RAMAN SPECTROSCOPIC STUDIES

OF ACYL CYSTEINE PROTEASES

Thesis submitted to the Department of Biochemistry
in partial fulfillment of the requirements for the
degree of Doctor of Philosophy

University of Ottawa
Ottawa, Ontario, Canada
April, 1996

ISBN 0-612-19949-5

Canada

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

Abstract

For a series of acyl cysteine proteases structure-reactivity relationships for the deacylation step have been established using a combination of Raman and absorption spectroscopies and enzyme kinetics. The simple chromophoric ligand 5-methylthienylacryloyl- (5MTA-) and three novel peptide based substrates labelled with the 5MTA- moiety, were used to create acyl enzyme adducts with papain, cathepsin B, cathepsin L, and protein engineered mutants of cathepsins B and L. The chromophoric specific substrates, 2-Ethoxycarbamido-3-(5-methylthienyl)acryloyl ethyl ester (NHCOOEt5MTAEt), 2-(N-acetyl-L-alanine)amino-3(5-methylthienyl)acryloyl ethyl ester (Ala5MTAEt), and 2-(N-acetyl-L-phenylalanine)amino-3-(5-methylthienyl) acryloyl ethyl ester (Phe5MTAEt), were designed to utilize hydrogen bonding and hydrophobic interactions which are known to promote catalysis in papain. For cathepsins B and L removing one of the hydrogen bonding groups making up the oxyanion hole reduces the deacylation rate 3-25 fold with the four substrates. The deacylation rate constants for the acyl cysteine protease series span a 214 fold range, from 0.07 to 15 x 10^{-3} sec^{-1}.

Using $^{13}C{=}O$ substitution it is possible to detect the acyl C=O frequency, ν_{C-O}, for each acyl cysteine protease in the Raman difference spectrum. A correlation between the ν_{C-O} frequency and the deacylation rate constant was established, where ν_{C-O}

increases with increasing reactivity. This trend is the opposite to that seen with acyl serine proteases. The opposite trend for acyl cysteine proteases is ascribed to the strong electron polarizing forces in the active site, due principally to an α-helix dipole, which change the hybridization about the carbonyl carbon atom.

A correlation was also established between the absorption maximum, λ_{max}, and the deacylation rate constant. As the deacylation rate increases, 214 fold across the series, λ_{max} red shifts from 367 to 384 nm. It is proposed that differential interactions between the α-helix dipole in the active site of the proteins and the π-electrons in the bound chromophoric substrate are responsible for these changes in λ_{max}, with increasing red shifts being caused by more favourable α-helix dipole-substrate interactions in the excited electronic state. It is also proposed that similar interactions occur between the α-helix dipole and the transition state of the acyl enzyme on the reaction pathway, giving rise to the observed differences in deacylation rates. These results demonstrate the importance of α-helix dipoles in cysteine protease active sites in modulating enzymatic activity, and provide the first experimental evidence for the role of α-helix dipoles in catalysis.

Great truths do not take hold of the hearts of the masses. And now, as all of the world is in error, how shall I, though I know the true path, how shall I guide? If I know that I cannot succeed and yet try to force success, this would be another source of error. Better then to desist and strive no more. But if I do not strive, who will?

Chuang Tzu

circa. fourth century B.C

Acknowledgements

I sincerely appreciate the numerous contributions that my supervisor, Dr. Paul R. Carey, made towards this thesis, especially when the adverse conditions to which our group had been subjected to is taken into consideration. I also appreciate the help provided by Dr. Peter J. Tonge during the first three years of my studies. For the organic chemistry portion of my studies I acknowledge the significant contributions made by Dr. Prabatt Ayra, Dr. Marianne Pustai, and Dr. Ross Williams. For the molecular biology work I acknowledge the significant contributions made by Dr. John Mort, Patty Mason, Sheryl Nymark, and fellow student Jonathon Latkins. Many thanks also to Dr. John Randell for running the NMR experiments for me.

Table of Contents

Derivatives With Wild Type and Mutant Acyl Cysteine Proteases

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

bicine	N,N-bis[2-hydroxyethyl]glycine
CBZ-Phe-Arg-MCA	carbobenzoxy-L-phenyl-alanyl-(7-amino-4-methylcoumarinyl)-L-arginine
CCD	charge coupled device
DCC	dicyclohexylcarbodiimide
DMSO	dimethyl sulfoxide
DTT	Dithiothreitol
E-64	1-(L-trans-epoxysuccinyl-L-leucylamino)-4-guanidinobutane
EDTA	ethylenediaminetetraacetic acid
FTIR	Fourier transform infra red
IPDA	intensified photodiode array
HPLC	high performance liquid chromatography
KPi	potassium phosphate
LDA	lithium diisopropyl amide
Mes	2-(N-Morpholino)ethanesulfonic acid
NMR	nuclear magnetic resonance
PCR	polymerase chain reaction
PDS	2,2'-dipyridyl disulfide
SCF-MO	self-consistent field-molecular orbital
RR	resonance Raman
5MTA	5-methylthienylacryloyl
Phe5MTAEt	2-(N-acetyl-L-phenylalanine)amino-3-(5-methylthienyl) acryloyl ethyl ester

Ala5MTAEt	2-(N-acetyl-L-alanine)amino-3-(5-methylthienyl) acryloyl ethyl ester
NHCOOEt5MTAEt	2-Ethoxycarbamido-3-(5-methylthienyl)acryloyl ethyl ester
5MTASC$_2$H$_5$	5-methylthienylacryloyl ethyl thiolester
5MTAOC$_2$H$_5$	5-methylthienyl acryloyl ethyl oxyester

CHAPTER 1

GENERAL INTRODUCTION

1.1 Introduction

Understanding how enzymes achieve such prodigious rate enhancements for chemical reactions, with great fidelity and specificity, is a primary goal of enzymologists. The ultimate goal of correlating the structure of the protein with its specific functions can only be achieved by using a variety of biophysical techniques. Vibrational spectroscopy, and in particular Raman spectroscopy, has proven to be a powerful tool for characterizing a variety of protein-ligand interactions. Resonance Raman, and more recently Raman difference, spectroscopies have provided quantitative insights into the functioning of a variety of serine protease-substrate complexes (Carey & Tonge, 1990). This thesis represents a logical extension to this work by applying a similar strategy to characterize the cysteine protease enzymes that reside in the papain super family.

1.2 The Raman Effect

An outline of the pertinent facts of Raman spectroscopy in relation to the study of biomolecules will be presented. A more detailed analysis of the origin of the Raman effect can be found in specialized texts about Raman spectroscopy (Herzberg, 1945, Long, 1977).

The vibrational frequencies of a molecule can be determined

by two different spectroscopic techniques, infrared and Raman spectroscopies. Both of these techniques depend upon the interaction of electromagnetic radiation with matter, but the resulting vibrational spectrum originates from different physical interactions for the two processes. In infrared spectroscopy a photon is absorbed causing the excitation of a vibration. In order for this absorption process to occur there must be a change in the permanent dipole moment of the molecule during the vibration. Raman spectroscopy, on the other hand, gives rise to vibrational spectra that result from the inelastic scattering of photons. In order for inelastic photon scattering to occur there must be a change in the dipole, induced by the electric field of the light wave, which in turn is a function of the molecular polarizability of a molecule, during a vibration.

Since the underlying physical origins of the two types of vibrational spectroscopies are different it follows that the selection rules governing the "allowedness" of a given vibrational transition are different. For example, the rule of mutual exclusion states that for molecules with a center of symmetry, vibrational transitions that are allowed in the infrared are forbidden in the Raman effect and vice versa. Thus, the complete vibrational analysis of any given molecule requires both IR and Raman spectroscopies.

As a consequence of these different selection rules the study of biomolecules via the Raman effect has a distinct advantage over using IR spectroscopy. This follows from the fact

that water, the biological solvent, has only a weak Raman spectrum in the key range of 0 - 2000 cm^{-1}. However, water absorbs IR radiation effectively in this region and this "intense" IR spectrum often obscures that of the biological solute. Thus, for studies in aqueous solutions Raman spectroscopy is usually the vibrational technique of choice.

1.3 Classical Theory of the Raman Effect

When a beam of monochromatic light is incident upon a sample a small portion of the radiation is scattered. Most of this radiation is scattered at the same frequency as the incident radiation, and is termed Rayleigh scattering. However a small portion of this radiation (typically 0.1 - 1 % of the total) will be scattered at different frequencies by the molecules making up the sample. This phenomenon is termed Raman scattering and is named after C.V. Raman who first observed this effect experimentally in 1928 (Raman & Krishnan, 1928).

When the electric field of a light wave encounters a molecule, it will tend to displace the negatively charged electrons from their average positions around the positively charged nucleus. In order for Raman scattering to occur the displacements must result in an induced dipole moment π in the molecule which is proportional to the electric field strength E. The following equation may therefore be written.

1.1 $$\pi = \alpha E$$

Where the proportionality factor α is called the electric polarizability of the molecule. A derivation from classical physics shows that the induced dipole moment, π, varies with three time dependent component frequencies ν', $\nu'-\nu_{vib}$, and $\nu'+\nu_{vib}$, where ν' is the frequency of the radiation that interacts with the molecule and ν_{vib} is a vibrational frequency of the molecule. The ν' component gives rise to Rayleigh scattering and comprises the majority (> 99%) of the scattered radiation. The $\nu'-\nu_{vib}$ component represents a frequency that is the difference between the incident radiation frequency and the molecular vibrational frequency, and is called Stokes Raman scattering. The $\nu'+\nu_{vib}$ component represents a frequency that is the sum of the incident radiation frequency and the molecular vibrational frequency, and is termed anti-Stokes Raman scattering.

Stokes and anti-Stokes scattering thus consist of scattered radiation that either loses or gains energy as a result of being inelastically scattered by a molecule. It follows from the conservation of energy that the energy gained or lost by this radiation reflect changes in the energy of the molecule that it interacts with. Hence, measurement of the energy (i.e. frequency) of the inelastically scattered radiation provides direct information on the changes in energy states of the molecule. These changes can be related to electronic, vibrational, or

rotational states of the molecule, although much of Raman spectroscopy is primarily concerned with changes in vibrational levels which occur in the range of 10 - 4000 cm^{-1}. Since the vibrational properties of a molecule are sensitive to its structure and its surrounding environment, Raman spectroscopy is a powerful tool for monitoring molecular chemistry.

1.4 The Resonance Raman Effect

Another advantage of Raman spectroscopy over IR spectroscopy concerns the resonance Raman (RR) effect, for which there is no parallel in IR spectroscopy. As described previously the Raman effect is very weak, however if the wavelength of the radiation used to elicit the Raman spectrum is coincident with an electronic absorption band present in the molecules making up the sample, then the Raman scattering process may be greatly amplified by a factor of 10^3 to 10^5. This amplification is the result of a large increase in the polarizability of the molecule associated with the vibration. This effect greatly facilitates the study of biomolecules that contain a chromophoric group. The main advantage is that the vibrational spectrum of the bound chromophore may be selectively obtained while the corresponding solvent and macromolecule vibrational bands are lost in the background. Another related advantage is that high quality spectra of the bound chromophore may be obtained at low

concentrations of sample (10^{-4} - 10^{-5} M). Furthermore, although the intensities of Raman bands are a function of both the ground and excited state, the peak positions are solely a function of the ground state (see section 1.5). Hence peak positions can be used to monitor ground state molecular chemistry, a useful property for analyzing enzyme catalyzed reactions which usually proceed exclusively through the ground state.

A disadvantage of RR spectroscopy is that the light used to elicit the RR spectrum may be energetic enough to cause undesirable photochemical reactions that either alter the structure or destroy the chromophore being studied. This can greatly complicate the interpretation of the resulting spectrum. Another disadvantage is that since the exciting wavelength is coincident with an electronic absorption band, it is possible for fluorescence to occur. If the emitted fluorescent photons have similar energies to the scattered Raman photons, then the Raman spectrum can be completely lost in the fluorescent emission spectrum, which may have 10^3 - 10^5 greater intensity.

1.5

Quantum Mechanical Picture of Raman Scattering

The classical theory discussed in section 1.3 provides a useful conceptual model for Raman scattering. However energy interactions with molecules are quantized, and hence quantum mechanics must be applied to achieve a rigorous description of

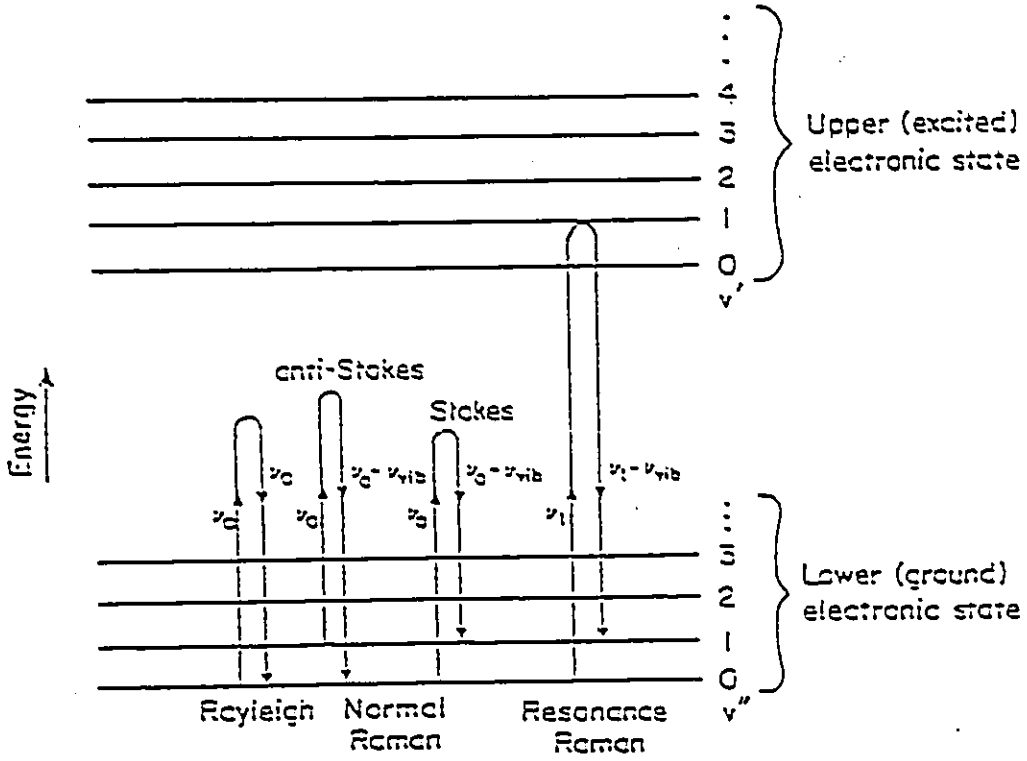
the Raman process. The quantum mechanical approach assumes that electromagnetic radiation consists of discrete packets of energy called photons, and that molecular energy levels between different vibrational states are quantized. Using this approach Raman spectroscopy can be viewed as a two photon process, the possible consequences of which are illustrated in Figure 1.1. The first step of the light scattering process is the combination of a photon and a molecule to raise the energy of the molecule for an extremely short life time. This is represented by the upward arrows in Figure 1.1. The second step is the release of the photon, depicted by the downward arrows of Figure 1.1. The length of the upward and downward arrows are proportional to the frequencies of the incoming light and subsequent scattered light, respectively.

If the photon begins at and returns to the lowest ground state vibrational level (v=0), then no energy is exchanged between the molecule and photon, resulting in Rayleigh scattering. Stokes Raman scattering occurs when the incoming light loses energy to the molecule, resulting in a molecule at the v=1 state. The energy lost by the photon, $\nu_0-\nu_{vib}$, is equal to the energy gained by the molecule as represented by the spacing between the v=0 and v=1 ground states. Conversely, anti-Stokes Raman scattering occurs when incoming light gains energy, $\nu_0+\nu_{vib}$, equal to the energy lost by the molecule in the v=1 to v=0 ground state transition. The results of the quantum mechanical treatment given here are consistent with the results from the classical treatment

Figure 1.1

Some of the Possible Consequences of

Photon-Molecule Interactions



given in Section 1.3.

A similar description may be used for the resonance Raman process. The main difference is that the incoming photon frequency in this case is coincident with an electronic absorption band of the molecule, resulting in absorption of the photon into the excited electronic state and giving rise to the RR effect as described in section 1.4.

The quantum mechanical model in Figure 1.1 illustrates a key generalization that the position of both Raman and resonance Raman peaks is a property solely of the ground electronic states. This follows from the fact that all of the molecular vibrational transitions occur in the ground state.

1.6

Assignments of Various Group Frequencies

When relatively simple molecules are being studied, complete mathematical procedures exist for calculating vibrational frequencies. However as the molecule becomes larger these methods become difficult to implement. Complete calculations, however, are not always necessary in order to identify some bonds or functional groups if they can be treated as independent oscillators. If the frequency of an oscillating group is sufficiently different from those of the rest of the molecule its vibrational coupling with other molecular frequencies may be minimal and that frequency can then be regarded as being

characteristic of the particular oscillating group. This permits the assignment of various vibrational bands to particular functional groups. For example, the C=O stretching mode has a relatively high vibrational frequency compared to the single bonds present in the rest of the molecule and hence its vibrational band can usually be discerned in its characteristic spectral region (1900 - 1550 cm^{-1}).

If a functional group can be treated as an independent oscillator then its frequency can be related to the strength (i.e the force constant) of the localized bond within the functional group. The utility of this is that the exact position of the frequency is a function of conformational and environmental factors which influence the strength of the functional group's bond. For example, the C=O bond is sensitive to mesomeric, inductive, and hydrogen bonding effects which are manifest in the position and intensity of the C=O stretching frequency band. Hence, vibrational spectroscopy can be used as a tool to monitor conformational and environmental changes via variations in group frequencies.

The group frequency approach is an approximation and caution must be exercised in determining to what extent the vibration is independent of the rest of the molecule. If a functional group frequency approaches the frequency of one or more groups in the molecule, then the frequencies may vibrationally couple and the resultant frequencies will be a function of multiple bonds and atoms. When this occurs it is not possible to view frequency

values as representing a localized bonding system and hence interpretation of the vibrational band position is more difficult.

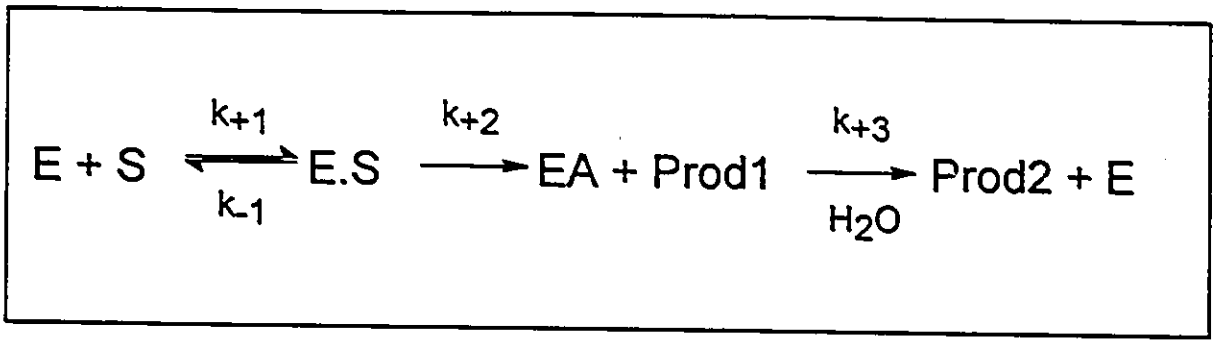
The use of isotopic substitution is a valuable aid for the assignment of vibrational bands. If it is assumed, as is usually the case, that force constants are unaltered by substitution, then shifts in frequency can be attributed to mass effects. Qualitatively this shift will aid in the assignment of frequencies simply by observing which bands are altered, and to what extent, as a result of the substitution.

1.7 Serine and Cysteine Protease Enzyme Systems

Serine and cysteine proteases are two classes of enzymes which break down proteins by hydrolyzing their peptide linkages. Both classes have been extensively studied over the past decades and a wealth of information has been obtained concerning structure and mechanism. Both categories are involved in key physiological processes, and they have been implicated in several disease states. Therefore research is still expanding so that a deeper insight into their mechanism of action can lead to mechanism-based drug design.

Serine and cysteine proteases share a similar mechanism of action in that hydrolysis of either an amide or ester substrate is achieved by the three step process shown in Scheme 1.1.

Scheme 1.1



The first step is the association of the free enzyme and substrate (E+S) to form a Michaelis complex E.S with a dissociation constant $K_s(=k_{-1}/k_{+1})$. The carbonyl carbon of the scissile ester or amide bond then undergoes nucleophilic attack from the enzyme's active site nucleophile resulting in the expulsion of the substrate leaving group (Prod1), and the formation of an acyl-enzyme complex, (EA). This is followed by the deacylation reaction which occurs via hydrolysis, resulting in the acid product of the substrate, (Prod2) and regeneration of the free enzyme, (E).

Although the overall three dimensional structures of serine and cysteine proteases are different, they share a common structural motif for catalysis. Serine proteases possess a catalytic triad composed of serine, histidine, and aspartic acid residues that has been termed the charge relay system. An analogous triad is present in cysteine proteases consisting of

cysteine, histidine and asparagine. The role of histidine in both systems is to increase the nucleophilicity of the acylating side chain (the hydroxyl group of serine for serine proteases, and the sulfhydryl group of cysteine in cysteine proteases), accomplished by proton transfer from the active site side chain to the imidazole. The imidazole also functions as a general base during deacylation by assisting in the removal of a proton from the nucleophilic water molecule.

The catalytic triad of serine proteases was originally thought to function as a charge relay system in which negative charge is transferred from the aspartate through the imidazole to the serine oxygen (Blow *et al.* 1969). However more recent experiments have demonstrated that it is the histidine which is protonated and not the aspartate (Bachovchin *et al.* 1981). It is now hypothesized that the aspartate functions by stabilizing the developing charges that occur in both the transition state and in the tetrahedral intermediate. The importance of the aspartate residue has been verified by site directed mutagenesis, which showed that replacement of the aspartic acid results in a 95% drop in activity (Kaiser, 1985).

Cysteine proteases have a similar catalytic mechanism although the nucleophile is a sulfhydryl side chain instead of the hydroxyl side chain of serine. A current hypothesis regarding cysteine protease mechanism is that the sulfhydryl side chain of the active site cysteine exists as an ion pair with the active site imidazole from histidine of the form S^----^+HIm, thus

accounting for the reactivity of cysteine proteases at acidic pH (Brocklehurst *et al.* 1987). The active site asparagine is believed to be hydrogen bonded to the imidazole, and its probable function is to orient the imidazole side chain (Baker & Drenth, 1987).

1.8 The Oxyanion Hole

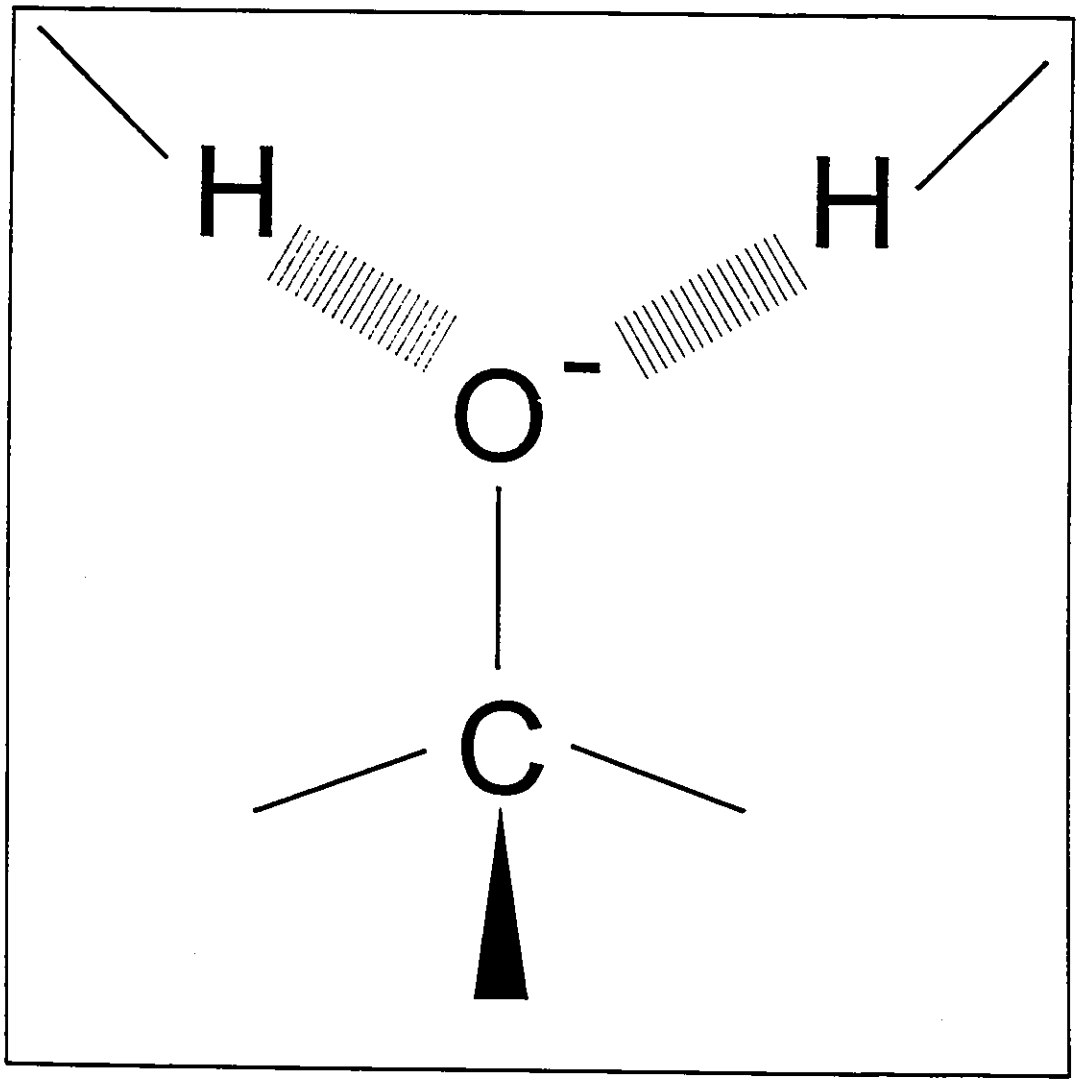
Another key mechanistic feature shared by serine and cysteine proteases is that they both possess an oxyanion hole. The oxyanion hole is a pair of hydrogen bonds strategically placed in the active site that can help stabilize the negative charge that builds up on the substrate carbonyl oxygen in the transition state. This is illustrated schematically in Figure 1.2.

According to transition state theory any chemical factors which help lower the energy of the transition state will lead to an increase in the rate of the reaction. A good test to evaluate the role of the oxyanion hole in serine and cysteine proteases therefore is to remove the hydrogen bonds, when feasible, through site directed mutagenesis and experimentally observe what effects this has on rate.

Site directed mutagenesis experiments were first performed on Asn 155, whose side chain contributes a hydrogen bond in the oxyanion hole, in the serine protease subtilisin, a bacterial

Figure 1.2

Schematic Representation of the Oxyanion Hole



protease isolated from *Bacillus subtilis*. It was found that the Asn side chain contributes between 2.2-4.8 kcal/mol towards transition state stabilization (Wells *et al.* 1986, Bryan *et al.* 1986, Rao *et al.* 1987, Carter & Wells, 1990).

An analogous pair of hydrogen bonds exists in the structures of cysteine proteases, and their role in catalysis has been determined by Ménard *et al.* (1991) via site directed mutagenesis with the cysteine protease papain. Examination of the X-ray structures of papain-inhibitor complexes indicate that the Gln 19 side chain contributes a hydrogen to the oxyanion hole (Drenth *et al.* 1976), and its removal by making the Gln19Ala substitution resulted in a $\Delta\Delta G^{\ddagger}$ value (i.e a reduction of the energy of stabilization of the transition state) of 2.4 kcal/mol, while the Asn155Gly substitution in subtilisin leads to a $\Delta\Delta G^{\ddagger}$ value of 3.2 kcal/mol. This study concluded that the extent of transition state stabilization via the oxyanion hole for papain is smaller than the analogous stabilization in subtilisin, although it is still a significant contribution.

It has been determined from the crystal structure of the cysteine protease cathepsin B that the Gln23 side chain is positioned in the oxyanion hole the same way as papain's Gln19 side chain is (Musil *et al.* 1991). Kinetic investigations of the oxyanion hole mutant Q23A by Fox *et al.* (1996) demonstrated that the oxyanion hole contributes 1.8 kcal/mol to transition state stabilization, similar to that exhibited by the same mutation in papain.

1.9

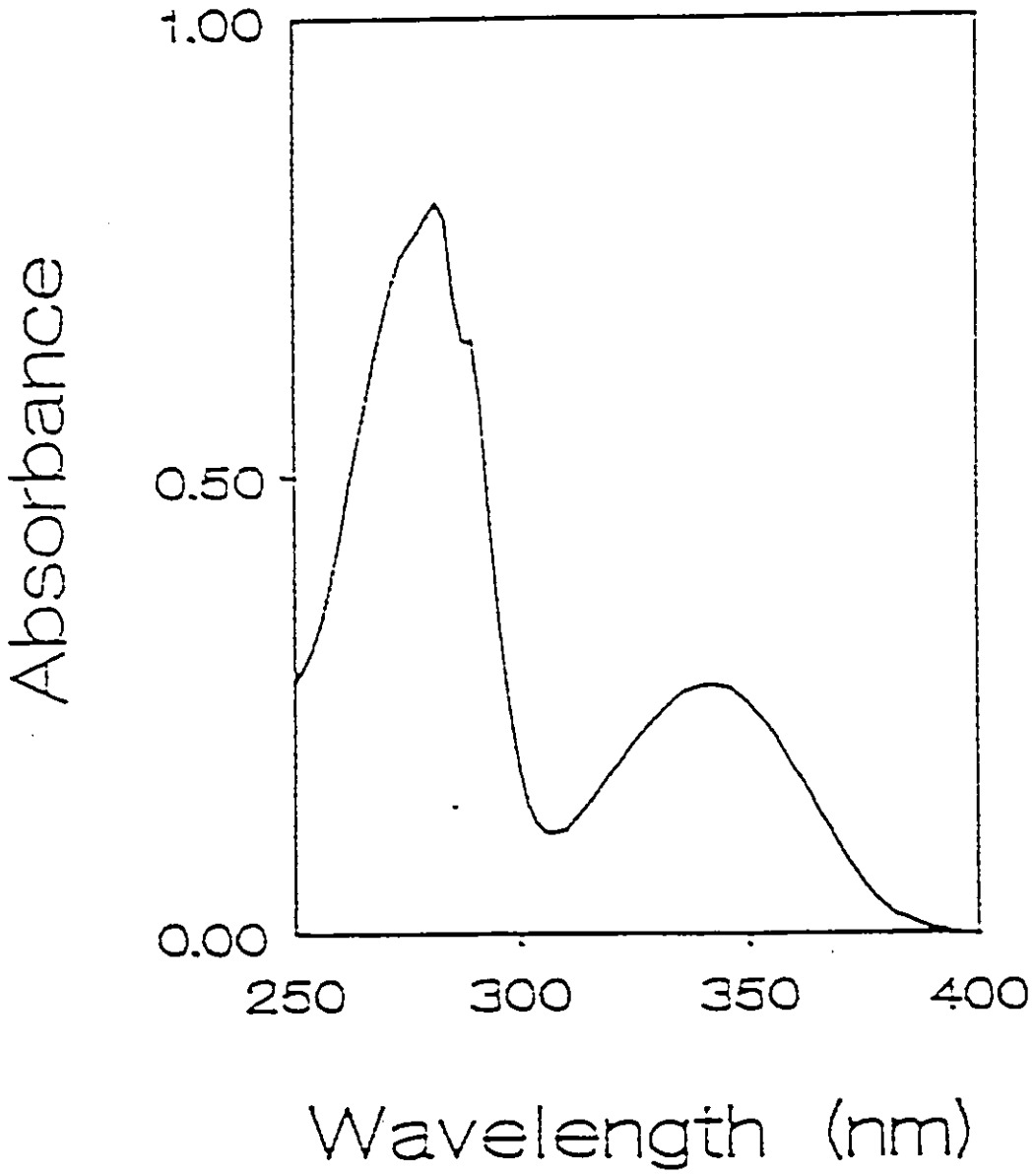
Application of the Resonance Raman Technique to the Study of Acyl-Serine Proteases.

Serine proteases catalyze the hydrolysis of amide bonds via a two step process of acylation and deacylation with the generation of a transient acyl-enzyme complex as discussed in section 1.7. With amide substrates acylation of the enzyme is rate limiting, thus making the study of the transient acyl-enzyme complex difficult. However, proteases can also hydrolyze ester substrates where deacylation is usually rate limiting by virtue of the fact that most esters have much better leaving groups than amides. Hence, it is possible to accumulate the acyl enzyme complex with ester substrates and study them by a variety of biophysical techniques. Serine and cysteine proteases have the added advantage that the deacylation reaction is greatly slowed down at low pH, (cysteine proteases typically have deacylation pKa's ≈ 4, serine proteases ≈ 7), thus permitting the purification of the acyl-enzyme which may then be subjected to direct spectroscopic and kinetic characterization.

In order to apply the resonance Raman technique to acyl-serine protease complexes it is necessary to have a chromophoric substrate. Arylacryloyl α,β-unsaturated compounds are non-specific substrates of proteases that may be used to generate chromophoric acyl enzymes. These compounds possess a conjugated π electron system which absorb light in the 300 - 400 nm region,

Figure 1.3

Absorption Spectrum of SMTACT



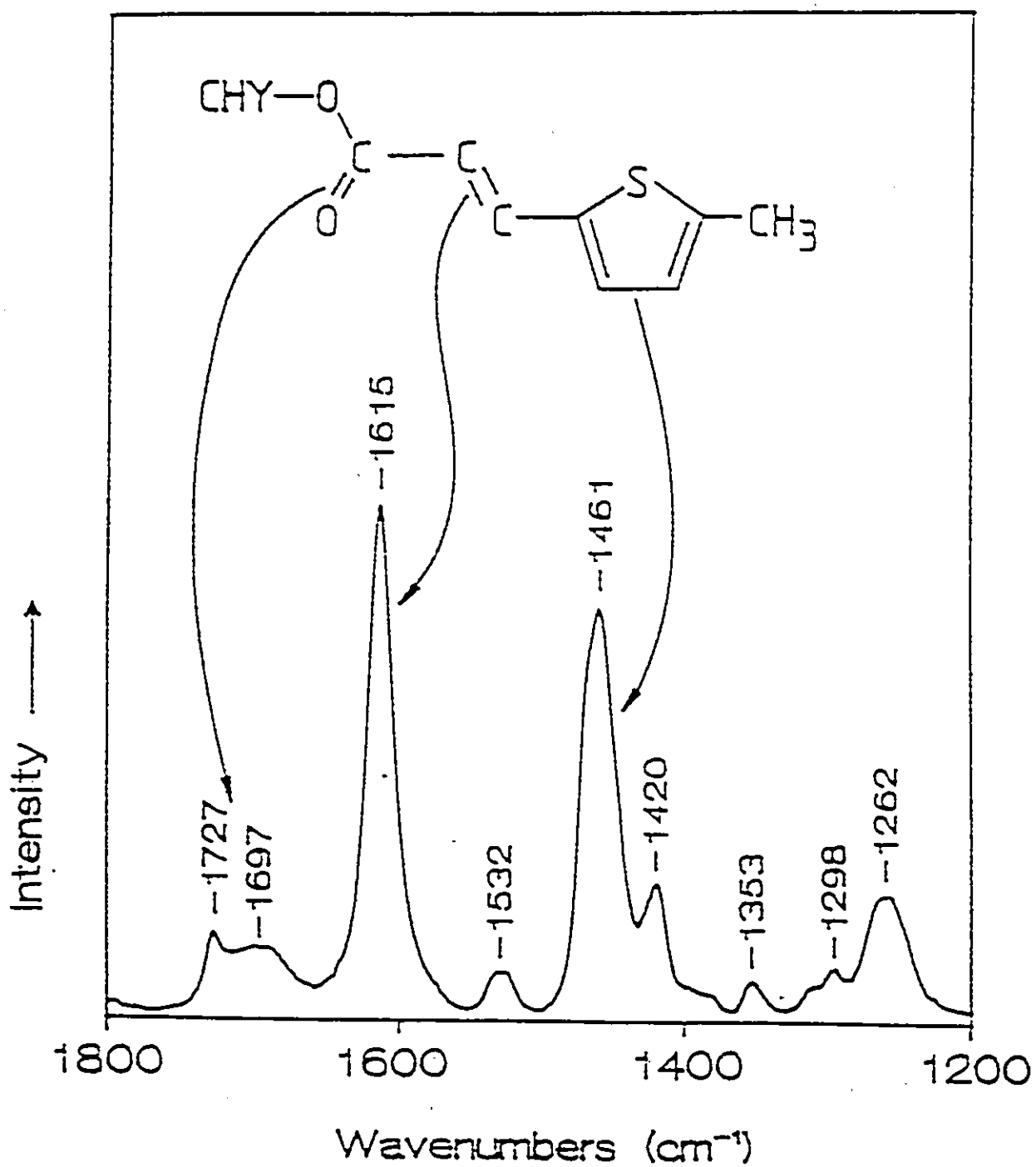
and they have been used extensively as active site directed spectroscopic probes for serine and thiol proteases (Bender *et al.* 1962, Bernhard *et al.* 1965, Hinkle & Kirsch, 1970, Carey & Tonge, 1990). Figure 1.3 shows the absorption spectrum of 5-methylthienylacryloyl-chymotrypsin (5MTACT),a typical acyl-enzyme. The absorption band with λ_{max} 340 nm is due to the substrate's acyl group bound to the active site serine-195 of chymotrypsin. By using an excitation wavelength that lies within this electronic absorption band, such as the 337.5 nm Kr^+ laser line, selective enhancement of the bound substrate's vibrational modes may be achieved free from interference from solvent or protein bands.

