## NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

## AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

# SECONDARY ENZYME-SUBSTRATE CONTACTS IN PAPAIN CATALYZED ESTER HYDROLYSIS

**Paul Berti**

This thesis is submitted to the Department of Biochemistry in partial fulfillment of the requirements for the M.Sc. degree.

University of Ottawa
Ottawa, Ontario

ISBN 0-315-60088-8

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

## ABSTRACT

Secondary enzyme-substrate contacts in the hydrolysis of ester substrates by the cysteine protease papain were investigated by systematically altering backbone hydrogen bonding and side chain hydrophobic contacts in the substrate and determining each substrate's $k_{cat}/K_M$. $k_{cat}/K_M$, the second order rate constant, directly reflects the equilibrium of free enzyme plus free substrate to the enzyme-substrate transition state. The incremental binding energies of the substrate backbone hydrogen bonds are estimated at -2.3 kcal/mol for the $P_2$ NH, -2.2 kcal/mol for the $P_1$ NH, and -2.3 kcal/mol for the $P_2'$ NH. The observed hydrogen bonding energy for the substrate backbone $P_1'$ C=O is -1.0 kcal/mol. The incremental binding energy of the Phe side chain in the $P_2$ position is estimated at -4.0 kcal/mol.

The contacts on the acyl side of the scissile bond (the $P_2$ NH, $P_2$ Phe side chain and $P_1$ NH) display a strong interdependence of binding energies that is characteristic of enzyme-substrate interactions. This interdependence arises largely from the entropic cost of forming the rate-determining enzyme-substrate transition state. As favourable contacts are added successively to a substrate, the entropic penalty associated with each decreases and the binding energy expressed approaches the incremental binding energy.

On the leaving group side, the observed -1.0 kcal/mol for the $P_1'$ C=O probably underestimates the incremental binding energy whilst the -2.3 kcal/mol for the $P_2'$ NH is a reasonable estimate.

Elucidation of favourable enzyme-substrate contacts remote from the catalytic site will assist in the design of highly specific cysteine protease inhibitors.

**ACKNOWLEDGEMENTS**

I gratefully acknowledge the assistance of R.G. Carriere for preparing much of the $Hg^{++}$-papain used in this study, H. Séguin for elemental analyses, Louis Sans Cartier and Tomoko Hirama for pH-stat kinetics and Ron Angus for the synthesis of many substrates.

I would like to thank Irene Ekiel for the NMR spectrum of $MocPheOCH_2COOH$ and Barbara Gour-Salin for invaluable advice on syntheses, Sadiq Hasnain for his unflagging encouragement, Peter Tonge for brilliantly brutal criticism, Harvey Kaplan and Andy Storer for advice, criticism and encouragement; and, of course, Ma and Pa.

**TABLE OF CONTENTS**

## LIST OF TABLES

## LIST OF FIGURES

## GLOSSARY OF TERMS AND ABBREVIATIONS

All amino acid abbreviations are the standard three letter abbreviations as recommended by the IUPAC-IUB Joint Commission on Biochemical Nomenclature (*J. Biol. Chem.* (1985) **260**,14-42). All substrates containing chiral amino acid residues contain the L-stereoisomer.

| | |
|---|---|
| -C(O)↑X- | - site of cleavage in ester or amide hydrolysis |
| $X^{\ddagger}$ | - the species, equilibrium or energy refers to a transition state |
| $\Phi$ | - dihedral angle defined by the rotation of the middle bond of $C_{(n-1)}$-$NH_n$-$C_{\alpha,n}$-$C_n$. |
| $\Psi$ | - dihedral angle defined by the rotation of the middle bond of $NH_n$-$C_{\alpha,n}$-$C_n$-$NH_{(n+1)}$. |
| $X_1$ | - dihedral angle defined by the rotation of the middle bond of NH-$C_\alpha$-$C_\beta$-$X_{\gamma 1}$, where X = C, O, or S. |
| Ac- | - acetyl group |
| AcPheGlyOMe | - *N*-acetyl-phenylalanylglycine methyl ester |
| $Ph(CH_2)_2COOH$ | - β-phenylpropionic acid |
| $Ph(CH_2)_2C(O)GlyOMe$ | - *N*-β-phenylpropionyl glycine methyl ester |
| $Ph(CH_2)_2C(O)OCH_2COOH$ | - *O*-β-phenylpropionyl glycolic acid |
| $Ph(CH_2)_2C(O)OCH_2C(O)OMe$ | - *O*-β-phenylpropionyl glycolic acid methyl ester |
| $Ph(CH_2)_2C(O)OCH_2C(O)NH_2$ | - *O*-β-phenylpropionyl glycolamide |
| -*n*-Bu | - *n*-butyl group |
| $\Delta\Delta G_{obs}$ | - change in free energy between the $k_{cat}/K_M$'s of two substrates: $\Delta\Delta G_{obs} = -RT * \ln\left(\frac{(k_{cat}/K_M)_A}{(k_{cat}/K_M)_B}\right)$ |
| $Et_2O$ | - diethyl ether |

| | |
|---|---|
| EtOAc | - ethyl acetate |
| Im/$ImH^+$ | - imidazole/imidazolium ring of histidine |
| -Me | - methyl group |
| MeCN | - acetonitrile |
| Moc- | - methyloxycarbonyl group |
| MocPheOH | - methyloxycarbonyl-L-phenylalanine |
| MocPheGlyOMe | - methyloxycarbonyl-L-phenylalanylglycine methyl ester |
| $MocPheOCH_2C(O)CH_2CH_3$ | - *O*-(methyloxycarbonyl-L-phenylalanyl)-(2-oxo)-butyl ester |
| $MocPheOCH_2C(O)CH_3$ | - *O*-(methyloxycarbonyl-L-phenylalanyl)-(2-oxo)-propyl ester |
| $MocPheOCH_2C(O)OMe$ | - *O*-(methyloxycarbonyl-L-phenylalanyl)glycolic acid methyl ester |
| $MocPheOCH_2C(O)NH_2$ | - *O*-(methyloxycarbonyl-L-phenylalanyl)-glycolamide |
| Ph- | - phenyl group |
| -*n*-Pr | - *n*-propyl group |
| $S^-/ImH^+$ | - thiolate/imidazolium ion pair of papain's active site |
| THF | - tetrahydrofuran |
| THI | - tetrahedral intermediate |
| Z- | - benzyloxycarbonyl |

## 1. INTRODUCTION

The objective of the work described herein is to elucidate certain secondary (*i.e.* non-covalent) enzyme-substrate interactions in papain and their role in the catalytic mechanism. This has been accomplished by studying the kinetic constants associated with the hydrolysis of a number of ester substrates of papain, which have defined specific interactions at sites remote from the site of catalysis. Attention has been focussed on the role of the amide NH groups of the $P_1$ and $P_2$ substrate residues, on the side chain of the $P_2$ residue and on the presence of hydrogen bonding partners for the $P_1'$-$P_2'$ amide group (Fig 1.A)

### 1.1 Cysteine proteases

Papain is the most studied and consequently best understood member of the papain superfamily of cysteine proteases. Members of this superfamily have been found in many plants, in molds (*Dictyostelium discoideum*), insects (dust mite) and many forms have been isolated from mammalian tissues (cathepsins B, H, L, S, and the calpains). Sequence and structural data indicate that these proteases have all descended from a single ancestral protease (1).

The cysteine proteases are so named because their catalytic mechanism

**Figure 1.A** Binding subsites in the active site cleft of papain

Following the work of Schechter and Berger (46), the substrate binding subsites of papain can accommodate four residues on the N-terminal side of the hydrolytic site, and three on the C-terminal side. The amino acid residues (or analogues) of the substrate are designated $P_4$ to $P_1$ on the acyl side of the active site and $P_1'$ to $P_3'$ on the leaving group side. The complementary binding subsites in the active site cleft of papain are designated $S_4$ to $S_3'$.

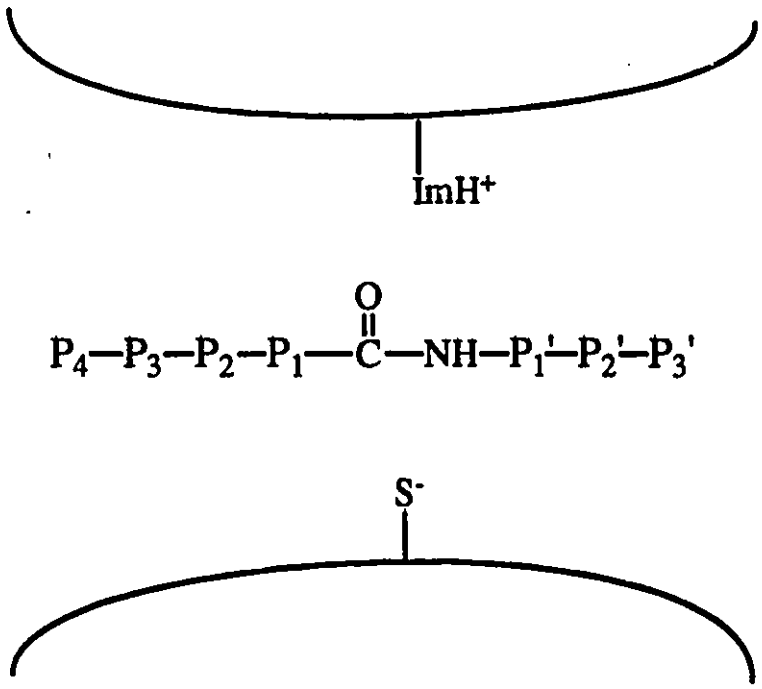



**Figure 1.A**

involves the nucleophilic attack of the catalytic site[1] cysteine's sulfur onto the scissile bond's carbonyl carbon. All the known cysteine proteases possess a pair of residues that are directly involved in the chemical steps of catalysis, the catalytic site cysteine and the catalytic site histidine, Cys 25 and His 159 (papain numbering).

### 1.1.1 Structure

The x-ray crystallographic structure of papain was first solved by Drenth *et al.* (2). It has since been refined at 1.65 Å (3), another crystal form has been solved (4) and a series of α-chloroketone inhibited papain structures have also been solved (5), as well as papain complexed with the irreversible inhibitor E-64 (6). Two other cysteine protease structures have been solved, both from plant sources, actinidin (7,8) and calotropain D1[2] (9,10). In all three cases, the structures are remarkably similar considering the 46% to 48% sequence identity between the three enzymes. The backbone atoms of all three can be superimposed with a root mean square deviation of less than 0.7 Å. The active sites are even more similar structurally; the atoms can be superimposed with a root mean square deviation of less than 0.4 Å (10,11).

---

[1] For clarity, the *catalytic site*, those residues undergoing chemical change (*i.e.* Cys 25, His 159, the scissile peptide or ester bond and the nucleophilic water in deacylation) are differentiated from other enzyme-substrate contacts that take place in the *active site* cleft.

[2] Previous reports have all referred to this enzyme as 'calotropin D1'. Since it was discovered that another enzyme of the same name exists, the cysteine protease has been changed to calotropain D1 (R. Hilgenfeld, personal communication).

The protein structures consist of two domains, which wrap 'arms' around each other (12). The N-terminus is located in the right (the 'second') domain, the chain crosses over to the left domain at residue 12. The left domain consists of residues 12-111 and 209-212 and contains three α-helices, with Cys 25 located at the N-terminus of the first. This domain adopts a 'standard' architecture for α-helical domains, with the three helices mutually perpendicular to each other (13). A long loop with no regular structure crosses from the left to the right domain, which is composed of residues 1-11 and 112-208 and consists of a bifurcated β-sheet, with two helices at the top and bottom of the β-sheet. The C-terminus of the enzyme crosses back to the left domain, becoming the 'arm' that wraps around the left domain. The active site of the cysteine proteases is located in the deep cleft separating the two domains. His 159 is in the right domain and located opposite from Cys 25 in the left domain.

The structures of the α-chloroketone inhibited papains have suggested the probable mode of binding of substrates in the active site on the acyl group side of the catalytic site; i.e. in the $S_1$ to $S_4$ subsites and will be discussed in Section 1.1.3.