Figure 1.4 depicts the RR spectrum of 5MTACT obtained at pH 3.0. The major peaks are due to vibrational modes that participate in the conjugated π bond system. Some of the major peaks are assigned in Figure 1.4. Of particular interest are the two peaks, 1727 and 1696 cm^{-1}, assigned to two C=O stretching modes. The vibrational signatures of this catalytically crucial carbonyl group can be used to relate subtle changes in its chemistry produced by its interactions with the enzyme with the deacylation rate constant, thus permitting the determination of structure-function correlations.

A key advantage of applying RR, and vibrational spectroscopy in general, to the study of enzyme-substrate complexes is that experimentally determined frequencies can be correlated with small changes in bond length. This is important since extremely

Figure 1.4

Resonance Raman Spectrum of 5MTACT



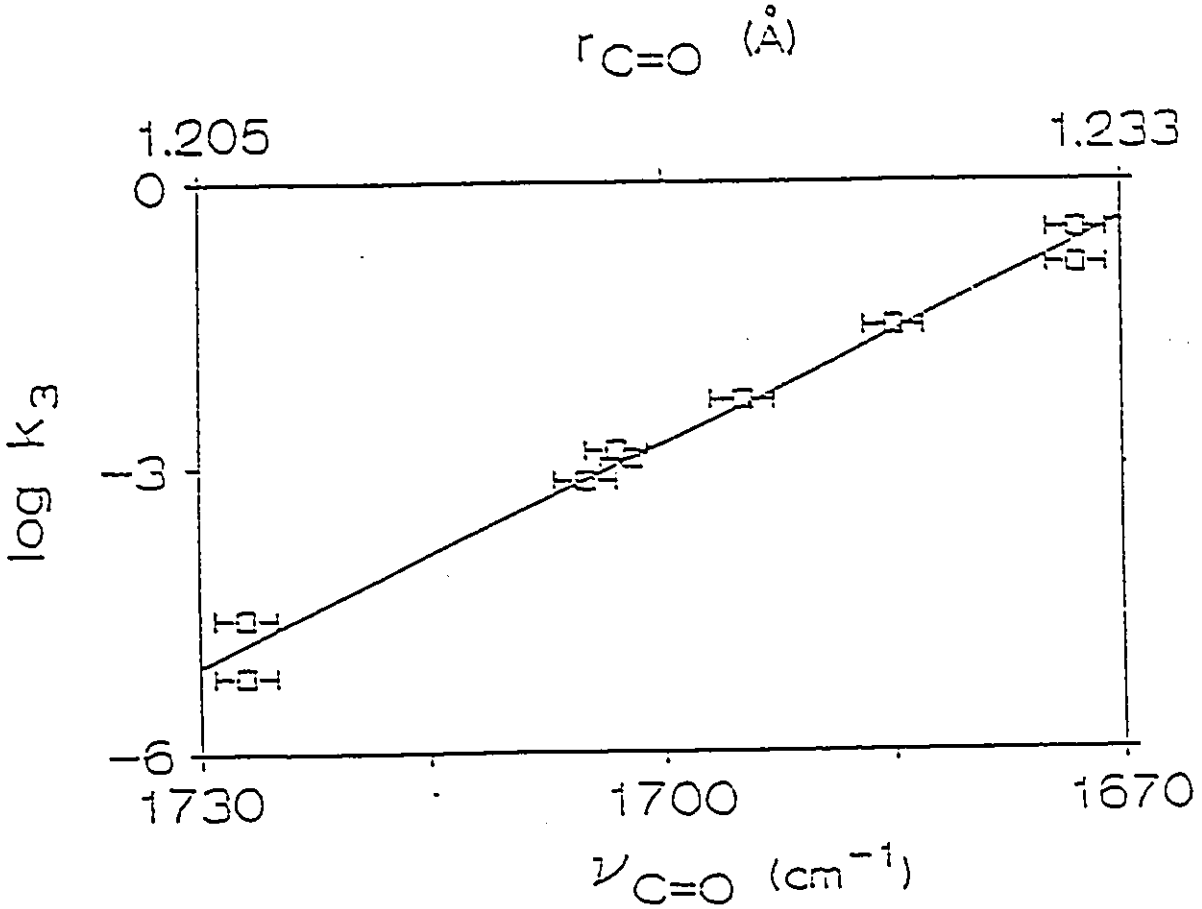
small changes in bond length can have major effects on chemical reactivity. For example, Jones & Kirby (1979, 1984) have shown for a series of alkyl aryl acetals that changes in the C-O bond lengths of the order of 0.03 Å can have a 10^6 fold or more effect on the rate constants for spontaneous hydrolysis.

Using RR spectroscopy to probe the acyl enzymes, Tonge and Carey (1990, 1992) were able to correlate the length of the acyl carbonyl bonds with the deacylation rates of a number of acyl-serine proteases generated from arylacryloyl substrates. Figure 1.5 is a reproduction of their plot correlating carbonyl stretching frequency with log k_3, where k_3 is the maximal deacylation rate at pH 10.0. A linear relationship is obtained between the two parameters over a 16,300 fold range of reactivity. It is possible to correlate the frequency of the carbonyl with its bond length through the use of an empirical relationship developed by Horvath *et al.* (1987) from infrared spectroscopic and crystallographic observations on cyclic and heterocyclic organic compounds. The C=O length is represented at the top portion of the graph. Using this empirical correlation it was determined that the length of the C=O bond increases by 0.025 Å as the deacylation rate increases 16,300 fold through the series of acyl enzymes.

Increasing the strength of hydrogen bonds to a carbonyl oxygen atom decreases its stretching frequency (Pimental & McClellan, 1960). Consistent with this, the data from Figure 1.5 show that the oxyanion hole mutants of subtilisin in which the H

Figure 1.5

Plot of log k_3 Against $\nu_{C=O}$ for Acyl Serine Proteases



bond donor Asn is substituted with Leu, Gln, or Arg result in higher carbonyl frequencies (> 1700 cm^{-1}) from the bound substrate; while the wild type subtilisin with an intact oxyanion hole gives lower carbonyl frequencies (1675 - 1680 cm^{-1}). The small variations observed in carbonyl bond length are thus ascribed to differential degrees of hydrogen bonding by the enzyme oxyanion hole to the carbonyl oxygen of the bound substrate. This leads to a significant generalization; stronger hydrogen bonding from the oxyanion hole gives rise to more rapid acyl enzyme deacylation. A current hypothesis which rationalizes this result is that the hydrogen bonding in the ground state distorts the structure of the carbonyl group towards the transition state structure, thus causing rate acceleration. Additionally, the hydrogen bonding can increase the polarization of the carbonyl group, leading to a reduction in electron density of the carbonyl carbon and thus increasing its susceptibility to nucleophilic attack.

The RR spectra of most α,β-unsaturated acyl-chymotrypsins are characterized by two carbonyl features; one at 1725 cm^{-1} which is characteristic of a non-hydrogen bonded carbonyl population and another in the 1700 - 1680 cm^{-1} range characteristic of a hydrogen bonded carbonyl environment (For example, see Figure 1.4, carbonyl feature of 5MTACT). Extensive studies by Tonge *et al.* (1993) using a combination of Raman difference, FTIR, and NMR spectroscopies have established that the 1725 cm^{-1} band is due to the formation of a *cis* isomer about

the ethylenic linkage of the bound α,β-unsaturated substrate. The *cis* isomer is caused by a photoisomerization reaction due to the laser beam used to collect the RR data. The creation of a mixture of *cis* and *trans* isomers complicates the interpretation of the RR vibrational spectra; however, in the case of the α,β-unsaturated acyl chymotrypsins it was possible to rigorously characterize the *cis* isomer. Purification of pure *cis* 5MTACT was possible due to a fortuitous large difference in deacylation rates between the two isomers. Characterization of the *cis* isomer for 5MTACT enabled its C=O band and log k_3 to be placed on the linear correlation shown in Figure 1.5, resulting in the extension of the reactivity range from 500 to 16300 fold.

These studies demonstrated unequivocally that laser lines in the 300 - 400 nm range are sufficiently energetic to create a photostationary mixture of *cis* and *trans* isomers with α,β unsaturated compounds. Therefore, caution must be employed when attempting to interpret α,β-unsaturated acyl-enzyme RR spectra when near UV laser lines are used as the excitation source.

1.10 Extension of the Carbonyl Bond Frequency - Reactivity Correlation to Cysteine Proteases.

In spite of the commonality of mechanistic features shared by cysteine and serine proteases, there are significant differences between the two classes of enzymes. As discussed

previously in section 1.8 the relative importance of the oxyanion hole appears to be different for the two classes. There are also intrinsic chemical differences between esters and thiolesters (discussed in detail in Chapter 3), and in the presence of strong electron polarizing forces in cysteine protease active sites which appear to be absent in serine proteases (discussed in detail in Chapter 4). It is these differences which provide the motivation to explore structure-reactivity relationships for cysteine protease acyl-enzymes in this thesis. A simple extension of the serine protease structure-reactivity correlation to cysteine proteases would provide powerful evidence that different classes of enzymes utilize the same structural motifs to achieve transition state stabilization.

In order to develop a carbonyl frequency - reactivity correlation for cysteine proteases, two problems had to circumvented. One problem is that the range of deacylation rates of acyl-cysteine proteases generated from simple α, β-arylacryloyl unsaturated compounds are limited. In order to extend the range of reactivity novel chromophoric substrates were synthesized based upon what is presently known about the specificity requirements of papain. These results are presented in Chapter 2.

The second problem is that thiol esters are intrinsically more photolabile than oxygen esters as determined by Tonge & Carey (1989(a)). The problem of photoisomerism as discussed in section 1.9 also occurs with both model compound and acyl-

cysteine protease data. Unfortunately it is not possible to isolate *cis* and *trans* isomers as was possible for model oxygen ester compounds and acyl serine proteases, since for thiol esters the chromophore is essentially destroyed during the photoisomerization process. This fact precludes any possibility of rigorously defining *cis* and *trans* isomer Raman bands as was possible for acyl-serine protease data, and hence interpretation of the RR band positions (including the crucial carbonyl) of acyl-cysteine proteases is rendered much more ambiguous. Recent technological advances in instrumentation permit this drawback to be overcome by obtaining high quality acyl-enzyme data under non-resonance conditions. The advances include the development of holographic notch filters and of extremely efficient photon detectors sensitive in the red region of the spectrum. An instrument incorporating these technologies is described by Kim *et al.* (1993)(see also section 3.2.4).

Holographic notch filters have an O.D. of greater than 6 at the wavelength of excitation and block undesirable Rayleigh photons from entering the spectrograph, while readily transmitting photons at other wavelengths. Figure 1.6 illustrates the performance of a notch filter tuned for 752.5 nm, in which the photons corresponding to the 200 cm^{-1} and beyond Raman shift are transmitted at greater than 80% efficiency. This represents a vast improvement over the use of a double monochromator to eliminate the Rayleigh line, in which the Raman photons are transmitted at less than 1% efficiency. The second technological

Figure 1.6

Transmission Efficiency of a Holographic Notch Filter

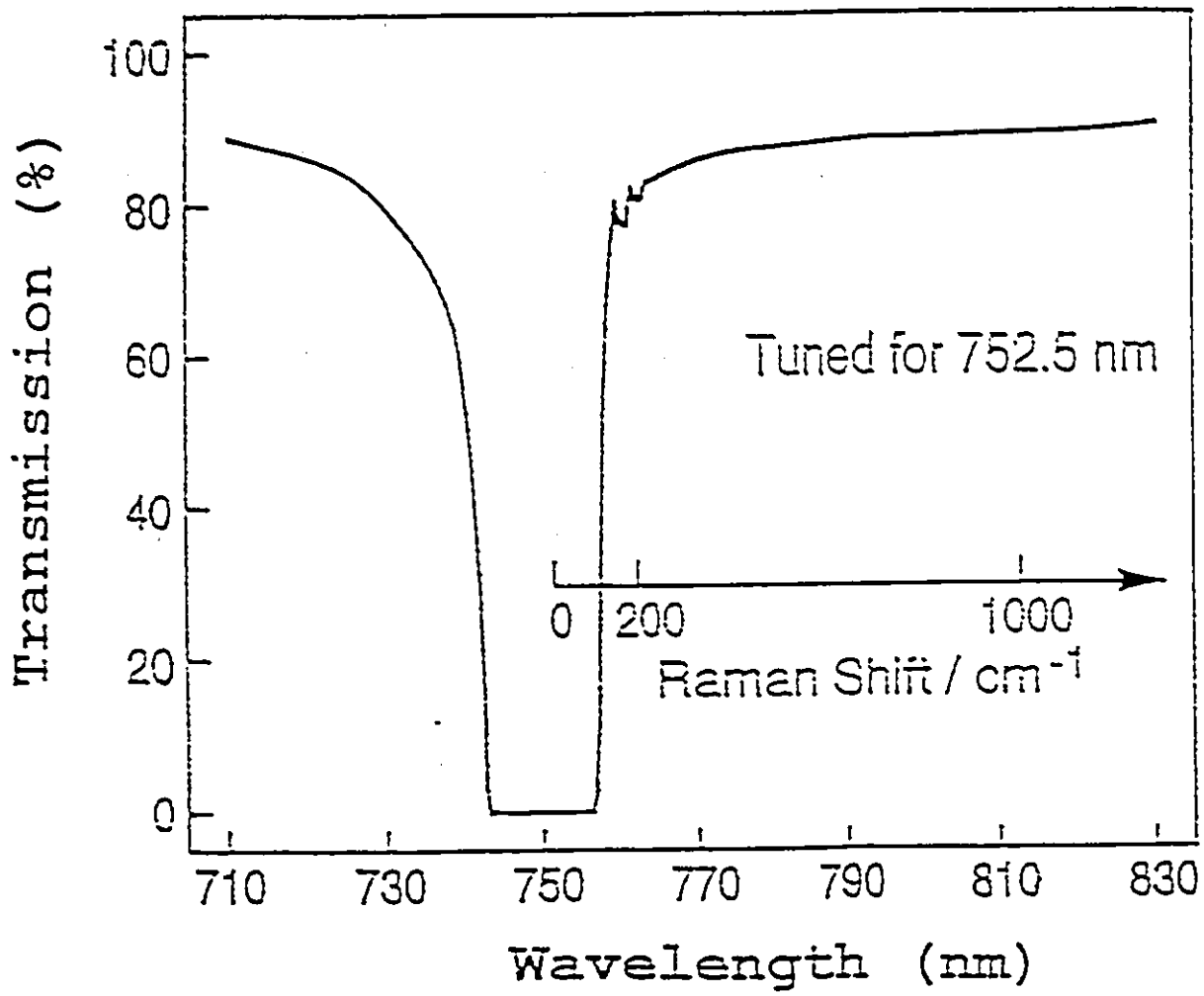
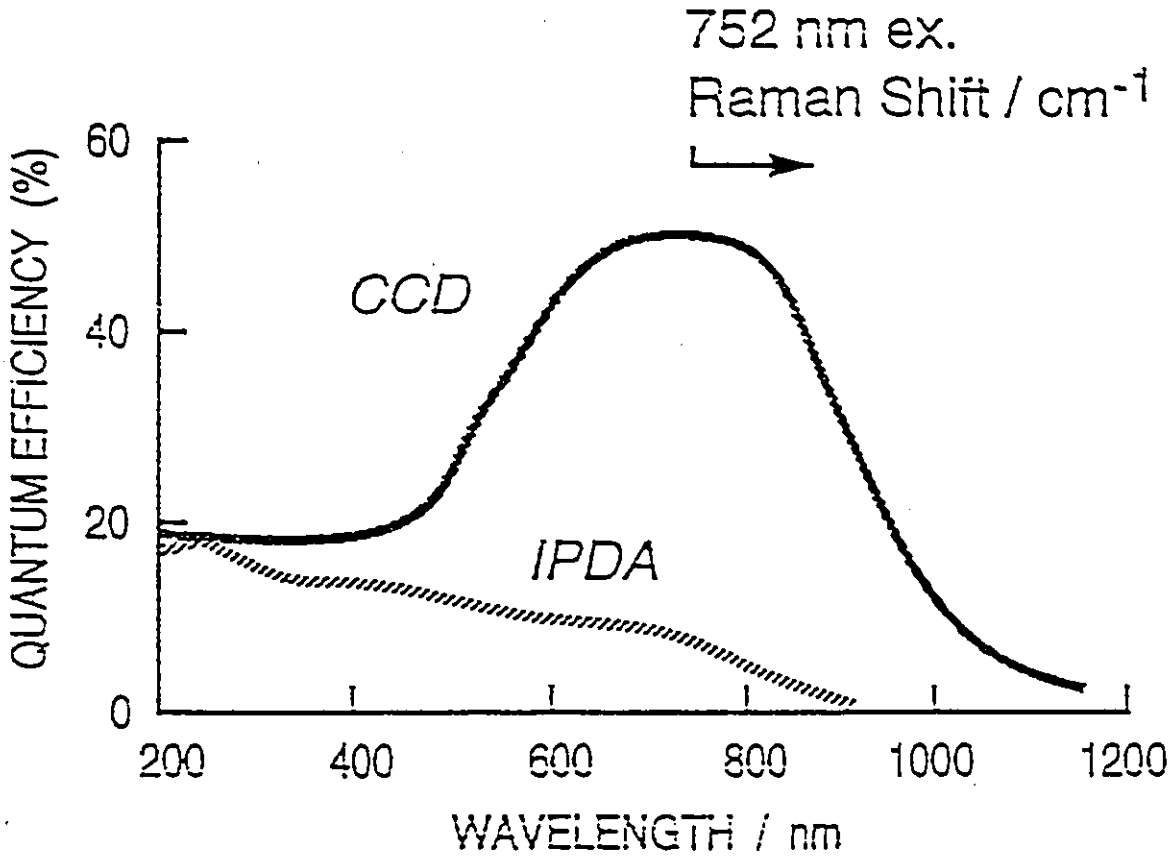


Figure 1.7

Efficiency of a CCD Detector Compared to an IPDA Detector



advance is the development of charge coupled device (CCD) detectors, which have enhanced sensitivity in the red region of the spectrum compared to older detectors. Figure 1.7 illustrates the quantum efficiency of a CCD detector versus an intensified photodiode array (IPDA) detector as a function of wavelength. It is apparent that the CCD is the more sensitive detector, and that this sensitivity increases as the wavelength increases up to a maximum around 700 nm. These two technological advances permit for the first time the acquisition of high quality acyl enzyme Raman spectra under non-resonance conditions.

CHAPTER 2

KINETIC CHARACTERIZATION OF 5-METHYLTHIENYLACRYLOYL SUBSTRATE DERIVATIVES WITH WILD TYPE AND MUTANT ACYL CYSTEINE PROTEASES

2.1 Introduction

Members of the papain super family of cysteine proteases are widely distributed among living organisms and participate in a number of physiological processes in plants, bacteria, and animals (Brocklehurst *et al.* 1987). There is a considerable amount of evidence implicating cysteine proteases in disease states. In particular, lysosomal cathepsins are believed to be involved in several pathological conditions such as muscular dystrophy (Katunuma, 1987), rheumatoid arthritis (Werb *et al.* 1989), and tumour invasiveness (Sloane, 1990). Thus, interest in the detailed mechanisms of cathepsins such as B and L is enhanced by their being potential targets for inhibitor-based drug design.

In the absorption and Raman difference spectroscopic studies of acyl cysteine proteases it was decided to focus on chromophoric acyl groups, in particular 5-methylthienyl acryloyl- (5MTA), because of the background in interpreting the vibrational spectra of this moiety bound to serine proteases (see Section 1.9 of General Introduction) and because in the initial studies the resonance Raman condition had to be used due to technical limitations. Arylacryloyl substrates like 5MTA are highly polarizable (see Section 3.4.3), thus they are strong Raman scatterers and yield high signal to noise spectra, even under non-resonance conditions. However, they are poor substrates for cysteine proteases, with deacylation rates for the corresponding acyl-enzymes that are 10 - 100,000 fold less than those obtained

with specific ester and amide substrates. For example, Zannis & Kirsch (1978) determined that deacylation rates for papain using a series of non-specific substituted cinnamoyl substrates ranged from 0.0005 to 0.05 sec^{-1}, while Brubacher & Zaher (1979) found deacylation rates ranging from 1.5 - 59 sec^{-1} for a series of specific peptide methyl ester substrates.

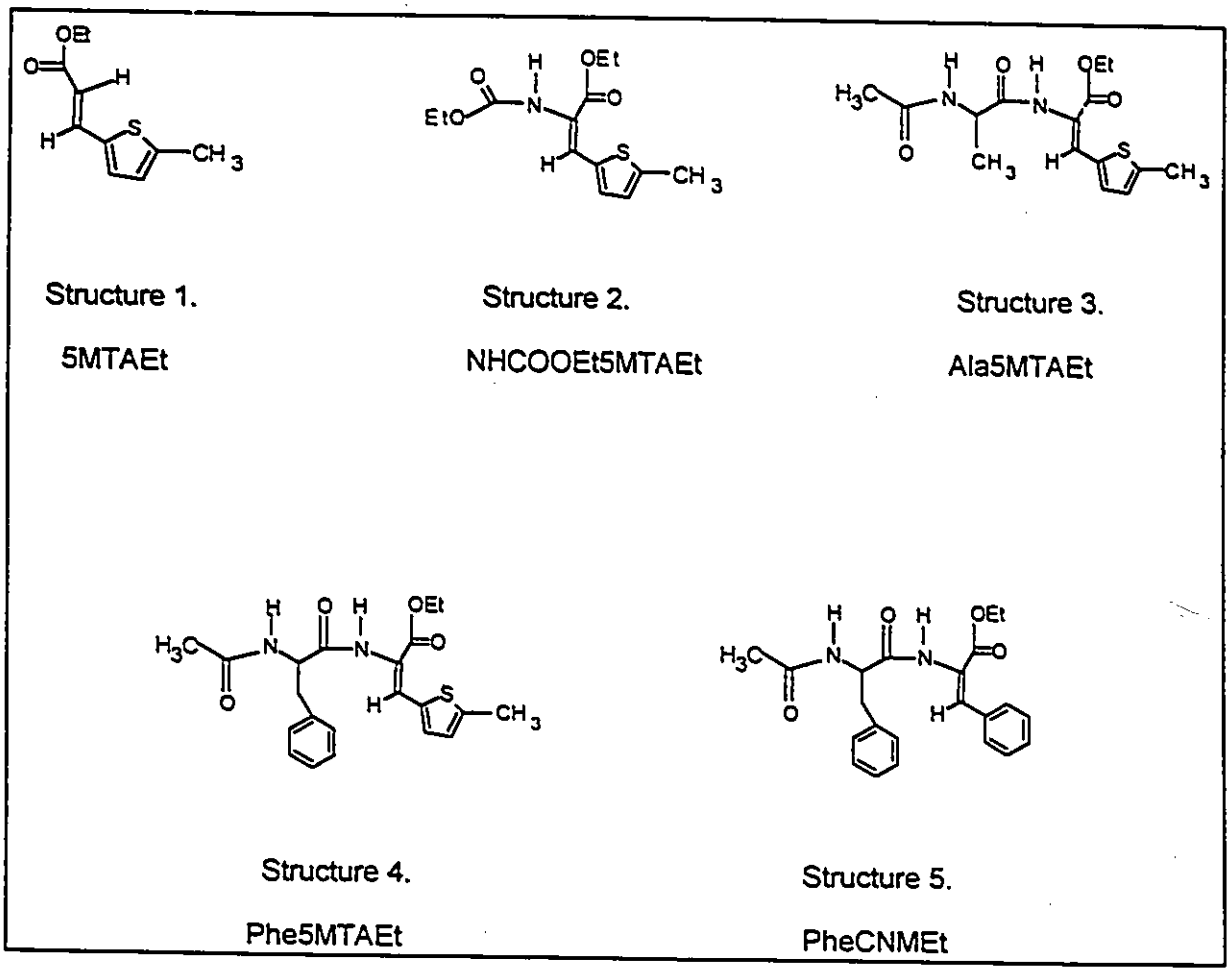
An attempt was therefore undertaken to synthesize more reactive chromophoric substrates containing the 5MTA moiety in order to extend the reactivity range. Early studies on the specificity requirements for papain by Schecter and Berger (1967,1968, see also Berger & Schecter, 1970) demonstrated that papain's active site is composed of 7 subsites which accommodate amino acid residues of the substrate. An important discovery from these studies is the strong preference for the phenylalanine residue at the S_2 position, a characteristic that is also shared by cathepsins B and L. Schecter and Berger (1967, 1968) originated the classification system using the P-S notation which is still currently in use. The basis of this system is that the amino acid residues of a polypeptide are designated P_1, P_2, etc. and P_1', P_2', etc. (counting from the bond cleaved towards the amino- and carboxyl termini, respectively), and the corresponding subsites in the enzyme are designated S_1, S_2, etc. and S_1', S_2', etc.

For papain-inhibitor complexes based on specific substrates, X-ray crystallographic experiments by Drenth and coworkers (1976) identified the importance of hydrogen bonds involving $P_1NH---O=C$

(Asp-158), P_2NH---O=C (Gly-66), and P_2C=O---HN (Gly-66). The energetic contributions of these interactions, as well as that for a Phe at the substrate's P_2 position, were estimated in the kinetic studies of Berti *et al.* (1991) to be in the range of 2.5 - 4.0 Kcal/mol per interaction. Given the importance of these contacts, an approach similar to Smolarsky (1978, 1979) was undertaken to generate specific chromophoric substrates involving the 5MTA group. By placing acetylPheNH- at the α carbon of cinnamoyl ethyl ester (Structure 5, Scheme 2.1), Smolarsky created a chromophoric compound capable of engaging in the specific active site interactions discussed above. In the present work, three substrates which incorporate 5MTA, shown in Scheme 2.1, were synthesized using a novel synthetic strategy.

Structure 1 is the non-specific 5MTA substrate which has been extensively characterized from previous work on serine proteases (Tonge & Carey, 1990,1992 & Tonge *et al.* 1991,1993). Structure 2 explores modulating reactivity which may occur by including a putative hydrogen bond (from NH) at the P_1 position. Structures 3 and 4 are capable of forming favourable hydrogen bonds at both P_1 and P_2 positions, but have different hydrophobic side chains at P_2. Using the above substrates to generate acyl-enzyme complexes with papain, cathepsins B and L, and oxyanion hole mutants of cathepsins B and L, it was possible to spread the rates of deacylation over a 200 fold range. This represents the range over which structure-reactivity correlations can be explored. This chapter presents some details of the kinetic

Scheme 2.1



properties of simple and specific (e.g. Structure 4) chromophoric substrates. While characterizing the kinetics of the Phe5MTA specific substrate, Structure 4, apparent anomalies in the deacylation kinetics were observed. These could be resolved by developing a model where the protonated form of the product acts as an acylating agent.

2.2 Materials and Methods

2.2.1 Materials and Instrumentation for Synthesis and Analysis of Substrates

The following chemicals were purchased from Aldrich Chemical Company (Milwaukee, Wi.). 5-methylthiophenecarboxaldehyde, malonic acid, 1,1'-carbonyldiimidazole, dicyclohexylcarbodiimide, acetyl-N-phenylalanine, glycine ethyl ester, lithium diisopropylamide, methanesulfonyl chloride, triethylamine. Silica gel F-254, and all solvents were from BDH Chemicals Canada Ltd (Toronto, Ontario).

GC-MS spectra were obtained using a Hewlett-Packard 5995 Gas Chromatograph/Mass Spectrometer equipped with a 18965A Flame Ionization Detector. LC-MS spectra were obtained using a Hewlett -Packard 5988 Liquid Chromatograph/Mass Spectrometer equipped with a thermospray interface. NMR spectra were recorded using a Varian Unity 400 NMR spectrometer.

2.2.2 Synthesis of 5MTAEt and 5MTAIm

5-methylthienylacryloyl acid was synthesized from 5-methylthiophenecarboxaldehyde and malonic acid via a Knoevenagel

condensation reaction according to the method of Koo *et al.* (1963). 5MTA imidazolide (5MTAIm) was obtained by reacting 5MTA acid with 1,1'-carbonyldiimidazole (1:1 molar ratio) in dry THF under N_2 for three hours at room temperature. The resultant solid was filtered, and then purified over silica gel using 50:50 acetonitrile/ether. Fractions were collected and analyzed spectrophotometrically using a λ_{max} of 378 nm for the purified 5MTAIm. The product bearing fractions were pooled and then dried with a rotary evaporator. The purified product was stored at -20° C until needed.

5MTAEt was synthesized and purified as described in Section 3.2.2.

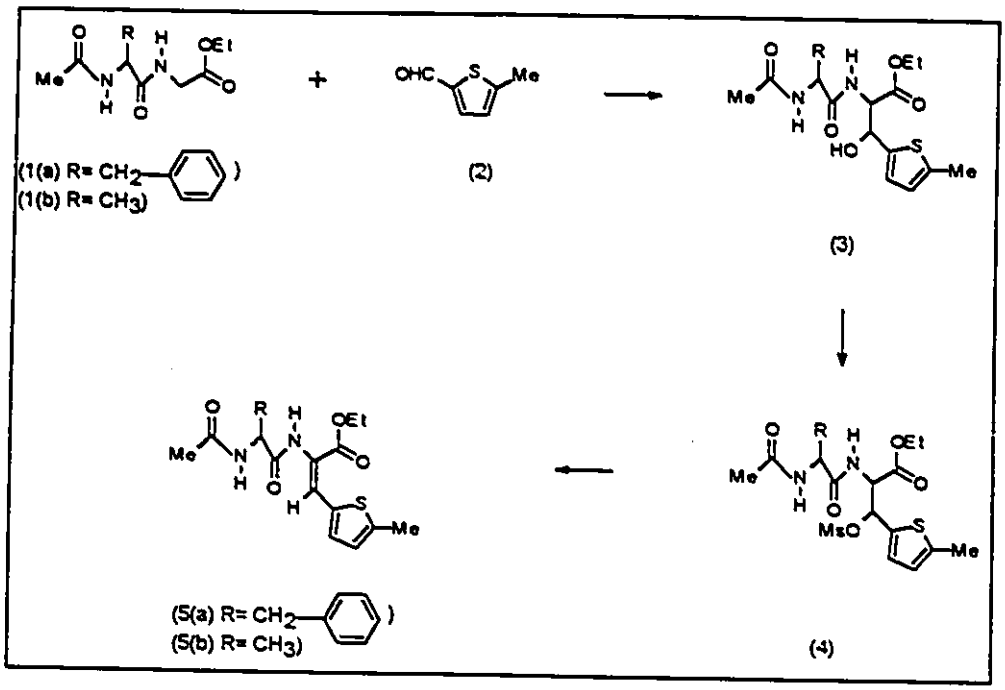
2.2.3

Synthesis of Phe5MTAEt and Ala5MTAEt

2-(N-acetyl-L-phenylalanine)amino-3-(5-methylthienyl) acryloyl ethyl ester (Phe5MTAEt, Structure 4 of Scheme 2.1) and 2-(N-acetyl-L-alanine)amino-3(5-methylthienyl)acryloyl ethyl ester (Ala5MTAEt, Structure 3 of Scheme 2.1) were synthesized using the methods described in Doran *et al.* (1995). Scheme 2.2 is a flowchart illustrating the synthetic route for these two compounds.

To prepare Phe5MTAEt, the protected dipeptide 1(a) was prepared from coupling glycine ethyl ester and N(acetyl)-L-phenylalanine by the standard dicyclohexylcarbodiimide (DCC)

Scheme 2.2



activation procedure and was purified over silica gel using flash chromatography (ethyl acetate:hexane, 2:1). The structure and purity of 1(a) were verified using GC-MS. The enolate of 1(a) was generated using 1.5 eq. of lithium diisopropyl amide (LDA) in dry THF at -78° C under N_2. It was then reacted with 1.5 eq. of 5-methylthiophenecarboxaldehyde (2) for 6 hours to give the corresponding secondary alcohol (3). A small amount (10% of total volume) of 0.1 M KPi buffer, pH 7.0 was added to quench the reaction, and the product (3) was then extracted with CH_2Cl_2 using a seperatory funnel. It was dried with anhydrous $MgSO_4$, filtered, and then redissolved in dry CH_2Cl_2. The secondary alcohol (3) was then derivatized to its mesityl derivative (4) by

treatment overnight with 3 eq. of triethylamine and 1.5 eq. of methane sulfonyl chloride (MsCl) under N_2. The presence of the excess base leads to the spontaneous conversion of (4) to the product Phe5MTAEt (5(a)) by elimination. The reaction was then quenched with pH 7.0 buffer, and the product was extracted with CH_2Cl_2, and dried with $MgSO_4$ as described previously. The product was then subjected to an initial purification step over silica gel using 2:1 ethyl acetate/hexane. Fractions containing Phe5MTAEt were collected, dried with a rotary evaporator, and then HPLC purified on a silica gel supelcosil LC-SI column (Pharmacia, Baie D'Urfé, Quebec, Canada). The column was eluted at 2.0 ml/min using the following hexane-ethyl acetate gradient; 0-5 min:100% hexane,5-50 min: linear gradient to 30% hexane, 70% ethyl acetate, 50-70 min: 30% hexane, 70% ethyl acetate: 70-71 minutes: linear gradient to 100% hexane: 71-80 minutes: 100% hexane. This profile results in Phe5MTAEt eluting at 53 minutes. The structure was confirmed by LC-MS, ^{1}H-NMR, and ^{13}C NMR.

LC-MS (Thermospray): 479 $(M+H)^+$.

^{1}H-NMR: 1.34 (t, 3H, J 7 Hz), 2.04 (s, 3H), 2.64 (s, 3H), 3.24 (ddd, 2H, J 6.7, 7.4, 14 Hz), 4.26 (q, 2H, 7 Hz), 4.94 (dd, 1H, 6.7, 7.4 Hz), 6.12 (bs, 1H), 6.32 (d, 1H, 3 Hz), 7.10 (d, 1H, 3.3 Hz), 7.32 (m, 5H), 7.68 (s, 1H).

^{13}C-NMR: 170.8, 170.5, 164.5, 146.5, 136.6, 134.0, 133.9, 130.3, 129.6, 129.5, 128.6, 128.3, 126.9, 125.8, 119.7, 61.2, 54.0, 37.2, 23.1, 17.8, 17.0.