## 1.1.2 Mechanism

### *1.1.2.1 Kinetics description*

The hydrolytic mechanism is minimally represented kinetically by a three step process (Scheme 1.1):

$$E + S \underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}} E{\cdot}S \xrightarrow{k_{+2}} EA \xrightarrow[+Prod2]{k_{+3}} E$$

(+ Prod1)

**Scheme 1.1**

Under steady state conditions with $[S]_0 \gg [E]_0$:

$$k_{cat} = k_{+2}k_{+3}/(k_{+2} + k_{+3})$$

$$K_M = (k_{+3}(k_{+2} + k_{-1}))/(k_{+1}(k_{+2} + k_{+3}))$$

$$k_{cat}/K_M = k_{+2}k_{+1}/(k_{+2} + k_{-1})$$
$$= k_{+2}/K_S \quad \text{(assuming } k_{-1} \gg k_{+2})$$

$$v_0/[E]_0 = k_{cat}[S]_0/([S]_0 + K_M)$$

The enzyme and substrate bind non-covalently in the first step to form the Michaelis complex with a dissociation constant $K_S = k_{-1}/k_{+1}$. The next step is acylation, occurring with a rate constant $k_{+2}$. The carbonyl carbon undergoes nucleophilic attack from the active site cysteine to form the acylenzyme intermediate and the first product (Prod1), an alcohol or amine. The acylenzyme intermediate then undergoes nucleophilic attack by water with a rate constant $k_{+3}$ to give the acid product (Prod2). The reverse of the acylation and deacylation steps are regarded as being negligible, particularly in determining $v_0/[E]_0$'s where [products] = 0, though under some conditions $k_{-2}$ and $k_{-3}$ can become significant. By appropriate manipulation of reaction conditions, papain can be used to synthesize ester and amide bonds (14-16).

The rate determining step of hydrolysis is dependent on the substrate. For amide substrates acylation is generally slower than deacylation ($k_{+2} < k_{+3}$). Frequently the difference in rates is great enough for $k_{+2}$ to become the rate determining step. Under these conditions $k_{cat} \approx k_{+2}$ and

$K_M \approx K_S$. For ester substrates deacylation is generally slower ($k_{+2} > k_{+3}$). For very specific ester substrates or for very good leaving groups such as *p*-nitrophenols, $k_{+3}$ is rate determining. Under these conditions $k_{cat} \approx k_{+3}$ but $K_M$ [$= k_{+3}(k_{-1}+k_{+2})/(k_{+1}k_{+2})$] remains a complicated constant that is not useful on its own.

*1.1.2.2 Chemical description*

Figure 1.1 provides a more detailed description of the catalytic mechanism. The enzyme and substrate bind non-covalently in the first step to form the Michaelis complex. The carbonyl carbon then undergoes nucleophilic attack by the active site Cys 25, and is believed to pass through a transient tetrahedral intermediate (THI). The His 159 imidazolium then acts as a general acid catalyst, donating a proton to the leaving group, forming the acylenzyme intermediate. Deacylation proceeds via an analogous pathway, using a water molecule as the nucleophile to hydrolyze the enzyme-substrate thiol ester. His 159 acts as a general base catalyst in this step, abstracting a proton from the water molecule.

*Thiolate/Imidazolium pair*

The sulfur of Cys 25 is believed to exist in an active form as a thiolate anion which is ion paired with the imidazolium of His 159. The $k_{cat}/K_M$ profile of papain catalyzed substrate hydrolysis is bell-shaped with $pK_a$'s ≈4.2 and ≈8.5 and the maximum near pH 6.5 (17). The acid $pK_a$ then represents ionization of the Cys 25 thiol to form the $S^-/ImH^+$ ion pair, with the basic $pK_a$ representing deprotonation of the imidazolium

**Figure 1.1** Schematic representation of cysteine protease catalyzed ester hydrolysis.

**E·S** ≡ Michaelis complex
**THI** ≡ Tetrahedral intermediates of acylation and deacylation
**EA** ≡ Acylenzyme intermediate
**E+P** ≡ Free enzyme plus acid product

$S^-$ is the thiolate anion of Cys 25; the imidazolium is contributed by His 159. Amide hydrolysis follows an identical mechanism.

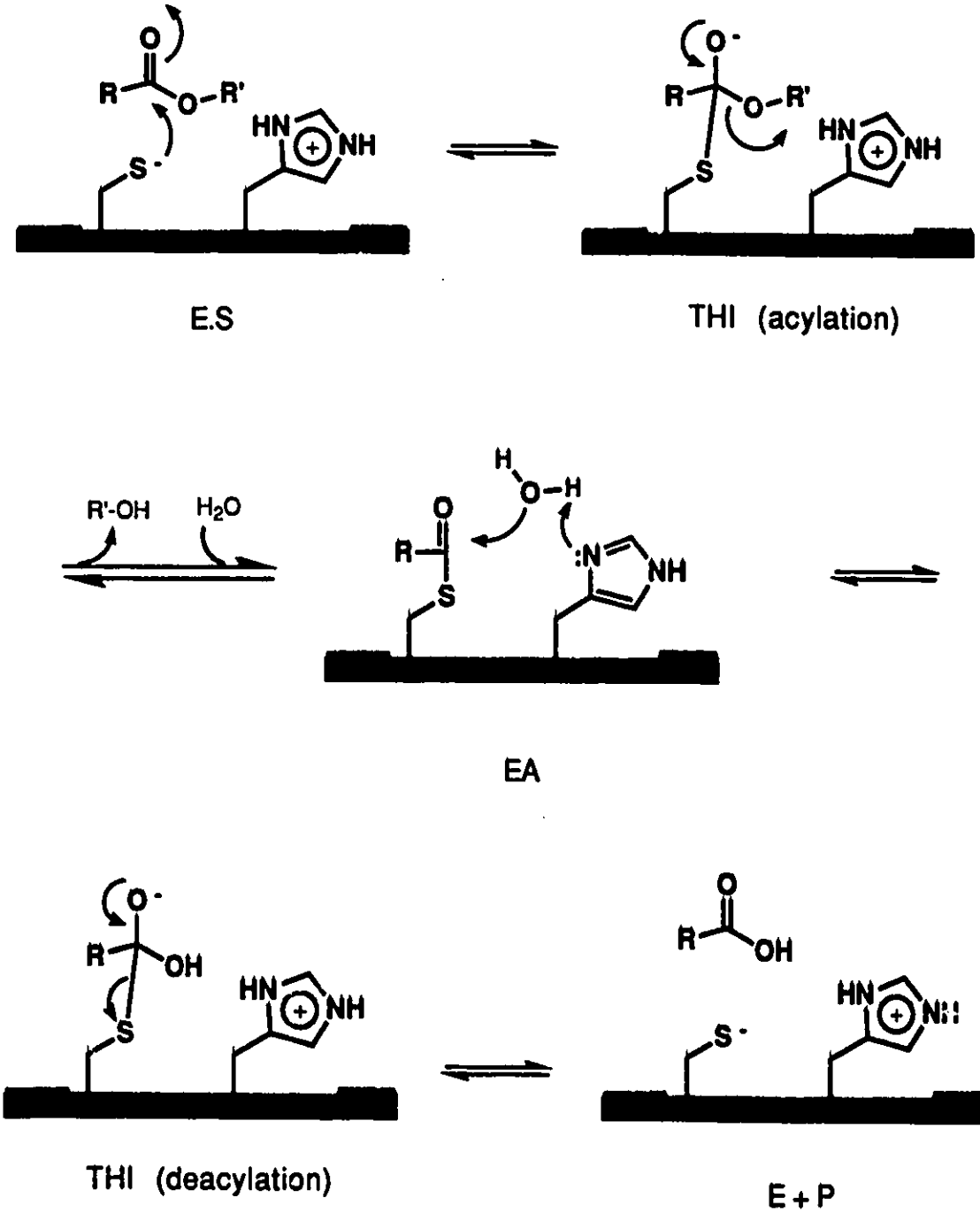



**Figure 1.1**

of His 159. There is some disagreement whether the existence of the $S^-/ImH^+$ pair has been rigorously demonstrated (18,19).

Evidence for the existence of the $S^-/ImH^+$ ion pair has included observation of the thiolate by UV absorption spectroscopy (20), potentiometric difference titration (21) and direct proton NMR observation (22,23). However, as discussed by Dixon (24), attempts to quantitate group ionization constants are unreliable unless it can be shown that the properties of the two ionizations under investigation are independent of each other, which is not the case here.

Additional evidence includes the pH dependence of the fluorescence of Trp 177, which is quenched by the His 159 imidazolium with a $pK_a = 8.6$ (25,26) and kinetic deuterium solvent isotope effects. Unlike deacylation in cysteine protease catalyzed hydrolyses, and unlike both acylation and deacylation by serine proteases, acylation by cysteine proteases is characterized by no effect or a small inverse kinetic isotope effect, indicating that there is no general base catalysis of acylation and that the thiolate is formed before the nucleophilic attack step (see 1). Finally, calculations on the papain (27,28) and actinidin (29) active sites using a variety of computational techniques indicate that the $S^-/ImH^+$ ion pair is more stable in the enzymes' active sites than the neutral SH/Im.

*Tetrahedral intermediate (THI)*

Nonenzymic nucleophilic substitution of esters and amides is believed to proceed through a THI (30). THI's are also generally believed to occur in cysteine protease catalyzed hydrolyses. Protease-bound THI's have

not been observed directly, consistent with the conclusion by Fastrez (31) that THI's are too unstable to accumulate measurably in cysteine or serine protease catalyzed hydrolyses. There is, however, considerable circumstantial evidence for their formation.

The effects of electron withdrawing groups on both acylation and deacylation suggest THI formation. A positive Hammett ρ value of 1.2 is observed for $k_{cat}/K_M$ with substituted aryl esters of *N*-benzoylGly (*90*), suggesting rate determining formation of the THI. The *para*-substitution of the anilides of *N*-benzoylGly and *N*-AcPheGly gives a Hammett ρ value of -1.04, suggesting the rate determining breakdown of the THI of acylation (33). O'Leary *et al.* (34) determined a nitrogen kinetic isotope effect ($k(^{14}N)/k(^{15}N)$ = 1.024 at pH = 6.0) close to the limit exptected for hydrolysis of *N*-benzoylArg$\hat{}$NH$_2$, indicating C-N bond cleavage (i.e. THI breakdown) in the rate determining step.

Modelling the THI of acylation into the active site based on the α-chloroketone inhibited papain structures (5) suggests that the negatively charged oxygen of the THI, the oxyanion, is placed within hydrogen bonding distance of two hydrogen bond donors, the Cys 25 backbone NH and Gln 19 $N_\varepsilon$. These two hydrogen bond donors will then stabilize the negative charge that accumulates on the oxyanion formed in the putative THI and therefore reduce the free energy of activation of acylation and deacylation. Substitution of the Gln 19 of papain to Ala significantly reduces $k_{cat}/K_M$ (A. Storer, J. Carrière, personal communication). Additionally, the Gln 19 side chain is absolutely conserved in all known cysteine protease sequences (unpublished data), another indication of the importance of the oxyanion hole.

Polgar *et al.* have noted that cysteine proteases can hydrolyze thiono esters (-C=S-O-) with $k_{cat}/K_M$'s similar to the analogous dioxygen esters whilst serine proteases cannot catalyze thiono ester hydrolysis (35,36). Since sulfur is larger than oxygen and inherently a poorer hydrogen bond acceptor, they have concluded that serine proteases cannot stabilize the sulfur anion in the transition state. The lack of effect then with cysteine proteases is seen as evidence that they do not possess an oxyanion hole. However, serine proteases have strong amino acid specificities which are expressed in the $S_1$ subsites and are therefore quite sterically restricted near the catalytic site. In contrast, cysteine proteases are not as sterically restricted at the catalytic site and should be able to compensate for the added size of the sulfur atom (37). If the steric constraints in the $S_1$ subsite are increased by changing the $P_1$ residue from Gly to Ala, the thiono ester substrates go from being better substrates than the dioxygen esters to being significantly worse (38). This implies that the additional steric constraints introduced by the side chain of Ala at $P_1$ makes hydrogen bonding to the sulfur less favourable, and by analogy suggests that thiono esters are poor substrates of serine proteases due to steric constraints.

*Acylenzyme intermediate*

Formation of the thiol ester intermediate in the hydrolysis of esters and amides was first hypothesized on the basis of kinetic evidence and has since been observed directly by: 1.) UV absorbance detection of chromophoric substrates with conjugated double bond systems (*trans*-cinnamoyl-papain intermediates) (39). 2.) UV absorbance

detection of dithioacyl papains formed during hydrolysis of thiono ester substrates (40). 3.) Resonance Raman spectroscopy of the dithioacylenzyme intermediates of thiono ester hydrolysis (41). 4.) Detection of the thiolester intermediate in the hydrolysis by papain of *N*-benzoyl im dazole labelled with $^{13}C$ at the carbonyl carbon by $^{13}C$ NMR at low temperature (42).

### 1.1.3 Specificity

Studies on sites of cleavage of naturally occurring polypeptides such as insulin B-chain (see 1) and glucagon (43) provide no strong indications of cysteine proteases' specificity. Schechter and Berger (44-46) 'mapped' out seven subsites in papain's active site, four on the acyl group side ($S_1$ to $S_4$) and three on the leaving group side ($S_1'$ to $S_3'$). They also identified a specificity for Phe at the $S_2$ subsite. In general, cysteine proteases have specificity for large hydrophobic residues at $S_2$. Papain has a specificity for large hydrophobic residues at $S_1'$ (39,47) and tends to prefer larger non-branched residues at $S_1$ as well (see 1). The importance of the $P_2$-$P_1$ peptide bond has also been recognized (32). While these individual contributions to specificity have been recognized, the inability to discern unambiguous specificity patterns from peptide cleavage indicates clearly that a number of favourable enzyme-substrate contacts at different sites contribute to specificity and no single contact is indispensable (*cf.* serine proteases such as trypsin which has a strong specificity for Lys or Arg at $S_1$ or chymotrypsin, which requires large hydrophobic residues such as Trp, Tyr or Phe at $S_1$ (48)).

Interpretation of kinetics results relating to specificity have relied heavily on the x-ray crystallographic structures of papain, particularly on the α-chloroketone inhibited papain structures (5). Figure 1.2 provides a schematic representation of enzyme-substrate contacts suggested by the Z-Phe-Ala-$CH_2$-S-papain structure. The $P_1$ NH···O=C $S_1$ (Asp 158) hydrogen bond is marked with a question mark to indicate that in the x-ray structure the distance between the nitrogen and oxygen atoms is too long for strong hydrogen bonding to occur. This is probably caused by the extra methylene group in the structure as compared with substrates. The importance of the $P_1$ NH has been demonstrated kinetically (32).

**Figure 1.2** Schematic representation of contacts in the Z-PheAla-$CH_2$-papain complex.

Starting from the top of the diagram, the benzyloxycarbonyl group occupies the $S_3$ subsite, with the C=O of the benzyloxycarbonyl group in the position of a $P_3$ C=O in a peptide substrate. The $P_2$ residue hydrogen bonds in an antiparallel fashion with the Gly 66 backbone C=O and NH. Aside from Val 133 and Val 157, the $P_2$ Phe side chain also makes contact with the side chains of Ala 160 and Pro 68. In the x-ray structure the $P_1$ NH is too far from the Asp 158 backbone C=O to form an hydrogen bond (indicated by a question mark); but this appears to be due to the presence of the extra methylene group. The $P_1$ NH has been shown previously to be kinetically important (32). Gln 19 $N_\varepsilon$ and the Cys 25 backbone NH form the oxyanion hole and hydrogen bond the oxygen in the tetrahedral intermediate.

The x-ray structure is from Drenth *et al.* (5); the figure is modified with permission from (6).

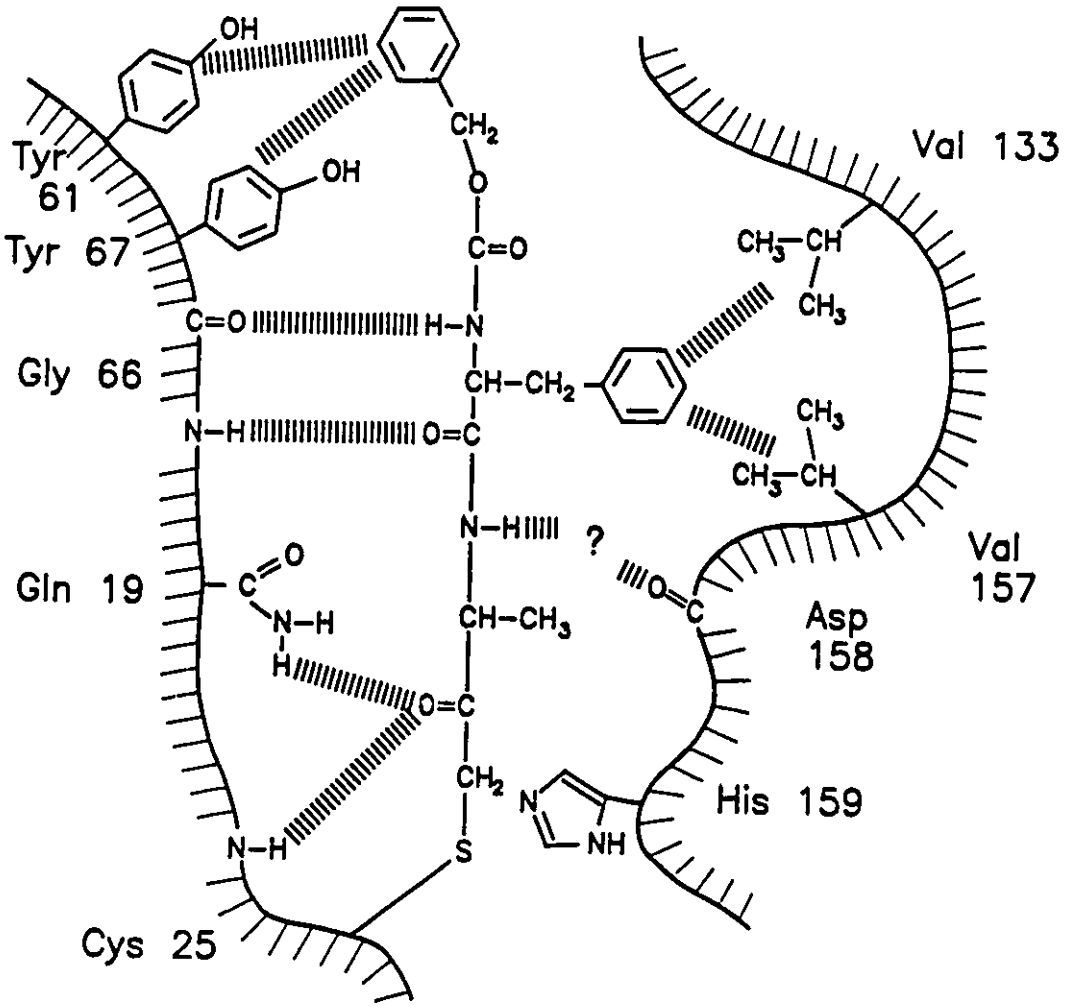

OH
Tyr
61
OH
Tyr 67
CH2
O
C=O
Val 133
CH3–CH
CH3
C=O
H–N
Gly 66
CH–CH2
N–H
O=C
CH3
CH3–CH
N–H
?
Gln 19
C=O
N–H
H
CH–CH3
O=C
Val
157
O=C
Asp
158
CH2
N
NH
His 159
N–H
S
Cys 25

**Figure 1.2**

The disulfide inhibitors of Brocklehurst *et al.* (49) have recently been used to examine cysteine protease specificity. Their use requires the assumption that the inhibitor-enzyme interactions parallel those between those of the substrate-enzyme complexes of which they are analogs. This assumption is made in spite of the fact that with many cysteine proteases this is clearly not true (50,51), and in spite of the large difference between the geometries of the transition states of disulfide exchange (trigonal bipyramid with the nucleophile and leaving groups as axial substituents) (52,53) and acyl hydrolysis (tetrahedral), and the postulated accumulation of negative charge delocalized over the three sulfurs in the disulfide transition state (54), which would explain the extreme sensitivity of these inhibitors to enzyme ionizations (55,56).

Brocklehurst *et al.* have attempted to demonstrate the effect of a Phe at $P_2$ by comparing the inhibition kinetics of disulfide reactivity probe analogs of the substrates *N*-AcGly and *N*-AcPheGly. By referring to these inhibitors as demonstrating the importance of the Phe side chain, they have assumed that the $P_3$-$P_2$ amide group that is also added has no affect on the kinetics. We show this not to be true. Further, they describe 'signalling' and 'non-signalling' interactions to explain their results; that the $P_2$ C=O, plus one of the $P_2$ Phe or the $P_1$ NH is required to 'signal' the catalytic site and enhance catalysis (57). They do this without thoroughly explaining the concept or rationalizing its physical basis.

Another recent specificity investigation has included probing the kinetic importance of the $P_2$ C=O by substituting a sulfur for the carbonyl oxygen (58). The decreased specificity upon addition of the

thionoamide was worth 4.4 kcal/mol for papain. This is compared with a difference of 1.6 kcal/mol observed for the same $P_2$ sulfur for oxygen exchange with a pair of poorer substrates (59). This difference in observed energies is consistent with the effects seen in this study at sites surrounding the $P_2$ C=O, that is at the $P_2$ NH, the $P_2$ side chain, and the $P_1$ NH.

Thus, in this study we have considerably augmented the existing knowledge of the enzyme-substrate interactions on the acyl group side of the catalytic site, have corrected some incorrect assumptions about the behaviour of substrates in the active site and have demonstrated that it is not necessary to introduce new or 'special' effects to explain our results.

## 1.2 Catalysis: mechanisms and kinetics

### 1.2.1 Transition state theory

Enzymes catalyze reactions by stabilizing the transition state (60). The transition state theory is less than 50 years old, and its application to enzymology is even more recent. As such, its application has sometimes been frustrated by the proliferation of terminologies and approaches (61) and sometimes misunderstanding of the theory's precepts, even by eminent workers in the field (62,63). Nonetheless, the power and broad applicability have made its use indispensable, and with wider employment it has become more easily understood and applied.

Transition state theory considers only two entities in a reaction, the ground state and the most unstable species on the reaction path, the

transition state. The transition state is the highest energy *point* on the lowest energy *path* for the reaction (Fig 1.3). If, for simplicity, one assumes a pseudo-equilibrium[3] between the ground state and the transition state:

$$A + B \overset{K^{\ddagger}}{\rightleftharpoons} AB^{\ddagger},$$

then the rate, $k$, is directly proportional to $[AB^{\ddagger}]$, which in turn is inversely proportional to dissociation constant, $K^{\ddagger}$. This is expressed as $k = \kappa\nu/K^{\ddagger}$, where $\kappa$ is transmission coefficient (which is usually approximately unity) and $\nu$ is the average frequency of barrier crossing (63). If the free energy of the transition state is decreased (*i.e.* stabilized) by $\Delta G$, then $K^{\ddagger}_{new} = K^{\ddagger}_{old} * \exp(\Delta G/RT)$, where R = the gas constant and T = temperature in Kelvin. The rate will increase by $K^{\ddagger}_{old}/K^{\ddagger}_{new}$.

For any chemical reaction there is at least one transition state between the reactants and products. For multistep reactions each step has associated with it a transition state. In cysteine protease catalyzed hydrolyses, there are transition states separating the reactants, the THI of acylation, the acylenzyme intermediate, the THI of deacylation and the products. If one chemical step is significantly slower than the others, it becomes the rate determining step. For good ester substrates of cysteine proteases, the rate determining step is often formation of

[3] It is possible to arrive at the same free energy relationships without the equilibrium assumption (98).

**Figure 1.3** Reaction profile for a reaction proceeding through one high energy intermediate.

Since Transition State 1 is higher energy than Transition State 2, it will represent the rate determining transition state. The population of the intermediate will be determined by the relative heights of the energy barriers on either side. The transition state represents the highest *point* on the lowest energy *path* from reactant to product. If plotted in three dimensions, the transition state would be a 'saddle point', like a pass between two mountains.