To prepare Ala5MTAEt, the protected dipeptide 1(b) was prepared from the coupling of glycine ethyl ester and N(acetyl)-L-alanine by the standard dicyclohexylcarbodiimide (DCC) activation procedure and was then purified over silica gel using flash chromatography. The product 1(b) was retained on the silica gel while 1:1 ethyl acetate/hexane was run through the column to elute impurities. The product 1(b) was then eluted from the column using 100% ethyl acetate. The structure and purity was verified using GC-MS, and it was then used to synthesize Ala5MTAEt (5(b)) using the same experimental conditions as were used for Phe5MTAEt. The crude product Ala5MTAEt was then subjected to an initial silica gel purification using 100% ethyl acetate, then HPLC purified on a silica gel supelcosil LC-SI column (Pharmacia). The column was eluted at 2.0 mls/min using the following hexane-CH_2Cl_2 gradient: 0-10 minutes: 100% hexane, 10-20 minutes: linear gradient to 80% hexane, 20% CH_2Cl_2, 20-35 minutes: 80% hexane, 20% CH_2Cl_2, 35-45 minutes: linear gradient to 100% CH_2Cl_2, 45-50 minutes: 100% CH_2Cl_2, 50-51 minutes: linear gradient to 100% hexane, 51-60 minutes: 100% hexane. This profile results in Ala5MTAEt eluting at 28 minutes. The structure was confirmed by LC-MS and ^{1}H-NMR.

LC-MS (ionspray): 402 $(M+H)^+$.

^{1}H-NMR: 1.32 (t, 3H, J 7 Hz), 1.52 (d, 3H, J 6.5 Hz), 2.07 (s, 3H), 2.50 (s, 3H), 4.24 (q, 2H, J 7 Hz), 4.90 (q, 1H, J 6.5 Hz), 6.33 (bs, 1H), 6.70 (d, 1H, J 3.3 Hz), 7.14 (d, 1H, 3.3 Hz), 7.47

(bs, 1H), 7.71 (s, 1H).

2.2.4

Synthesis of NHCOOEt5MTAEt

2-Ethoxycarbamido-3-(5-methylthienyl)acryloyl ethyl ester (NHCOOEt5MTAEt, Structure 2 of Scheme 2.1) was synthesized by Dr. P. J. Tonge according to the method in Doran *et al.* (1995).

2.2.5

Purification of Papain

Papain (from Sigma, St. Louis, Mo.) was purified on an agarose mercurial column and then by ion exchange HPLC as described previously (Kim & Carey, 1991). Preparations of papain were determined to be 99-100% active by active site titration with 2,2'-dipyridyl disulfide (PDS) from Sigma. At pH 4.0 the rapid reaction of PDS with intact active sites produces pyridine-2-thione (λ_{max}= 343 nm, ϵ_{343}= 8080 $M^{-1}cm^{-1}$) in an amount stoichiometric with the catalytic site thiol groups (Brocklehurst *et al.* 1987). Calculation of the amount of released product in conjunction with the known concentration of enzyme determined at 280 nm (for papain, ϵ_{280}= 56,000 $M^{-1}cm^{-1}$, Brocklehurst & Little, 1973) permits the determination of the concentration of active papain.

2.2.6 Cloning of Cathepsins B and L and Oxyanion Hole Mutants of Cathepsins B and L

2.2.6.1 Materials for Molecular Biology

Ultracompetent XL2-Blue *E. coli* cells were purchased from Stratagene (La Jolla, California). Restriction enzymes and alkaline phosphatase were from Boehringer Mannheim (Laval, Quebec). DNA purification was accomplished using either a Prep-A-Gene DNA purification kit from BioRad (Mississauga, Ontario) or the Magic Minipreps DNA purification system from Promega. Sheep anti-human cathepsin B (Mort *et al.* 1981) and rabbit anti-human procathepsin L (Iwata *et al.* in preparation) were used for immune detection of expressed proteins on Western blots using the Protein G gold Immun-Blot assay kit from BioRad. Oligonucleotides were synthesized with a Beckman oligo-1000 synthesizer. Polymerase chain reactions were prepared using the GeneAmp DNA amplification Reagent kit (Perkin-Elmer Cetus, Ottawa, Ontario). PCR was performed using a Hybaid Omni Gene thermal cycler.

2.2.6.2 Molecular Biology Methods

The cDNA for human procathepsin L was amplified by PCR from a reverse transcribed RNA preparation obtained from human

cartilage and ligated, in frame, into the *Hin*dIII site of the yeast α-factor cassette which had been previously inserted into the multiple purpose shuttle vector pVT105, a variant of the pVT100U series (Vernet *et al.* 1987) lacking the alcohol dehydrogenase promoter. A similar construct for the expression of rat cathepsin B was used as described previously (Rowan *et al.* 1992). The mutations Q19S for human cathepsin L and Q23A and Q23S for rat cathepsin B were prepared by site directed mutagenesis using the method of Kunkel *et al.* (1987). Initially, protein was expressed in *the Saccharomyces cerevisiae* strain BJ3501, however yields from this system were low for the oxyanion hole mutants (0.01 - 0.1 mg/L) so the constructs were transferred to the *Pichia pastoris* expression system (Invitrogen, San Diego, California, USA). This system uses the AOX1 promoter and methanol as an inducer and is reported to provide 10-100 fold higher heterologous protein expression levels than can be obtained using *S. cerevisiae* (Cregg *et al.* 1993).

The prepro-α-factor-procathepsin B and prepro-α-factor-procathepsin L constructs were transferred from the original plasmids to the pHIL-D2 *P. pastoris* vector using a variation of the *in vivo* recombination strategy of Jones and Howard (1992). This strategy places the α-factor construct upstream of the AOX1 promoter leading to secretion of the procathepsins B and L. The desired inserts were amplified by PCR using the appropriate primers so that they could be recombined by *E. coli in vivo* with pHIL-D2 cut at the single *Eco*RI site at position 957. Primers

were designed to be complementary to both a section of the pHIL-D2 sequence adjacent to the single *Eco*RI cut site and to the relevant DNA segment of the insert to be amplified. Primer design was aided through the use of the program PC/Gene (Intelligenetics Inc. University of Geneva, Switzerland), which evaluates the suitability of a given pair of primers for PCR. As an example, the following two primers were synthesized for amplification of the prepro-α-factor-procathepsin L construct. For the 3' primer the oligonucleotide 5'-CAGTCATGTCTAAGGCTCACACAGTGGGGTAGCTGG -3' was synthesized in which the first 17 nucleotides are complementary to base pairs 961 to 977 of the sense strand of pHIL-D2 while the remaining nucleotides are complementary to the 3' end of the antisense strand of the cathepsin L cDNA. For the 5' primer the oligonucleotide 5'-ACTAATTATTCGAAACGATGAGATTTCCTTCAATTTTTAC-3' was synthesized in which the first 17 base pairs are complementary to nucleotides 937 to 953 of the antisense strand of pHIL-D2 while the remaining nucleotides are complementary to the 5' end of the sense strand of the prepro-α-factor-procathepsin L construct representing the beginning of the α factor pre-sequence. These primers were used to amplify the prepro-α-factor-procathepsin L construct by PCR using the original plasmid as the template. The resultant fragment was mixed with pHIL-D2 plasmid cut at the *Eco*RI site and treated with alkaline phosphatase in order to minimize religation of the plasmid *in vivo* without incorporation of the required insert. The DNA mixture was then used to transform competent *E.*

coli following the Stratagene procedure. Colonies were selected for ampicillin resistance and the resultant clones screened for incorporation for the appropriate insert by agarose gel electrophoresis. On average about 25 % of the clones contained the required insert. The validity of all constructs used for *P. pastoris* transformation was confirmed by DNA sequencing.

The pHIL-D2 plasmids containing the desired inserts were transformed into *P. pastoris* strain GS115 according to the manual's instructions. Protein expression by yeast transformants was gauged qualitatively by SDS acrylamide electrophoresis and Western blotting. The transformants giving the best expression were then used for large scale preparations, which typically gave yields of 1-50 mg/L. The large scale protocol consisted of 2 days growth in 2 litres of minimal glycerol media. The cells were then separated by centrifugation and suspended in 1 L of minimal methanol media for 3 days of growth. Yields of enzyme ranged from 50 mg/L for the wild type enzymes to 5 mg/L for the mutants.

2.2.7

Protein Purification

After completion of the growth in minimal methanol media, the cells were harvested by centrifugation and the media solution containing the secreted recombinant protein was concentrated to 50 mL using a Millipore Pellicon Cassette concentrator followed by an Amicon stirred cell concentrator (YM-10 membrane). The

concentrated media were then dialysed in 20 mM citrate buffer, pH 4.7. Acidification of the media during the dialysis permits both wild type cathepsin B and L to autoprocess themselves from the proenzyme form to the catalytically active enzyme (Rowan *et al.* 1992). The oxyanion hole mutants were processed to mature enzyme using pepsin as described by Rowen *et al.* (1992).

A two step column purification procedure was used to purify the enzymes to homogeneity. An initial purification to remove the majority of the media components was accomplished by binding the enzyme to a CM-Sepharose fast flow resin (Pharmacia) equilibrated with 20 mM acetate, pH 5.0 buffer. The buffer was run through the column until all traces of the coloured media components were removed. The bound enzyme was then batch eluted with buffer containing 1 M NaCl.

For wild type and mutant cathepsin B enzymes the eluent was concentrated to 10 mL using an Amicon concentrator. During concentration the buffer was exchanged to 50 mM phosphate, 1 mM EDTA, pH 6.0. The concentrated solution was then purified using an Agarose-Ahx-Gly-Phe-Gly-semicarbazone resin as described by Fox *et al.* (1992). The resin was synthesized by coupling H-gly-phe-gly-semicarbazone salt (Bachem, Philadelphia, PA) with activated CH Sepharose 4B (Pharmacia) as described by Rich *et al.* (1986).

For cathepsin L and the Q19S mutant the eluent was concentrated as for the cathepsin B enzymes except that the buffer was exchanged with 50 mM formate, 1 mM EDTA, pH 4.0. The

concentrated solution was then applied to a thiolpropyl Sepharose 6B column (Pharmacia) equilibrated with the pH 4.0 buffer. The column was washed with 10 volumes of the pH 4.0 buffer and then the bound enzyme was eluted using 50 mM phosphate buffer, pH 6.0 containing 5 mM dithiothreitol (DTT). The eluted enzyme was then exchanged with three times 40 mL of 50 mM formate, pH 4.0 buffer with an Amicon concentrator in order to remove excess DTT, and was then concentrated to a final volume of 2 mL. This was then dialysed overnight against the pH 4.0 buffer containing 1.5 mM 2,2'-dipyridyl disulfide (PDS). The protected enzyme was then stored at 4° C until needed.

2.2.8

Preparation of Acyl Enzymes

200 μL of a 0.54 mM solution of papain, pH 4.0, were activated with DTT (final concentration 2 mM) for 15 minutes. To prepare Phe5MTApapain, 30 μL of a 35 mM solution of Phe5MTA ethyl ester or Phe5MTA acid in DMSO were added to activated papain solution and allowed to react for 3 minutes. The excess substrate was then removed by centrifugation through a 1 mL G-25 column equilibrated with 5 mM formate, 1 mM EDTA, pH 4.0. To prepare 5MTApapain and Ala5MTApapain, the more reactive imidazolide derivatives had to be employed as acylating agents. In this case activated papain had to be centrifugated through the same aforementioned G-25 column in order to remove excess DTT, (which

reacts with the imidazolide derivatives). The resulting DTT free activated papain (approximately 180 μL) was then reacted with either 30 μL of a 35 mM solution of 5MTA imidazolide or Ala5MTA imidazolide in DMSO. The Ala5MTA imidazolide was found to be unstable and therefore it had to be generated *in situ*. This was done by adding a stoichiometric amount of 1,1 dicarbonylimidazole to Ala5MTA acid just prior to the acylation reaction with enzyme. After 3 minutes of acylation the resultant acyl enzyme was purified via G-25 column centrifugation as described previously. 5MTApapain has a λ_{max} of 378 nm for the bound chromophore, while both Phe5MTApapain and Ala5MTApapain have λ_{max} at 384 nm. Similar procedures were used to prepare acyl cathepsin B complexes.

2.2.9

Deacylation Kinetics

Deacylation kinetics were determined using a Cary 3 UV - visible spectrophotometer. Deacylation reactions were initiated by diluting the desired acyl enzyme into 1 mL of buffer. Buffer solutions were : 1) pH range 3 - 3.75 and 5.0 - 5.5, 50 mM citrate; 2) pH range 4.0 - 4.5, 50 mM formate; 3) pH 6.0, 50 mM 2-(N-Morpholino)ethanesulfonic acid (Mes); 4) pH range 6.5 - 7.5, 50 mM phosphate; 5) pH 8.5, 50 mM N,N-bis[2-hydroxyethyl]glycine (bicine). All deacylation buffer solutions also contained 0.2 M NaCl (in order to maintain constant ionic

strength), and 1 mM EDTA. For some experiments involving Phe5MTApapain, the buffer solution also contained 1-(L-trans-epoxysuccinyl-L-leucylamino)-4-guanidinobutane (E-64) purchased from Boehringer Mannheim. Deacylation kinetics were monitored at 390 nm at 25° C and first order rate constants were evaluated by importing the data into the Enzfitter program (Elsevier-Biosoft) and fitting them to a first order exponential equation. The pK_a for deacylation of 5MTApapain, Ala5MTApapain, and Phe5MTApapain in the presence of E-64 was evaluated by fitting the data to the equation $y = (C*(10^{(pH-pKa)}))/(10^{(pH-pKa)} + 1)$, where C = upper limit for the rate constant. All experiments were performed in duplicate, with each data pair giving an error of less than 10%.

2.2.10

Steady State Kinetics

Steady state kinetic assays were performed in 50 mM bicine, 0.2 M NaCl, 1mM EDTA at pH 8.5 containing 20 % (v/v) acetonitrile at 25° C. Lucas and Williams (1969) have shown that in 20% v/v acetonitrile - water there is little change in the kinetic parameters for the papain catalyzed hydrolysis of N-benzoylglycine methyl ester compared to water alone. Initial velocities for hydrolysis of either Phe5MTA or Ala5MTA ethyl esters were determined at a constant papain concentration of 6.8 μM, while the substrate concentrations were varied. The reactions were monitored at 350 nm so that the initial rate of

disappearance of reactant could be monitored without interference from the product absorption band. Michaelis-Menton parameters k_{cat} and K_m were determined from the initial velocities by Hanes-Woolf plots.

2.2.11

Dissociation Constant Evaluation

The following procedure was carried out to evaluate K_d (i.e K_i) for Phe5MTA papain at various pH values. The ratio of the 280 nm protein band to 384 nm bound chromophore band was determined to be 2.5 at pH 4.0 for an 11 μM solution. The acyl enzyme complex is sufficiently stable under these conditions so that this ratio may be used to represent 100% acylation. Concentrated Phe5MTApapain at pH 4.0 was diluted to a given concentration into a buffer of the desired pH. The variation with time of the 384 nm band was used to verify that the system had reached equilibrium. By using the 280 nm/384 nm ratio of 2.5 and the measured concentration of protein in the sample, it was possible to calculate that the percentage of acyl enzyme remaining as equal to (100 x A_{384}/A°_{384}), where $A^{\circ}_{384} = A_{280}/2.5$, the magnitude of the 384 nm band before deacylation (i.e. representing 100% acylation), and A_{384} is the absorbance at 384 nm at equilibrium. Once the amount of deacylated enzyme has been determined it is possible to calculate the concentration of acyl enzyme, free enzyme, and free product due to the 1 : 1 : 1 stoichiometry of

the reaction. K_i^{app} may then be evaluated using Equation 2.1.

2.1) K_i^{app} = [free enzyme] [free product] / [acyl enzyme]

In order to calculate the true K_i corrected for product ionization, the Henderson-Hasselbach equation was used to calculate the degree of product ionization at the desired pH. A pK_a of 3.8 was determined spectrophotometrically for Phe5MTA acid and this value was used for all calculations to determine the degree of product ionization. The calculated protonated product concentration was then substituted into the K_i equation for [free product] as shown in Equation 2.2.

2.2) K_i = [free enzyme][protonated product] / [acyl enzyme]

2.3 Results

Table 2.1 lists the deacylation rates that were determined experimentally for the acyl cysteine proteases generated from the substrates shown in Scheme 2.1. A pH of 8.5 was used to monitor deacylation for the acyl papain complexes, while a pH of 6.0 was used for the acyl cathepsin B and L complexes. The cathepsins are measured at the lower pH value because both are irreversibly inactivated at pH values greater than 6.0 (Kirschke & Barrett, 1981). Since the acyl enzyme complexes display a characteristic absorption band near 380 nm, deacylation rates can be

Table 2.1

Deacylation Rates of Acyl Cysteine Proteases

Acyl Enzyme	Deacylation rate at active pH ($sec^{-1} \times 10^{-3}$)
5MTApapain	1.1
NHCOOEt5MTApapain	1.0
Phe5MTApapain	3.3
Ala5MTApapain	3.3
5MTAcatBwt	0.5
Phe5MTAcatBwt	2.8
5MTAcatBQ23A	0.18
5MTAcatBQ23S	0.07
Phe5MTAcatBQ23S	0.6
5MTAcatLwt	7.7
Phe5MTAcatLwt	15
5MTAcatLQ19S	0.3

Acyl enzymes were prepared and their deacylation rates measured as described in Sections 2.2.8 and 2.2.9 in the Materials and Methods. All experiments were performed in duplicate, with each data pair giving an error of less than 10%.

conveniently measured by following changes in absorbance as a function of time.

The base catalyzed hydrolysis rates of the ethyl ester substrates in Scheme 2.1 are tabulated in Table 2.2, along with the rate constant of the corresponding acyl papain generated from the substrate. The ratio of the deacylation rate constant of the acyl papain to the base catalyzed hydrolysis rate of the substrate is also calculated to provide a gauge of the degree of rate enhancement catalyzed by papain relative to the corresponding non-catalyzed hydrolysis reaction.

The reactions of the 5MTA, NHCOOEt5MTA, Ala5MTA, and PheMTA substrates with papain were subjected to a more detailed kinetic characterization. Table 2.3 contains values of k_{cat}, K_m, and k_{cat}/K_m determined from steady state kinetics for the above four substrates with papain. Figure 2.1 illustrates the variation of the deacylation rate constant with pH for 5MTApapain and Ala5MTApapain, from which pK_a's of 4.9 and 3.5, respectively, were determined.

For the Phe5MTApapain complex, complex deacylation profiles were obtained. At low pH little or no deacylation occurs. As the pH is raised, successively larger proportions of the total acyl papain deacylates. This is illustrated in Figure 2.2, where the final concentration of Phe5MTApapain at equilibrium is plotted as a function of pH. Figure 2.3 is a plot of the apparent pK_1, which may be readily calculated from the data in Figure 2.2, vs. pH. This plot is a straight line with a slope of -1.

Table 2.2

Rate Constants for the Hydrolysis of Ethyl Ester Substrates and Deacylation Rates for Their Corresponding Acyl Papains

Substrate	$k_{[OH^-]}$ (Mol^{-1} sec^{-1} x 10^{-3})	k_{obs} for acyl papain (sec^{-1} x 10^{-3})	Ratio k_{obs} / $k_{[OH^-]}$
5MTAEt	23	1.1	0.05
NHCOOET5MTAEt	2.9	1.0	0.34
Phe5MTAEt	0.50	3.3	6.6
Ala5MTAEt	0.65	3.3	5.1

Hydrolysis rates of the substrates were determined at pH 14.0, 25° C. This was accomplished by adding 5 μL of a stock solution of substrate in DMSO to 995 uL of a 1.0 N solution of NaOH, giving a final concentration of approximately 50 μM. Deacylation kinetics were monitored at 340 nm, and the rate constants were evaluated as described in Section 2.2.9 of the Materials and Methods. All experiments were done in triplicate with a standard deviation of error of less than 5%. Values of k_{obs} for the corresponding acyl papain were obtained from Table 2.1.

A deacylation rate - pH profile similar to the one for Ala5MTApapain in Figure 2.1 could be obtained for Phe5MTApapain if the deacylation reaction was carried out in the presence of E-64, a potent irreversible inhibitor of proteases from the papain super family. This plot is shown in Figure 2.4.

Table 2.3

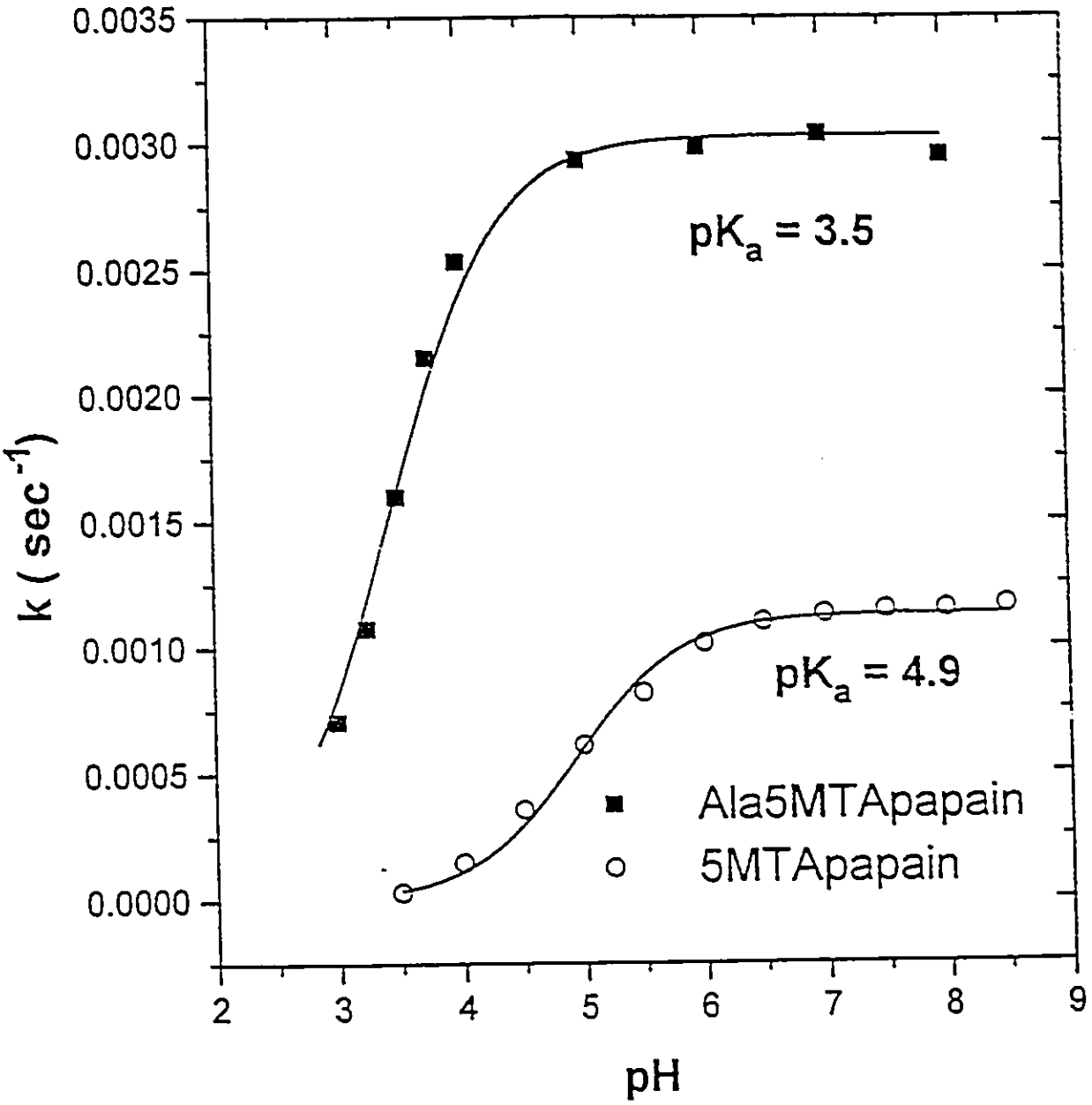
Kinetic Constants for Papain Catalyzed Hydrolysis Determined by Steady State Kinetics.

Substrate	k_{cat} (sec^{-1} x 10^{-3})	K_m (M x 10^{-3})	k_{cat}/K_m
5MTA ethyl ester	-	-	< 0.07
NHCOOEt5MTA ethyl ester	-	-	< 0.07
Ala5MTA ethyl ester	3.3	3.9	0.85
Phe5MTA ethyl ester	3.3	0.062	53

Values of k_{cat} and K_m were evaluated as described in Section 2.2.10 of the Materials and Methods. The steady state kinetic parameters for NHCOOEt5MTAOEt and 5MTAOEt could not be obtained since the reaction was too slow to measure even at concentrations of 3 mM, which approaches their solubility limits in the assay buffer, and hence an upper limit to k_{cat}/K_m is presented in Table 2.3.

Figure 2.1

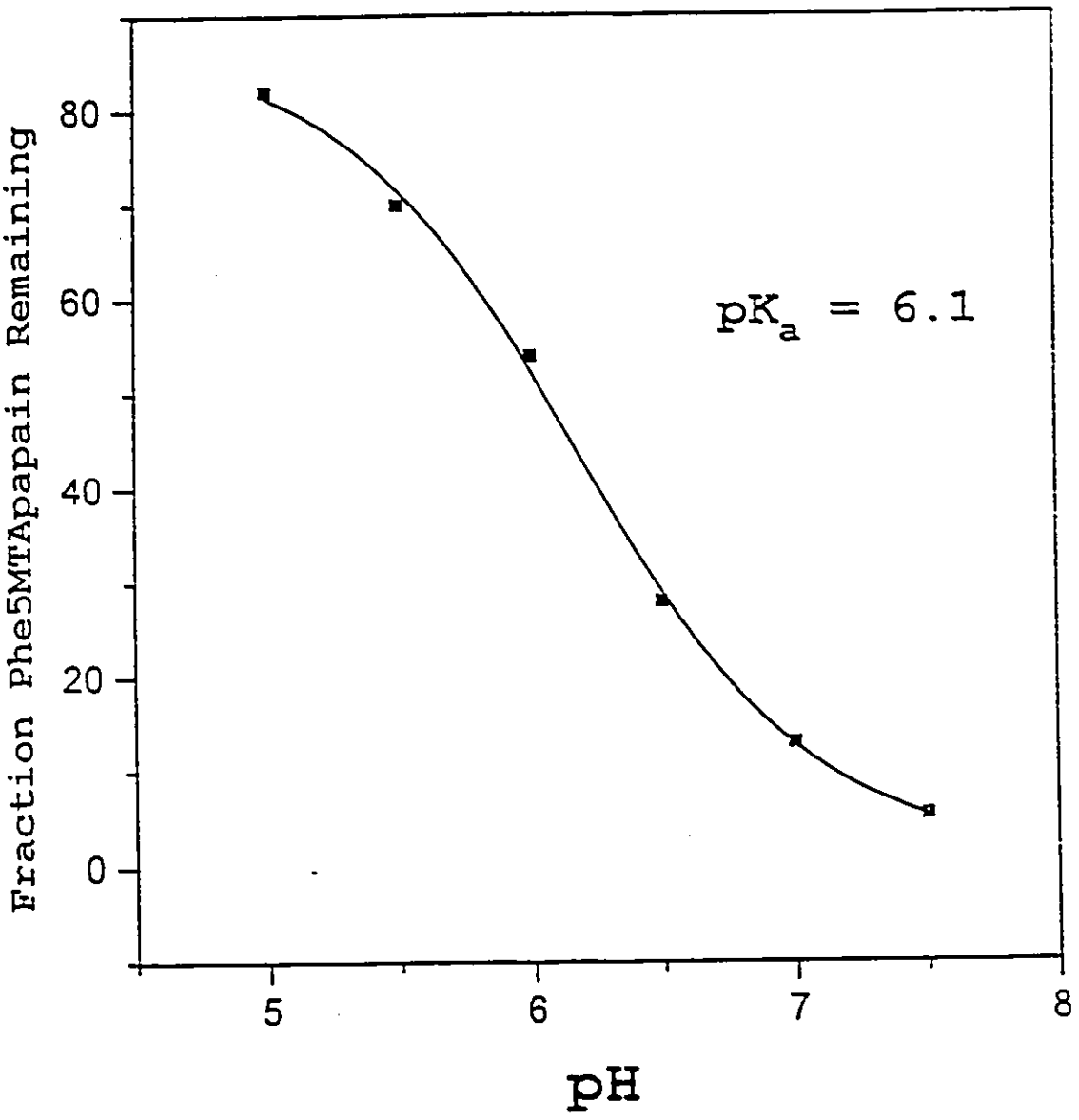
5MTApapain and Ala5MTApapain pK, Profiles



Plots of deacylation rate constant vs. pH for 5MTApapain and Ala5MTApapain. Deacylation kinetics were measured and the rate constant was evaluated as described in Section 2.2.9 of the Materials and Methods. The pK_a's were evaluated as described in Section 2.2.9 in the Materials and Methods.

Figure 2.2

Effect of pH on the Final Equilibrium Concentration of Phe5MTApapain

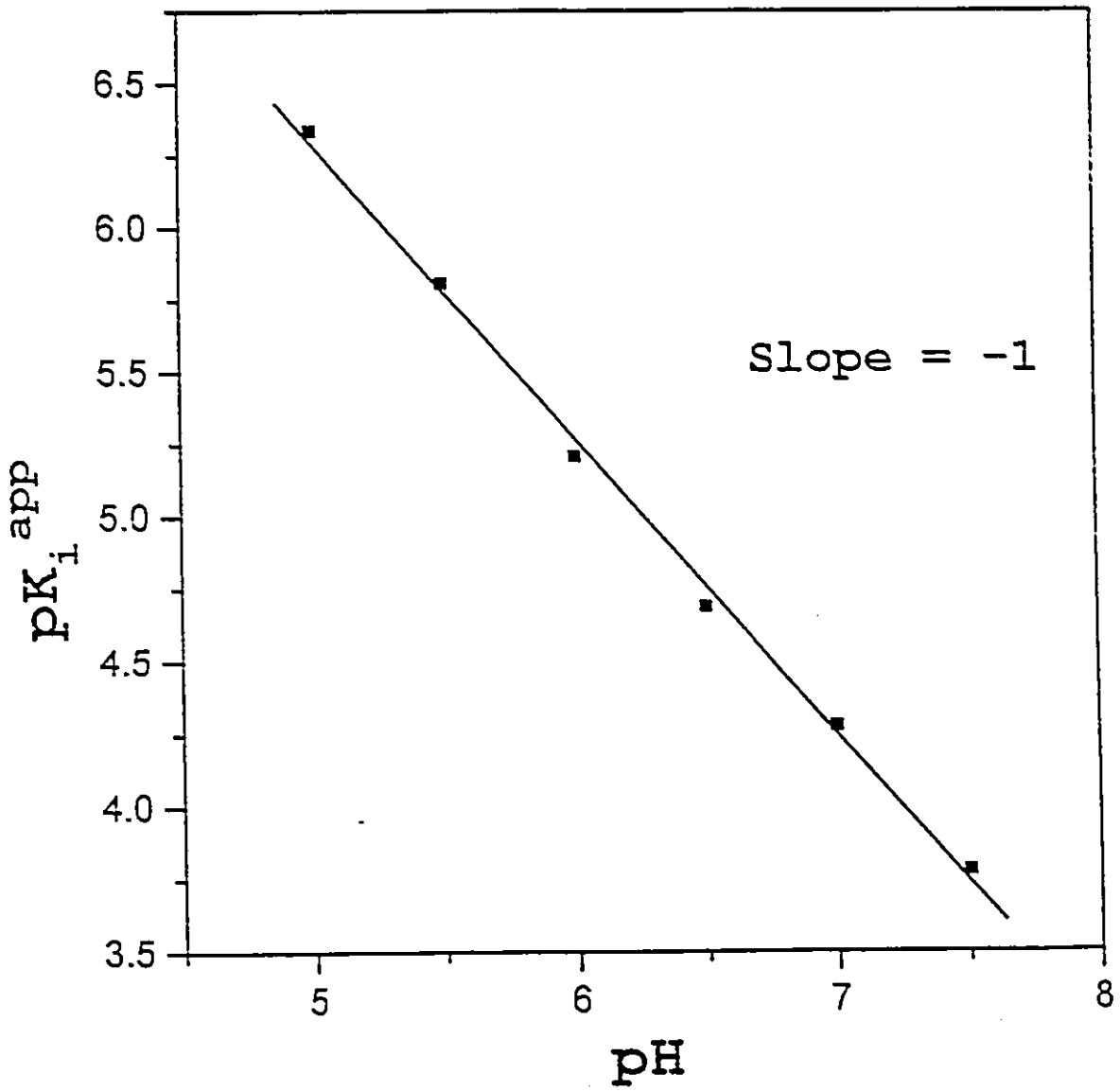


Plot of the fraction of Phe5MTApapain remaining at equilibrium *vs.* pH. The fraction remaining at any given pH value was calculated as described in Section 2.2.11 in the Materials and Methods. The pK_a was evaluated using the equation $y = C*(10^{(pH-pKa)}))/(10^{(pH-pKa)} + 1)$ as described in Section 2.2.9 of the Materials and Methods.

Figure 2.3

Effect of pH on the Apparent pKi for the Equilibrium System

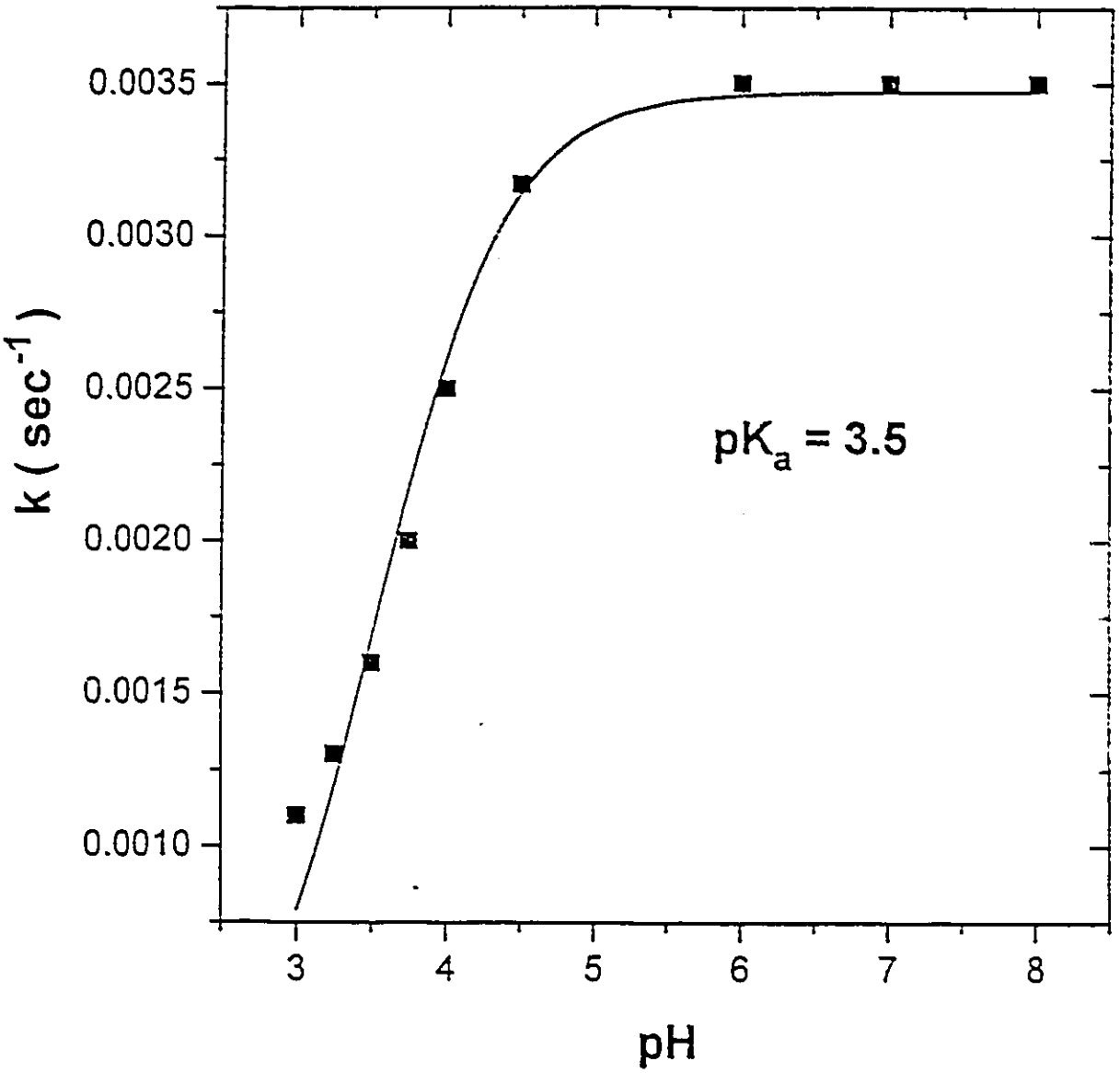
Phe5MTApapain $\rightleftharpoons$ Phe5MTACOOH + Papain



Plot of pK_i^{app} *vs.* pH for the Phe5MTApapain equilibrium system. The apparent pK_i was evaluated using Equation 2.1 as described in Section 2.2.11 of the Materials and Methods.