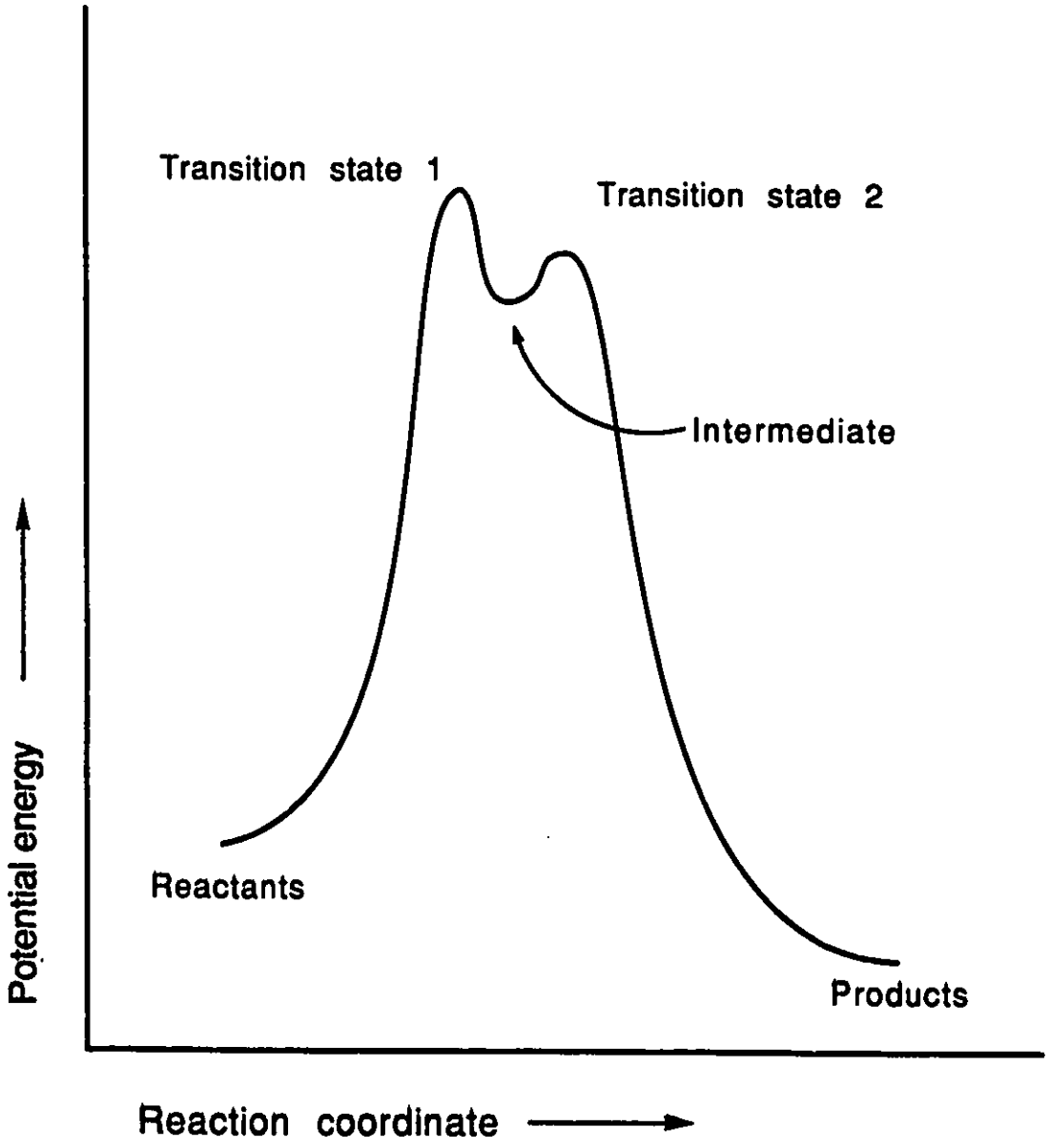

Transition state 1
Transition state 2
Intermediate
Reactants
Products
Potential energy
Reaction coordinate

**Figure 1.3**

the THI of deacylation (32). For amides, breakdown of the THI of acylation is often rate determining (33).

### 1.2.2 Enzymic catalysis

The most notable features of enzymic catalysis are the rate enhancement and specificity over non-enzymic reactions. Substrate molecules generally consist of a reactive functionality (*e.g.* an ester group) and many non-reacting atoms. Enzymes use binding energy provided by interactions with non-reacting parts of the substrate to stabilize the transition state; thus simultaneously providing selectivity against non-specific substrates and reaction rate enhancement. Enzymes also directly interact with the substrates' reactive functionality to stabilize the transition state, providing rate enhancement.

Free energy diagrams can be used to illustrate catalysis (Fig. 1.4). The rate enhancement is equal to $\Delta G_T^{\ddagger}$, the free energy of transition state stabilization. In the thermodynamic cycle (Fig. 1.4a), $k_{+2} \propto 1/K_E^{\ddagger}$ and $K_S$ is the dissociation constant as used in Michaelis-Menten kinetics.

**Figure 1.4a** Thermodynamic cycle comparing an uncatalyzed (top row) *versus* an enzymically catalyzed reaction.

$E + S$ ≡ Free enzyme plus free substrate in the ground state
$E + S^{\ddagger}$ ≡ Free enzyme plus free substrate transition state
$E \cdot S$ ≡ Michaelis complex
$ES^{\ddagger}$ ≡ Enzyme-substrate transition state.
$E + P$ ≡ Free enzyme plus free product in the ground state

$K_S \equiv [E]*[S]/[E \cdot S]$
$K_T^{\ddagger} \equiv [E]*[S^{\ddagger}]/[ES^{\ddagger}]$
$K_N^{\ddagger} \equiv [S]/[S^{\ddagger}]$
$K_E^{\ddagger} \equiv [E \cdot S]/[ES^{\ddagger}]$

$\Delta G_X = RT*\ln(K_X)$

It is a necessary consequence of the statement that enzymes catalyze reactions by stabilizing the transition state that $K_E^{\ddagger} < K_N^{\ddagger}$ and therefore that $K_T^{\ddagger} < K_S$, that is, the enzyme has a higher affinity for the transition state of the catalyzed reaction than for the reactants' ground states.

**Figure 1.4b** Free energy diagram for an uncatalyzed *versus* an enzymically catalyzed reaction.

The specificity of papain for a particular substrate is measured as $k_{cat}/K_M$. As demonstrated in Section 1.2.2.2, $\Delta G_{spec}^{\ddagger} = RT*\ln(k_{cat}/K_M)$.

Because $k_{cat}/K_M$'s in this study are expressed in units of $\underline{M}^{-1} \cdot s^{-1}$, the standard state for the $k_{cat}/K_M$'s is for all species to be at 1 $\underline{M}$ concentration.

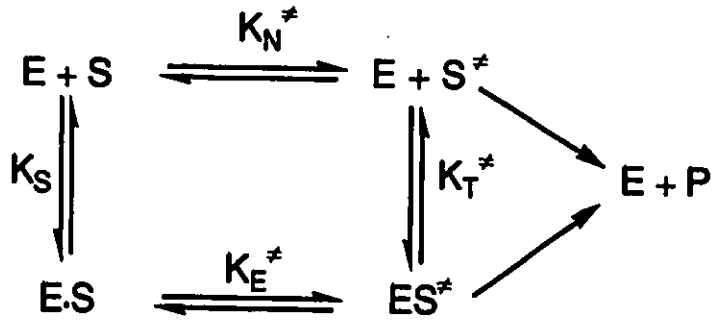



**Figure 1.4a**

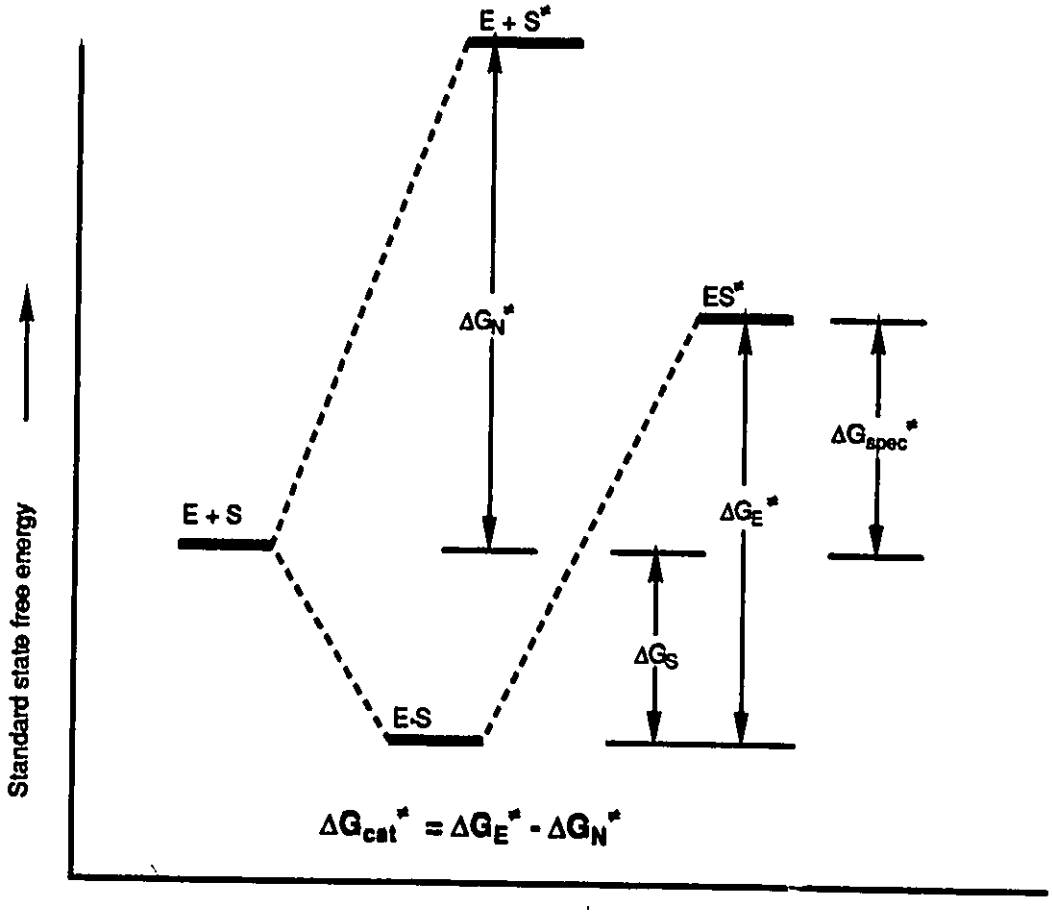



**Figure 1.4b**

*1,2,2,1 Enzymic transition state stabilization*

Enzymes can stabilize the transition state by entropy compensation or by destabilization of the ground state at the catalytic site towards the transition state.

*Entropy compensation*

All enzymically catalyzed reactions undergo some degree of entropy compensation. In enzymic catalysis the first step is the bringing together of the enzyme and the substrate to form the Michaelis complex, E·S. This entails a loss of free energy in the form of entropies of translation and overall rotation of the two molecules relative to each other. Attempts to quantify the free energy loss in this process have centred on the study of the relative rates of inter- and intra-molecular reactions. Intramolecular reactions (*e.g.* cyclizations), in the absence of conformational strain are faster than the equivalent intermolecular reactions by a factors of up to $10^8$ (64,65). This energy of approximately 10 kcal/mol at 300 K represents the loss of entropy on forming the transition state in the intermolecular reaction.

Alternately, the loss of entropy on formation of the enzyme-substrate transition state may be estimated as follows. Since the translational and overall rotational entropy of a molecule are only weakly dependent on the molecule's size (62), for all the substrates in this study the translational and overall rotational entropies will be similar. The change in translational and overall rotational entropy for the enzyme molecule on binding the substrate will be negligible. Thus, the loss of

entropy on formation of the enzyme-substrate transition state is approximately equal to the translational and overall rotational entropy of the substrate molecule in solution, which can be estimated to be approximately 40 to 50 cal/K/mol (65), less the internal entropy of the transition state. Since acyl transfer reactions proceed through 'tight' transition states (i.e. there is a large amount of covalent bond formation in the transition state and therefore its internal entropy is small) (65), the internal entropy about the atoms undergoing chemical transition will be low, on the order of 15 cal/K/mol (66). Therefore the entropy change on forming $ES^{\ddagger}$ is on the order of -35 to -45 cal/K/mol and is only weakly dependent on the identity of the substrate. This corresponds to a binding energy of -10 to -13 kcal/mol.

If an enzyme does nothing more in catalyzing a bimolecular reaction than to bind the two reactants in the active site, then the free energy of activation will be reduced by ≈10 kcal/mol; indeed this appears to be the primary mechanism of catalysis of β-galactopyranosyl pyridinium salt hydrolysis by β-galactosidase (see 66).

*Ground state destabilization*

Destabilization of the ground state at the catalytic site is a nebulous concept that entails a number of different possible mechanisms. This destabilization can be geometrical or electrostatic in nature. Neither the enzyme nor the substrate contains immovable constituent groups nor can they develop irresistible forces to exert on the other. Invoking strain in the catalytic mechanism implies strain and movement on the part of both the enzyme and the substrate, the distinction lies in whether it is the distortion of the substrate or the enzyme that

enhances catalysis. In induced fit, binding energy used to alter the conformation the enzyme enhances its reactivity. In geometrical (substrate) destabilization, binding energy is used to distort the substrate, making the distorted bond weaker and more reactive. Electrostatic destabilization may take the form of desolvating a charged reactant into a hydrophobic environment, destabilizing the substrate towards a neutral transition state. Alternatively, placing charged or hydrogen bonding groups in the active site to stabilize charges formed in the transition state will destabilize the substrate towards the transition state. The oxyanion hole of cysteine (19) and serine (48) proteases is an example of this.

*1.2.2.2 Kinetic constants and* $\Delta\Delta G_{obs}$

The conventional Michaelis-Menten kinetic constants can be related to the energies defined in figure 1.4:

$$K_S = \exp(\Delta G_S/RT)$$

$$k_{+2} \propto 1/K_E^{\ddagger} \propto \exp(\Delta G_E^{\ddagger}/RT)$$

$$k_{+2}/K_S \propto \exp(-(\Delta G_E^{\ddagger}+\Delta G_S)/RT) = k_{cat}/K_M \text{ (Section 1.1.2.1)}$$

Thus it is seen that $k_{cat}/K_M$, the specificity constant, represents the equilibrium between the *free* enzyme plus *free* substrate to the enzyme-substrate transition state:

$$E + S \overset{K_{spec}}{\rightleftharpoons} ES^{\ddagger},$$

where $k_{cat}/K_M \propto 1/K_{spec}$, $\Delta G_{spec} = RT * \ln(K_{spec})$

Since in the equation $k_{+2} = \kappa\nu/K_E^{\ddagger}$, $\kappa$ and $\nu$ are not known, and since the equation itself is only approximate, $\Delta G_{spec}$ cannot be determined

directly. However, if $k_{cat}/K_M$'s are determined for two substrates, *A* and *B*, $\Delta\Delta G_{obs}$, the difference between $\Delta G_{spec,A}$ and $\Delta G_{spec,B}$ can be determined:

$$\Delta\Delta G_{obs} = -RT * \ln\left(\frac{(k_{cat}/K_M)_A}{(k_{cat}/K_M)_B}\right)$$

Thus, if substrates *A* and *B* differ by the presence of a hydrogen bond donor (*e.g.* an amide NH) and if *A* is a better substrate, then $\Delta\Delta G_{obs}$ will be negative, indicating better transition state stabilization in *A*. In this way it is possible to determine the free energy of interaction of specific portions of the substrate molecule.

$k_{cat}/K_M$ reduces to $k_{+2}/K_S$ for the papain catalyzed hydrolysis of esters and amides. An improvement in $k_{cat}/K_M$ can be achieved by tighter binding on forming the Michaelis complex (*i.e.* lower $K_S$) or by lowering the activation energy of the reaction (*i.e.* higher $k_{+i}$). Comparison of a single pair of $k_{cat}/K_M$'s does not permit distinction between the two. With substrates where acylation ($k_{+2}$) is known to be rate determining, a simple comparison of $k_{cat}$ between two substrates will indicate whether the observed change in $k_{cat}/K_M$ is due to changes in $K_S$ or $k_{+2}$. With substrates where both $k_{+2}$ and $k_{+3}$ contribute significantly to $k_{cat}$, or where $k_{+3}$ is rate determining, it is not possible to determine which effect is responsible for a change in $k_{cat}/K_M$.

A decreased $K_S$ alone does not lead to catalytic enhancement. When [S] > $K_M$, the accumulation of the Michaelis complex simply prevents competitive substrates from binding to the enzyme. This is not a useful strategy for an enzyme to adopt *in vivo*. Indeed, it has been argued that a perfectly evolved enzyme, for a given $k_{cat}/K_M$, will have as high

a $K_M$ as possible, with as much binding energy as possible going into lowering the activation energy (62,67).

*1.2.2.3 Incremental and intrinsic binding energy*

In adding a particular substituent to a substrate, one seeks to determine how favourable the interaction is of this substituent alone with the enzyme, that is, to determine the incremental binding energy (56,68). The incremental binding energy of a substituent is the energy of interaction with the enzyme in the absence of extraneous effects such as steric or electrostatic interference by other parts of the substrate or by differences in entropy between the enzyme-substrate complexes of the substrates with and without the substituent in question. The incremental binding energy of a substituent is the maximum strength of *that* substituent specifically interacting at *that* site on *that* enzyme. The intrinsic binding energy is defined as the maximum binding energy that can be expressed by that substituent's chemical group interacting with perfect complementarity with any protein (56).

As pointed out by Fersht (68,69), several criteria must be met in order to determine the incremental binding energy of a substituent on a substrate. Firstly, it is necessary to *delete* the substituent under investigation since substitution with another group that can interact with the enzyme will lead to a strict specificity effect. Secondly, the deletion must be such that solvent has free access to the binding site on the enzyme, otherwise the energy cost of excluding solvent will be included in the observed energies. Thirdly, there must be no loss of entropy in the enzyme-substrate complexes between substrates in the

presence and absence of the substituent. This becomes only approximately true for very good substrates (see Section 4.4.1); so one may only determine, under ideal conditions, the lower limit of a group's incremental binding energy.

Clearly, deletion of a group located in the middle of a substrate molecule, such as the $P_1$ NH, is impossible. If that group is substituted with another, then potential interactions of the new substituent with the enzyme will interfere with determination of the incremental binding energy; that is, $\Delta\Delta G_{obs}$ is purely a specificity energy rather than an incremental binding energy. At best, when considering a substitution one can determine a quantity that will be referred to here as the *incremental specificity energy*, by analogy to the incremental binding energy. Use of the phrase 'incremental specificity energy' makes the same assumptions as determination of incremental binding energy except that the interaction with the enzyme of the group that was substituted in is included. If the energy of interaction of the group substituted in can be estimated independently, then the incremental binding energy can be estimated indirectly.

### 1.2.3 Limitations of kinetics determinations

Changes in $k_{cat}/K_M$ between two substrates indicates a change in $\Delta G_{spec}$; no structural information is provided *per se*. Structural information must derived by comparison of $\Delta\Delta G_{obs}$ for a given substituent with other kinetics results and by consideration of other, structure-related data such as x-ray structures or spectroscopic results. As such, any conclusions regarding the nature and structure of enzyme-substrate

interactions that are reached in this study are extrapolations from the available data and do not arise directly from that data. That is, the favourable interaction of, for example, a $P_1$ NH can be demonstrated directly by comparing $k_{cat}/K_M$'s of substrates in the presence and absence of this interaction. The deduction that there is hydrogen bonding between the $P_1$ NH and Asp 158 C=O is an interpretation of the available kinetic and structural data.

## 1.3 Secondary papain-substrate contacts

Much of the recent work on cysteine proteases has focussed on their inhibition, by naturally occurring or synthetic small inhibitors (6,70), or by protein inhibitors from the cystatin superfamily of cysteine protease inhibitors (71,72). This focus on inhibition arises from the implication of cysteine proteases in many disease states, including muscular dystrophy (defective inhibitor allows degradation of muscle tissue), myocardial infarction ($[Ca^{++}]$ is released, activating a $Ca^{++}$ dependent cysteine protease and causing tissue damage), cancer metastasis (tumour invasiveness), and parasitic and bacterial pathogenicity (see 70). Recently it was demonstrated that a synthetic cysteine protease inhibitor was effective in protecting mice after they were injected with a lethal dose of group A *streptococci* (73). Clearly, designing more effective cysteine protease inhibitors requires a knowledge of what potential interactions may be exploited to increase inhibitor-protease affinity.

We have set out to map interactions within the active site of papain by systematically varying parts of the substrate molecule and noting the

effect on $k_{cat}/K_M$ of these changes. One of the substrates used, MocPheOCH$_2$C(O)OMe, contains two ester functionalities and may be cleaved at both sites (Fig. 1.5). High pressure liquid chromatography (HPLC) was used to determine if cleavage was occurring at both sites and to quantitate the products formed. Whilst the potential hydrogen bonding of the $P_2$ NH to Gly 66 C=O was shown by the x-ray structures of α-chloroketone inhibited papain, its kinetic importance have never been demonstrated. Information on binding of the substrate's leaving group is scarce due to a paucity of kinetic studies (39,47) and lack of structural information. It had been hypothesized by several workers beforehand that E-64 would bind on the leaving group side of the catalytic site of papain and thus provide structural information on binding in this region of the active site. Solution of the x-ray structure of the papain—E-64 complex showed that it binds on the acyl group side of the catalytic site (6). The substrates used in this thesis with glycolamide as a leaving group, along with structurally related leaving groups has provided evidence for hydrogen bonding partners for the backbone $P_1'$ C=O and $P_2'$ NH groups. The strength of these interactions have been compared with those of the $P_1$ NH and the $P_2$ side chain. The results indicate that all these interactions are important and could all be utilized in the design of more effective cysteine protease inhibitors.

**Figure 1.5** Structure of MocPheOCH$_2$C(O)OMe indicating the sites and products of hydrolysis at the two ester functionalities.

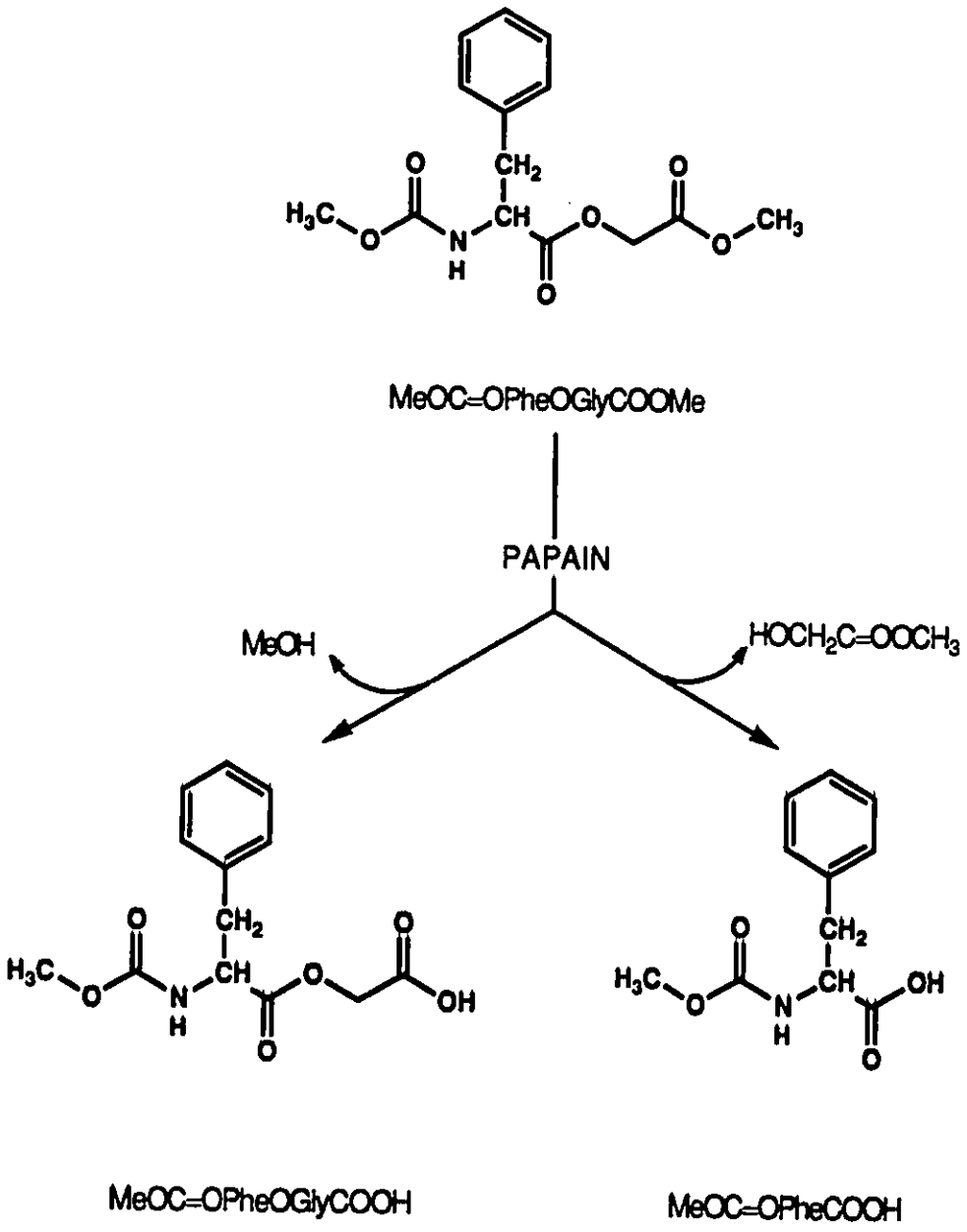

MeOC=OPheOGlyCOOMe
PAPAIN
MeOH
HOCH2C=OOCH3
MeOC=OPheOGlyCOOH
MeOC=OPheCOOH

**Figure 1.5**

## 2. MATERIALS AND METHODS

Except as noted, curve fittings were done using a non-linear least squares software package, Enzfitter (Elsevier Biosoft).