Figure 2.4

Phe5MTApapain pK, Profile



Plot of the deacylation rate constant *vs.* pH for Phe5MTApapain. Deacylation kinetics were measured and the rate constant was evaluated as described in Section 2.2.9 of the Materials and Methods. All measurements were performed in the presence of 118 μM E-64. The pK_a was evaluated as described in Section 2.2.9 of the Materials and Methods.

2.4 Discussion

2.4.1 Comparison of Deacylation Rate Constants for Acyl Enzymes Derived From Specific and Non-Specific Substrates

Each of the substrates shown in Scheme 2.1 can be used to generate an acyl enzyme complex, consistent with deacylation being rate limiting for esters. These acyl enzymes are most stable at pH 4.0 and will deacylate rapidly and completely when the pH is jumped to a higher value, (pH 8.5 for papain, pH 6.0 for the cathepsin B & L enzymes). The results for a series of such acyl enzymes are listed in Table 2.1. The 5MTApapain acyl enzyme deacylates with a rate constant of 0.0011 sec^{-1} and attachment of the NHCOOEt moiety to the α carbon of 5MTA has no effect on the deacylation rate of papain. There is however a 3 fold increase in deacylation with either the Phe5MTApapain or Ala5MTApapain acyl enzymes. The increase in deacylation rate on going from 5MTA to Phe5MTA based acyl enzymes is also observed with cathepsins B and L, and the cathepsin B Q23S mutant, where k_3 increases by 6 fold, 2 fold, and 9 fold respectively. Even though these increases are modest, they suggest that the presence of P_1-S_1 and P_2-S_2 contacts can contribute to accelerate deacylation.

Table 2.2 contains the second order rate constants for base catalyzed hydrolysis of the 5MTA, NHCOOEt5MTA, Phe5MTA, and

Ala5MTA ethyl esters and their ratios to the hydrolysis rates of the corresponding acyl papain complexes. Contrary to the effect observed on enzyme catalyzed deacylation, the attachment of a peptide moiety to the α-carbon of 5MTA results in a decrease in the base catalyzed hydrolysis of the peptide 5MTA esters compared to the non-specific 5MTA ester. Phe5MTAOEt is 46 fold less reactive than 5MTAOEt towards base catalyzed hydrolysis, and yet Phe5MTApapain deacylates faster than 5MTApapain by a factor of 3. Assuming that the corresponding thiol ethyl esters show the same base catalyzed hydrolysis trend, the data in Table 2.1 show that papain catalyzed deacylation is accelerated by approximately 100 fold compared to the non-specific 5MTApapain. Thus, the interactions of the dipeptide moiety with papain's active site contribute significantly to increase the deacylation rate and compensate for the lower susceptibility to hydrolysis inherent within the structures of Phe5MTA and Ala5MTA. The incremental effects on $k_{[OH^-]}/k_{obs}$ observed upon attachment of NHCOOEt and of either of the dipeptide moieties to 5MTA are reminiscent of those observed by Berti *et al.* (1992) for a variety of substrates that measure the different contributions made by P_1-S_1 and P_2-S_2 hydrogen bond contacts to the catalytic rate. Hence, it may be concluded that the P_1 amide (as evinced by the NHCOOEt substrate), the P_2 amide and the P_2 C=O (evinced by Phe5MTA and Ala5MTA) are making significant contributions to deacylation rate enhancements. The presence of phenylalanine in the S_2 active site pocket does not make a contribution to deacylation rate

enhancement, since exchanging Phe for Ala in the dipeptide 5MTA ester has only a small effect on the base catalyzed hydrolysis rate and no effect on the values of k_3 for the acyl papain complex.

For comparison, the parameters k_{cat} and K_m for papain catalyzed hydrolysis of the ethyl ester substrates were determined by steady state kinetics and are shown in Table 2.3. As mentioned previously, deacylation is rate limiting for the ethyl ester substrates and, accordingly, k_{cat} was found to equal k_3, the deacylation rate constant obtained through deacylation kinetic measurements (see Table 2.1). The presence of the dipeptide moieties in Ala5MTAOEt and Phe5MTAOEt results in a major increase in the specificity constant k_{cat}/K_m compared to the NHCOOEt5MTAOEt and 5MTAOEt substrates. This illustrates the significant contribution the P_2-S_2 interactions make for improving the overall reactivity relative to the two substrates that lack this interaction. The data in Table 2.3 also show the preference for phenylalanine compared to alanine in the S_2 pocket, since k_{cat}/K_m is 62 fold greater for Phe5MTA over Ala5MTA. This is similar to the 74 fold difference in k_{cat}/K_m's determined by Brubacher and Zaher (1979) for the substrates Z-Phe-Ala-OMe *vs.* Z-Ala-Ala-OMe. Since the k_{cat} values for the Phe5MTA and Ala5MTA substrates are identical, the increased specificity of the Phe5MTA substrate relative to the Ala5MTA is manifested exclusively in the acylation step. This finding is consistent with other work demonstrating the S_2 subsite's preference for Phe

over Ala is predominantly seen in the acylation process (Lowe & Yuthavong, (1971), Brubacher & Zaher, (1979)).

2.4.2

Extending the Range of Deacylation Rate Constants by the Use of Oxyanion Hole Mutants

To provide a basis for the development of structure-activity relationships using Raman and absorption spectroscopies, it is important to have acyl enzyme complexes with as wide a range of rate constants as possible. Using 5MTA as the acyl group, the deacylation rate constants for oxyanion hole mutants of cathepsins B and L were measured and are listed in Table 2.1. Replacing Gln23 by Ala or Ser in cathepsin B results in a 3 to 7 fold decrease in the deacylation rate while the mutation of Gln19 to Ser in cathepsin L causes a 26 fold decrease. These mutations have been shown to decrease k_{cat}/K_m for hydrolysis of the specific substrate CBZ-Phe-Arg-MCA by approximately 60 to 600 fold in papain (Ménard *et al.*, 1991) and 12 to 100 fold in cathepsin B (Ménard *et al.*, unpublished data). Hence, it can be concluded that even with a non-specific substrate the oxyanion hole is operative in catalysis, although not to the same degree as seen with a specific substrate. The use of these mutants permits the expansion of the deacylation rate constant range from 15 x 10^{-3} sec^{-1} with Phe5MTAcatL down to 0.07 x 10^{-3} sec^{-1} with 5MTAcatB-Q23S.

2.4.3

Variation of Deacylation with pH

The influence of pH on deacylation rates was determined for 5MTA, Ala5MTA, and Phe5MTApapain. Results for the first two acyl enzymes are shown in Figure 2.1. The non-specific 5MTApapain has a deacylation pK_a of 4.9, which is similar to the pK_a value of 4.7 obtained for cinnamoylpapain (Bender and Brubacher, 1964). The deacylation pK_a of Ala5MTApapain, however, is 3.5, more than one pH unit lower than that for 5MTApapain. Assuming this pK_a represents ionization of the active site His159 residue, this result indicates that the presence of the dipeptide moiety giving rise to S_1-P_1 and S_2-P_2 contacts lowers the pK_a of His159 in the acylenzyme. This implies that the environment about the His159 is rendered more hydrophobic in the case of the dipeptide substrate.

For Phe5MTApapain, complex deacylation profiles were obtained in which successively larger proportions of the total acyl enzyme concentration deacylates as the pH is increased. The pK of this equilibrium process may be calculated by plotting the fraction of acyl papain population that does not deacylate *vs.* pH as shown in Figure 2.2, where the fraction of acyl enzyme population is derived as discussed in Section 2.2.11 of the Materials and Methods. The pK of 6.1 (Figure 2.2) represents the pH at which 50 % of the acyl papain deacylates. This result closely resembles Smolarsky's determination of the extent of acyl enzyme formation as a function of pH for PheCNM acid. He

determined that the extent of acyl papain formation was maximal at pH 4.0, and that it decreases sigmoidally with a transition point at pH 5-5.5, indicating that 50% acylation is achieved in this pH range (Smolarsky 1978).

Smolarsky's interpretation of these results was that the K_i of the PheCNM acid is increasing as a function of pH. The pK_a of protonation for PheCNM acid is 3.5, which is 1.5 - 2 pH units below the pH where 50 % acylation is achieved. This pK differential lead Smolarsky to conclude that a protonation step involving one or more amino acid side chains of papain is required for the enzyme to become acylated by the acid product. He suggested that protonation of the enzyme provides some of the free energy required to shift the equilibrium toward acylation.

Other workers have also observed inhibition of papain activity by acid product molecules. Sluyterman (1964) found that the inhibition of papain by benzoylarginine is strongly pH dependent. Berger and Schechter (1970) determined that a variety of di- and tri-peptides, with both a free carboxylic acid group and an aromatic amino acid capable of interacting at the S_2 subsite of papain, gave rise to competitive inhibitors whose K_i would increase with increasing pH. The latter authors hypothesized that the pH dependence of inhibition could be explained, at least in part, by assuming that only the protonated form of the acid product (i.e. -COOH) was in equilibrium with the enzyme. A method for calculating a "true" binding constant pK_i' was developed by assuming that only the protonated form of the

inhibitor was in equilibrium with the enzyme, leading to a constant pK_i' value of 6.8 in the pH range 4.3 - 7.0 for tert-Boc-p-iodo-Phe-Arg (Berger & Schechter, 1970). These authors however could not eliminate the possibility that ionization of one or more groups in papain could also be a factor in the observed pH dependence of inhibition.

The apparent pK_i's of the Phe5MTApapain complex were calculated as a function of pH using Equation 2.1 as described in Section 2.2.11, and the data are plotted in Figure 2.3. A straight line is obtained with a slope of -1, which illustrates the increase in K_i^{APP} as pH increases. According to Dixon & Webb (1979(a)) the slope of any straight line section of a pK *vs.* pH graph is numerically equal to the change in charge occurring in that pH range when the complex dissociates into free enzyme and ligand. For Phe5MTApapain in the pH range 5 - 7.5 the product is negatively charged and suggests that it is responsible for the slope of -1.

A calculation, similar to that of Berger & Schecter (1970), to evaluate the true binding constant pK_i' was performed using Equation 2.2 in Section 2.2.11, and the results are given in Table 2.4. It is apparent from Table 2.4 that once it is assumed that only the protonated inhibitor reacts with the enzyme the pK_i' value is invariant, with a standard error of deviation of less than 2 % (7.42 ± 0.13). This result implies that the increasing K_i values in Figure 2.3 are due to the protonation state of the inhibitor, however, it does not rule out the

Table 2.4

Values of pK_i' Determined by Assuming Only the Protonated Form of the Product Binds to Papain

pH	pK_i'
5	7.6
5.5	7.2
6.0	7.4
6.5	7.4
7.0	7.5
7.5	7.4

Values of pK_i' were calculated using Equation 2.2 in Section 2.2.12. The average value of pK_i' is 7.42 ± 0.13.

possibility of enzyme group ionization also playing a role.

A simple mathematical model, which expands on the ideas inherent in Equation 2.2, has been developed that predicts the deacylation behaviour of systems such as Phe5MTApapain. The model makes two assumptions, that only the protonated form of the product will acylate the enzyme and that, in the pH range of concern, binding is unaffected by possible ionizations of active site amino acid side chains.

2.4.4

Model for pH Dependence of Inhibitor - Papain Interaction

Deacylation is represented as :

$$A \underset{k'}{\overset{k}{=}} B + C$$

where A = acyl enzyme , B = free enzyme, C = free product, k = deacylation rate constant (sec^{-1}), k' = acylation rate constant ($M^{-1}\ sec^{-1}$).

At equilibrium, we assume an X amount of acyl enzyme has deacylated. The equilibrium constant expression is:

1) $$(X)(X') / (A_o - X) = K$$

where A_o = starting concentration of acyl enzyme.

X = amount of free enzyme.

X' = amount of free product, and X = X'.

K = equilibrium constant.

the total amount of product, X', is the sum of the protonated, X_H, and unprotonated forms, X_U.

2) $X' = X_U + X_H$ and

$$X_H \overset{K_a}{=} X_U + H^+$$

3) Where $K_a = [X_U][H^+] / [X_H]$ and therefore

$$[X_U] = K_a[X_H] / [H^+]$$

Substitution of Equation 3 into Equation 2 gives:

4) $X' = K_a [X_H]/[H^+] + [X_H]$

Rearrangement gives:

5) $[X_H] = X' / (1 + K_a / [H^+])$

At equilibrium, we can rearrange Equation 1 to give Equation 6.

6) $K(A_o - X) = (X) (X')$

We assume that only the protonated product [X_H] participates in the equilibrium, therefore we substitute $X'/(1+K_a/[H^+])$ from Equation 5 into X' in Equation 6.

7) $K(A_o - X) = (X) (X' / (1 + K_a / [H^+]))$

We now expand Equation 7 and disregard the X' notation:

8) $KA_o - KX = (X^2) / (1 + K_a/[H^+])$

Rearrangement leads to:

9) $X^2 + K(1+K_a/[H^+])X - KA_o(1+K_a/[H^+]) = 0$

Solve for X using the quadratic formula.

10) $$X=\frac{-K(1+K_a/[H^+])\pm\sqrt{(K(1+K_a/[H^+]))^2-4(-KA_o(1+K_a/[H^+])}}{2}$$

Using Equation 10 we can determine the equilibrium concentration of any species provided that the equilibrium constant K (which is equal to K_i', the "true" equilibrium constant), initial starting concentration, and K_a of the acid

product are known. For Phe5MTApapain K is equal to 3.8 x 10^{-8} M (i.e the inverse negative logarithm of 7.42 as evaluated from Table 2.4).

Using the above model we can make predictions of the behaviour of this type of system. It is important to note that the derivation assumes that the ionization of the product is the only factor in the observed kinetics and that any effect of ionization by enzyme groups is ignored.

The underlying hypothesis of the model is that rebinding of the product occurs only for the protonated form. It follows that if the rebinding step is blocked then one should obtain complete deacylation at all pH values. This may be tested by carrying out the deacylation in the presence of E-64, a powerful cysteine protease inhibitor. These experiments were performed and the results are shown in Figure 2.4. Complete deacylation was obtained at all pH's and a pK_a of 3.5 was determined. This pK_a is identical to that of the Ala5MTApapain, and the deacylation rate constants of the two acyl papains are very similar at any given pH. This is consistent with the conclusions drawn from the data in Table 2.2 that the effects of the Phe side chain operate in the acylation process. Other workers have reached similar conclusions. Lowe and Yuthavong (1971) determined that the papain catalyzed deacylation rates for a variety of N-acylglycines are relatively constant, while the k_{cat}/K_m values range over a factor of 107, with substrates with Phe in the P_2 position giving the highest values. Brubacher and Zaher (1979) found the same

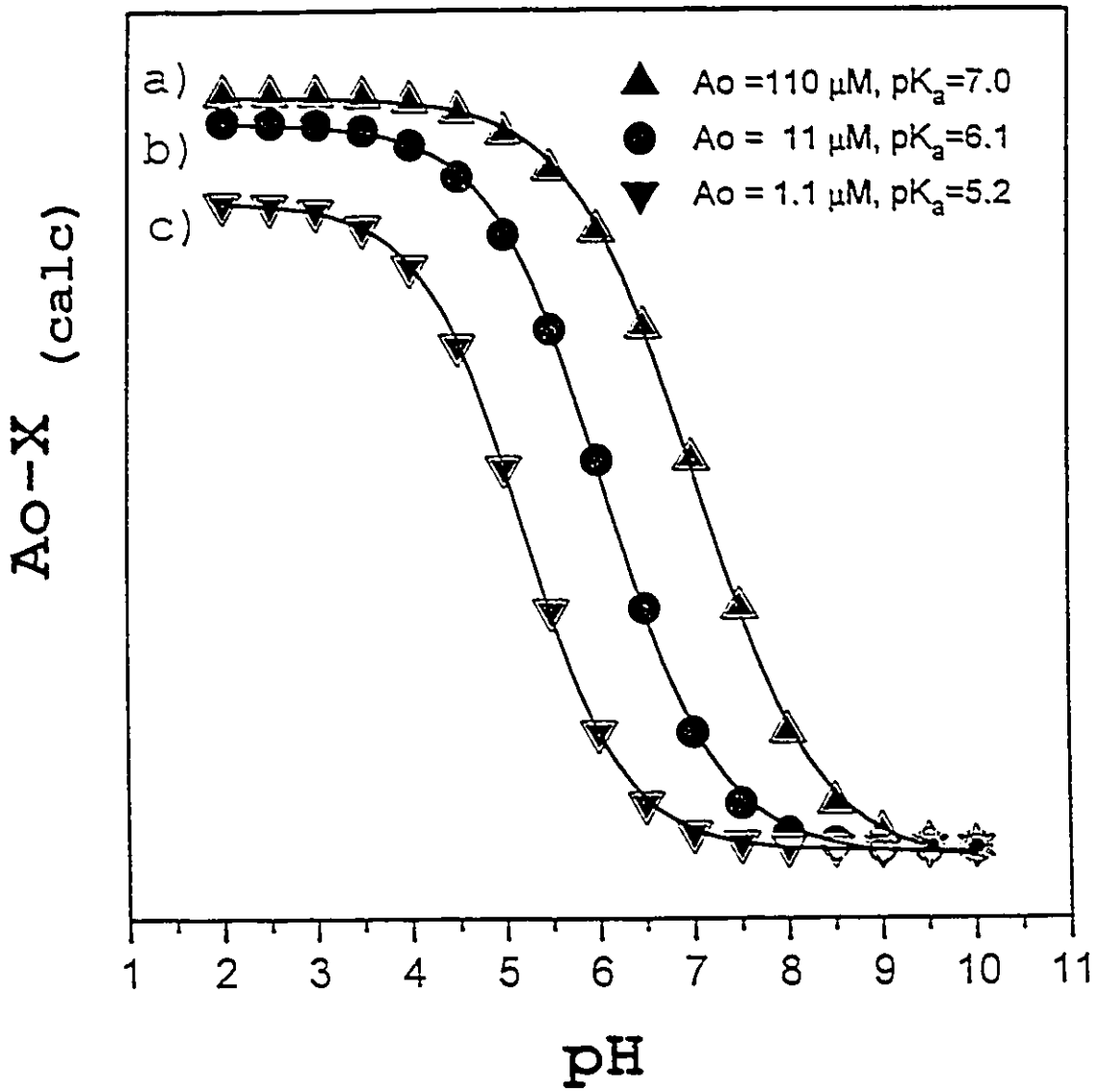
behaviour for a series of methyl esters, and they concluded that the S_2 site preference for phenylalanine is apparent only in terms of overall reactivity.

An important conclusion stemming from the experiments with E-64 is that the model's assumption that the product is rebinding to the enzyme is correct. It is apparent that the Phe-S_2 subsite interaction results in extremely efficient acylation. It is so efficient that even the protonated product will bind to papain, which explains why this substrate becomes an inhibitor at low pH. It is only at higher pH values (pH > 7.5) that the concentration of protonated product becomes so insignificant that normal deacylation may proceed. Thus whether Phe5MTA ethyl ester behaves as a substrate or as an inhibitor is a function of the degree of ionization of the released acid product.

The model may also be used to predict the effects of initial acyl enzyme concentration on the final equilibrium concentrations at any given pH. Figure 2.5 contains simulations illustrating the effect changing the initial acyl enzyme concentration has while holding all other parameters constant. The apparent pK shifts to lower values as the acyl enzyme concentration decreases. This means that as the initial concentration of acyl enzyme gets lower the reacylation process becomes less significant. This process was verified experimentally. An 11 μM solution of Phe5MTApapain which was stable at pH 4.0 was diluted ten fold in the same pH 4.0 buffer. The acyl enzyme proceeded to deacylate, indicating that the inhibitor binding is a function of the starting acyl

Figure 2.5

Effect of Initial Acyl Papain Concentration on the Final Acyl Papain Concentration at Equilibrium



Simulation curves illustrating the effect the initial acyl enzyme concentration, A_o, has on the final equilibrium concentration of Phe5MTApapain. Equation 10 from page 71 was used to generate the curves. The following values were used for all of the simulations. Phe5MTA acid product K_a = 0.00016 M, K = 3.8 x 10^{-9} M (from Table 2.4). For the sake of clarity curves b and c were multiplied by factors of 10 and 100, respectively, so that they appear on the same scale.

enzyme concentration as well as the pH. Thus, the pK of 6.1 derived from Figure 2.2 indicates that this is the pH where half of the acyl enzyme deacylates, but this value is only applicable for an 11 μM solution of Phe5MTApapain. Equation 10 may be used to accurately predict the pK that will result from an initial starting Phe5MTApapain concentration of 11 μM. Simulation b in Figure 2.5 was generated by using an initial acyl papain concentration of 11 μM, which is the same concentration that was used in all experiments in Figure 2.2 to determine the pK at which 50% of the acyl papain deacylates. Both the simulation (Figure 2.5, curve b) and the experimentally determined curve (Figure 2.2) give a pK of 6.1, thus Equation 10 from the model predicts precisely the effects of pH on the degree of inhibitor binding to papain.

These results indicate that the apparent pK is in fact dependent on the initial starting acyl enzyme concentration, and therefore should not be interpreted as being indicative of the ionization of an active site group. Furthermore, the model may be used as a general tool to understand the pH dependence of enzyme binding. If Equation 10 successfully reproduces the experimentally determined equilibrium concentrations of a given enzyme-ligand system as a function of pH, then it is probable that the degree of binding is solely a function of the protonation state of the ligand molecule. If Equation 10 fails to satisfactorily reproduce the experimental data, then it can be concluded that ionization of one or more enzyme groups is also

involved in the binding process.

2.5 Summary

In order to investigate structure-reactivity relationships within a series of acyl cysteine proteases, deacylation kinetics have been measured for a number of acyl intermediates involving members of the papain super family. A derivative of the simple chromophoric ligand 5-methyl-thienylacryloyl- (5MTA) and those based on three specific chromophorically labelled derivatives, vis. NHCOOEt5MTA, Ala5MTA, and Phe5MTA, were used to create acyl enzyme adducts with papain, cathepsin B and cathepsin L. The chromophoric specific substrates were designed to utilize hydrogen bonding and hydrophobic interactions which are known to promote catalysis by papain. For cathepsins B and L removing one of the hydrogen bonding groups making up the oxyanion hole has a modest effect on catalysis by these substrates. With these substrates and the wild type and oxyanion hole mutants the values of the deacylation rate constants stretch over a 214 fold range, from 0.07 to 15 x 10^{-3} sec^{-1}.

The pK_a for deacylation of 5MTApapain is 4.9, close to that reported for similar papain intermediates, while that for Ala5MTApapain is at 3.5, which in the latter likely represents the effect of P_1-S_1 and P_2-S_2 interactions on the pK_a of His159. For Phe5MTApapain the extent of deacylation was found to depend

on the pH and the starting acyl enzyme concentration. A simple model has been derived which accounts for this behaviour, using the assumptions that the protonated form of the acyl product reacylates the enzyme, and that in the pH range 5.0 to 7.5 the ionization of active site groups has no effect on the reacylation. The validity of the first assumption was demonstrated by following the deacylation of Phe5MTApapain in the presence of the potent inhibitor E-64, whereupon complete deacylation occurred at all pH's with a pK_a identical to that for Ala5MTApapain, *viz.* pH 3.5. The assumption that ionization of groups in the enzyme's active site has no effect on the reacylation process was demonstrated by the fact that the model could quantitatively account for the observed kinetics without invoking any variables related to enzyme group ionization.

CHAPTER 3

THE SPECTROSCOPIC AND CONFORMATIONAL PROPERTIES OF $5MTASC_2H_5$ AND $5MTAOC_2H_5$

3.1 Introduction

Thiol esters play a key role in many biochemical processes. For example, a transient thiol ester is generated during hydrolysis by cysteine proteases (Brocklehurst *et al.* 1987); and enzyme systems requiring coenzyme A, which serves as the acyl carrier for a wide variety of enzyme catalyzed reactions, involves the formation of thiol esters (Dixon & Webb, 1979(b)). In order to understand the role of enzyme bound thiolesters in biochemical systems it is necessary to first understand the chemistry of simple thiolester compounds. In particular, with respect to this thesis, the vibrational spectra of model α,β-unsaturated thiol esters must be characterized before a meaningful interpretation of the vibrational spectra of an enzyme-bound α,β-unsaturated thiol derivatives can be achieved. In spite of the importance of thiolesters in biochemistry, there is a paucity of information in the literature concerning thiolesters when compared to that available for the corresponding oxygen esters. What little literature there is concerns comparing thiolesters to oxygen esters, and hence this introduction will focus on this comparative aspect. An examination of the spectroscopic differences between thiol and oxy esters will also be valuable for understanding the basis of the differing chemistry between cysteine and serine proteases.

It was determined by Barnes *et al.* (1944), and verified more comprehensively by Nyquist and Potts (1959), that in the

infrared spectra of thiolesters of the type R-C(=O)-S-R', where R and R' are saturated hydrocarbons, the carbonyl stretching frequencies occur some 30-40 cm^{-1} lower than in the spectra of the analogous oxygen ester. Collings *et al.* (1970) devised simple vibrational models of methyl acetate and methyl thioacetate, and their calculations revealed that the larger mass of the sulphur atom and the properties of the C-S bond are partially responsible for the reduction in thiolester carbonyl frequency relative to oxyesters. More recently, extensive *ab initio* SCF-MO calculations for formic and thioformic acid have been performed by Fausto (1994), and the results were used to perform a detailed charge density analysis. Upon substituting sulphur for oxygen, the total electron population on the carbonyl carbon increased, while the corresponding total electron population of the carbonyl oxygen decreased. This is consistent with sulphur, which is less electronegative than oxygen, being less able to draw charge away from the carbonyl carbon. The distribution of charges is shown in Table 3.1.

The calculations also revealed that the substitution of sulphur for oxygen leads to a substantial reduction in the total overlap electron population in the carbonyl double bond. The overlap electron population may be viewed as a measure of the strength of a given chemical bond. A greater electron overlap between two atoms corresponds to a stronger bond. Fausto (1994) calculated the overlap populations of the σ and π bonds, and came to the following important conclusions. Substitution of sulphur

Table 3.1

Distribution Of Charges in Formic and Thioformic Acid

Molecule	HCOOH			HCOSH		
Atom	C	=O	O	C	=O	S
Charge[A]	0.524	-0.479	-0.652	0.161	-0.439	0.011

[A] Charge was determined by taking the gross electron population calculated by Fausto (1994) for each atom and subtracting the atom's nuclear charge. Units are given in terms of e, where e=1.6021892 x 10^{-19} Coulombs.

for oxygen causes the total and σ C=O overlap electron population to decrease, while the π C=O overlap population increases. Since the total electron overlap population decreases overall in the carbonyl bond, it is apparent that σ effects predominate in weakening the carbonyl bond. The weakening of the carbonyl bond results in a reduction in the carbonyl bond force constant and hence a lower characteristic C=O stretching frequency for molecules containing the -C(=O)-S- linkage.

The effects on the C-X bond (where X=O or S) are opposite to that seen in the carbonyl bond. In this case replacing oxygen by sulphur causes the total and σ C-X electron overlap populations to increase, while the π C-X electron overlap population decreases. This result indicates that mesomerism in the C(=O)-X moiety is more prevalent when X is O than when X is S (see Scheme 3.4, Section 3.4.5).

Another important thiolester parameter is the basicity of the carbonyl group. The stronger basicity of any given functional group is reflected by its stronger hydrogen bonding acceptor strength. Hence a comparison of basicity between the carbonyl oxygens of thiol and oxyesters will provide a gauge of their relative hydrogen bonding strength. In solution thiol esters are less basic than oxyesters (Baker & Harris, 1960; Grunwell *et al.* 1977). Using ion cyclotron resonance techniques Grunwell *et al.* (1977) determined that in the gas phase thiolacetate and acetate have similar basicities. On the basis of CNDO/2 calculations that determined the total charge of protonated thiolacetate and

protonated acetate, Grunwell et al. (1977) concluded that in solution thiolacetate is less basic than acetate because it is less well solvated. The veracity of Grunwell's conclusion that differential solvation effects are responsible for the lower basicity of thiolesters in solution was confirmed by the FTIR studies of Smolders et al.(1988). It was determined that in CCl_4 solutions the ΔH of hydrogen bonding of methyl thiolacetate with a variety of phenols is 63 to 87% of the analogous values for methyl acetate. The $\Delta\nu$ OH between free phenol derivatives and of that complexed with methyl thiolacetate likewise is 83 to 93% of the $\Delta\nu$ OH values determined for acetate. On the other hand, the $\Delta\nu$ OH shifts for the same interactions in inert low temperature Ar matrices are the same for both methyl thiol acetate and acetate. These results are consistent with the lower basicity of thiolactetate being due to poorer solvation.

To summarize, the lower frequency of the C=O stretching mode of thiolesters compared to oxyesters is due to the predominance of σ orbital effects in weakening the C=O bond and the C=O constant. The lower basicity of the C=O group in thiol esters compared to oxy esters can be attributed to the poorer solvation of the protonated thiolester. Similar effects may result in weaker hydrogen bonds to thiol ester carbonyls in solution. Another factor that can contribute to the slightly smaller hydrogen bonding strength of thiolesters is the smaller negative charge on the carbonyl oxygen (see Table 3.1).

3.2 Materials and Methods

3.2.1

Synthesis of 5-methylthienylacryloyl ethyl thiolester.

50 mg of 5MTA-Im were added to 1 ml of ethanethiol and stirred overnight at room temperature. The next day 300 ml of 50 mM KPi, pH 7.0 were added, and the product 5MTASC$_2$H$_5$ was then extracted using CH_2Cl_2. The crude product was then HPLC purified on a silica gel supelcosil LC-SI column (Pharmacia). The column was eluted at 2.0 mls/min with a hexane-CH_2Cl_2 gradient; 0-10 min:100% hexane,10-20 min; linear gradient to 80% hexane, 20% CH_2Cl_2; 20-35 min, 80% hexane, 20% CH_2Cl_2; 35-45 minutes; linear gradient to 100% CH_2Cl_2. This profile results in the 5MTASC$_2$H$_5$ product eluting at 27 minutes. The purified product was authenticated by GC-MS. ^{13}C thiolester derivatives were synthesized from ^{13}C labelled 5MTA-Im, which in turn was synthesized as described in section 2.2.1 using ^{13}C substituted malonic acid (^{13}C; C_1,C_3 and ^{13}C; C_2) purchased from MSD Isotopes, Canada.

3.2.2

Synthesis of 5-methylthienylacryloyl ethyl oxyester

50 mg of 5MTA-Im were added to 1 ml of ethanol and stirred overnight at room temperature. The sample was then dried using a

rotary evaporator and resuspended in acetonitrile. The crude product was then HPLC purified on a Delta Pak C18 Reverse Phase column using an isocratic 90% acetonitrile/10% H_2O gradient over 60 minutes. The column was eluted at 2 mls/min, with the product eluting at 36 minutes. The purified sample was then dried with a rotary evaporator to remove acetonitrile, and the water was removed by freeze drying. The purified product was authenticated by GC-MS.

3.2.3

FTIR Instrumentation and Data Collection.

FTIR data were recorded with a FTS-40A Biorad instrument. The sample cell had either KBr or CaF_2 windows and a path length of 0.05 mm was used for all experiments. All solvents used were spectroscopic grade from BDH and were dried with molecular sieves before use. Phenol used for the hydrogen bonding study was from Aldrich. Spectral manipulations were performed with the program Spectracalc (Galactic Industries, NH).

3.2.4

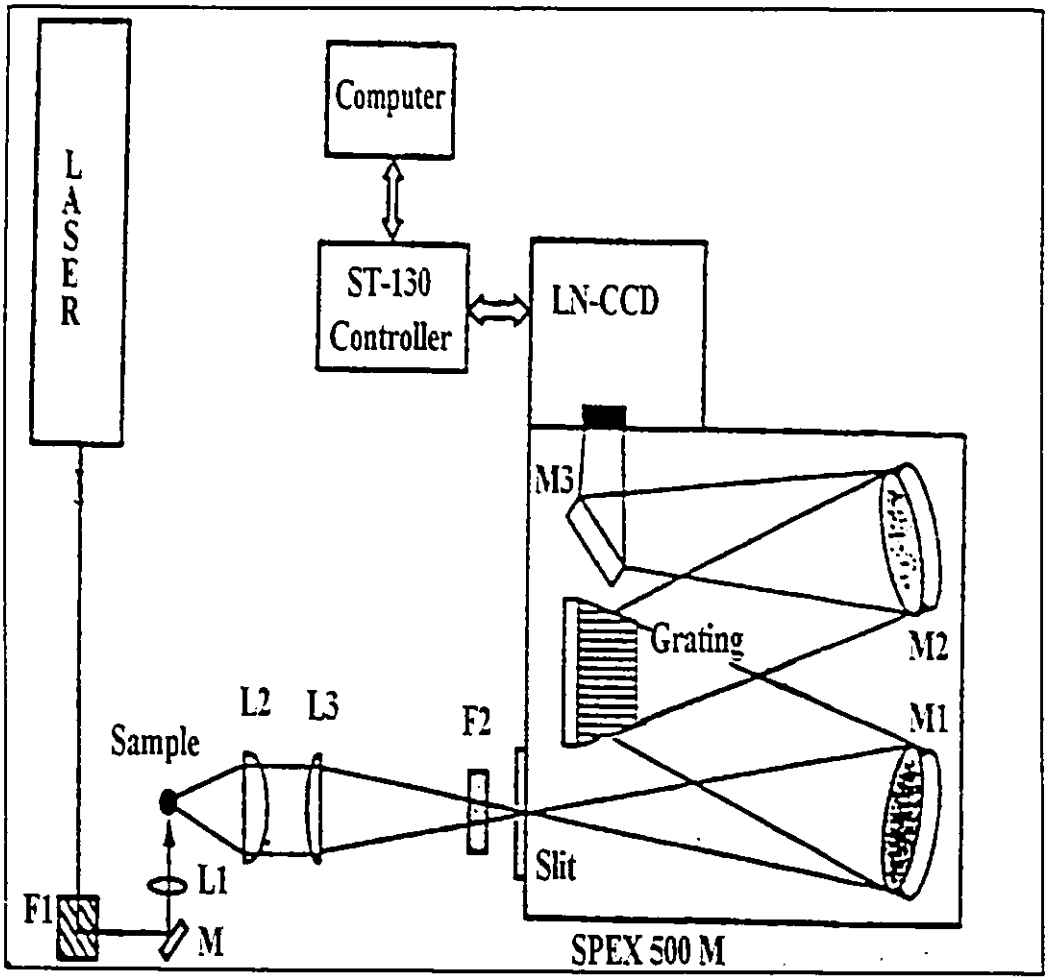
Raman Instrumentation and Data Collection

A schematic diagram of the Raman spectroscopic system employed is shown in Figure 3.1. The excitation wavelengths were either a 647.1 nm line obtained from a Coherent CR-2000K Kr^+

laser or a 488 nm line obtained from a Spectra Physics 164 argon ion laser. For 647.1 nm excitation, the laser beam was passed through a Pellin-Broca prism and a holographic laser bandpass filter (Kaiser Optical Systems) in order to separate this line from other plasma radiation. The beam was then introduced into a rectangular Suprasil cell (5x5x30 mm) containing the desired sample. The standard 90° excitation/detection geometry, resulted in the laser beam passing vertically in front of the spectrometer slit. The beam was focused prior to its entry in the sample cell with a quartz biconvex spherical lens (F2: f=100 mm). The scattered photons from the sample were then collected using two plano-convex lens (L2: 65 mm, f/1.3; and L3: 200 mm, f/4.0), and then passed through a holographic super-notch filter tuned for 647.1 nm to eliminate Raleigh photons. For experiments using the 488 nm laser line a similar light train was used. The Raman photons were then dispersed with an ion-etched holographic grating (1800 grooves/mm) within a Spex 500M single stage spectrograph. The dispersed spectrum was imaged onto a Princeton Instruments model LN-CCD detector via mirrors M1, M2, and M3. The CCD detector is operated by a ST-130 controller, and is equipped with an UV-coated EEV Model 88130 array consisting of 298(row) x 1152(column) pixels. The data were collected through the use of CSMA software from Princeton Instruments. The collected data was then exported into the program Spectracalc for spectral manipulations.