$Hg^{++}$-papain was prepared from papain (Sigma, Type III, 2 × crystallized in suspension in 50 mM NaOAc, pH 4.5) using the method of Sluyterman and Wijdenes (74).

BDH Omnisolv grade MeCN and ultrapure $H_2O$ were used for HPLC. MocPheCOOH was synthesized by R. Angus from L-phenylalanine using methyl chloroformate (75). *N*-AcPheGlyOMe was synthesized by A. Storer. Other substrates whose synthesis is not described below were synthesized by R. Angus with the assistance of R.G. Carriere (37,38). All other reagents and solvents were the highest available grade.

### 2.1 Syntheses

#### 2.1.1 $Ph(CH_2)_2COOH$

$Ph(CH_2)_2COOH$ was prepared from 3 g (19 mmol) $Ph(CH_2)_2C(O)Cl$ by stirring 1 hr. in 50 ml 5% (w/v) $NaHCO_3$. The solution was made to pH 10 with 1 N NaOH, up to 400 ml with $H_2O$, extracted with *ca.* 100 ml hexane, brought to pH 1 with 2 N HCl, extracted with 3 × 300 ml EtOAc. The second and third cuts, which were purer by TLC than the first, were pooled, dried under vacuum and allowed to solidify overnight, then dried thoroughly. This was used for the synthesis of $Ph(CH_2)_2C(O)OCH_2C(O)OMe$, $Ph(CH_2)_2C(O)OCH_2C(O)NH_2$ and $Ph(CH_2)_2C(O)OCH_2COOH$. After extensive drying

under vacuum, it was redissolved in EtOAc, and the insoluble material filtered off. This was submitted for elemental analysis and was used for the HPLC standard curve with PB6 (see below).

## 2.1.2 Glycolic Acid Methyl Esters and Glycolamides

MocPheOH or $Ph(CH_2)_2COOH$ (1.1 equiv.) was dissolved in *ca.* 25 ml acetone with 1 equiv. methyl bromoacetate or iodoacetamide and 1 equiv. anhydrous $K_2CO_3$, and refluxed for 5 hr. The acetone was removed under vacuum. Remaining material was dissolved in *ca.* 30 ml water and *ca.* 30 ml $CH_2Cl_2$. The organic layer was washed with 1 N HCl, 1 N NaOH and saturated NaCl solution, stood overnight over $Na_2SO_4$, then was filtered and the solvent removed under vacuum. The remaining material was dissolved in MeOH, filtered, dried under an $N_2$ stream, then extensively under vacuum. Compounds gave acceptable elemental analyses [$\pm 0.02 \times$ *(calculated percentage)*], except %N, which was [*(calculated percentage)*$\pm 0.2\%$] (Table 2.1) and gave only one peak by HPLC using solvent programs PB3, PB5 or PB6 as described below.

**Table 2.1** Elemental analyses for synthesized compounds.

| Compound | | %C | %H | %N |
|---|---|---|---|---|
| MocPheOCH$_2$C(O)OMe | *obs.* | 57.09 | 5.79 | 4.84 |
| | *calc.* | 56.95 | 5.80 | 4.74 |
| MocPheOCH$_2$C(O)NH$_2$ | *obs.* | 55.19 | 5.81 | 10.20 |
| | *calc.* | 55.69 | 5.75 | 10.03 |
| MocPheOCH$_2$C=OCH$_3$ | *obs.* | 59.89 | 6.11 | 5.20 |
| | *calc.* | 60.19 | 6.13 | 5.04 |
| MocPheOCH$_2$C=OCH$_2$CH$_3$ | *obs.* | 61.21 | 6.28 | 4.96 |
| | *calc.* | 61.41 | 6.53 | 4.79 |
| Ph(CH$_2$)$_2$COOH | *obs.* | 72.11 | 6.74 | - |
| | *calc.* | 71.98 | 6.71 | - |
| Ph(CH$_2$)$_2$C(O)OCH$_2$C(O)OMe | *obs.* | 64.57 | 6.44 | - |
| | *calc.* | 64.85 | 6.35 | - |
| Ph(CH$_2$)$_2$C(O)OCH$_2$C(O)NH$_2$ | *obs.* | 63.47 | 6.34 | 6.65 |
| | *calc.* | 63.74 | 6.32 | 6.79 |
| Ph(CH$_2$)$_2$C(O)OCH$_2$COOH | *obs.* | 63.21 | 5.86 | - |
| | *calc.* | 63.45 | 5.81 | - |

### 2.1.3 MocPheOCH$_2$COOH

MocPheOCH$_2$C(O)NH$_2$ was deamidated by the method of Nefkens and Nivard (76) by dissolving 180 mg (0.64 mmol) in *ca.* 7 ml glacial acetic acid, then stirring in an 8°C water bath as 0.3 g (2.4 mmol) nitrosylsulfuric acid ($HO_3SONO$) was added in 4 parts over 10 minutes. This mixture was stirred for 1 hr. at room temperature, then *ca.* 8 ml EtOAc was added. It was boiled for 1 minute to remove nitrogen oxides, then reduced to a small volume under vacuum, added into *ca.* 30 ml $H_2O$, extracted with 2 × *ca.* 40 ml EtOAc, and dried under vacuum. It was purified by flash chromatography on silica gel in 7 EtOAc/3 hexane, dried under an $N_2$ stream, then under vacuum (yield = 57 mg). HPLC analysis indicated a ratio of MocPheOCH$_2$COOH:MocPheOH of 32:1. It was further purified by preparative HPLC, using 70% (0.1% TFA/$H_2O$)/30% MeCN. The pure material was dried under vacuum and characterized by NMR (run by Dr. Irene Ekiel). HPLC analysis indicated a ratio MocPheOCH$_2$COOH:MocPheOH of 500:1.

### 2.1.4 Ph(CH$_2$)$_2$C(O)OCH$_2$COOH

Ph(CH$_2$)$_2$C(O)OCH$_2$C(O)NH$_2$ (620 mg, 3 mmol) was deamidated (76) by dissolving in *ca.* 30 ml glacial acetic acid, adding 1.5 g (12 mmol) $HO_3SONO$ in 4 parts over 10 minutes while stirring in a 10 °C water bath. It was allowed to warm to room temperature while stirring for 1 hour. EtOAc (*ca.* 40 ml) was added and the solution boiled for 1 minute to remove nitrogen oxides. The solution was reduced to a small volume under vacuum. Water was added to form a white precipitate. The solution was evaporated under vacuum at room temperature until more

solid precipitated. This was filtered and washed with cold $H_2O$, then dried overnight under vacuum (yield = 110 mg). The material was further purified by HPLC eluting with 40% (0.1%TFA/$H_2O$)/60% MeCN, then rotary evaporated dry.

### 2.1.5 MocPheOCH$_2$C(O)CH$_3$, MocPheOCH$_2$C(O)CH$_2$CH$_3$

MocPheOH (1.12 g, 5 mmol) was dissolved with stirring in THF with 0.44 g (6 mmol) 1-hydroxypropanone or 0.53 g (6 mmol) 1-hydroxy-2-butanone, plus 0.74 g (5.5 mmol) 1-hydroxybenzotriazole and 3 drops N-methyl-morpholine at -10°C. Dicyclohexylcarbodiimide (1.13 g, 5.5 mmol) was added. Stirring was continued overnight, allowing the mixture to warm to room temperature. The solution was rotary evaporated to dryness, then ca. 30 ml $CH_2Cl_2$ was added. After sitting overnight, the precipitated dicyclohexylurea and 1-hydroxybenzotriazole were filtered off. The material was purified by silica gel flash chromatography in 95 $CH_2Cl_2$/5 MeOH. MocPheOCH$_2$C(O)CH$_3$ was further purified by HPLC eluting with 55% (0.1% TFA/$H_2O$)/45% MeCN. MocPheOCH$_2$C(O)CH$_2$CH$_3$ was purified by HPLC with solvent program PB7 (Table 2.2). Both were white solids (yield = 240 mg for MocPheOCH$_2$C(O)CH$_3$, approximately 150 mg for MocPheOCH$_2$C(O)CH$_2$CH$_3$).

**Table 2.2** HPLC solvent programs.

Detection at 205 nm
Solvent A ≡ 0.1% TFA/$H_2O$
Solvent B ≡ MeCN
Solvent C ≡ $H_2O$

***a.)* Solvent program PB3**
Flow = 2.0 ml/min
End time = 11.0 min

| Time (min) | %A | %B |
|---|---|---|
| 0.0 | 80 | 20 |
| 6.0 | 0 | 100 |
| 7.5 | 80 | 20 |

***b.)* Solvent program PB5**
Flow = 2.5 ml/min
End time = 11.0 min

| Time (min) | %A | %B |
|---|---|---|
| 0.0 | 90 | 10 |
| 6.0 | 30 | 70 |
| 7.5 | 90 | 10 |

***c.)* Solvent program PB6**
Flow = 2.5 ml/min
End time = 11.0 min

| Time (min) | %A | %B |
|---|---|---|
| 0.0 | 75 | 25 |
| 6.0 | 35 | 65 |
| 7.5 | 75 | 25 |

***c.)* Solvent program PB7**
Flow = 5.0 ml/min
End time = 60 min

| Time (min) | %A | %B | %C |
|---|---|---|---|
| 0.0 | 0 | 50 | 50 |
| 4.0 | 50 | 50 | 0 |
| 28.0 | 50 | 50 | 0 |
| 38.0 | 20 | 80 | 0 |
| 48.0 | 0 | 50 | 50 |

**2.1.6 O-Ac-phenyllactylGlyOMe**

L-β-phenyllactic acid was synthesized from L-phenylalanine by treatment with nitrous acid (77). L-phenylalanine (12.5 g, 7.6 mmol) was dissolved in 250 ml 1 M $H_2SO_4$, then cooled to -5°C. $NaNO_2$ (7.66 g, 11 mmol) was dissolved in 40 ml $H_2O$ and cooled to 0°C. It was added to the L-phenylalanine solution over 3 hr with stirring and maintaining the temperature below 0°C. The solution was stirred overnight, allowing it to warm to room temperature. The reaction mixture was extracted with 4 × 100 ml $Et_2O$, saturating the aqueous layer with NaCl between the third and fourth extractions. The extracts were pooled, rotary evaporated to 200 ml and let stand over $Na_2SO_4$ overnight. Charcoal was added, then filtered and the $Et_2O$ rotary evaporated to dryness. It was recrystallized from $CHCl_3$/light petroleum ether to yield pure white crystals (3.07 g).

O-acetyl-L-β-phenyllactic acid was synthesized by refluxing with acetic anhydride (78). The L-phenyllactic acid was added to 20 ml acetic anhydride and heated at 100°C for 3 hr. Water (200 ml) was added, then solid $NaHCO_3$ to pH >9. The solution was extracted with 100 ml $Et_2O$, concentrated HCl was added to pH <3 and then extracted with 3 × 150 ml $Et_2O$. This was dried overnight over $Na_2SO_4$. The solvent was rotary evaporated to give a yield of 1.24 g.

O-Ac-phenyllactlyGlyOMe was made by the mixed anhydride method. The O-acetyl-L-β-phenyllactic acid (5.9 mmol) was dissolved in *ca.* 15 ml THF with 1 equiv. N-methylmorpholine (0.65 g) and cooled to -15°C with stirring. *iso*-Butyl chloroformate (1 equiv., 0.81 g) was dissolved in *ca.* 10 ml THF and added dropwise. After 1 to 2 min glycine methyl ester

hydrochloride (5.9 mmol, 0.74 g) and N-methylmorpholine (0.65 g) in *ca.* 10 ml N,N-dimethylformamide were added dropwise. After 2 hr at -15°C, the solution was allowed to warm to room temperature. The solids were filtered off, the THF rotary evaporated and the remaining solids taken up in *ca.* 200 ml EtOAc. This was washed with 150 ml $H_2O$, 100 ml 5% $NaHCO_3$, 200 ml $H_2O$, 100 ml 1 M HCl, 100 ml $H_2O$ and 75 ml saturated NaCl. The solution was let stand overnight over $Na_2SO_4$. It was then purified by silica gel chromatography using 9 $CH_2Cl_2$/1 MeOH. The material was rotary evaporated to give a yield of 0.91 g of a yellow oil. The compound was shown to be pure by NMR and the stoichiometry of the hydrolysis by papain showed 97.8% purity.