Figure 3.1

Schematic Diagram of Raman Instrumentation



Abbreviations: F1; holographic laser band pass filter and Pellin Broca prism. M, M1, M2, M3; mirrors. L1; quartz biconvex spherical lens. L2, L3; plano-convex lenses. F2; holographic super-notch filter.

3.2.5

Absorption Spectroscopy

Absorption spectroscopic data were measured using a Cary 3 UV-visible spectrophotometer.

3.3 Results

3.3.1

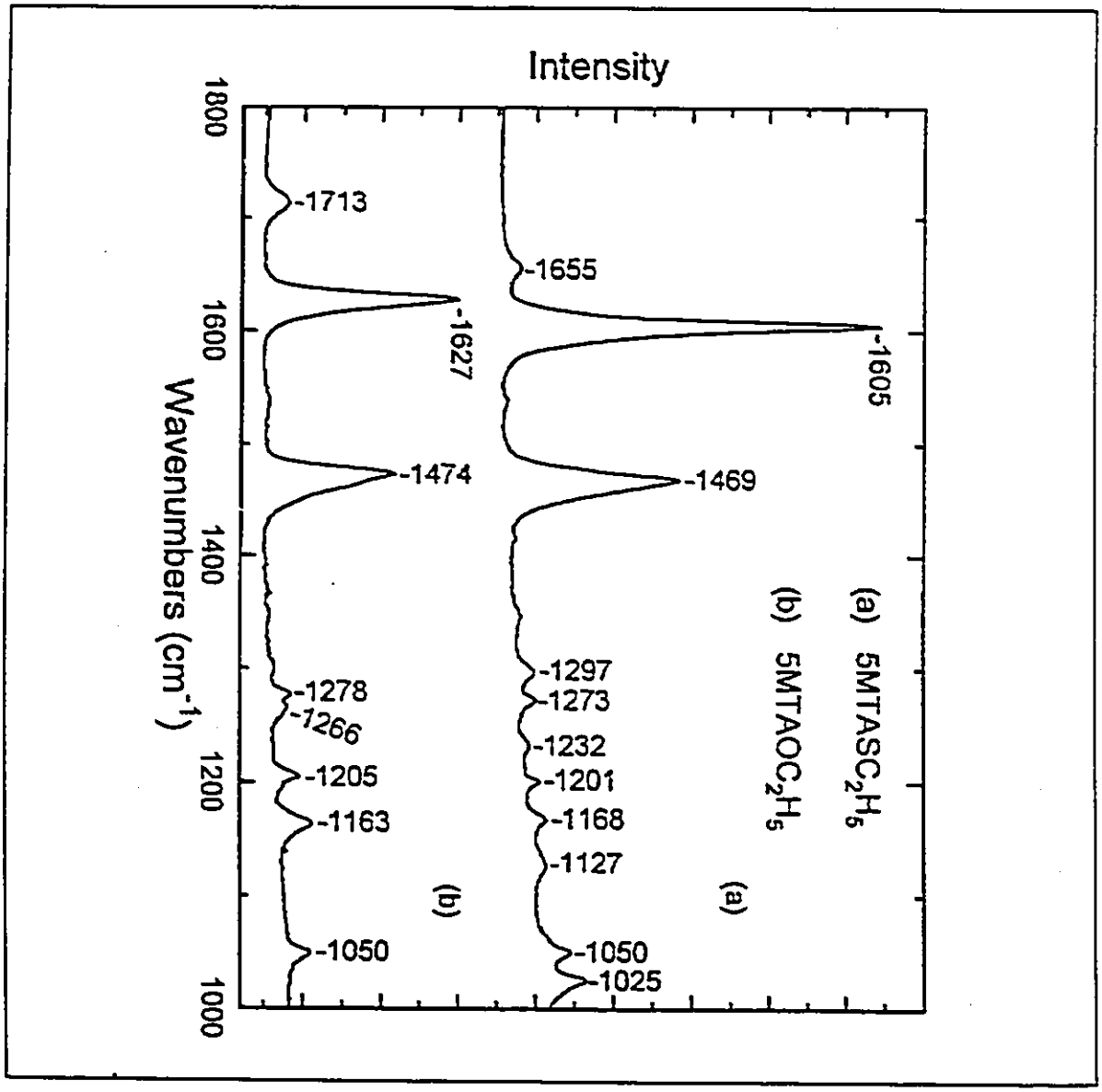
Spectroscopic Characterization of 5MTASC$_2$H and 5MTAOC$_2$H

Figure 3.2 shows the Raman spectra of 5MTASC$_2$H and 5MTAOC$_2$H in CCl_4 superimposed upon the same scale. Figures 3.3 and 3.4 are the FTIR spectra for 5MTASC$_2$H and 5MTAOC$_2$H, respectively, in CCl_4. Figures 3.5 and 3.6 are the FTIR and Raman spectra of C=O 5MTASC$_2$H$_2$ in CCl_4. All samples shown in the Figures are approximately 25 mM, and the CCl_4 solvent bands were interactively subtracted using the program Spectracalc. Experimental parameters for the Raman experiments were λ_{ex}=647 nm at 300 mW power, 8 sec. x 40 scans data acquisition time. FTIR data were collected using 64 scans.

Table 3.2 contains values of $\nu_{C=O}$, $\nu_{C=C}$, and λ_{max} for 5MTASC$_2$H$_2$ and 5MTAOC$_2$H$_2$ in solvents of varying polarity. Table 3.3 summarizes the effects ^{13}C substitution at the C_1 position (i.e. the carbonyl carbon) and the C_2 position (i.e the ethylenic carbon adjacent to the carbonyl group) have on the carbonyl and

Figure 3.2

Raman Spectra of 5MTASC$_2$H$_5$ and 5MTAOC$_2$H$_5$; 647.1 nm Excitation



ethylenic vibrational frequencies for $5MTASC_2H_5$ in CCl_4. Unpublished data by Dr. P. J. Tonge for the corresponding ^{13}C substitutions in 5MTAOMe (i.e. the methyl ester), have been included in Table 3.3. The effect of ^{13}C substitution on the *s-cis* and *s-trans* C_1-C_2 stretching modes is shown in Table 3.4.

The resonance Raman spectrum of 5MTA methyl ester has been extensively characterized by MacClement *et al.* (1981), and Tonge and Carey (1989(b)). As well, Faria *et al.* have performed detailed *ab initio* SCF-MO calculations to determine the complete vibrational frequencies of methyl *trans*-crotonate (1991(a)) and methyl *trans*-cinnamate (1991(b)). These calculations were shown to be in close agreement with both the Raman and FTIR spectra of these compounds. This background literature permits the assignment, listed in Table 3.5, of the vibrational bands of $5MTAOC_2H_5$ seen in Figure 3.4. These data, in conjunction with the vibrational assignments of ethyl thiocrotonate by Fausto (1994), may in turn be used to make the assignments, listed in Table 3.6, of the vibrational bands of $5MTASC_2H_5$ which appear in Figure 3.5. The assignments, although not specifically referred to in the discussion, provide part of the foundation required for the analysis of the properties of the oxy- and thiol-ester.

Figure 3.3

FTIR Spectrum of 5MTASC-H

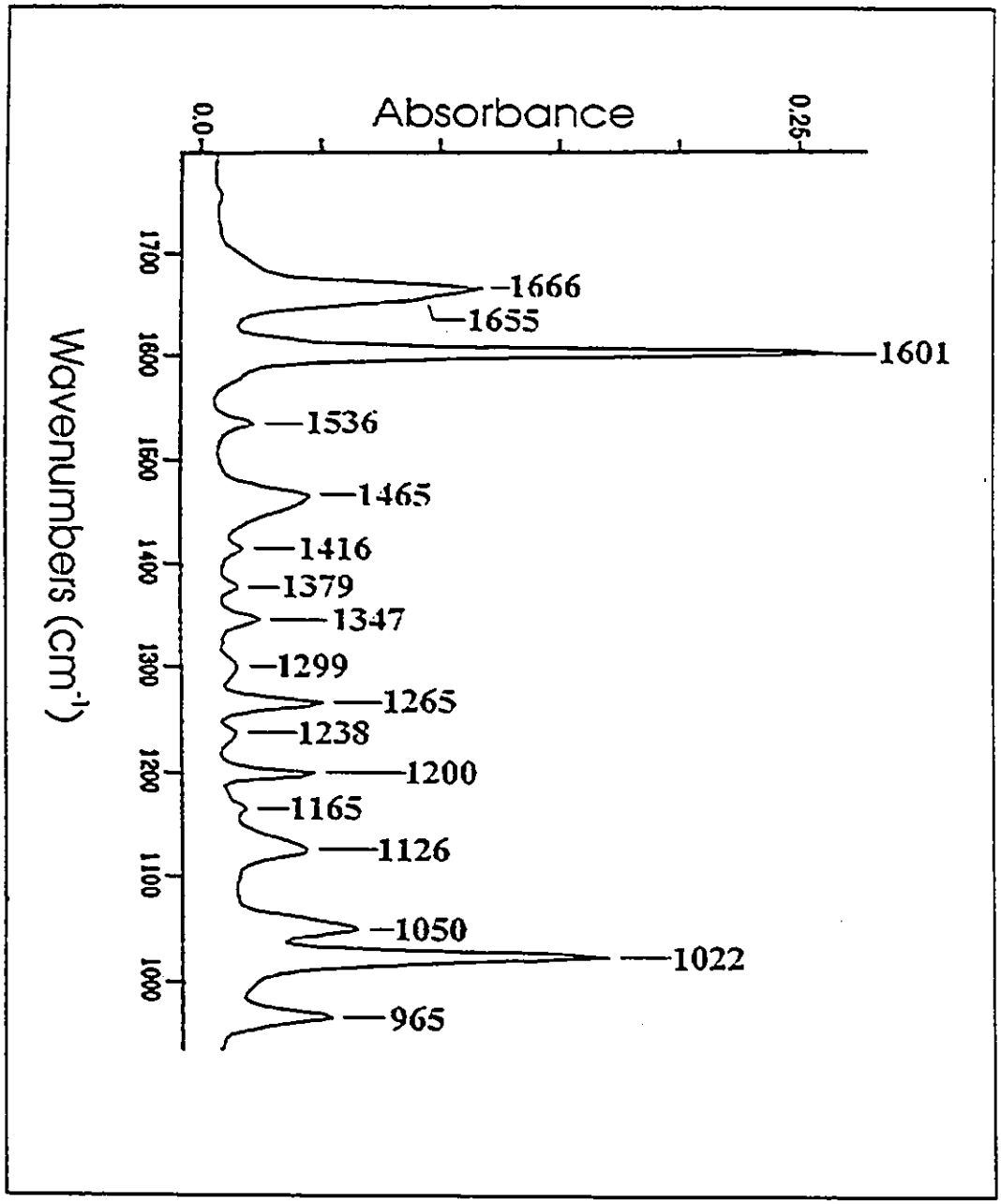


Figure 3.4

FTIR Spectrum of 5MTAOC.H.

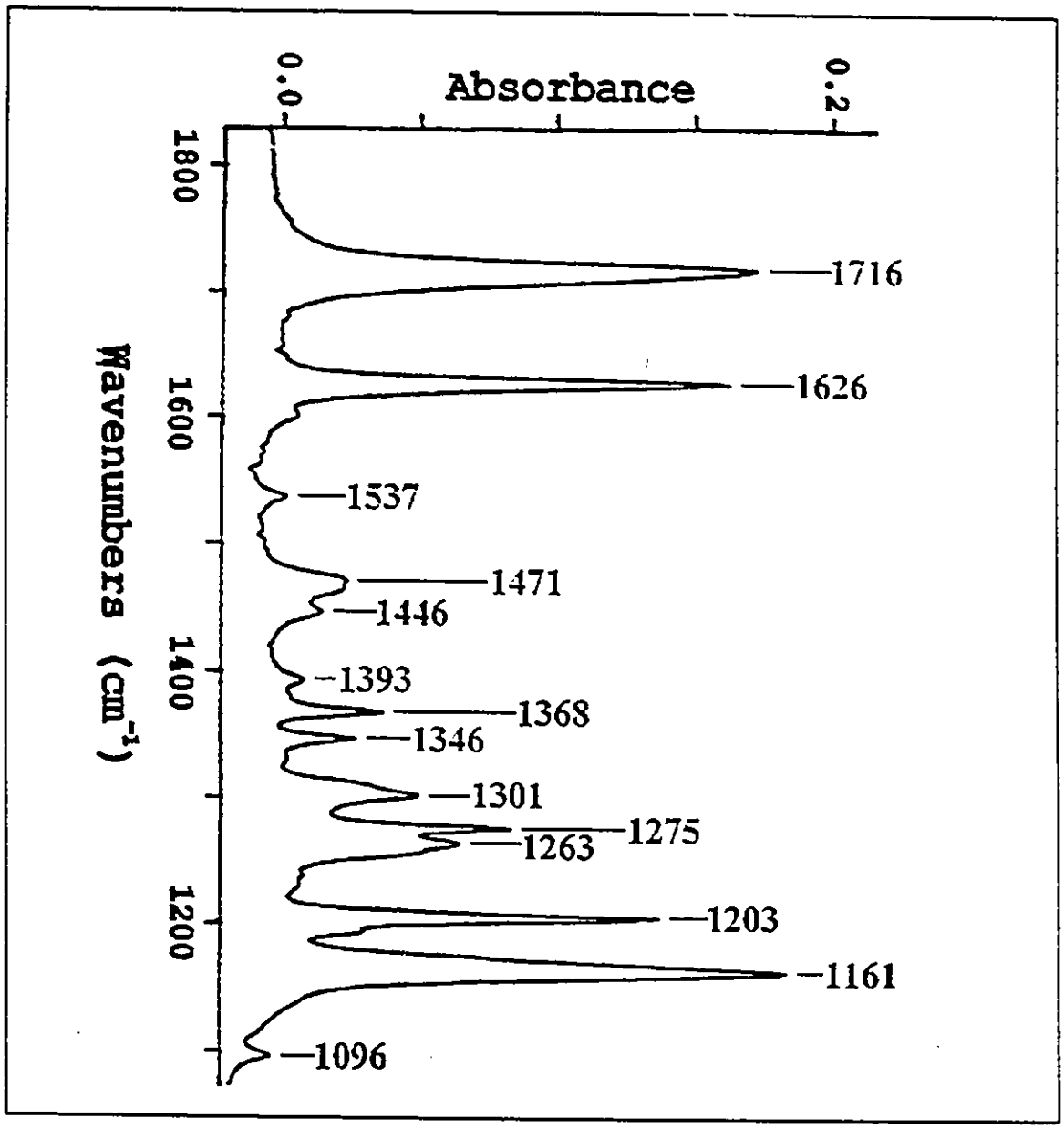
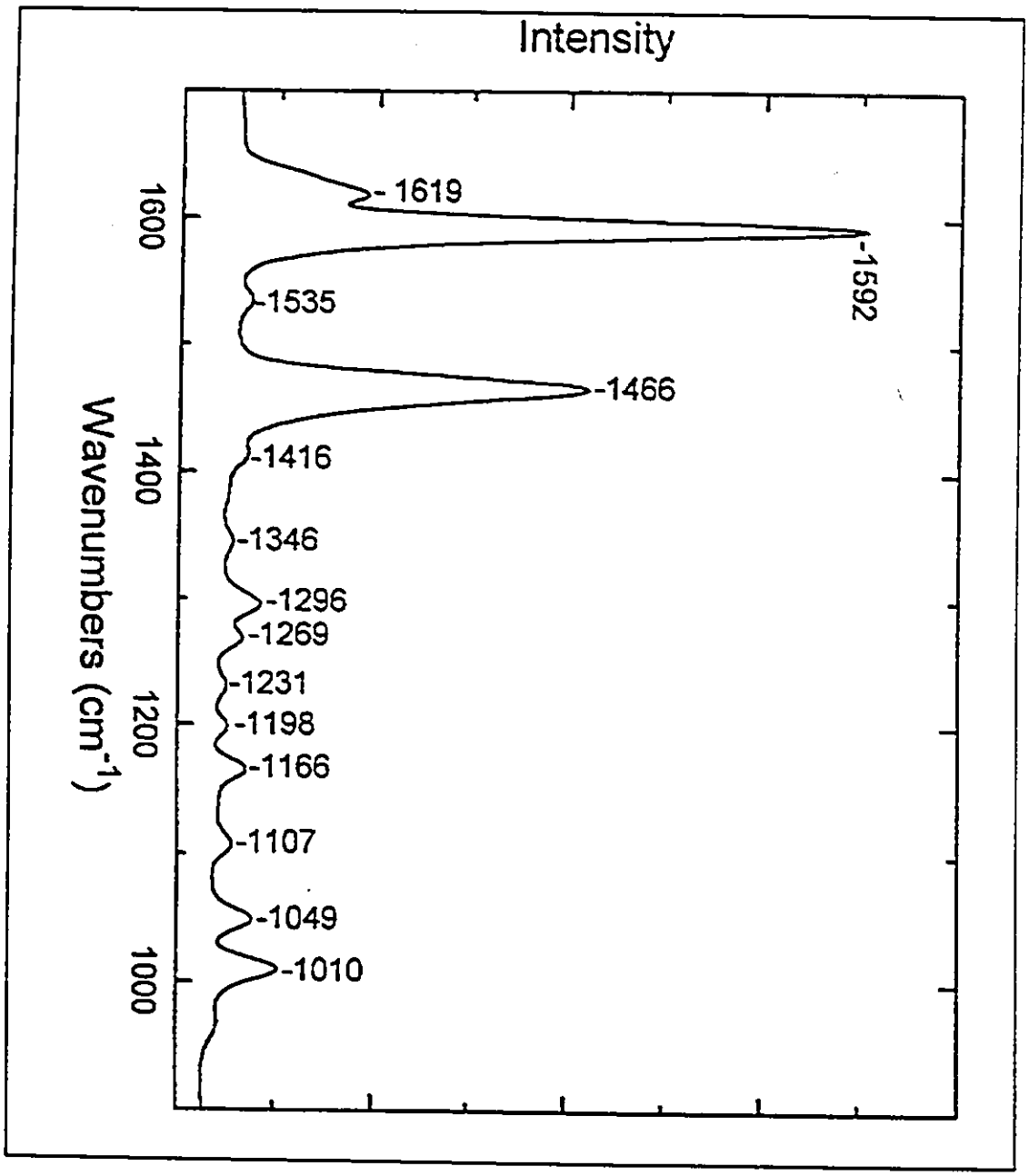


Figure 3.5

FTIR Spectrum of 5MTASC-H, Labelled at $^{13}C=O$



S = Solvent Feature

Figure 3.6

Raman Spectrum of 5MTASC-H, Labelled at $^{12}C=O$; 647.1 nm Excitation

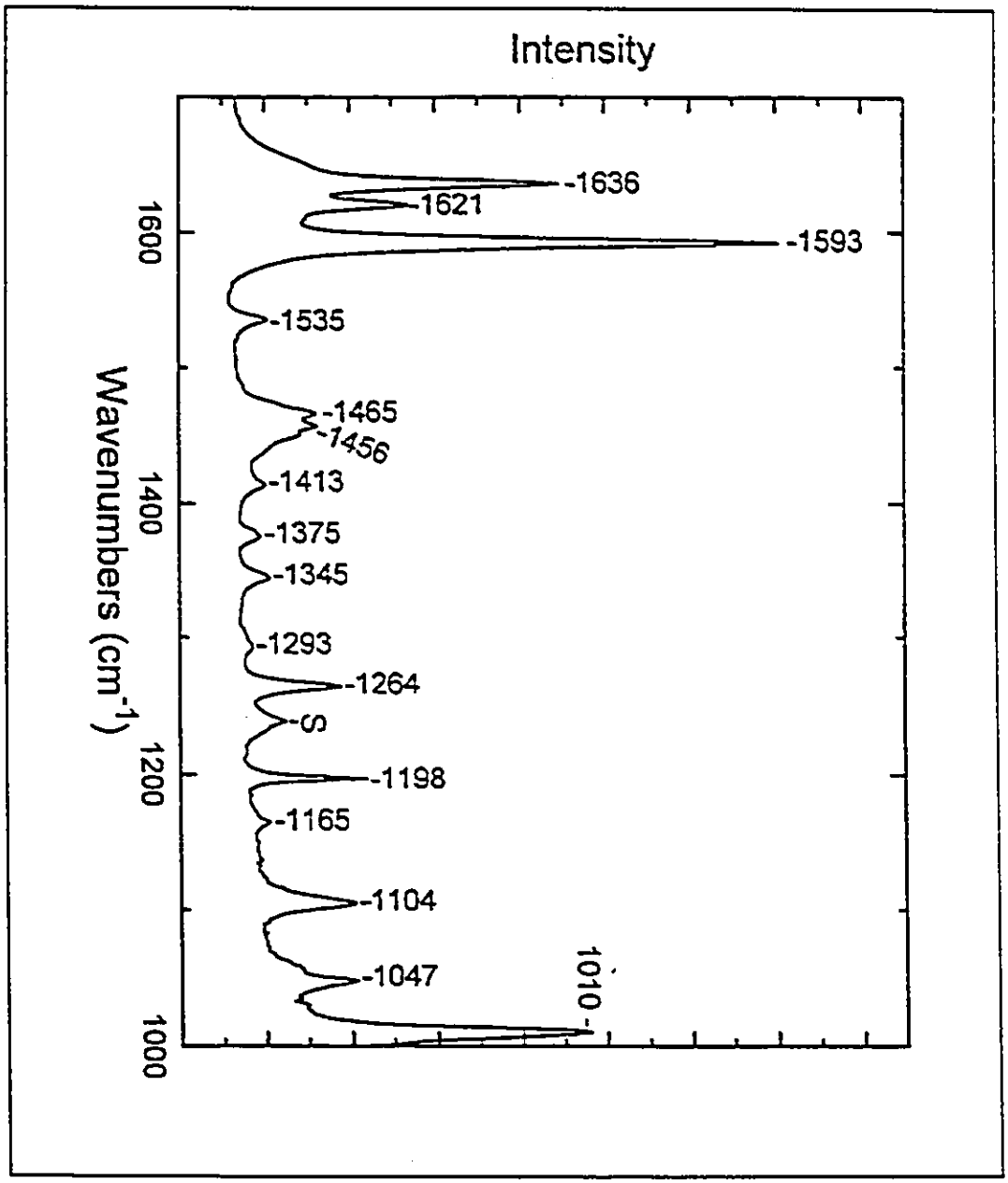


Table 3.2

Spectroscopic Data for $5MTASC_2H_5$ and $5MTAOC_2H_5$

Compound	Solvent	λ_{max} (nm)	$\nu_{C=C}$ (cm^{-1})	$\nu_{C=O}$ (cm^{-1})
$5MTASC_2H_5$	CCl_4	338	1603	1666,1658[S]
"	CH_3CN	340	1601	1661,1652[S]
"	H_2O	350	1595[R]	ND
$5MTAOC_2H_5$	CCl_4	318	1626	1716
"	CH_3CN	319	1624	1705
"	H_2O	325	1621[D]	1686[D]

[R] = Peak position determined via Raman spectroscopy using 647.1 nm excitation in 90% H_2O/10% CH_3CN. Acquisition parameters; 8 sec. x 40 scans.
[D] = FTIR peak position determined in 90% D_2O/10% CH_3CN.
[S] = Shoulder.
ND = not determined.

Unless stated otherwise, all peak position data were determined by FTIR. $5MTASC_2H_5$ and $5MTAOC_2H_5$ were prepared as 20 mM solutions in either CH_3CN, CCl_4, or D_2O. Spectra were recorded (64 scans) for both the solvent and the solvent + sample, and the solvent was subtracted from the sample spectrum using the subtraction feature in Spectracalc.

Absorption spectra were obtained by diluting the sample in the desired solvent to a final concentration of 50 μM. Prior to the data collection a baseline was recorded with the solvent alone.

Table 3.3

Effect of ^{13}C substitution on the IR Ethylenic and Carbonyl Stretching Frequencies of $5MTASC_2H_5$ and $5MTAOCH_3$

Compound (in CCl_4)	$\nu_{C=C}$ (cm^{-1})	$\nu_{C=^{13}C}$ (cm^{-1})	$\nu_{C=O}$ (cm^{-1})	$\nu_{^{13}C=O}$ (cm^{-1})	$\nu_{C=C}$ from $\nu_{^{13}C=O}$ data (cm^{-1})
$5MTASC_2H_5$	1603	1585	1666,1658[S]	1636,1621	1593 (Δ=10)[A]
$5MTAOCH_3$	1627	1604	1722,1707[S]	1676	1624 (Δ=3)[A]

$5MTASC_2H_5$ and its ^{13}C labelled derivatives at the C_1 and the C_2 positions were prepared as CCl_4 solutions with a final concentration of approximately 20 mM. FTIR Spectra were recorded (64 scans) for both CCl_4 and for CCl_4 + sample. The CCl_4 spectrum was then subtracted from the sample spectrum using the subtraction feature in the Spectracalc program. The $5MTAOCH_3$ unpublished data from Dr. P. J. Tonge were obtained under similar conditions.

[A] = The Δ value represents the magnitude of the change in wavenumber for $\nu_{C=C}$ between unlabelled $5MTASC_2H_5$ and $5MTASC_2H_5$ labelled at the carbonyl carbon.

[S] = Shoulder

Table 3.4

Effect of ^{13}C substitution on the *s-cis* and *s-trans* $\nu_{C_1-C_2}$ Stretching Frequencies of 5MTASC-H_5

Compound	$\nu_{C_1-C_2}$ (*s-cis*) (cm^{-1})	$\nu_{C_1-C_2}$ (*s-trans*) (cm^{-1})
Unlabelled 5MTASC-H_5	1126	1022
^{13}C=O 5MTASC-H_5	1104	1010
C=^{13}C 5MTASC-H_5	1115	1017

Data are from FTIR spectra obtained according to the procedure described in the legend for Table 3.2.

Table 3.5

Assignment of Vibrational Frequencies in 5MTAOC$_4$H$_3$

Vibrational Frequency (cm^{-1})	Assignment
1716	$\nu_{C=O}$
1626	$\nu_{C=C}$
1537	thiophene ring mode
1471	thiophene ring mode
1446	δ C-H from $-CH_2CH_3$ (*s-cis*)
1393	thiophene ring mode
1368	δ C-H from $-CH_2CH_3$ (*s-trans*)
1346	thiophene ring mode
1301	δ C_β-H (*s-cis*)
1275	δ C_β-H (*s-trans*)
1263	ν C-O (*s-trans*)
1203	thiophene ring mode
1160	ν C-O (*s-trans*)

ν = bond stretching mode.
δ = bond bending mode.

Table 3.6

Assignment of Vibrational Frequencies of 5MTASC$_5$H$_5$

Vibrational Frequency (cm^{-1})	Assignment
1666	$\nu_{C=C}$ (*s-cis*)
1655	$\nu_{C=C}$ (*s-trans*)
1601	ν_{C-C}
1537	Thiophene ring mode
1465	Thiophene ring mode
1347	Thiophene ring mode
1265	δ C_β-H (*s-cis*)
1200	Thiophene ring mode
1165	Thiophene ring mode
1126	$\nu_{C_1-C_2}$ (*s-trans*)
1050	Thiophene ring mode
1022	$\nu_{C_1-C_2}$ (*s-cis*)

ν = bond stretching mode.
δ = bond bending mode.

3.3.2

Determination of the Decrease in $\nu_{C=O}$ of $5MTASC_2H_5$ upon Hydrogen Bond Formation.

Thijs & Zeegers-Huyskens (1984) have shown for saturated oxy esters and ketones that a relationship exists between the enthalpy of hydrogen bonding, ΔH involving a C=O group, and the shift in the C=O's stretching frequency, $\Delta\nu_{C=O}$. When the enthalpy of hydrogen bond formation decreases, (i.e. the strength of hydrogen bond increases), $\Delta\nu_{C=O}$ decreases. The magnitude of $\Delta\nu_{C=O}$ upon hydrogen bond formation may be used as a measurement of the strength of hydrogen bonding. The larger the value of $\Delta\ \nu_{C=O}$ is, the larger the value of $-\Delta H$ and the stronger the hydrogen bond is. Table 3.7 lists for $5MTASC_2H_5$ the frequency values of free $\nu_{C=O}$ in CCl_4, and of $\nu_{C=O}$ when it is hydrogen bonded to phenol. For comparison similar values are included for $5MTAOCH_3$.

Table 3.7

$\nu_{C=O}$ Values for "Free" (Non-Hydrogen Bonded), and Hydrogen bonded Carbonyls for $5MTASC_2H_5$ and $5MTAOCH_3$.

Compound	Free $\nu_{C=O}$ (cm^{-1})	$\nu_{C=O}$ + phenol (cm^{-1})	$\Delta\nu_{C=O}$ (cm^{-1})
$5MTASC_2H_5$	1666	1645	21
$5MTAOCH_3$[A]	1722	1697	25

The FTIR spectrum of a 20 mM solution of $5MTASC_2H_5$ in CCl_4 was recorded (64 scans). A second spectrum was obtained under the same experimental conditions except that the solution also contained 10 mM phenol. The phenol concentration is less than that of the thiol ester to ensure that the carbonyl groups only had a single hydrogen bond donor.The decrease in carbonyl frequency upon hydrogen bond formation was determined by interactively subtracting an FTIR spectrum of free acceptor (i.e. thiolester) from a solution of acceptor and donor (i.e. thiolester and phenol) using the program Spectracalc.

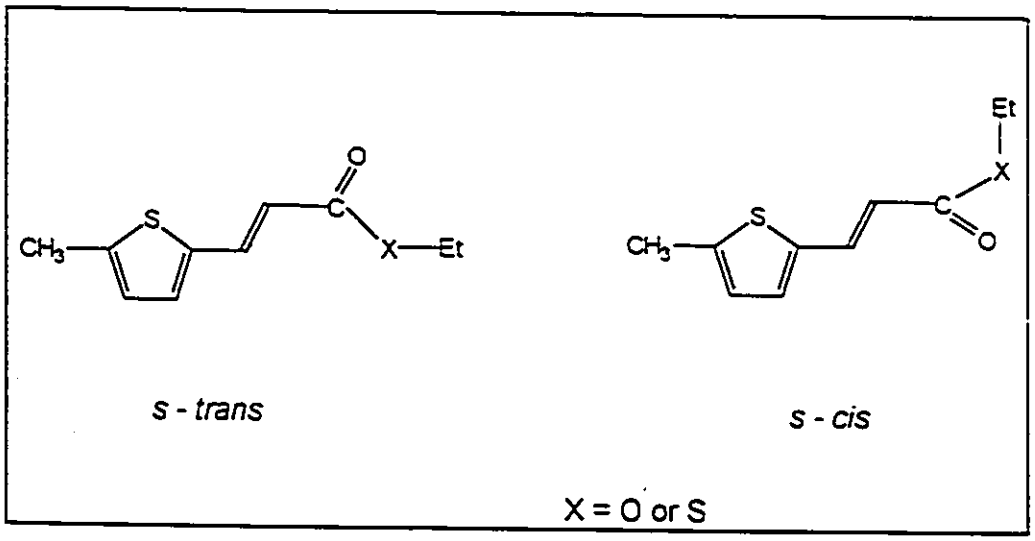
[A] = Unpublished Data from Peter J. Tonge. These experiments were done under similar conditions as described above.

3.4 Discussion

3.4.1 Evidence for *s*-*cis*/*s*-*trans* Conformational Heterogeneity in the 5MTASCH Molecule

The C_1-C_2 bond in molecules containing an α,β-unsaturated carbonyl moiety has partial double bond character due to π electron delocalization between the ethylenic bond and the carbonyl group. This causes restricted rotational freedom about this bond and gives rise to *s*-*cis* and *s*-*trans* conformers as shown in scheme 3.2.

Scheme 3.2



The carbonyl group experiences different environments in the two conformations, leading to an observed carbonyl vibrational

profile having contributions from both conformers. Fausto *et al.* (1994) have determined for ethyl thiocrotonate that the *s-cis* conformer is more stable than the *s-trans* conformer by *ca* 7 kJ/mol for the isolated molecule, and that the *s-cis* carbonyl absorbs at *ca* 8 cm^{-1} higher frequency. Hence, it is hypothesized that the carbonyl profile in the FTIR spectrum of $5MTASC_2H_5$ (Figure 3.3) contains contributions from both *s-cis* (1666 cm^{-1}) and *s-trans* (1655 cm^{-1} shoulder) conformers. The *s-cis* $\nu_{C=O}$ band predominates since the *s-cis* conformer is the more stable of the two. These two conformers are more discernible in the FTIR spectrum of $5MTASC_2H_5$ with ^{13}C substitution in the carbonyl, with the more intense *s-cis* band and the less intense *s-trans* band moving down to 1636 cm^{-1} and 1621 cm^{-1}, respectively (Figure 3.4).

The $5MTAOC_2H_5$ molecule is also capable of *s-cis/s-trans* isomerism, however it is not possible to differentiate the bands in the carbonyl profile of the FTIR spectrum (Figure 3.4), presumably due to band overlap. However, the $5MTAOCH_3$ FTIR data (Table 3.2) show two peaks ascribable to *s-cis/s-trans* isomerism. Faria *et al.* have performed detailed *ab initio* SCF-MO calculations to theoretically determine the complete vibrational frequencies of methyl *trans*-crotonate (1991(a)), methyl *trans*-cinnamate (1991(b)), and methyl acrylate (1991 (c)). As is the case with thiocrotonate the *s-cis* conformer was calculated to be the more stable conformer. With the oxygen esters, however, it is the *s-trans* $\nu_{C=O}$ which absorbs infrared radiation at the higher frequency. Hence, Tonge *et al.*(1996) assign the 1722 cm^{-1} band

(Table 3.3) to the *s-trans* conformer and the shoulder band to the *s-cis* conformer.

Fausto *et al.* (1994) also determined that the 1165 cm^{-1} and 1035 cm^{-1} IR bands for ethyl thiocrotonate are predominantly comprised of C_1-C_2 stretch of the *s-trans* and *s-cis* forms, respectively, and concluded that these features would serve as useful marker bands for the presence of these two conformers. It is therefore proposed that the 1127 and 1025 cm^{-1} bands in the FTIR spectrum of 5MTASC$_2$H$_5$ (Figure 3.2) contain a significant degree of C_1-C_2 stretch from the *s-trans* and *s-cis* forms, respectively. The effect of ^{13}C substitution shown in Table 3.4 provides evidence for the validity of these assignments. ^{13}C substitution in the carbonyl (i.e. at the C_1 carbon) and in the ethylenic bond at the C_2 carbon both cause significant shifts in the frequency of the 1126 and 1022 cm^{-1} bands. These shifts are expected since both substitutions directly effect the total mass of C_1-C_2 bonding system, and hence will perturb the C_1-C_2 stretching frequency.

3.4.2

Comparison of the Carbonyl Profiles of 5MTASC$_2$H$_5$ by FTIR and Raman Spectroscopy

As was discussed in section 1.2 of Chapter 1, IR and Raman spectroscopies are based upon different physical processes, leading to different selection rules and hence to the possibility

of different intensities for any given vibrational mode. A key difference between the FTIR and Raman spectra of $5MTASC_2H_5$ is that the carbonyl profile in the FTIR spectrum is a mixture of at least two bands, an intense 1666 cm^{-1} band and a 1655 cm^{-1} shoulder (Figure 3.3), while the $5MTASC_2H_5$ carbonyl band in the Raman spectrum at 1655 cm^{-1} is relatively weak and its profile cannot be simply analyzed by inspection (Figure 3.2). An explanation for the observed spectra is that both the *s-cis* and *s-trans* carbonyls are infrared active, giving frequencies of 1666 and 1655 cm^{-1}, respectively. In the Raman spectrum on the other hand the *s-cis* mode has low intensity due to the lesser polarizability of the C=O bond at the terminus of the *s-cis* π-electron chain, resulting in a predominantly *s-trans* carbonyl band at 1655 cm^{-1}. The low intensity is due to the smaller population of *s-trans* conformers present (as demonstrated in the FTIR data), and to the low scattering cross section associated with the carbonyl.

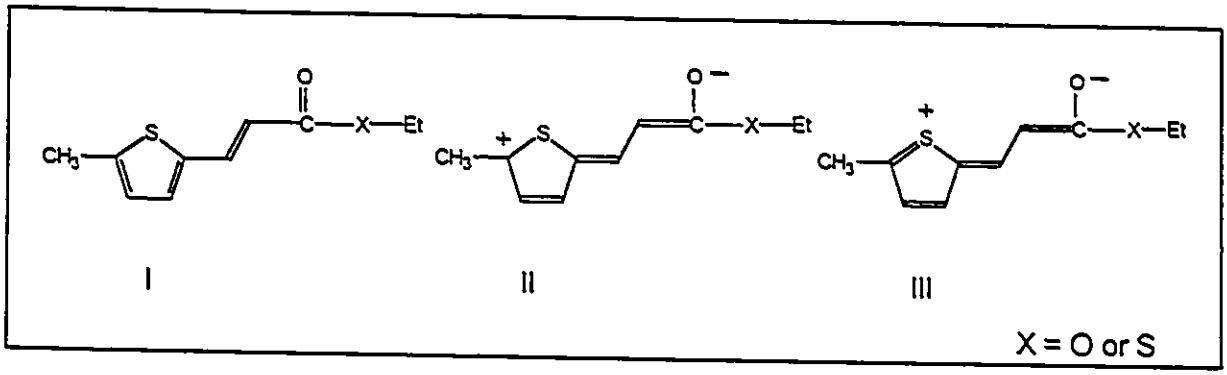
The above interpretations are consistent with what is observed in the FTIR and Raman spectra of $5MTASC_2H_5$ labelled at the carbonyl carbon (Figures 3.5 and 3.6, respectively), where the more intense *s-cis* carbonyl band at 1636 cm^{-1} in the FTIR spectrum is absent in the Raman spectrum, and only the less intense *s-trans* band is visible in both spectra in the vicinity of 1620 cm^{-1}.

3.4.3

Correlation of the Spectroscopic and Structural Properties of $5MTASC_2H_5$ and $5MTAOC_2H_5$

Both $5MTASC_2H_5$ and $5MTAOC_2H_5$ have extended π conjugation systems. Scheme 3.3 illustrates possible canonical structures for these two compounds.

Scheme 3.3



The spectroscopic data in Table 3.1 may be rationalized through the use of the structures in Scheme 3.3. In their absorption spectra, both $5MTASC_2H_5$ and $5MTAOC_2H_5$ show a marked red shift in λ_{max}, of 12 and 7 nm respectively, in going from the non-polar CCl_4 solvent to polar H_2O. In general, for molecules that have excited state structures that are more polar than those that are found in the ground state, an increase in solvent polarity will

stabilize the excited state more than the ground state, leading to a decrease in the energy required for the electronic transition and hence a red shift in λ_{max} (Barltrop and Coyle,(1975)). A red shift in λ_{max} with increasing solvent polarity is also characteristic of π-π^* transitions, which in general have more charged excited states than ground states. The absorption data therefore indicate a π-π^* transition. Compared to the ground state, it is likely that the excited state has a more significant contribution from structures resembling canonical forms II and III in Scheme 3.3, giving the excited state molecule a more dipolar character.

5MTASC_2H_5 is more red shifted than 5MTAOC_2H_5 (338 nm vs 318 nm in CCl_4), and it also shows a larger red shift when the solvent system is changed from CCl_4 to H_2O (12 nm vs 7 nm). This indicates that the 5MTASC_2H_5 molecule in the excited state has more charged character, and hence a greater contribution of structures resembling canonical forms II and III, than the analogous 5MTAOC_2H_5 excited state. The energy required for an electronic transition in absorption spectroscopy is a function of both ground and excited states. However, as discussed in Section 1.4 of Chapter 1, peak positions obtained from vibrational spectroscopy are solely a property of the ground state. Hence vibrational peak positions are useful for monitoring ground state effects. There are a number of studies emphasizing the value of using ν_{C-C} as a probe of electron delocalization/ polarization in the ground state (Carey & Storer, 1984; Storer *et al.* 1981). The

Raman spectra of $5MTASC_2H_5$ and $5MTAOC_2H_5$ (Figure 3.2) show a substantial reduction of 22 cm^{-1} in $\nu_{C=C}$, and a small reduction of 5 cm^{-1} in the thiophene ring mode of $5MTASC_2H_5$ compared to $5MTAOC_2H_5$. This indicates that for $5MTASC_2H_5$ there is increased electron polarization in the ground state relative to $5MTAOC_2H_5$. Thus, canonical structures II and III in Scheme 3.3 are making a significantly larger contribution to the overall structure of $5MTASC_2H_5$ in the ground state. Further evidence to support this conclusion is the fact that there is a larger reduction in the C=C frequency on going from CCl_4 to H_2O for $5MTASC_2H_5$ (8 cm^{-1}) than there is for $5MTAOC_2H_5$ (5 cm^{-1}, see Table 3.1). This indicates that $5MTASC_2H_5$ is experiencing greater stabilization of the more charged canonical structures II and III on going to the polar solvent H_2O relative to $5MTAOC_2H_5$.

Thus, the data in Table 3.1 demonstrate that the $5MTASC_2H_5$ molecule is more polarizable than $5MTAOC_2H_5$ in both the ground and excited states; and thus in the ground state in the presence of polarizing forces, has a greater dipole moment by virtue of the greater separation of charges on the molecule.

3.4.4

Evidence for Vibrational Coupling Between $\nu_{C=O}$ and $\nu_{C=C}$ of $5MTASC_2H_5$

Functional groups such as C=C and C=O are sometimes treated as independent oscillators in vibrational spectroscopy (see Section 1.5 of Chapter 1). However, this approximation can break

down if there are molecular vibrational frequencies that approach the C=C and/or C=O frequencies, which cause vibrational coupling. The vibrational frequency of a thiolester carbonyl occurs at a relatively low frequency due to the lower force constant of $\nu_{C=O}$ in $5MTASC_2H_5$ (see section 3.1). This provides the opportunity for $\nu_{C=O}$ to approach the frequency value of $\nu_{C=C}$, resulting in the potential for the two modes to couple with one another. Thus, for the thiolester $\nu_{C=O}$ and $\nu_{C=C}$ are closer in frequency (63 cm^{-1}, Table 3.2), than that found in $5MTAOC_2H_5$ (95 cm^{-1}, Table 3.2). This gives rise to greater vibrational coupling in $5MTASC_2H_5$. The ^{13}C substitution data in Table 3.2 show that there is a relatively large decrease of 10 cm^{-1} in the ethylenic frequency when ^{13}C is substituted at the carbonyl carbon of $5MTASC_2H_5$. This is strong evidence for vibrational coupling between the two vibrational modes, since if they were not coupled very little change in $\nu_{C=C}$ would be expected. The oxygen ester on the other hand shows only a small decrease of 3 cm^{-1} in $\nu_{C=C}$ with ^{13}C substitution at the carbonyl carbon, indicating only a small degree of coupling.

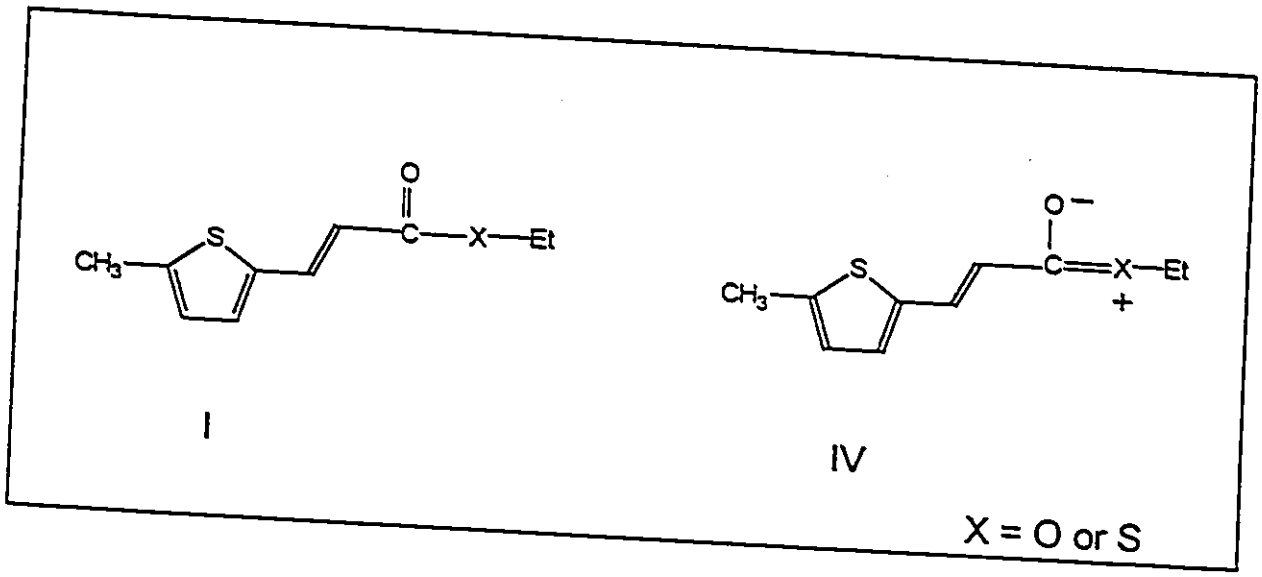
3.4.5

A Tentative Hypothesis to Explain the Greater Polarization of $5MTASC_2H_5$

As was discussed in Section 3.4.3, α,β-unsaturated thiolesters such as $5MTASC_2H_5$ have a more polarized π-electron system relative to the corresponding oxygen ester. This can be

explained at least in part by a more significant resonance interaction between the carbonyl bond and the ester oxygen than occurs for the thiol ester. Scheme 3.4 illustrates the relevant canonical structures.

Scheme 3.4



The *ab initio* calculations by Fausto (1994) demonstrated that canonical structure IV makes a more significant contribution for oxygen esters than for thiolesters. This result is due to the efficient $2p_z$-$2p_z$ bond overlap that occurs in the -C(=O)-O- moiety in oxygen esters. The $2p_z$-$3p_z$ overlap between the carbonyl carbon and the sulphur in thiol esters is less efficient and hence canonical structure IV is less significant in thiolesters.

The greater polarization of α,β-unsaturated thiolesters

relative to their corresponding oxygen esters could be due in part to canonical structure IV making a more significant contribution to the overall structure in oxygen esters. The greater significance of structure IV means that the ester oxygen can be viewed as a "competing source" to donate electrons to the carbonyl oxygen, competing with the π electrons present in the conjugated system of the thiophene ring and ethylenic bond. Therefore, the resonance interaction with the ester oxygen will reduce the significance of canonical structures II and III (Scheme 3.3) in α,β-unsaturated oxyesters, leading to a reduction in the overall polarization in the molecule.

3.4.6

Comparison of the Hydrogen Bonding Capacities of 5MTASC$_2$H$_5$ and 5MTAOC$_2$H$_5$

Quantitative data on the relationship between the $\nu_{C=O}$ shift and the enthalpy of hydrogen bonding are not yet available for α,β-unsaturated thiolesters; however, the shifts in $\Delta\nu_{C=O}$ may be used as a semi-quantitative estimate of the enthalpy of hydrogen bonding. Table 3.5 shows that for 5MTASC$_2$H$_5$, $\Delta\nu_{C=O}$, upon hydrogen bond formation with phenol is 84% of the corresponding value for 5MTAOCH$_3$. Data from Smolders *et al.*(1988) show that methyl thiolacetate (MTA) complexing with phenol in CCl_4, the $\Delta\nu_{OH}$ for phenol is smaller (83-93%) and -ΔH is less (67-87%) relative to methyl acetate (MA). It was also shown that relative decreases in

$\Delta\nu_{CH}$ between MTA and MA are reflected in a smaller corresponding relative decrease in $\Delta\nu_{C=O}$. For example, using 3,4,5 trichlorophenol as the hydrogen bond donor, Smolders *et al.* (1988) determined that both $\Delta\nu_{C=O}$ and $\Delta\nu_{CH}$ for MTA were 93% of that obtained for MA, and are accompanied by a decrease of 63% in the $-\Delta H$ of hydrogen bond formation for MTA. It is therefore assumed that the slightly smaller decreases in $\Delta\nu_{C=O}$ for a thiolester compared to its corresponding oxygen analogue are accompanied by smaller decreases in $-\Delta H$. This leads to the conclusion that the hydrogen bonding strength of $5MTASC_2H_5$ is about 80% as strong as that found in $5MTAOCH_3$.

3.5 Summary

The $5MTASC_2H_5$ molecule is more polarizable than the corresponding $5MTAOC_2H_5$ molecule since upon going to a more polar solvent it shows a greater red shift in the absorption spectrum and a greater reduction in its $\nu_{C=C}$ frequency position in the Raman spectrum. The carbonyl stretching frequency of $5MTASC_2H_5$ is coupled to a larger extent with the ethylenic stretching frequency than is the case for $5MTAOC_2H_5$, primarily due to the intrinsically lower carbonyl frequency exhibited by thiolesters compared to oxyesters. The hydrogen bonding capacity of the carbonyl oxygen in $5MTASC_2H_5$ is significant, but is somewhat smaller (approximately 80%) than that of $5MTAOC_2H_5$.

For $5MTASC_2H_5$, the presence of two features in the carbonyl

stretching region of the FTIR spectrum at 1666 and 1655 cm^{-1} is ascribed to the presence of *s-cis* and *s-trans* conformers. Similarly, the presence of bands near 1126 and 1022 cm^{-1} in the FTIR and Raman spectra is ascribed to the presence of C_1-C_2 stretching features for the *s-cis* and *s-trans* isomers, respectively. It is proposed that the 1126 and 1022 cm^{-1} bands may be used as marker bands for the degree of conformational *s-cis*/*s-trans* heterogeneity, and these bands will be used to detect *s-cis* and *s-trans* isomers in the Raman spectra of acyl cysteine proteases in Chapter 4.

CHAPTER 4

ABSORPTION AND RAMAN SPECTROSCOPIC STUDIES OF ACYL CYSTEINE PROTEASES

4.1 Introduction

Structure-reactivity relationships have been established for acyl serine proteases by using a combination of resonance Raman spectroscopy and kinetics (Tonge & Carey, 1992; Carey & Tonge, 1995). Based on the resonance Raman spectra of a series of acyl chymotrypsins and acyl subtilisins, a strong correlation was found between the carbonyl stretching frequency, $\nu_{C=O}$, and log k_3, where k_3 is the deacylation rate constant (see Figure 1.5 in Chapter 1).

Extension of structure-reactivity relationships to acyl cysteine proteases is of interest because although there are mechanistic similarities, significant known differences exist between the properties of acyl serine and acyl cysteine proteases. The relative importance of the oxyanion hole appears to be different (see Section 1.8) and there are intrinsic chemical differences between thiol esters and oxy esters (see Chapter 3). One of the significant differences not yet discussed is that, with respect to acyl enzymes derived from α,β-unsaturated substrates, it appears that there are strong electrostatic polarizing forces in the active sites of acyl cysteine proteases which appear to be absent in acyl serine proteases (Carey, 1982). Using acyl papains derived from furylacryloyl, 4-dimethylamino-3-nitro-cinnamoyl, and 4-dimethylamino-3-nitro-(α-benzamido)cinnamoyl substrates, Carey *et al.* (1976, 1978) demonstrated that binding of these substrates to

papain resulted in dramatic differences in the spectroscopic properties compared to those of either the substrate or product. For each of these substrates, upon forming an acyl papain, a large red shift occurs in the absorption spectrum, accompanied by the appearance of an intense ethylenic bond frequency at an unusually low frequency. It was proposed that these results could be due to the papain active site setting up a highly polarized π-electron system within the substrate's acyl group.

Acyl serine proteases derived from α,β-unsaturated substrates, however, do not show as large a red shift in the absorption spectrum or as low a $\nu_{C=C}$ frequency in the resonance Raman spectrum (MacClement *et al.* 1981, Phelps *et al.* 1981). It was also shown by Tonge & Carey (1989(a)) that acyl thiolsubtilisins (i.e. subtilisin that has had its active site serine chemically converted to a cysteine) derived from α,β-unsaturated substrates do not in general give rise to highly polarized π-electron systems as is seen with papain. This result suggests that the source of polarization for members in the papain super family is not simply due to the acyl linkage being a sulphur atom instead of oxygen.

It has been proposed that the principal cause of polarization in papain's active site is due to the α-helix dipole present in the papain active site (Carey, 1982, Tonge & Carey, 1989(a)). The active site Cys25 is the terminal residue of papain's α-helix, hence acyl groups bound at Cys25 are capable of interacting with this helix.

α-Helical segments in proteins possess a substantial dipole moment due to the electrostatic field that results from the dipole moments of the individual peptide groups. It has been proposed that the α-helix dipole has important structural and catalytic roles, although the latter have not been experimentally characterized (Wada, 1976, Hol *et al.* 1978). Independently, Nicholson *et al.* (1988) and Sali *et al.* (1988) demonstrated that interactions between α-helix dipoles and charged side chains can increase protein stability. For example, in the small ribonuclease barnase the interaction between the protonated form of His-18 and an α-helix stabilizes the native structure by 2.1 kcal/mol at neutral pH (Sali *et al.* 1988). Aqvist *et al.* (1991) have pointed out the importance of using a correct value for the effective dielectric constant when calculating the interactions of an α-helix dipole. The conclusions of the latter authors are in agreement with the findings of Nicholson *et al.* (1991) that the α-helix dipole is best considered not as a giant macrodipole but rather as a manifestation of the concentration of local unsatisfied charges on the amides and carbonyls within the first and last turns of the helix. These findings are also in accord with the Hartree-Fock SCF calculations on the active site of papain by Rullman *et al.* (1989). It was found that the protein matrix stabilizes the zwitterionic form of the active site Cys(25)-His(159) pair, with the most important contribution to the stabilization coming from the α-helix and more than half of this effect coming from the backbone atoms of Cys25 itself.

Although the structural importance of α-helix dipoles has been established by studies of the kind outlined above there appears to be no experimental evidence relating changes in enzymatic reactivity to the effects of an α-helix dipole.

In order to investigate how α-helix-substrate interactions might affect structure-reactivity relationships for the papain superfamily the results of the kinetic characterization from Chapter 2 have been combined with a Raman and electronic absorption analysis in this chapter. Initially, resonance Raman studies of acyl cysteine proteases were attempted, however, due to the highly photolabile nature of the thiolester intermediates (Tonge & Carey, 1989(a)), it was not possible to interpret the data with a high degree of confidence. This drawback was overcome by obtaining high quality Raman difference spectra under non-resonance conditions (see Sections 1.10 and 3.2.4). In view of the correlations that have been observed for acyl serine proteases, the results of the Raman and kinetic studies were initially unexpected and surprising. However, when taken with the data from absorption spectroscopy, they point to the importance of α-helix dipoles in cysteine protease active sites and pave the way for an understanding of the role of these dipoles in modulating enzymatic activity.

4.2 Materials and Methods

4.2.1 Synthesis of Substrates

All unlabelled substrates were synthesized as described in Sections 2.2.2 and 2.2.3. 5MTAIm labelled with ^{13}C at either the C_1 or the C_2 carbon was synthesized as described in Section 3.2.1.

Phe5MTAEt labelled with ^{13}C at either the C_1 or the C_2 carbon was synthesized with the following procedure. Glycine-1-^{13}C or glycine-2-^{13}C, purchased from Cambridge Isotope Laboratories, were esterified by adding 1.2 eq. of toluene-4-sulphonic acid in absolute ethanol and refluxing overnight. The sample was then dried with a rotary evaporator followed by a vacuum pump. The glycine ethyl ester labelled at the desired position was then used to synthesize Phe5MTAEt as described in Section 2.2.3.

4.2.2 Preparation of Acyl Enzymes and Data Collection

Enzymes and their protein engineered variants were purified and their respective acyl enzymes were prepared as described in the Materials and Methods of Chapter 2. For the Raman experiments, buffers were prepared using 99% D_2O (Aldrich) to remove the weak H-O-H bending vibrational band in the 1600 - 1700

cm^{-1} spectral region. The reported pD values for D_2O solutions were calculated by adding 0.4 to the observed pH meter reading (Glascoe & Long, 1960). Raman difference spectra were recorded using the spectrometer described in Section 3.2.4. The following experimental conditions were used for the collection of all acyl enzyme Raman data; 350 mW of 488 nm laser excitation, acquisition parameters 8 sec. x 40 accumulations, acyl enzyme concentration 200 ± 5 μM. A spectrum of the acyl enzyme was recorded using these conditions, and then a spectrum of enzyme alone at the same concentration was recorded. The spectrum of the enzyme was then subtracted from that of the acyl enzyme using the program Spectracalc.

Absorption data were measured using a Cary 3 UV-visible spectrophotometer.

4.2.3

Calculations used to Determine $\Delta\Delta E^{Electronic}$ and $\Delta\Delta E^{Catalysis}$

In order to calculate the electronic energy differences associated with the absorption spectra for two different acyl enzymes, $\Delta\Delta E^{Electronic}$, Planck's equation was employed in the following manner.

4.1) $$\Delta\Delta E^{Electronic} = (Lhc)/\lambda_1 - (Lhc)/\lambda_2$$

where L = Avogadro's constant, h = Planck's constant,

c = speed of light, λ_1 and $\lambda_2 = \lambda_{max}$ of the two acyl enzymes.

In order to calculate the difference in activation energies, $\Delta\Delta E^{catalysis}$, between two acyl enzymes, the Arrhenius equation was employed in the following manner.

4.2) $$\Delta\Delta E^{catalysis} = \ln(k_1/k_2) \times RT$$

where R = gas constant, T = temperature, k_1 and k_2 = deacylation rate constants of the two acyl enzymes.

4.3 Results

Figures 4.1, 4.2, and 4.3 are the Raman spectra of unlabelled, $^{13}C_2$ labelled, and $^{13}C_1$ labelled 5MTApapain at pD 4.0. Figure 4.1 also contains the spectrum of $5MTASC_2H_5$ in CCl_4 to illustrate the changes occurring within the 5MTA moiety upon binding to papain. Figure 4.4 contains superimposed spectra of unlabelled, $^{13}C_2$ labelled, and $^{13}C_1$ labelled Phe5MTApapain at pD 4.0.

In order to investigate the effects on the vibrational bands upon going to active pH, spectra were recorded at pDs 4.0 and 8.5 for all acyl papains and pDs 4.0 and 6.0 for acyl cathepsins B and L. A typical result is shown in Figure 4.5, which demonstrates that there is a small (2-4) increase in cm^{-1} in the

ethylenic and carbonyl vibrational bands, accompanied by a reduction in the overall intensity of the spectrum, upon increasing the pD from 4.0 to 8.5 for 5MTApapain. Acyl enzymes derived from the Phe5MTA- substrate did not show this pH effect in the pD 4.0 - 8.5 range. A unique result, shown in Figure 4.6, was obtained with 5MTAcatL, where upon going to active pH a relatively intense carbonyl band appears at 1643 cm^{-1}. This is the only acyl enzyme which clearly shows a carbonyl feature without the aid of isotopic substitution.

The carbonyl band can only be clearly discerned in all acyl enzyme data when ^{13}C is substituted at the carbonyl carbon. These peak positions are listed in Table 4.1 along with values of log k_3, where k_3 are the deacylation rate constants at active pH for the acyl enzymes derived in Chapter 2. Included in Table 4.1 are values of λ_{max} obtained from absorption spectroscopic measurements at pH 4.0 and at either pH 6.0 or 8.5.

Figure 4.7 is a plot of ν_{C-C} frequencies versus λ_{max} obtained from unlabelled acyl enzyme complexes. Included in Figure 4.7 are values of ν_{C-C} and λ_{max}, obtained from Chapter 3, Table 3.2, for the model compound $5MTASC_2H_5$ in various solvent systems. Figures 4.8 and 4.9 are plots of log k_3 versus $^{13}C{=}O$ frequencies and λ_{max}, respectively. The data for these graphs were obtained from Table 4.1. Included for comparative purposes in Figure 4.8 is the log k_3 *vs.* ν_{C-O} data for the acyl serine protease complexes reported by Tonge & Carey (1992).

Figure 4.1

Raman Spectra of 5MTAPapain at pD 4.0 and 5MTASC$_2$H$_5$ in CCl$_4$

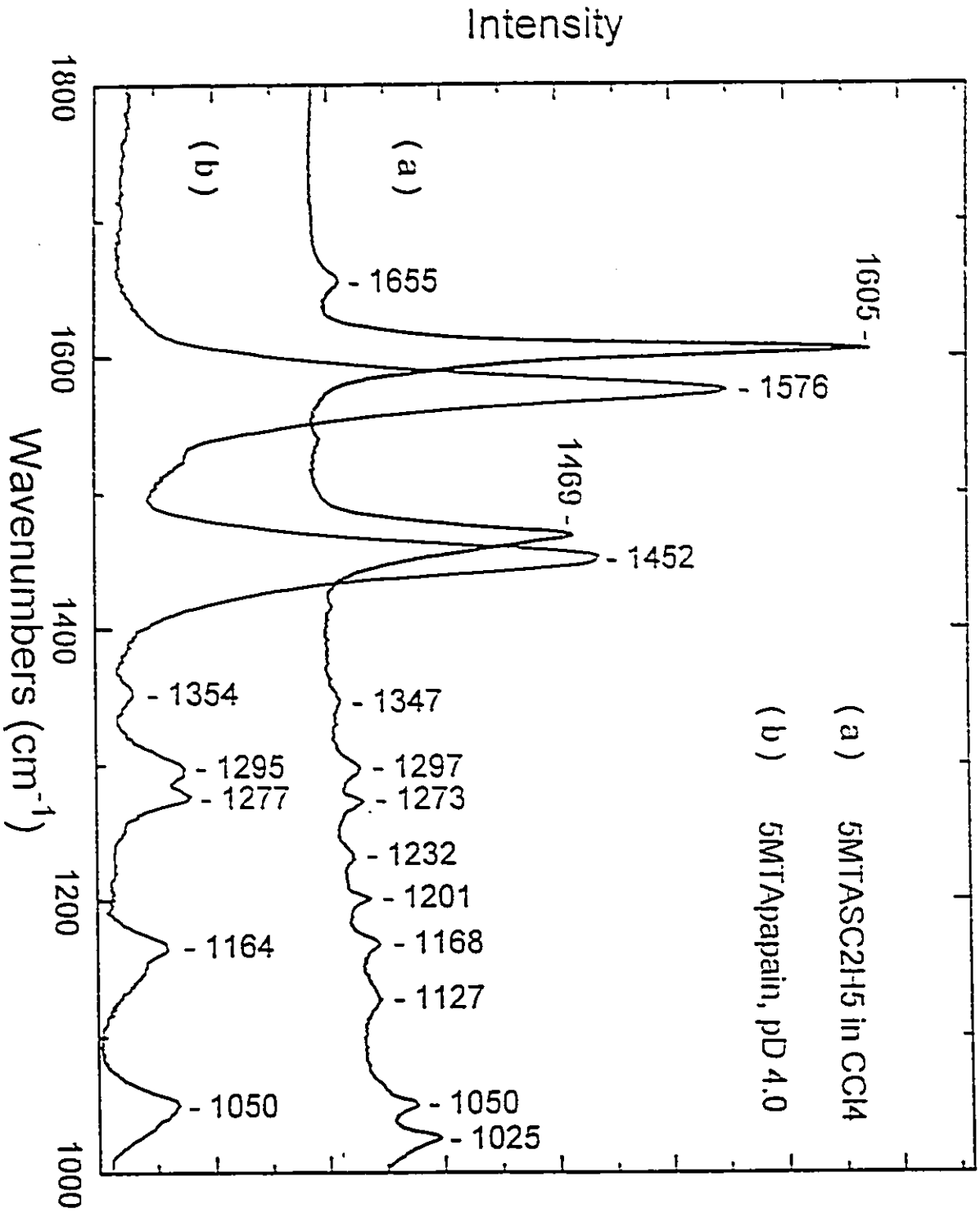


Figure 4.2

Raman Spectrum of ^{13}C Labelled 5MTApapain at pD 4.0

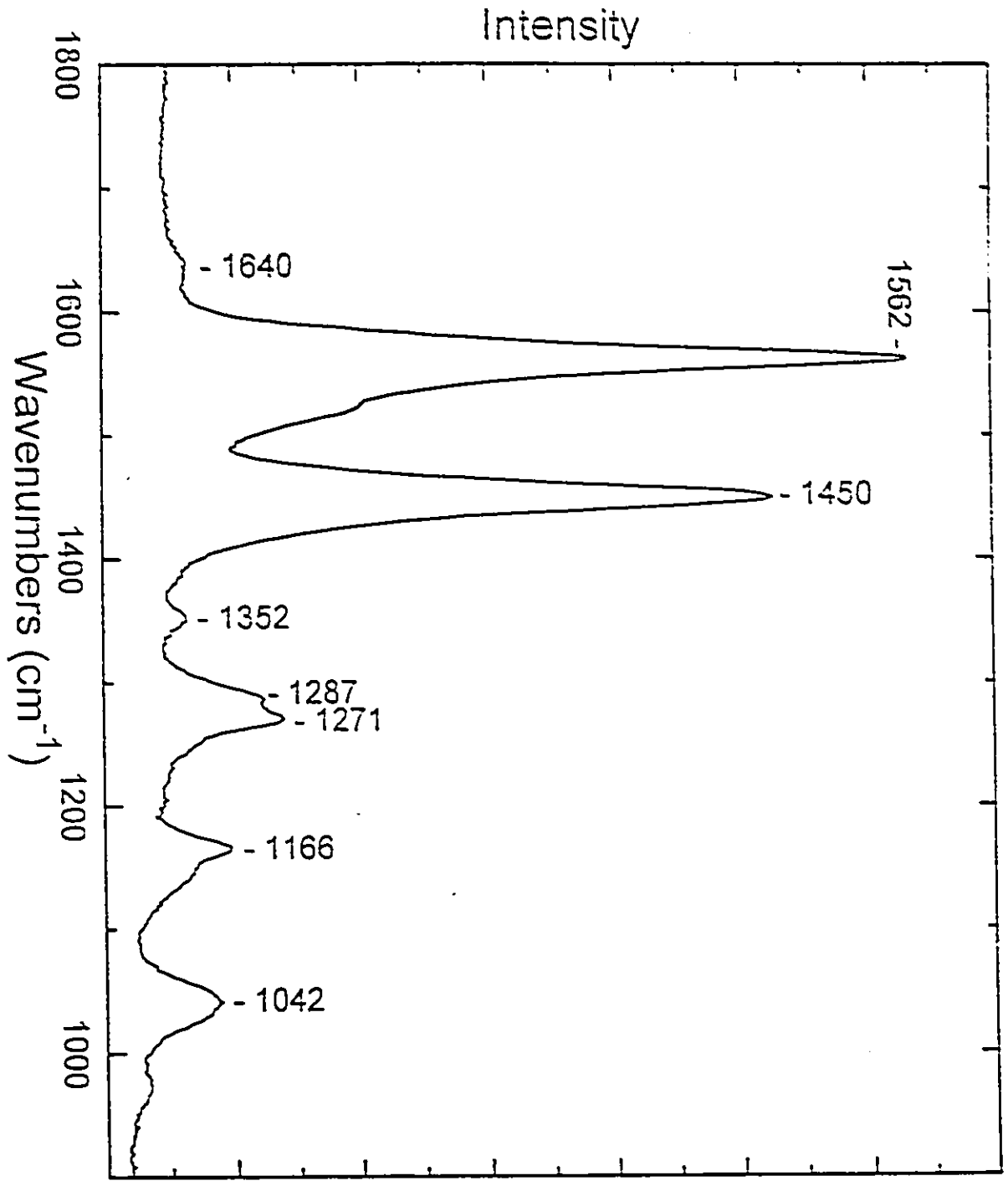


Figure 4.3

Raman Spectrum of $^{13}C_1$ labelled 5MTApapain at pD 4.0

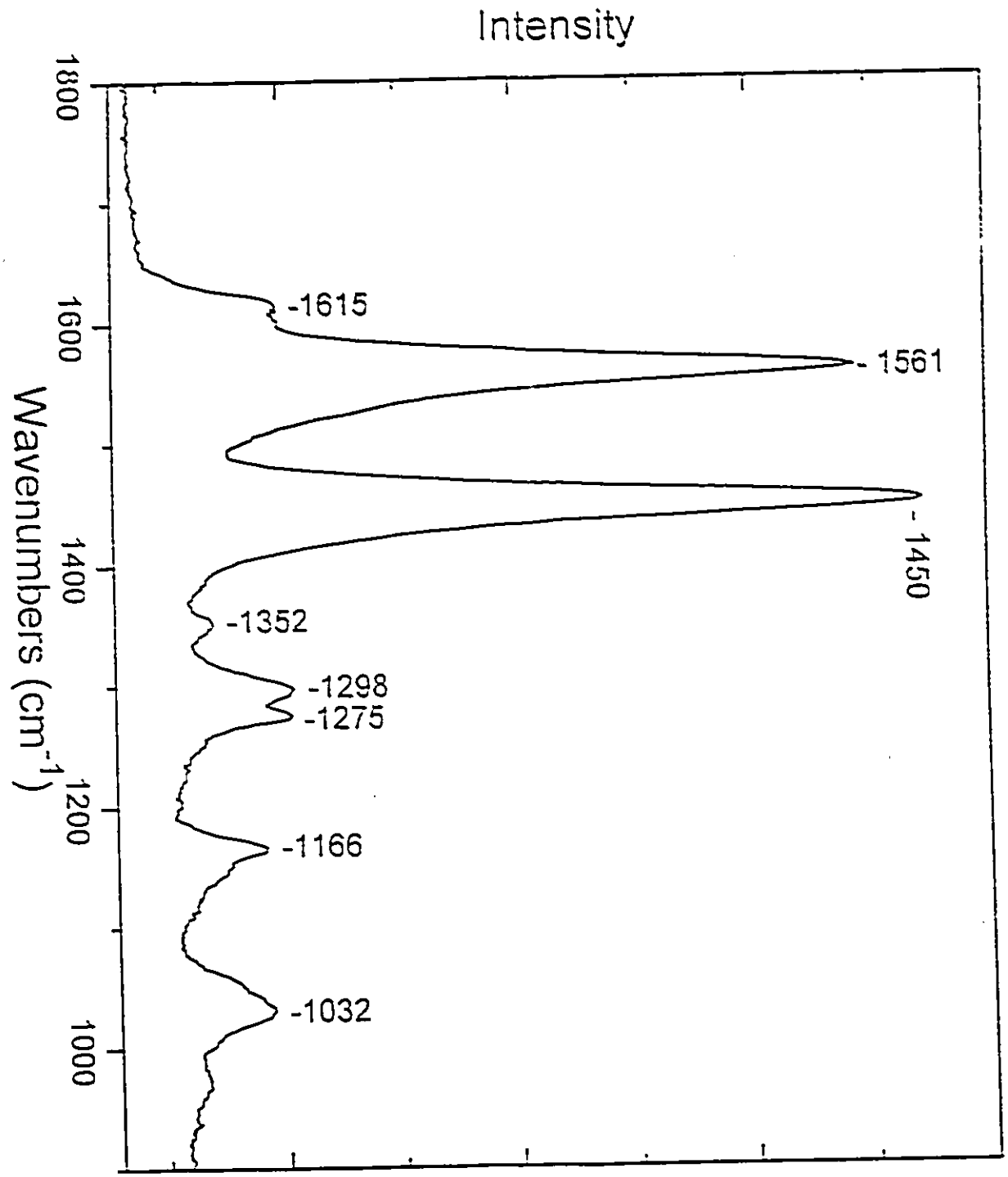


Figure 4.4

Raman Spectra of Unlabelled, $^{13}C_2$ Labelled, and $^{13}C_1$ Labelled Phe5MTApapain at pD 4.0