## 2.2 Papain Kinetics

$Hg^{++}$-papain was activated by stirring it with 3 µl β-mercaptoethanol/ml $Hg^{++}$-papain for 30 minutes, then separating papain from the $Hg^{++}$ and β-mercaptoethanol by G-15 Sephadex chromatography in 1 mM EDTA. Total [protein] was determined from the $A_{280}$ using a molar absorptivity of $\varepsilon_{280} = 5.6 \times 10^4$ $M^{-1} \cdot cm^{-1}$ (79). Total [thiol], which equals the amount of active enzyme present, was determined using 5,5′-dithiobis(2-nitrobenzoic acid) (80). Freshly activated papain contained 1.0 mol SH/mol protein and was stored on ice until used.

### 2.2.1 pH-stat

pH-stat kinetics were performed using a Radiometer RTS822 recording titration system, with a PHM 84 pH-meter. The carboxylic acid product of ester hydrolysis was titrated with 20 mM NaOH, which was prepared

fresh daily and standardized against 1 μmol HCl. Reaction solvent was pH 6.0, 20% (v/v) MeCN, 300 mM NaCl and 1 mM EDTA. Ideally substrate concentrations of 0.1 to 10 × $K_M$ were used, though high $K_M$ and/or low solubility of some substrates made this impossible to achieve. Active enzyme concentrations of 0.14 to 6 μM were used. Initial velocities were fitted by eye from a chart of volume NaOH added *versus* time. For some substrates, $v_0/[E]_0$ was determined using a computer program interfaced with the pH-stat that used linear least squares regression, or if there was significant curvature in the plot, Cornish-Bowden's direct linear plot method of $v_0/[E]_0$ determination (81).

The kinetics constants were determined directly from $v_0/[E]_0$'s by the algorithm of Cornish-Bowden (82). This method assumes that the errors in the observed $v_0/[E]_0$'s is proportional to the true $v_0/[E]_0$. This has been shown to be the case for pH-stat kinetic determinations (37). In this situation there is an exact solution for the determination of $k_{cat}/K_M$, $k_{cat}$ and $K_M$ from $v_0/[E]_0$'s (83). Cornish-Bowden's method also provides an estimate of the standard errors in $k_{cat}/K_M$, $k_{cat}$ and $K_M$.

With several substrates the plots of $v_0/[E]_0$ *versus* $[S]_0$ tended to have significant negative y-intercepts; that is, that the rate reached zero at significant substrate concentrations. By fitting the poorer substrates' data directly to the Michaelis-Menten equation plus offset using the program Enzfitter, better standard errors were obtained without materially changing the constants calculated. Iterative applications of offset resulted in a smooth approach to zero offset in typically one or two, occasionally three iterations. The new $v_0/[E]_0$'s

were then used with Cornish-Bowden's method as with other substrates to determine the kinetic constants.

### 2.2.2 HPLC Methods

A Varian LC5560 with a UV200 detector and DS654 data station was used for all HPLC work. All analytical runs were done using C18, end capped, reversed phase Varian MHC-5N-CAP columns, 4 mm by 15 cm with 5 μm irregular particles with absorbance detection at 205 nm. Solvent programs PB3 and PB5 were used with MocPheOCH$_2$C(O)OMe, and PB6 with Ph(CH$_2$)$_2$C(O)OCH$_2$C(O)OMe (Table 2.2). A Waters Delta-Pak C18-100Å, 19 mm by 30 cm column, with absorbance detection at 205 nm and isocratic elution was used for preparative HPLC.

#### *2.2.2.1 Dialysis Setup*

In quantitating substrate and products in the reaction mixtures, it was necessary to quickly and gently separate them from papain, which irreversibly bound to and destroyed the HPLC column.

This was accomplished by a dialysis arrangement using 2 nested test tubes separated by a dialysis membrane (Fig. 2.1). The dialysis tubing was by SPECTRA/POR, with a cylindrical diameter of 1.6 cm (2.0 ml/cm) and a molecular weight cutoff of 3500. The outer test tube was a 30 ml Corex centrifuge tube which was silanized (see below) and contained 20 ml reaction mixture. The inner test tube was a 13 mm by 100 mm

**Figure 2.1** Dialysis setup for HPLC kinetics.

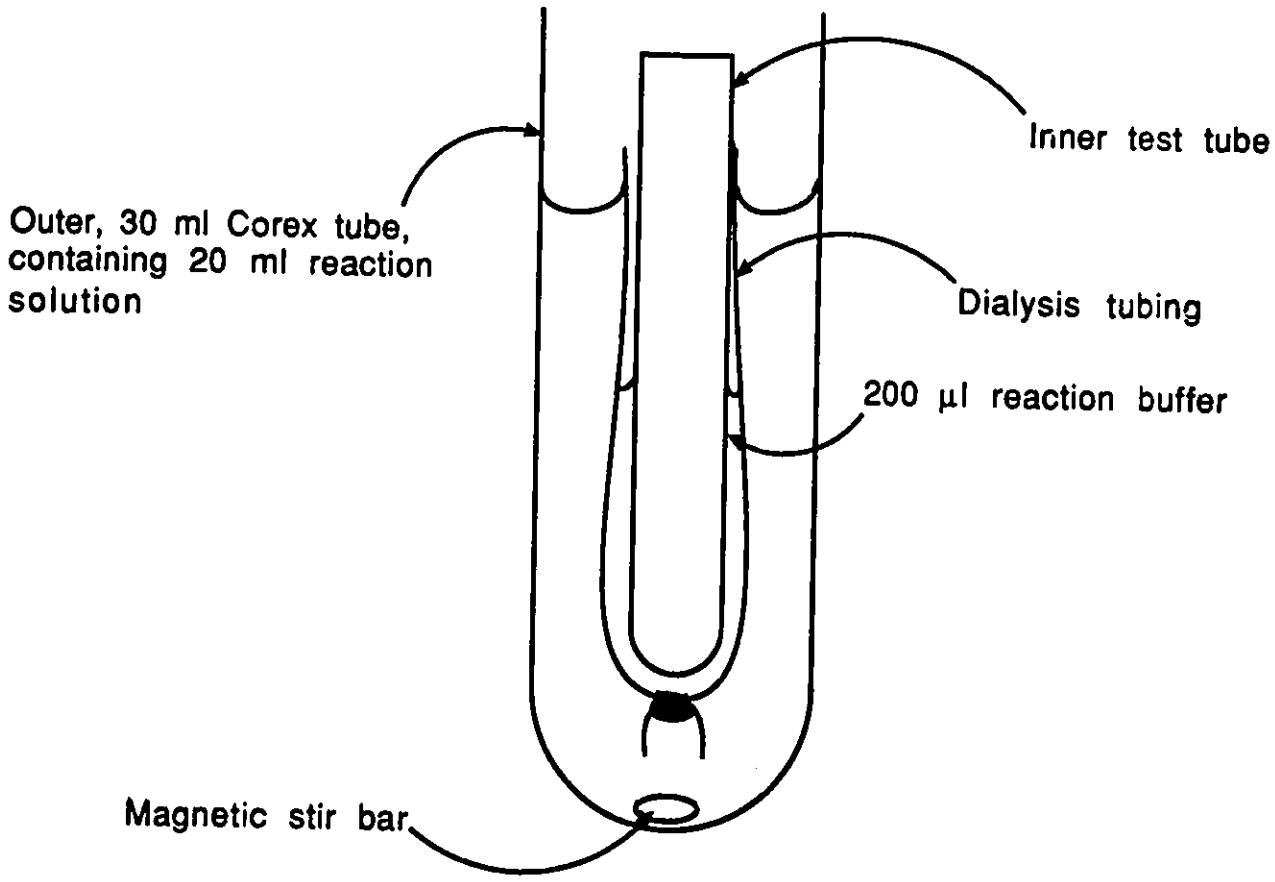

Inner test tube
Outer, 30 ml Corex tube, containing 20 ml reaction solution
Dialysis tubing
200 µl reaction buffer
Magnetic stir bar

**Figure 2.1**

disposable culture tube. Using an inner test tube dramatically increased the surface area and therefore the speed of dialysis. A small volume (200 µl) of reaction buffer was placed inside the dialysis membrane and allowed to equilibrate for 5 to 10 minutes before being removed and injected directly onto the HPLC column. The reaction buffer was 100 mN sodium phosphate, 200 mM NaCl, pH 6.0, 20% (v/v) MeCN. This is the same ionic strength, percentage MeCN and pH as the pH-stat reaction solution. Phosphate ions have been shown not to have an affect on papain activity (37,84).

The rate of dialysis was tested with $Ph(CH_2)_2C(O)OCH_2C(O)OMe$ by using the dialysis setup as usual, using 2.5 mM $Ph(CH_2)_2C(O)OCH_2C(O)OMe$ in the outer test tube. Successive samples of 200 µl reaction buffer were placed in the inner test tube and withdrawn at times from 1 to 15 min afterward. $[Ph(CH_2)_2C(O)OCH_2C(O)OMe]$ was determined from the $A_{216}$ of the inner solution. The $[Ph(CH_2)_2C(O)OCH_2C(O)OMe]_{obs}$'s were fit to a first order rate equation, where:

$$\frac{[Ph(CH_2)_2C(O)OCH_2C(O)OMe]_{in}}{[Ph(CH_2)_2C(O)OCH_2C(O)OMe]_{out}} = e^{-kt}$$

The rate of dialysis was found to be have a rate constant of $k = 0.333 \pm 0.025$ min$^{-1}$, corresponding to $t_{1/2} = 2.03$ min (Fig. 2.2).

*Silanization*

Papain at 1 µM in untreated centrifuge tubes became inactivated over time, presumably due to adsorption onto the glass. By increasing [papain] to 2 µM and silanizing the centrifuge tubes, this problem was eliminated.

**Figure 2.2** Rate of dialysis.

$[Ph(CH_2)_2C(O)OCH_2C(O)OMe]_{in}/[Ph(CH_2)_2C(O)OCH_2C(O)OMe]_{out}$ *vs.* time (min), with curve fitted to:
$[Ph(CH_2)_2C(O)OCH_2C(O)OMe]_{in}/[Ph(CH_2)_2C(O)OCH_2C(O)OMe]_{out} = e^{(-0.333*t)}$.

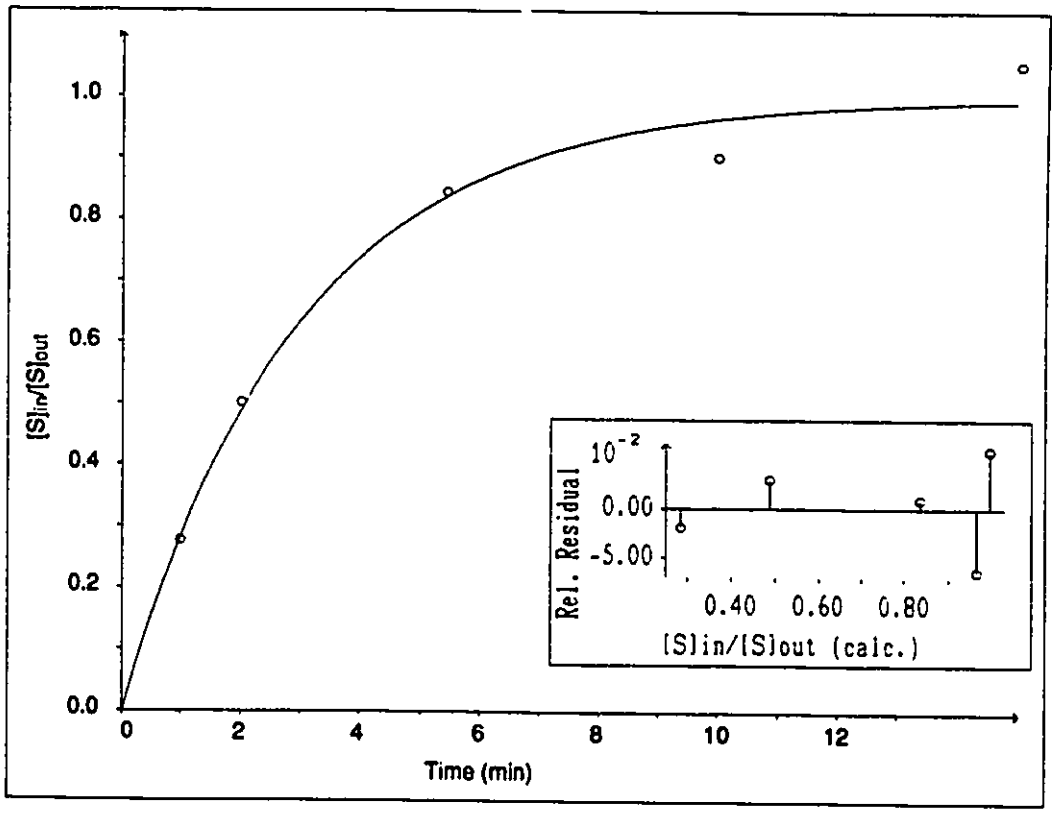

1.0
0.8
0.6
0.4
0.2
0.0
[S]in/[S]out
0
2
4
6
8
10
12
Time (min)
Rel. Residual
10⁻²
0.00
-5.00
0.40
0.60
0.80
[S]in/[S]out (calc.)