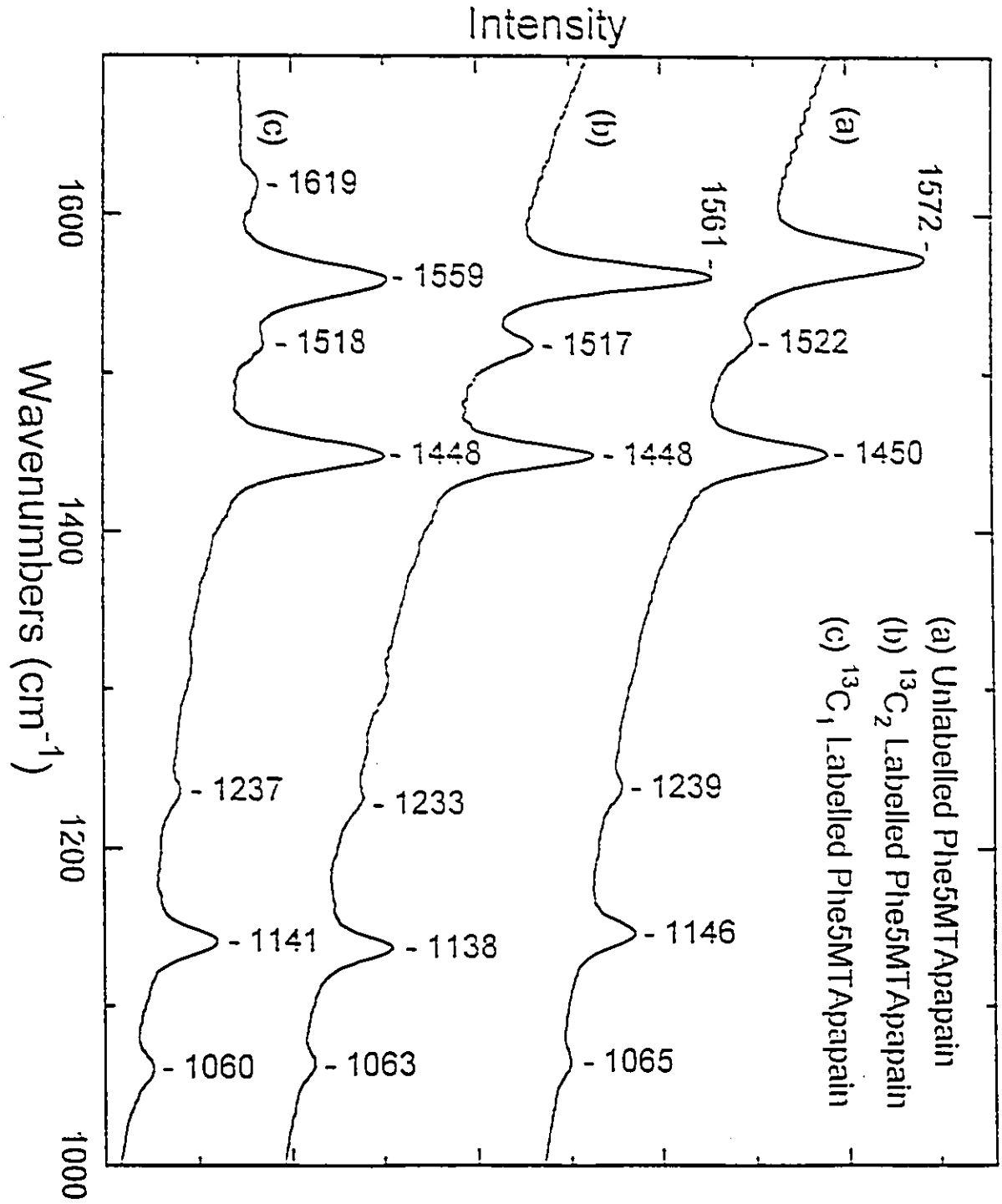


Figure 4.5

Raman Spectra of SMTAnapain at pD 4.0 and 8.5

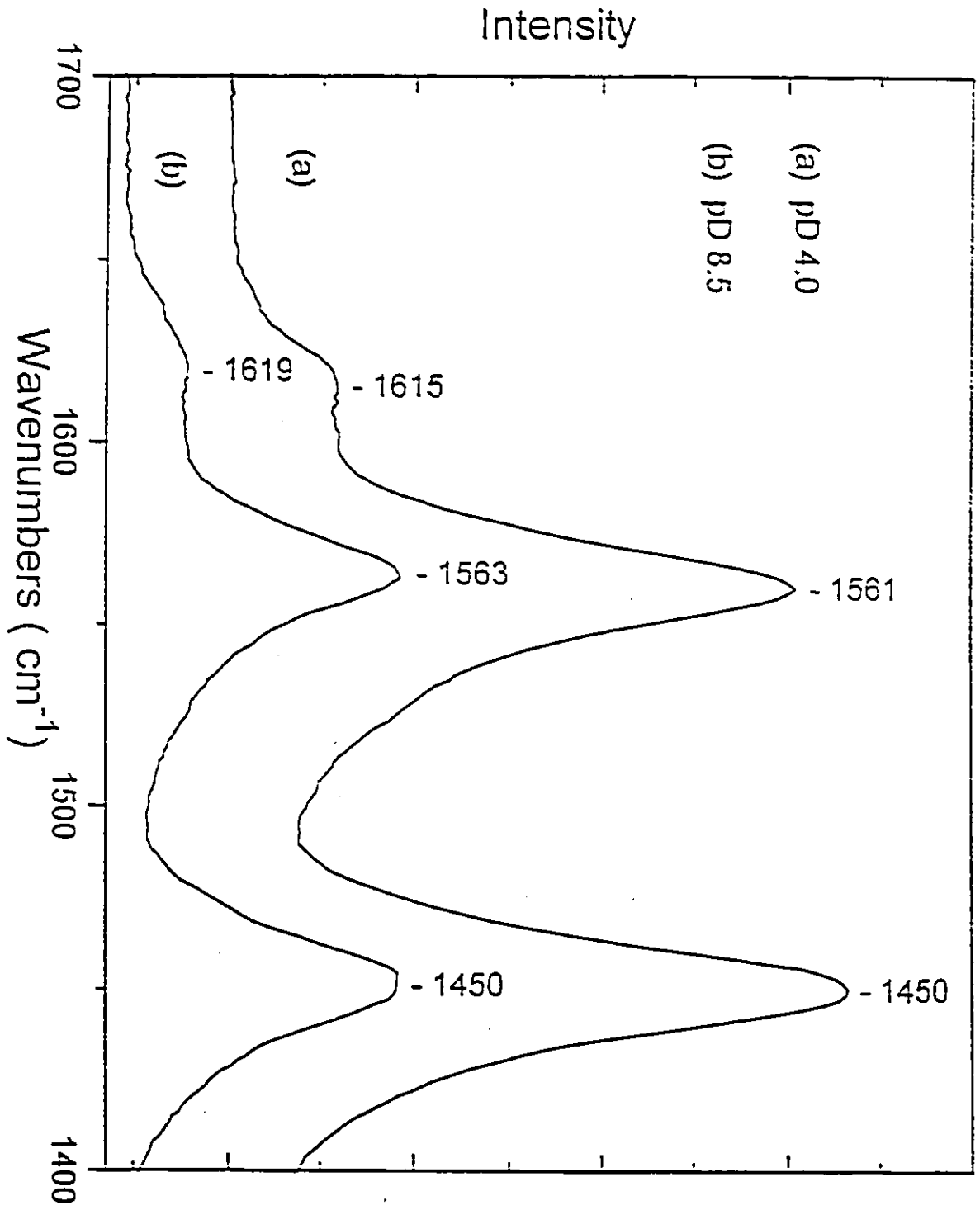


Figure 4.6

Raman Spectrum of SMTAcetL at pD 6.0

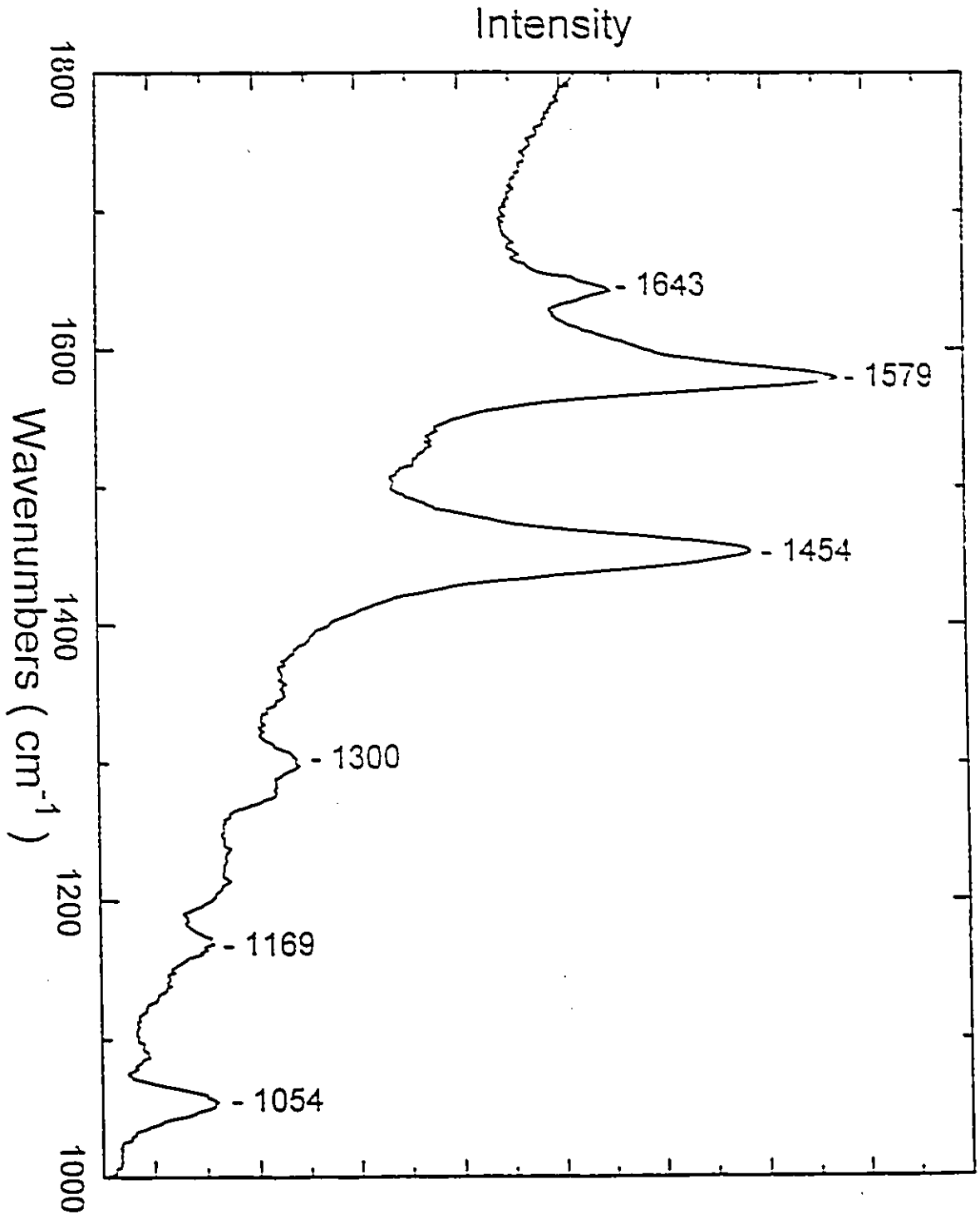


Table 4.1

Deacylation, Absorption, and Raman Spectroscopic Data for Ten Acyl Cysteine Proteases

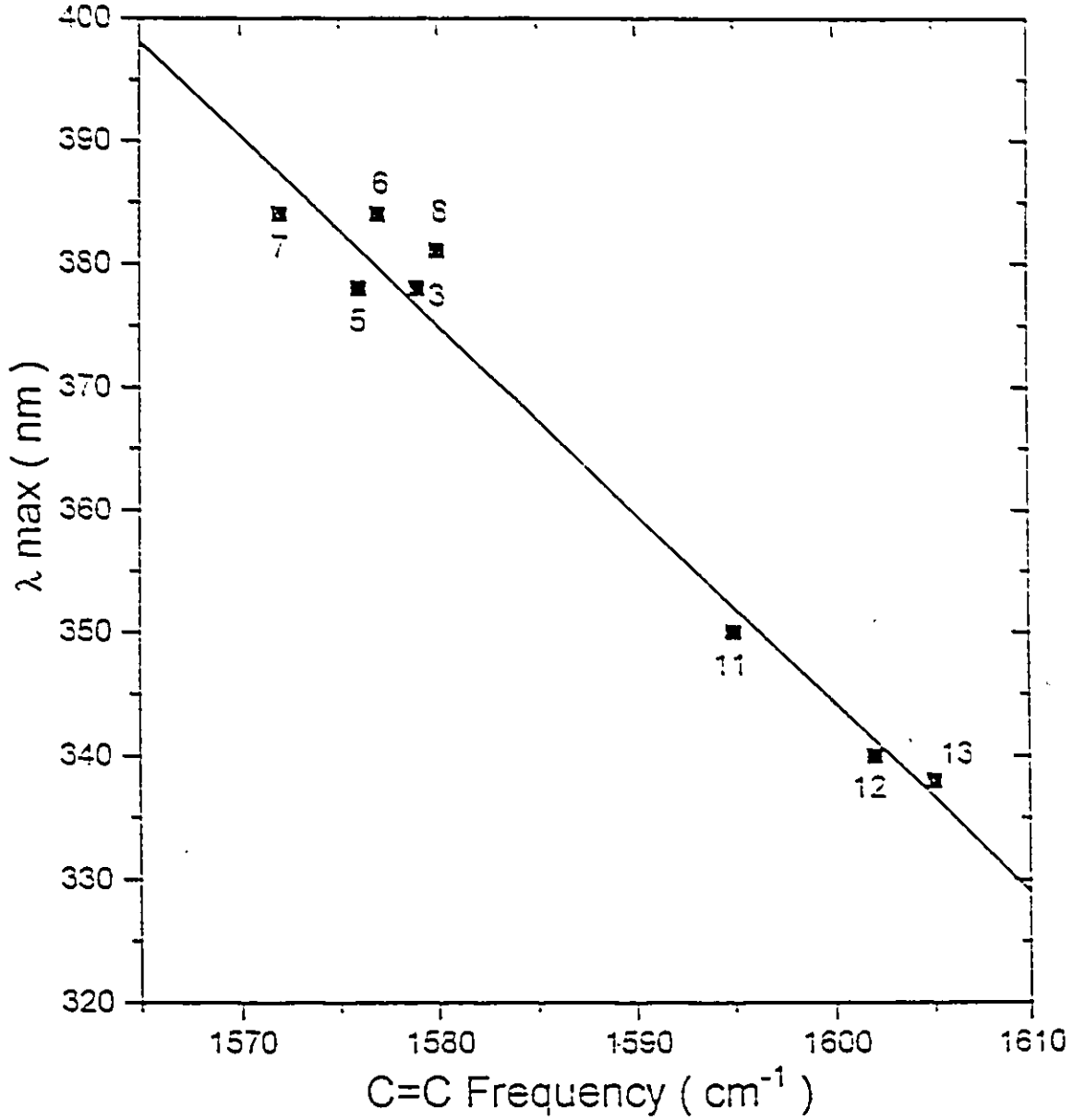
Acyl Enzyme	log k_3 deacylation	λ max low pH (nm)	λ max high pH (nm)	C=C high pD (cm^{-1})	^{13}C=O high pD (cm^{-1})
5MTApapain	-3.0	378	374	1579	1619
Phe5MTApapain	-2.5	384	384	1573	1624
5MTAcatB	-3.3	378	374	1579	1620
Phe5MATcatB	-2.5	384	384	1576	1623
5MTAcatBQ23A	-3.7	374	369	ND	1618
5MTAcatBQ23S	-4.15	374	369	ND	1612
Phe5MTAcatBQ23S	-3.2	375	375	ND	1624
5MTAcatL	-2.1	381	378	1579	1635
5MTAcatLQ19S	-3.55	ND	367	ND	ND
Phe5MTAcatL	-1.8	384	384	ND	ND

ND = Not determined.

Values of k_3 are from the data in Table 2.1, Chapter 2. Low pH and pD values are 4.0, and high pH and pD values are 8.5 for acyl papains and 6.0 for acyl cathepsins B and L.

Figure 4.7

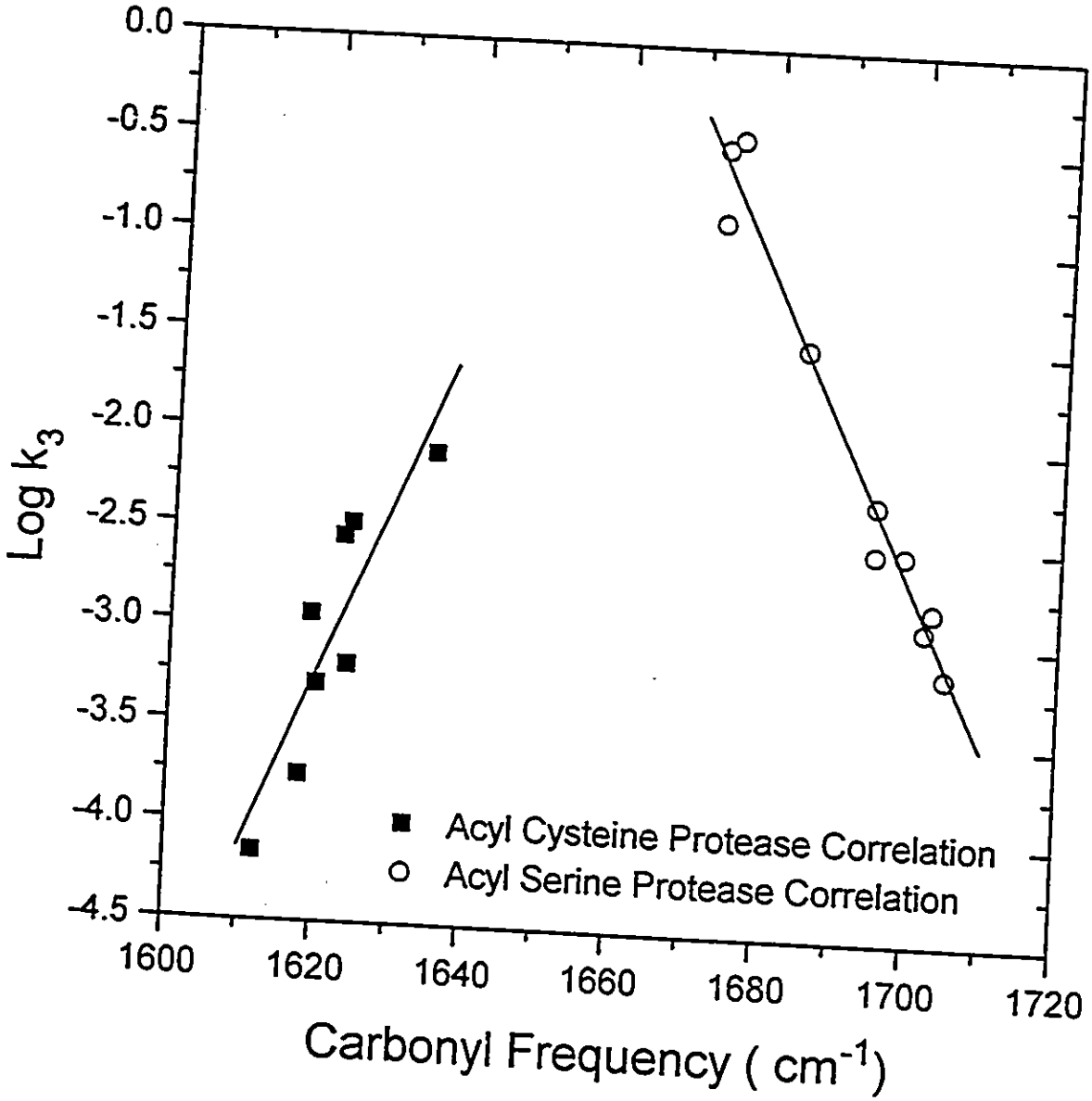
Plot of λ_{max} vs. Ethylenic Stretching Frequencies for Acyl Enzymes and for $5MTASC_2H_5$ in Different Solvents



Compound numbering:
3. 5MTAcatB; 5. 5MTApapain; 6. PheSMTAcatB; 7. PheSMTApapain;
8. 5MTAcatL; 11. $5MTASC_2H_5$ (in H_2O); 12. $5MTASC_2H_5$ (in CH_3CN);
13. $5MTASC_2H_5$ (in CCl_4)

Figure 4.8

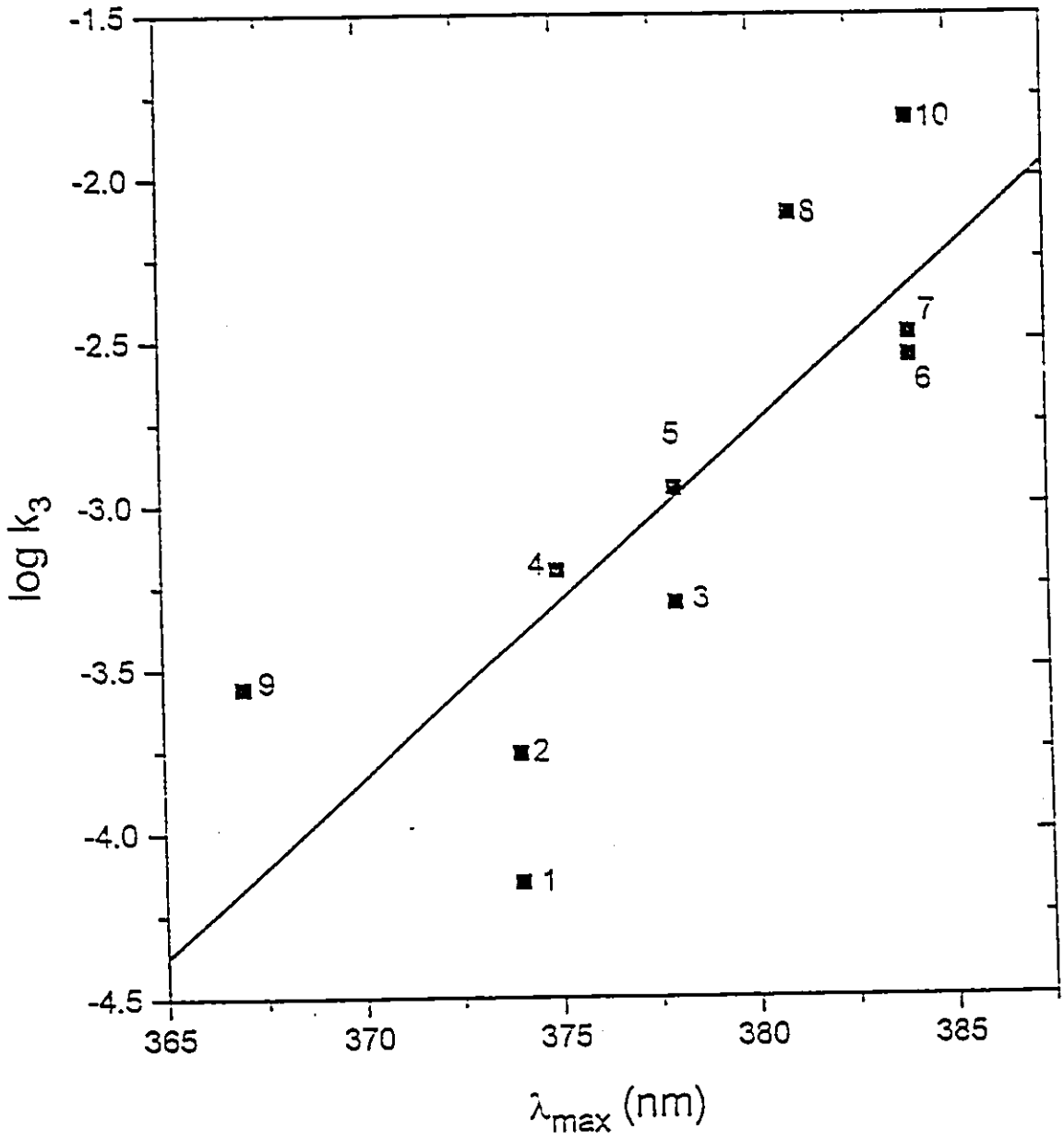
Plots of log k_3 vs. Carbonyl Stretching Frequency For Acyl Cysteine Proteases and Acyl Serine Proteases



The values for acyl cysteine proteases are for the $^{13}C{=}O$ substituted isotopomers.
Compound Numbering:
1. 5MTAcatBQ23S; 2. 5MTAcatBQ23A; 3. 5MTAcatB;
4. Phe5MTAcatBQ23S; 5. 5MTApapain; 6. Phe5MTAcatB;
7. Phe5MTApapain; 8. 5MTAcatL.

Figure 4.9

Plot of log k_3 vs. λ_{max} for Acyl Cysteine Proteases



Compound Numbering:
1. 5MTAcatBQ23S; 2. 5MTAcatBQ23A; 3. 5MTAcatB;
4. Phe5MTAcatBQ23S; 5. 5MTApapain; 6. Phe5MTAcatB;
7. Phe5MTApapain; 8. 5MTAcatL; 9. 5MTAcatLQ19S; 10. Phe5MTAcatL.

4.4 Discussion

4.4.1 Evidence for *s-cis*/*s-trans* Conformational Heterogeneity in Acyl Cysteine Proteases

α,β-Unsaturated thiolesters usually exist as a mixture of *s-cis* and *s-trans* conformational isomers about the C_1-C_2 bond. The 1126 and 1025 cm^{-1} Raman bands of $5MTASC_2H_5$ have a high degree of C_1-C_2 stretch from the *s-trans* and *s-cis* conformers, respectively (see Section 3.4.1). These two conformers are present in the 5MTApapain spectrum shown in Figure 4.1. The 1050 cm^{-1} and 1164 cm^{-1} bands of 5MTApapain are composed of overlapping thiophene ring and C_1-C_2 stretching vibrational modes. The 1025 C_1-C_2 *s-cis* stretching mode and 1050 cm^{-1} thiophene ring mode in the Raman spectrum of $5MTASC_2H_5$ (Figure 4.1 (a), see also $5MTASC_2H_5$ assignments in Table 3.6) overlap in the 5MTApapain Raman spectrum, resulting in the asymmetricallly shaped 1050 cm^{-1} band. A similar situation exists for the 1164 cm^{-1} 5MTApapain band, where a 1126 cm^{-1} C_1-C_2 *s-trans* stretching mode and 1168 cm^{-1} thiophene ring mode of $5MTASC_2H_5$ appear unresolved in the 5MTApapain spectrum as the asymmetrically shaped 1164 cm^{-1} band. Isotopic substitutions at the C_1 or the C_2 positions provide evidence for these assignments. For example, substitution of ^{13}C at the C_1 or C_2 positions causes the 1050 cm^{-1} band to shift to 1042 cm^{-1} (Figure 4.2) and 1032 cm^{-1} (Figure 4.3). It may

therefore be speculated that the bound 5MTA moiety has sufficient conformational freedom in the active site to interchange between these conformers.

For the specific substrate Phe5MTA bound to the active sites the conformation about the C_1-C_2 bond cannot be deduced. It is unlikely that the marker bands discussed above will be of value since the 5MTA moiety is now carrying the phenylalanine-based substituent on C_2. Close scrutiny of the Raman difference spectra for Phe5MTA acyl enzymes does not reveal new putative candidates for marker bands (see Figure 4.4). As expected, the attachment of the Phe dipeptide to C_2 causes substantial changes in the 1000-1400 cm^{-1} region, but it is impossible to interpret precisely the nature of these changes without a more detailed analysis.

4.4.2

The Carbonyl Stretching Frequency

The Raman carbonyl stretching frequency, ν_{C-O}, of thiol esters is considerably less intense than that of the corresponding oxy esters (see Section 3.4.2). This leads to difficulties in detecting ν_{C-O}. All of the unlabelled acyl cysteine proteases, with the exception of 5MTAcatL at active pH, show no band in the carbonyl region of the spectrum (e.g. Figure 4.1, 5MTApapain spectrum). Substitution of ^{13}C at the C_2 position shifts ν_{C-C} to a lower frequency, making the carbonyl somewhat easier to detect. For example, Figure 4.2 shows that for C_2

labelled 5MTApapain the carbonyl band is an extremely weak feature at 1640 cm^{-1}. This weak feature cannot be seen for the majority of the intermediates. In order to detect ν_{C-O} it was necessary to label the carbonyl carbon with ^{13}C, whereupon ν_{C-O} could be seen in the Raman difference spectrum of every acyl enzyme. The intensity enhancement of ν_{C-O} labelled at the carbonyl carbon is evident in Figures 4.3 and 4.4 (c), and is similar to the increase in ν_{C-O} intensity seen in the $^{13}C_1$ labelled $5MTASC_2H_5$ Raman spectrum (see Figure 3.5).

Labelling the carbonyl carbon however introduces the complication of increased vibrational coupling between the carbonyl and ethylenic stretching modes. Evidence for increased coupling is demonstrated by the decrease in ν_{C-C} by 15 cm^{-1} upon $^{13}C_1$ labelling relative to the unlabelled 5MTApapain. (compare Figures 4.1 and 4.3). A similar result was obtained with $5MTASC_2H_5$, where ν_{C-C} decreased by 10 cm^{-1} upon $^{13}C_1$ labelling (see Section 3.4.4). Hence α,β-unsaturated thiol esters in general give rise to coupled vibrations between the carbonyl and ethylenic stretching frequencies, predominantly due to the relatively low frequency of thiol ester carbonyls. It is interesting that binding of the 5MTA moiety to papain causes an even larger decrease in ν_{C-C} upon $^{13}C_1$ labelling than in $5MTASC_2H_5$. This is probably due to the strong polarizing forces present in the papain active site (see Section 4.4.5) causing greater coupling of ν_{C-C} and ν_{C-O}.

Due to the increased vibrational coupling between ν_{C-C} and

$\nu_{C=O}$ it is not possible to view the $\nu_{C=O}$ band in the presence of $^{13}C_1$ labelling as a pure carbonyl stretching vibration. In fact, the increased relative intensity of $\nu_{C=O}$ may be due to intensity being borrowed from the ethylenic stretch. It is therefore best to view $^{13}C_1$ labelled $\nu_{C=O}$ as a mixed vibrational mode containing contributions from both $\nu_{C=O}$ and $\nu_{C=C}$ stretching frequencies. For the sake of simplicity this band will be referred to as $\nu_{C=O}$ throughout the course of this thesis, with the understanding that it is in fact a coupled vibrational mode.

4.4.3

Spectroscopic Effects Upon Going to Active pH

All of the acyl enzymes, with the exception of those based upon the Phe5MTA substrate, show a 2 - 4 cm^{-1} increase in the $\nu_{C=C}$ and when it can be discerned, $\nu_{C=O}$, upon going to active pH (for example, see Figure 4.5). This small shift is accompanied by a decrease in the intensity of the bands. As well there is a 3 - 5 nm blue shift in λ_{max} upon going to active pH (Table 4.1).

An explanation for this phenomenon is found from earlier studies by Carey *et al.* (1976, 1978) and Phelps *et al.* (1981). Carey *et al.* (1978) found that for 4-dimethylamino-3-nitro-(α-benzamido)cinnamoylpapain, $\nu_{C=C}$ increased dramatically from 1570 cm^{-1} to 1582 cm^{-1} between pH's 3.0 and 7.0, accompanied by a 90% drop in intensity. Studies by Phelps *et al.* (1981) on 4-NH_2-3-NO_2-cinnamoyl- and 5MTA-chymotrypsins found similar, although not

as dramatic, effects upon taking the acyl enzyme to active pH. In this case there was a 3 - 4 cm^{-1} increase in ν_{C-O} and a 5 nm decrease in λ_{max}.

These results were explained by assuming that deprotonation of the active site histidine at active pH in both papain and chymotrypsin brings about a reduction in polarization of the bound substrate. This hypothesis is consistent with the fact that the Phe5MTApapain, which has a pK_a of 3.5 (see Figure 2.4), shows no variation in Raman peak position or λ_{max} between pH's 4.0 and 8.5. The concept of electron polarization will be explored in greater depth in the following sections.

4.4.4

Hydrogen Bonding in the Active Site of Cysteine Proteases

As was discussed in Section 4.4.2, in order to detect ν_{C-O} in all acyl enzymes it is necessary to label the carbonyl carbon, resulting in a mixed vibrational mode. Fortunately, however, the intermediate 5MTAcatL is unique in that, for reasons which are not fully understood, the carbonyl is more intense and it is possible to detect ν_{C-O} at active pH in the absence of labelling. It occurs at 1643 cm^{-1} at pD 6.0 (Figure 4.6), and since it may be regarded as a relatively pure C=O stretching mode, it may be used to gauge the strength of hydrogen bonding in the active site.

For the model compound 5MTASC$_2$H$_5$ in CCl_4, $\nu_{C=C}$ occurs in the Raman spectrum at 1655 cm^{-1} (see Figure 3.2 (a)). The 12 cm^{-1} shift upon going from a totally non-hydrogen bonding environment to the active site represents the formation of relatively weak hydrogen bonding to the carbonyl oxygen. Quantitative data on the relationship between the $\nu_{C=C}$ shift and the enthalpy of hydrogen bonding are not yet available for α,β-unsaturated thiol esters, but a semi-quantitative estimate can be deduced from the corresponding data for α,β-unsaturated oxy esters (Tonge *et al.* 1996). Tonge & Carey (1992) determined that $\nu_{C=C}$ shifts 32 cm^{-1} between 5MTAOMe in CCl_4 and 5MTAchymotrypsin at pH 9.0, which corresponds to a change in hydrogen bonding strength of 16 kJ mol^{-1} (Tonge *et al.* 1996). Based on literature comparisons of simple oxy esters and thiol esters (Grunwell *et al.* 1977, Smolders *et al.* 1988) it is expected that an equivalent hydrogen bonding situation to the acyl C=O in active site thiol esters would correspond to a similar shift of $\nu_{C=O}$. Since the observed shift in $\nu_{C=O}$ is only 12 cm^{-1}, it is concluded that hydrogen bonding is substantially weaker to the 5MTA acyl carbonyl oxygen bound to the cathepsin L active site than it is when bound to the chymotrypsin active site. The weaker hydrogen bonding may be one of the reasons why the contribution to transition state stabilization by the oxyanion hole in papain is somewhat less than that found in subtilisin (Ménard *et al.* 1991, see also Section 1.8).

4.4.5

π-Electron Polarization in Cysteine Protease Active Sites

The first resonance Raman studies on substituted cinnamoyl papains revealed the presence of strong electron polarization forces in papain's active site (Carey *et al.* 1976, 1978). The consequent π-electron reorganization is so large that the resonance Raman spectrum of the bound acyl group is totally distinct from that of either the substrate or product. The major changes in the Raman spectra are accompanied by large perturbations to the electronic absorption spectrum. The early resonance Raman studies focused mainly on the intense double bond feature which could be ascribed to $\nu_{C=C}$ or a coupled mode involving $\nu_{C=C}$ (Carey *et al.* 1976, 1978). The double bond stretching feature moves to lower cm^{-1} with increasing polarization, and there is a correlation between $\nu_{C=C}$ and λ_{max}, with a reduction on $\nu_{C=C}$ being accompanied by a red shift (Carey & Storer, 1984, Storer *et al.* 1981).

The absorption spectral data of 5MTA and Phe5MTA acyl cysteine proteases, and of the model compound $5MTASC_2H_5$ are presented in Table 4.1. For the acyl enzymes, binding to the active site is accompanied by a red shift of approximately 30 nm relative to $5MTASC_2H_5$. Thus the intrinsic polarizability of thiol esters such as $5MTASC_2H_5$ is even more pronounced when bound to cysteine protease active sites. This large red shift could be due

to predominantly electronic ground state or excited state effects, or a mixture of both. However, the major changes in peak positions in the Raman spectra indicate that substantial electron rearrangement is occurring in the ground state. Figure 4.1 compares a typical acyl cysteine protease spectrum, that of 5MTApapain, with the Raman spectrum of the model compound $5MTASC_2H_5$ in CCl_4. Going from the model compound to papain's active site, polarization of the π-electrons in the ground electronic state is evidenced by the drop in frequency of the intense C=C stretching feature from 1605 to 1576 cm^{-1}, along with a drop of 16-18 cm^{-1} in the position of thienyl ring modes at 1537 and 1468 cm^{-1}. The more intense bands appear to be broader in the case of the acyl enzyme, and this may be due to the fact that each band profile is made up of unresolved contributions from *s-cis* and *s-trans* conformers. These conformers may have increasingly large cm^{-1} differences due to differential interactions with the polarization field.

In Figure 4.7, the values of $\nu_{C=C}$ for 5 acyl cysteine proteases at "high" pH (where deacylation is maximal) and the three values for the model compound, $5MTASC_2H_5$, in different solvents, are plotted against the corresponding values for λ_{max}. The λ_{max} values for the acyl enzymes are taken at low pH so that the Phe5MTA- based substrates may be included. This will not have a major effect on the plot since the blue shift upon going to active pH is of the same magnitude for all acyl enzymes. A general trend of a red shift in λ_{max} accompanied by a drop in $\nu_{C=C}$

is observed in Figure 4.7. The straight line indicates that the change in the ground state electronic state π-electron system, as probed by ν_{C-C}, is accompanied by a proportional change in the difference between the ground and excited state energy levels as monitored by λ_{max}.

4.4.6

Correlation Between ν_{C-O} and Reactivity

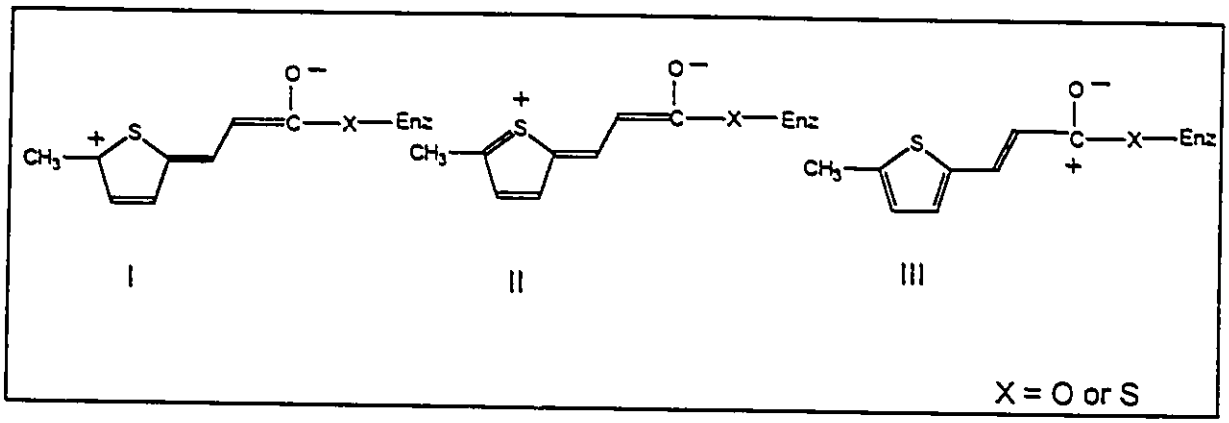
For acyl serine proteases, which included chymotrypsin, subtilisin, and oxyanion hole mutants of subtilisin, a striking correlation was developed between ν_{C-O} and log k_3, over a range of 16,000 fold in k_3 (Tonge & Carey, 1992, see also Figure 1.5). As the acyl enzymes become more reactive ν_{C-O} is found at lower values. The data could also be used to evaluate quantitatively the changes in C=O bond length and hydrogen bonding strengths to the C=O in the active site. The more reactive intermediates have a more polarized C=O group and stronger active site-to-C=O hydrogen bonds.

Figure 4.8 shows a plot of ν_{C-O} against log k_3 for 8 acyl cysteine proteases. Included in the plot for comparative purposes is the acyl serine protease data from Tonge & Carey (1992). Surprisingly, for the cysteine protease intermediates exactly the opposite behaviour to that for the serine proteases is found. As reactivity increases over the 214 fold range there is an increase in ν_{C-O}.

As was discussed in Section 4.4.2, in order to detect ν_{C-C} in every acyl enzyme it was necessary to ^{13}C label the carbonyl carbon, leading to increased coupling between the ν_{C-C} and ν_{C-C} stretching modes. It is therefore not possible to interpret the increase in ν_{C-C} with reactivity for acyl cysteine proteases as being indicative of simple changes in the length of the C=O bond, as was possible with acyl serine proteases. Nevertheless, the increases in ν_{C-C} demonstrate a change in the ground state π-electron system with increasing reactivity that does not occur with acyl serine proteases derived from similar substrates.

The explanation for the extraordinary reversal of slope seen in Figure 4.8 lies in ideas that were developed some time ago to explain the changes in the electronic absorption and resonance Raman spectra of furylacryloyl and thienylacryloyl chymotrypsins upon raising the pH and activating the charge relay system (Phelps *et al.* 1981). As pH is raised these acyl serine proteases undergo a blue shift in λ_{max} and in the RR spectrum a modest increase in ν_{C-C}, both of which indicate a minor reduction in the polarization in the π-electron system. This effect is akin to that seen for the acyl cysteine proteases upon raising pH as was discussed in Section 4.4.3. The changes can be explained using the same valence bond terms that was used to explain the greater polarizability of $5MTASC_2H_5$ compared to $5MTAOC_2H_5$ (Section 3.4.3). The canonical structures I and II in Scheme 4.1, where X is oxygen, contribute to the true structure of the acyl enzyme, and give rise to a small permanent dipole

Scheme 4.1



moment for the 5MTA group in the active site. If prior to hydrolysis the carbonyl is activated by, for example, a water molecule, resonance forms such as III in Scheme 4.1 become important. Canonical structure III has the positive charge localized on the carbonyl carbon, and canonical structures I and II, giving rise to the permanent dipole, are reduced in importance. Thus, interactions associated with carbonyl activation favour structure III at the expense of structures I and II, giving rise to a blue shift in λ_{max} and an increase in $\nu_{C=C}$.

The values of $\nu_{C=C}$ and λ_{max} for each acyl cysteine protease (Table 4.1) demonstrate that the π-electron system for the 5MTA group is highly polarized. That means that canonical forms I and II, where X is sulphur, make a major contribution to the true structure. This will perturb $\nu_{C=C}$ and $\nu_{C=O}$ strongly and, as we have observed, contribute to the effect of increasing vibrational

coupling between the two groups, as evidenced by the higher sensitivity of $\nu_{C=O}$ to ^{13}C labelling of the carbonyl carbon for acyl cysteine proteases compared to acyl serine proteases. Increased polarization resulting from increased contributions of structures I and II also explains why the degree of coupling is larger for acyl cysteine proteases than in the model compound $5MTASC_2H_5$ (see Section 4.4.2). The predominance of canonical forms I and II in cysteine protease bound acyl groups may reduce the reactivity of the carbonyl group because high susceptibility to nucleophilic attack depends on canonical form III making a significant contribution to the true structure, and this is suppressed in the highly polarized 5MTA acyl group. This is in contrast to acyl serine proteases where there is relatively little polarization throughout the acyl group, so that increasing hydrogen bonding to the C=O group increases the reactivity by increasing the importance of forms equivalent to structure III.

There are three possible causes giving rise to the observed polarization in the 5MTA group in the active site. These are, hydrogen bonding between the oxyanion hole and the acyl C=O group, the presence of the positively charged imidazole side chain from the active site histidine in the vicinity of acyl C=O group which is present below the deacylation pK_a of the acyl enzyme, and the presence of an α-helix dipole due to the α-helix which terminates at the active site cysteine. The fact that strong polarization exists for 5MTAcatL, where there is only weak hydrogen bonding (see Section 4.4.4), suggests that hydrogen

bonding at the acyl carbonyl is not a major contributor to polarization. This is confirmed by cathepsin B oxyanion hole data, where removal of a putative oxyanion hole hydrogen bond donor blue shifts λ_{max} by only 4 nm and increases ν_{C-C} by 2-4 cm^{-1} (see Table 4.1). Similarly, the modest changes seen in ν_{C-C} and λ_{max} upon raising pH where the charge effects of the protonated imidazole are removed shows that the protonated imidazole is not a major contributor (see Section 4.4.3). These data demonstrate that hydrogen bonding to the acyl carbonyl and the electric field due to the active site histidine's positively charged side chain make a small contribution to the observed polarization of the π-electrons in the 5MTA moiety. However, these data cannot account for the observed 28 nm red shift in λ_{max} or the 22 cm^{-1} decrease in ν_{C-C} observed in 5MTA based acyl enzymes relative to $5MTASC_2H_5$ in H_2O (see Table 3.2 for $5MTASC_2H_5$ spectroscopic data), therefore the major factor giving rise to polarization appears to be the α-helix. The helix dipole has been put forward as a key part of the mechanistic machinery of active sites (Hol *et al.* 1978, Wada, 1976), but experimental evidence for its effects have been lacking.

Small modulations of the α-helix dipole 5MTA interaction could result in variations in the contribution of canonical forms I and II to the overall structure. Since this will probably change the hybridization at the C=O carbon the effect on ν_{C-O} is difficult to predict based on a simple model.

4.4.7

Correlation Between λ_{max} and Reactivity

For the acyl serine proteases there is little or no variation in λ_{max} as reactivity varies over 16,000 fold (Tonge & Carey, 1992). However, across the acyl cysteine protease series with its 214 fold reactivity change, λ_{max} varies by 17 nm and as Figure 4.9 shows there is a linear correlation between λ_{max} and log k_3, with the most red shifted intermediates having the fastest deacylation rates. The implications of this relationship are important. The electronic absorption maximum represents a π-π^* transition between the ground and first excited states, and a red shift in λ_{max} represents a narrowing of this energy gap. In chromophores such as 5MTA the wave function for the excited electronic state contains major contributions from canonical forms I and II, where there is charge separation. That is, the excited state has a larger dipole moment than the ground state (Barltrop and Coyle, 1975). That is why there is a red shift for 5MTASC_2H_5 upon going from a solvent of low dielectric constant to a solvent with a high one (see Section 3.4.3 and Table 3.2). It is therefore hypothesized that the progressive red shift in Figure 4.9 is due to increasingly favourable α-helix dipole interactions which preferentially stabilize the excited electronic state. These changes undoubtedly involve minor variations in the α-helix-chromophore alignment and interactions. Thus the data in Figure 4.9 can be interpreted as the most

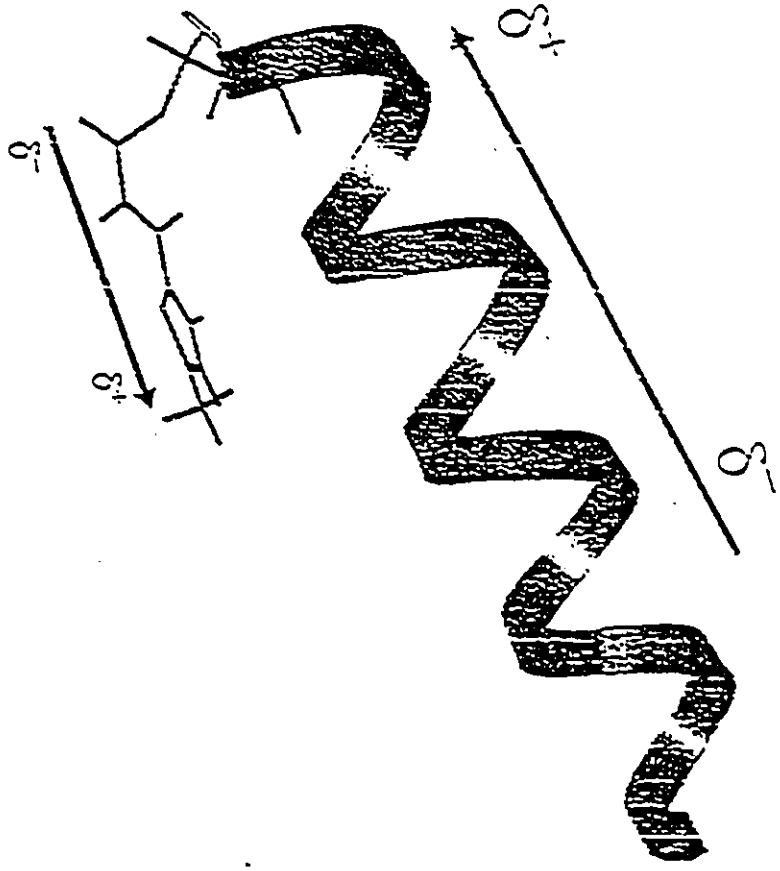
reactive acyl enzymes having the most effective α-helix dipole-chromophore interactions in terms of stabilizing the dipolar nature of the excited electronic state.

In Figure 4.10 a representation of the 5MTA acyl group bound to papain's cysteine-25 is shown. The orientation and conformation of the 5MTA group and the cysteine linkage were obtained via energy minimization. Due to the open cleft-like nature of papain's active site it is not possible to state unequivocally that the results shown represent a unique "true" conformation. However, the energy minimization clearly demonstrates that the axis of the 5MTA group can lie along that of the α-helix, such that the direction of the α-helix dipole and the dipole induced in the 5MTA group are anti-parallel. This type of interaction demonstrates that the α-helix dipole is capable of stabilizing the dipole in the 5MTA group, which in turn explains the observed large red shift in λ_{max} and the large drop in ν_{C-C} upon 5MTA binding to papain.

The connection to reactivity is made by remembering that the hydrolysis of the acyl enzyme likely goes through a transition state which resembles a tetrahedral intermediate (Ménard *et al.* 1995). In the latter species the acyl C=O is converted to a negatively charged C-O$^-$ group. Thus, it is proposed that α-helix dipole interactions that stabilize the charge separation in the excited electronic state and bring about a red shift will also stabilize negative charge build up in the transition state and bring about enhanced reactivity. Of course, the enzyme reaction

Figure 4.10

Schematic Representation of the Orientation of the SMTA moiety in the Active Site of Papain



Molecular modeling was performed by Dinakarpandian & Carey (unpublished data) using the InsightII (95.0) and Discover (2.9.7 & 95.0/3.00) software packages from Biosym/Molecular Simulations. SMTA- was built manually and minimized independently before being bonded to Cys25 of a relaxed crystal structure of papain. The ligand was placed initially in a position parallel to the axis of the α-helix and energy minimization carried out within a 10 Å radius of the acyl carbonyl oxygen. Subsequently, all restraints were removed and the entire molecule minimized until a maximum derivative of 0.1 kcal mol^{-1} $Å^{-1}$ was achieved. The only part of the active site shown above is a portion of the α-helix, residues Ile40 to cys25, pointing towards the active site.

proceeds solely through the ground electronic state, but the commonality involves the ability of the α-helix dipole to stabilize charge build up be it in the transition state of the reaction pathway or in the excited electronic state which determines electronic spectral properties.

4.4.8

Quantification of the Catalytic Effect of an α-Helix Dipole

The deacylation rate constants for the present series of acyl cysteine proteases vary by a factor of 214. On the basis of a simple Arrhenius equation calculation (Equation *4.2*), this means that the activation energies $\Delta\Delta E^{catalysis}$ vary by 3.2 kcal/mole throughout the series. In the previous section it was postulated that the interaction between the α-helix dipole and the negative charge on the carbonyl oxygen of the tetrahedral intermediate helps stabilize the charge build up, and thus contributes to rate acceleration. It was further postulated that there are differences in the helix-acyl group interactions throughout the series, possibly due to subtle changes in helix-chromophore alignment, that lead to the observed changes in λ_{max}. The origin of these changes lies in the differential contributions of canonical structures I and II (Scheme 4.1), with the more dipolar, more polarized acyl groups having the largest red shift. It is of interest to note that throughout the series

of acyl cysteine proteases λ_{max} varies from 367 to 384 nm, and if the differences in the electronic energy level transitions for these extremes of λ_{max} are calculated using Planck's relationship (Equation *4.1)* it is found that $\Delta\Delta E^{electronic}$ is 3.2 kcal/mole across the series. The identity of the values for $\Delta\Delta E^{catalysis}$ and $\Delta\Delta E^{electronic}$ must be fortuitous. However, it may be an indication that the underlying postulates have some validity and, crucially, that the observed variation in deacylation rate is caused by differential helix-acyl group interactions. If this is the case, then this work is the first experimental outline of the effects of α-helices on catalytic rate.

4.5 Summary

Raman and absorption spectroscopic data were combined with the deacylation rate constants for a series of acyl cysteine proteases to provide insight into the role of α-helix dipoles in rate acceleration.The Raman spectra, obtained by Raman difference spectroscopy, of 5-methylthienylacryloyl-papain, cathepsins B and L, and two oxyanion hole mutants of cathepsin B (Q23S and Q23A), show that extensive polarization throughout the π-electron chain occurs for the bound acyl group in the active sites. A similar result was obtained using the specific chromophoric substrate 2-(N-acetyl-L-phenylalanine) amino-3(5-methyl thienyl) acrylate methyl ester.

Using $^{13}C{=}O$ substitution it was possible to detect the acyl C=O stretching frequency for each acyl enzyme. This mode was found to be strongly coupled with the C=C stretching frequency of the acyl group. A correlation between ν_{C-O} and log k_3, where k_3 is the deacylation rate constant, was found where ν_{C-O} increases with increasing reactivity. This is exactly the opposite behaviour to the relationship found for a series of acyl serine proteases. The opposite trend in the direction of the correlation for the acyl cysteine proteases is ascribed to the strong polarizing forces in the active site, due principally to the α-helix dipole, giving rise to canonical forms of the acyl group which change the hybridization about the C=O carbon atom. A correlation was also observed between the absorption maximum, λ_{max}, and log k_3 for the

series of acyl cysteine proteases. As the deacylation rate increases, 214 fold across the series, λ_{max} red shifts from 367 to 384 nm. It is proposed that differential interactions between the active site's α-helix dipole and the acyl chromophore give rise to the observed changes in λ_{max}, with the red shift being caused principally by interactions with the excited electronic state, which has a high degree of charge separation, and the dipole. Similar interactions between the dipole and the transition state of the acyl enzyme in the ground state, are proposed as the source of the differential rates in deacylation.

CHAPTER 5

GENERAL SUMMARY AND CONCLUSIONS

Structure-reactivity relationships for a series of acyl cysteine proteases have been established using a combination of Raman and absorption spectroscopies and kinetic techniques. The simple chromophoric ligand 5-methylthienylacryloyl- (5MTA-) and three novel, more specific, chromophorically labelled substrates including the 5MTA- moiety, were used to create acyl enzyme adducts with papain, cathepsins B and L, and oxyanion hole mutants of cathepsins B and L. The chromophoric specific adduct NHCOOEt5MTA- gave no increase in overall reactivity relative to 5MTA-, while the Ala5MTA- and Phe5MTA- adducts gave modest increases of 3 - 9 fold in deacylation rates relative to 5MTA-. Ala5MTApapain and Phe5MTApapain had k_{cat}/K_m values of 0.85 and 53 sec^{-1} M^{-1}, respectively, and an upper limit of 0.07 sec^{-1} M^{-1} was determined for NHCOOEt5MTApapain and 5MTApapain. These results demonstrate that the Ala5MTA and Phe5MTA substrates are more specific than the NHCOOEt5MTA and 5MTA substrates, and most of the specificity is manifested in the acylation step. The Phe5MTA- adduct was the most specific substrate, thus reflecting the well known propensity of papain for a Phe side chain in the S_2 pocket. It was observed that the extent of deacylation of Phe5MTApapain was pH dependent, with successively larger proportions of the total acyl papain concentration deacylating as the pH was increased. A simple model was derived which quantitatively accounted for this behaviour, based on the assumptions that the protonated product rebinds to the enzyme and that in the pH range of 5.0 to 7.5, ionization of enzyme groups is not a factor in

deacylation or reacylation.

For cathepsins B and L removing one of the hydrogen bonding groups making up the oxyanion hole reduces the catalytic rate 3-25 fold with the above substrates. Using these substrates to make acyl enzymes with papain, wild type cathepsins B and L and oxyanion hole mutants of cathepsin B, resulted in deacylation rate constants stretching over a 214 fold range, from 0.07 to 15 x 10^{-3} sec^{-1}.

Using $^{13}C=O$ substitution it is possible to detect the acyl C=O frequency, $\nu_{C=O}$, for each acyl cysteine protease by Raman difference spectroscopy. The C=O stretching frequency is vibrationally coupled with the C=C stretching frequency of the adjacent ethylenic bond in α,β-unsaturated substrates. A correlation between the $\nu_{C=O}$ frequency and the deacylation rate constant was established, where $\nu_{C=O}$ increases with increasing reactivity. This trend is the opposite to that seen with acyl serine proteases derived from similar α,β-unsaturated substrates. The opposite trend is ascribed to the strong electron polarizing forces in the active site, due principally to an α-helix dipole, which change the hybridization about the carbonyl carbon atom. This strong polarizing effect was demonstrated to be a property of cysteine protease active sites, since comparative spectroscopic studies of the model compounds $5MTASC_2H_5$ and $5MTAOC_2H_5$ revealed that the polarizing effects due to the substitution of sulphur for oxygen were insufficient to explain the large red shifts in λ_{max} or the large decreases in $\nu_{C=C}$

observed with acyl cysteine proteases. It was also determined that hydrogen bonding and protonation of the active site histidine in the active site have only minor polarization effects, leading to the hypothesis that the enzyme's α-helix dipole is primarily responsible for the polarization. Molecular modelling of 5MTA- bound to papain obtained via an energy minimization resulted in a parallel alignment of the bound chromophore with the α-helix dipole. This configuration stabilizes charge separation in the 5MTA moiety, providing support for the hypothesis that the α-helix is the polarizing source.

A correlation was also established between the absorption maximum, λ_{max}, and the deacylation rate constant. As the deacylation rate increases, 214 fold across the series, λ_{max} red shifts from 367 to 384 nm. It is proposed that differential interactions between the α-helix dipole of the protein and the π-electrons in the bound chromophoric substrate is responsible for these changes in λ_{max}, with the red shift being due to more favourable α-helix dipole interactions with the substrate's dipolar excited electronic state. It is also proposed that similar interactions occur between the α-helix dipole and the transition state of the acyl enzyme in the ground electronic state, giving rise to the observed differences in deacylation rates. These results demonstrate the importance of α-helix dipoles in cysteine protease active sites in modulating enzymatic activity, providing the first experimental verification of

numerous theoretical studies implicating a role for α-helix dipoles in catalysis.

In this thesis, the chromophoric acyl group 5MTA- and its more specific derivatives was used for a variety of technical reasons based on the development of Raman spectroscopy to study acyl enzyme intermediates. For this type of chromophore, interactions with enzyme dipoles give rise to a complementary dipole in the acyl group's π-electron system, which extends along the long axis of the acyl group. It is thus likely that the reverse slope relationship between ν_{C-O} and log k_3 observed for acyl cysteine proteases is a consequence of the extended π-electron systems characteristic of the substrates used in this study. For a natural substrate, which lacks an extended π-electron chain, the polarization will be limited to the C=O group itself. It may therefore be predicted that an acyl serine protease type relationship between ν_{C-O} and log k_3 will assert itself if substrates are chosen which lack an extended π-electron system. In this case there can be no possibility for the stabilization of resonance structures with extensive charge separation (such as canonical structures I and II in Scheme 4.1) which reduce the nucleophilic susceptibility of the carbonyl carbon. It is recommended that future experiments be planned to confirm or refute this prediction.

References

Aqvist, J., Luecke, H., Quiocho, F. A., & Warshel, A. (1991) *Proc. Natl. Acad. Sci. 88*, 2026

Asboth, B., Stokum, E., Khan, I.U., & Polgar, L. (1985) *Biochemistry 24*, 606-609.

Bachovchin, W. W., Kaiser, R., Richards, J. H., & Roberts, J. D. (1981) *Proc. Natl. Acad. Sci. USA 78*, 7323.

Baker, A.W. & Harris, G. H. (1960) *J. Am. Chem. Soc. 82*, 1923-1928.

Baker, E. N. & Drenth, J. (1987) *Biological Macromolecules and Assemblies: Vol. 3 - Active Sites of Enzymes*. pp 334, John Wiley & Sons, Inc.

Barltrop, J. A. & Coyle, J. D. (1975) *Excited States in Organic Chemistry*. pp 59-62, John Wiley & Sons, London.

Barnes, R. B., Gore, R. C., Liddel, V., & Williams, V. Z., (1944) *Infrared Spectroscopy*. Reinhold, New York.

Bender, M. L., Schonbaum, G. R., & Zerner, B. (1962) *J. Am. Chem. Soc. 84*, 2540.

Bender, M. L. & Brubacher, L. J. (1964) *J. Am. Chem. Soc. 86*, 5333-53334.

Berger, A. & Schechter, I. (1970) *Phil. Trans. Roy. Soc. Lond. B. 257*, 249-264.

Bernhard, S. A., Lau, S. J., & Noller, H. (1965) *Biochemistry 4*, 1108.

Berti, P. J., Faerman, C. H., & Storer, A. C. (1991) *Biochemistry*

30, 1394-1402.

Blow, D. M., Birktoft, J. J., & Hartley, B. S. (1969) *Nature* 221, 337.

Brocklehurst, K. & Little, G. (1973) *Biochem. J.* 133, 67-80.

Brocklehurst, K., Willenbrock,F., & Salih,E. (1987) *Hydrolytic Enzymes* (Neuberger,A. & Brocklehurst,K., Eds.) Chapter 2, Elsevier.

Brubacher, L. J. & Zaher, M. R. (1979) *Can. J. Biochem. 57*, 1064-1072.

Bryan, P., Pantoliano, M. W., Quill, S. G., Hsiao, H., & Poulos, T. (1986) *Proc. Natl. Acad. Sci. USA 83*, 3743.

Callender, R. & Deng, H. (1994) *Annu. Rev. Biophys. Biomol. Struc. 23*, 215

Carey, P. R., Carriere, R. G., Lynn, K. R., and Schneider, H. (1976) *Biochemistry 15*, 2387

Carey, P. R., Carriere, R. G., Phelps, D. J., and Schneider, H. (1978) *Biochemistry 17*, 1081-1087.

Carey, P. R. (1982) *Biochemical Applications of Raman and Resonance Raman Spectroscopies*. pp 169-176. Academic Press, New York.

Carey, P. R. & Storer, A. C. (1984) *Ann. Rev. Biophys. Bioeng. 13*, 25.

Carey, P. R. & Tonge, P. J. (1990) *Chem. Soc. Rev. 19*, 293-316.

Carter, P. & Wells, J. A. (1990) *Proteins: Struct., Funct., Genet. 7*, 335.

Collings, A. J., Jackson, P. F., & Morgan, K. J. (1970) *J. Chem.*

Soc. (B), 581

Cregg, J. M., Vedvick, T. S., & Raschke, W. C. (1993) *Biotechnology 11*, 905-910.

Dixon, M. & Webb, E. C. (1979(a)) *Enzymes 3rd Ed.* pp 154-157 Longman Group Limited, London.

Dixon, M. & Webb, E. C. (1979(b)) *Enzymes 3rd Ed.* pp 511-512 Longman Group Limited, London.

Doran, J. D., Tonge, P. J., Carey, P. R., & Arya, P. (1995) *Bioorg. Med. Chem. Lett. 5*, 2381-2384.

Drenth, J., Kalk, K. H., & Swen, H. M. (1976) *Biochemistry 15*, 3731-3738.

Faria, M. D. G., Teixeira-Dias, J. J. C., & Fausto, R. (1991(a)) *Vibrat. Spectros. 2*, 107-123.

Faria, M. D. G., Teixeira-Dias, J. J. C., & Fausto, R. (1991(b)) *Vibrat. Spectros. 2*, 519-523.

Fausto, R. (1994) *J. Mol. Struct. 315*, 123.

Fausto, R., Tonge, P. J., & Carey, P. R. (1994) *J. Chem. Soc. Faraday Trans. 90*, 3491.

Fox, T., De Miguel, E., Mort, J. S., & Storer, A. C. (1992) *Biochemistry 31*, 12571-12576.

Fox, T., Mason, P., Mort, J. S., & Storer, A. C. (1996) *Biochemstry* (submitted).

Glascoe, P. K. & Long, F. A. (1960) *J. Phys. Chem. 64*, 188-191.

Grunwell, J. R., Forest, D. L., Kaplan, F., & Siddigui, J. (1977) *Tetrahedron 33*, 2781-2784.

Herzberg, G. (1945) *Infrared and Raman Spectra of Polyatomic*

Molecules. Van Nostrand, New York.

Hinkle, P. M. & Kirsch, J. F. (1970) *Biochemistry 9,* 4633.

Hol, W. G. I., van Duijen, P. T., & Berendsen, H. J. C. (1978) *Nature 273,* 443.

Horvath, G., Illenyi, J., Pusztay, L., & Simon, K. (1987) *Acta Chim. Hung. 124,* 819.

Jones, P. G. & Kirby, A. J. (1979) *J. Chem. Soc., Chem. Commun.* 288.

Jones, P. G. & Kirby, A. J. (1984) *J. Am. Chem. Soc. 106,* 6207.

Jones, D. H. & Howard, B. H. (1992) *Biotechniques 12,* 528-553.

Kaiser, E. T. (1985) *Nature* 313, 630.

Katunuma, N., & Kominami, E. (1987) *Rev. Physiol. Biochem. Pharmacol. 108,* 1-20.

Kim, M. & Carey, P. R. (1991) *J. Mol. Struct. 242,* 421 - 431

Kim, M., Owen, H., & Carey, P. R. (1993) *Appl. Spectrosc. 47,* 1780-1783.

Kirschke, H. & Barrett, A. J. (1981) *Methods Enzymol.* 80, 535-560.

Koo, J., Fish, M. S., Walker, G. N., & Blake, J. (1963) *Organic Syntheses, Collect. vol. IV,* Wiley, New York. p. 327.

Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154,* 367-382.

Long, D. A. (1977) *Raman Spectroscopy*. McGraw Hill, England.

Lowe, G. & Yuthavong, Y. (1971) *Biochem. J. 124,* 107-115.

Lucas, E. C. & Williams, A. (1969) *Biochemistry 8,* 5125-5135.

MacClement, B. A., Carriere, R. G., Phelps, D. J., & Carey, P. R.

(1981) *Biochemistry 20,* 3438.

Mackenzie, N. E., Grant, S. K.,Scott, A. I., & Malthouse, J. P. G. (1986) *Biochemistry 25,* 2293-2298.

Ménard, R., Carriere, J., Laflamme, P., Plouffe, C., Khouri, H. E., Vernet, T., Tessier, D. C., Thomas, D. Y., & Storer, A. C. (1991) *Biochemistry 30,* 8924-8928.

Ménard, R., Plouffe, C., Laflamme, P., Vernet, T., Tessier, D. C., Thomas, D. Y., & Storer, A. C. (1995) *Biochemistry 34,* 464-47

Mort, J. S., Poole, A. R., & Decker, R. S. (1981) *J. Histochem. Cytochem. 29,* 649-657.

Musil, D., Zucic, D., Turk, D., Engh, R. A., Huber, R., Popovic, T., Turk, V., Towatari, T., Katunuma, N., & Bode, W. (1991) *EMBO 10,* 2321-2330.

Nicholson, H., Becktel, W. J., & Mathews, B. W. (1988) *Nature 336,* 651.

Nicholson, H., Anderson, D. E., Dao-pin, S., & Matthews, B. W. (1991) *Biochemistry 30,* 9816.

Nyquist, R. A. & Potts, W. J. (1959) *Spectrochimica Acta 7,* 514.

Phelps, D. J., Schneider, H., & Carey, P. R. (1981) *Biochemistry 30,* 3447.

Pimental, G. C. & McClellan, A. L. (1960) *The Hydrogen Bond,* Freeman, London.

Raman, C. V. & Krishnan, K. S. (1928) *Nature (London) 121,* 501.

Rao, S. N., Singh, U. C., Bash, P. A., & Kollman, P. A. (1987) *Nature 328,* 551.

Rich, D. H., Brown, M. A.,& Barrett, A.J. (1986) *Biochem J.* 235, 731-734.

Rowan, A. D., Mason, P., Mach, L., & Mort, J. S. (1992) *J. Biol. Chem. 267,* 15993-15999.

Rullman, J. A. C., Bellido, M. N., & van Duijnen, P. T. (1989) *J. Mol. Biol. 206,* 101.

Sali, D., Bycroft, M., & Fersht, A. R. (1988) *Nature 335,* 740.

Schechter, I. & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27,* 157.

Schechter, I. & Berger, A. (1968) *Biochem. Biophys. Res. Commun. 32,* 898.

Sloane, B. F., Moin, K., Krepela, E., & Rhozhin, J. (1990) *Cancer Metastasis Rev. 9,* 333-352.

Sluyterman, L. A. A. (1964) *Biochim. Biophys. Acta 85,* 316-321.

Smolarsky, M. (1978) *Biochemistry 17,* 4606-4615.

Smolarsky, M. (1979) *Biochemistry 19,* 478-484.

Smolders, A., Maes, G., & Zeegers-Huyskens, T. (1988) *J. Mol. Struct. 172,* 23-40.

Storer, A. C., Phelps, D. J., & Carey, P. R. (1981) *Biochemistry 20,* 3454.

Thijs, R. & Zeegers-Huyskens, T. (1984) *Spectrochim. Acta 40A,* 307-313.

Tonge, P. J. & Carey, P. R. (1989(a)) *J. Mol. Liq. 42,* 195-212.

Tonge, P. J. & Carey, P. R. (1989(b)) *Biochemistry 28,* 6701-6709.

Tonge, P. J. & Carey, P. R. (1990) *Biochemistry 29,* 10723-10727.

Tonge, P. J., Pusztai, M., White, A. J., Wharton, C. W., & Carey,

P. R. (1991) *Biochemistry 30*, 4790-4795.

Tonge, P. J. & Carey, P. R., (1992) *Biochemistry 31*, 9122-9125.

Tonge, P. J., Carey, P. R., Callender, R., Deng, H., Ekiel, I., Muhandiram, D. R. (1993) *J. Am. Chem. Soc. 115*, 8757-8762.

Tonge, P. J., Fausto, R., & Carey, P.R. (1996) *J. Mol. Struct. (Submitted)*.

Vernet, T., Dignard, D., & Thomas, D. Y. (1987) *Gene 52*, 225-233.

Wada, A. (1976) *Adv. Biophys. 9*, 1.

Wells, J. A., Cunningham, B. C., Graycar, T. P., & Estell, D. A. (1986) *Phil. Trans. R. Soc. Lond. A 317*, 415.

Werb, Z. (1989) Proteinases and matrix degradation, in *Textbook of Rheumatology* (Keller, W.N., Harris, E.D., Ruddy, S., & Sledge, C.S., Eds.) pp 300-321, W. B. Saunder Co., Philadelphia, PA.

Whiting, A. K. & Peticolas, W. L. (1994) *Biochemistry 33*, 552-561.

Zannis, V. I. & Kirsch, J. F. (1978) *Biochemistry 17*, 2669-2674.