**Figure 2.2**

Corex centrifuge tubes (30 ml) were washed and rinsed with 5% $HNO_3$, then rinsed thoroughly with $H_2O$. They were dried, then soaked in 5% dichlorodimethylsilane (v/v) in $CH_2Cl_2$ for 15 min to 1 hr. They were then rinsed with MeOH, and soaked in MeOH for 15 min before rinsing thoroughly with $H_2O$ and drying overnight at 60°C.

*2.2.2.2 Standard Curves*

Standard curves were determined for MocPheO$CH_2$C(O)OMe, MocPheO$CH_2$COOH and MocPheOH with both solvent programs PB3 and PB5. Standard curves were determined for Ph$(CH_2)_2$C(O)O$CH_2$C(O)OMe, Ph$(CH_2)_2$C(O)O$CH_2$COOH and Ph$(CH_2)_2$COOH with PB6.

*2.2.2.3 Reaction Runs*

Samples taken at various times from the dialysis setup in the reaction of papain with substrate were directly injected onto the column. For the reactions with MocPheO$CH_2$C(O)OMe, the resulting Area Counts for each component were converted to nmol injected, the initial substrate concentration was then used to calculate the concentration of each component present in the reaction mixture from the relative amounts present, as well as the ratio of [MocPheO$CH_2$COOH]/[MocPheOH]. Initial velocities for substrate depletion and product formation were determined by the direct linear plot method of Cornish-Bowden (81).

## 2.3 Kinetics derivations

In order to determine the constants reflecting the hydrolysis of MocPheOCH$_2$C(O)OMe at the two ester functionalities, the King-Altman method of deriving kinetic constants was used (82), given the master pattern in Figure 2.3 and:

$e_0$ ≡ initial enzyme concentration

s ≡ substrate concentration at time = t

$v_{0,T}$ ≡ total initial rate (-ds/dt)

$v_{0,B}$ ≡ initial rate of hydrolysis of substrate B

$v_{0,C}$ ≡ initial rate of hydrolysis of substrate C

$k_{cat,B}$, $k_{cat,C}$, $k_{cat,T}$ ≡ $k_{cat}$ for substrates B, C and the total reaction respectively.

$K_{M,B}$, $K_{M,C}$, $K_{M,T}$ ≡ $K_M$ for substrates B, C and the total reaction, respectively.

Substrates B and C are treated as being separate, though they are the same compound, since the binding of the compound to the enzyme in a productive mode for hydrolysis of the methyl ester precludes hydrolysis of the glycolic acid methyl ester without dissociation of the enzyme and substrate and subsequent re-association in a different complex, effectively the binding of a separate molecule. To simplify the derivation of the equations, it was assumed that the rates of the reverse reactions for the covalent steps ($k_{-2}$, $k_{-3}$, $k_{-5}$, & $k_{-6}$) are negligible. This assumption was justified by allowing the reaction of

**Figure 2.3** Master King-Altman pattern for hydrolysis of $MocPheOCH_2C(O)OMe$ at two sites.

**E + S** ≡ Free enzyme plus free substrate in the ground state
$\mathbf{E \cdot S^B}$, $\mathbf{E \cdot S^C}$ ≡ Michaelis complexes for $MocPheOCH_2C(O)OMe$ productively bound to be hydrolyzed at each ester functionality.
$\mathbf{ES^B}$, $\mathbf{ES^C}$ ≡ Acylenzyme intermediates for hydrolysis at both sites.

**Figure 2.4** King-Altman subpatterns derived from the master pattern in Fig. 2.3.

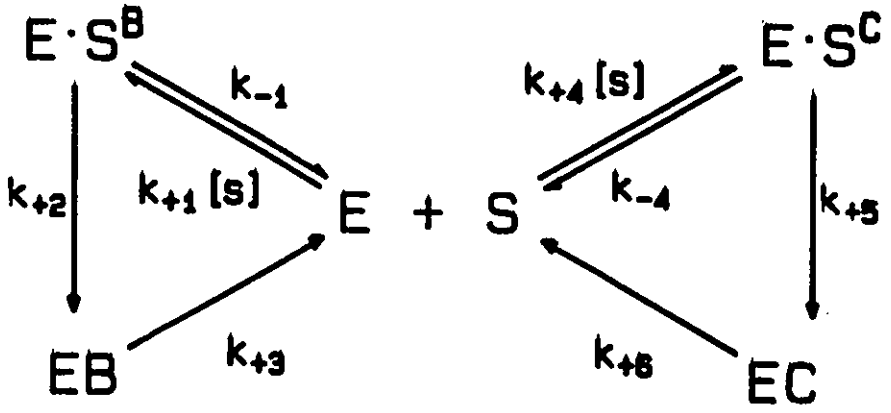



**Figure 2.3**

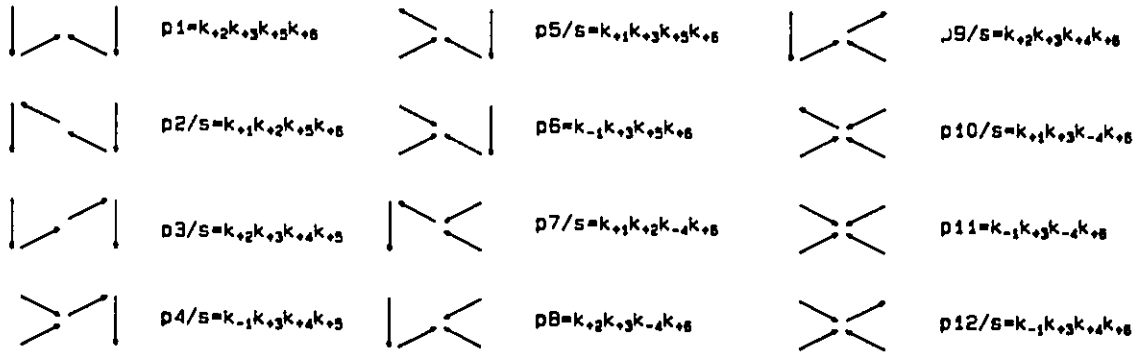



**Figure 2.4**

MocPheOCH$_2$C(O)OMe with papain to go to completion, where no remaining MocPheOCH$_2$C(O)OMe was detectable.

**(I)** $v_{0,T} = v_{0,B} + v_{0,C}$

$= [EB]k_{+3} + [EC]k_{+6}$

Given the master pattern of Figure 2.3, subpatterns were generated (Fig. 2.4). The number of subpatterns was checked using the method of Huang (85):

$\pi = a*b - l_{AB}^2$, where: a,b = number of lines in subfigures A and B (in this instance, each subfigure being one half of the master pattern).

$l_{AB}^2$ = number of common boundaries between A and B.

$\therefore \pi = 3*3 - 0^2$

$= 9$

From the subpatterns:

**(II)** $$[EB]/e_0 = \frac{k_{+1}k_{+2}k_{+6}s(k_{-4}+k_{+5})}{\Sigma}$$

**(II.1)** $$[EC]/e_0 = \frac{k_{+3}k_{+4}k_{+5}s(k_{-1}+k_{+2})}{\Sigma}$$

Where (from Fig. 2.4):

(V) $\Sigma = (p1+p6+p8+p11) + s*(p2+p3+p4+p5+p7+p9+p10+p12)$

Since substrates B and C are equivalent, it is sufficient to solve for one only, in this case B.

From I,II:

$$\text{(VII)} \qquad v_{0,B}/e_0 = \frac{k_{+1}k_{+2}k_{+3}k_{+6}s(k_{+5}+k_{-4})}{\Sigma}$$

$$= \frac{k_{+2}k_{+3}s}{\left[\frac{\Sigma}{k_{+1}k_{+6}(k_{+5} + k_{-4})}\right]}$$

Divide by: $(k_{+2}+k_{+3})$

$$= \frac{\left[\frac{k_{+2}k_{+3}*s}{(k_{+2} + k_{+3})}\right]}{\left[\frac{\Sigma}{k_{+1}k_{+6}(k_{+5}+k_{-4})\,(k_{+2}+k_{+3})}\right]}$$

$$= \frac{k_{cat,B}*s}{\left[\frac{\Sigma}{k_{+1}k_{+6}(k_{+5}+k_{-4})\,(k_{+2}+k_{+3})}\right]}$$

Where $k_{cat,B}$ is the rate constant for B in the absence of substrate C.

Thus the denominator of this equation becomes (from V):

(IV)

$$\text{Denominator} = \frac{\left[s(p2+p5+p10+p7) + s(p3+p4+p9+p12) + (p1+p6+p8+p11)\right]}{k+1k+6(k+5+k-4)\,(k+2+k+3)}$$

$$= s + \left[\frac{s(p3+p4+p9+p12) + (p1+p6+p8+p11)}{k_{+1}k_{+6}(k_{+5}+k_{-4})\,(k_{+2}+k_{+3})}\right]$$

$$= s + \left[\frac{s(k_{+3}k_{+4}(k_{+2}+k_{-1})\,(k_{+5}+k_{+6})) + k_{+3}k_{+6}(k_{+2}+k_{-1})\,(k_{+5}+k_{-4})}{k_{+1}k_{+6}(k_{+5}+k_{-4})\,(k_{+2}+k_{+3})}\right]$$

Divide by: $k_{+3}k_{+6}(k+2+k-1)\,(k+5+k-4)$

$$= s + \left[ \frac{1 + \dfrac{s(k_{+3}k_{+4}(k_{+2}+k_{+1})(k_{+5}+k_{+6}))}{k_{+3}k_{+6}(k_{+2}+k_{-1})(k_{+5}+k_{-4})}}{\dfrac{k_{+1}k_{+6}(k_{+5}+k_{-4})(k_{+2}+k_{+3})}{k_{+3}k_{+6}(k_{+2}+k_{-1})(k_{+5}+k_{-4})}} \right]$$

$$= s + \left[ \frac{1 + \dfrac{sk_{+4}(k_{+5}+k_{+6})}{k_{+6}(k_{+5}+k_{-4})}}{\dfrac{k_{+1}(k_{+2}+k_{+3})}{k_{+3}(k_{+2}+k_{-1})}} \right]$$

$$= s + \left[ \frac{1 + \dfrac{s}{K_{M,C}}}{\dfrac{1}{K_{M,B}}} \right]$$

$K_{M,B}$, $K_{M,C}$ refer to the Michaelis constants for substrates B and C respectively, each in the absence of the other.

$$\therefore \textit{Denominator} = s + K_{M,B}^{*}(1 + s/K_{M,C})$$

$$\text{(III.1)} \qquad \therefore v_{0,B}/e_0 = \frac{k_{cat,B}^{*}s}{K_{M,B}^{*}(1 + s/K_{M,C}) + s}$$

and (III.2),

$$v_{0,C}/e_0 = \frac{k_{cat,C}{}^{*}s}{K_{M,C}{}^{*}(1 + s/K_{M,B}) + s}$$

These are exactly the equations expected for the rates of hydrolysis of competitive substrates (82). If two different compounds B and C were used, $s/K_{M,B}$ and $s/K_{M,C}$ would become $b/K_{M,B}$ and $c/K_{M,C}$, respectively.

For the overall rate equation, from I:

$$v_{0,T} = v_{0,B} + v_{0,C}$$
$$= [EB]k+3 + [EC]k+6$$

∴ from II and II.1:

(VIII)

$$v_{0,T}/e_0 = \frac{k_{+1}k_{+2}k_{+3}k_{+6}s(k_{+5}+k_{-4}) + k_{+3}k_{+4}k_{+5}k_{+6}s(k_{+2}+k_{-1})}{\Sigma}$$

From V, $\Sigma$ may be rearranged to:

(IX) $$\Sigma = k_{+3}k_{+6}(k_{+5}+k_{-4})(k_{+2}+k_{-1}) + s[k_{+1}k_{+6}(k_{+2}+k_{+3})(k_{+5}+k_{-4}) + k_{+3}k_{+4}(k_{+5}+k_{+6})(k_{+2}+k_{-1})]$$

From VIII and IX, divide by: $(k_{+5}+k_{-4})(k_{+2}+k_{-1})$

$$v_{0,T}/e_0 = \frac{\frac{k_{+1}k_{+2}k_{+3}k_{+6}}{(k_{+2}+k_{-1})} + \frac{k_{+3}k_{+4}k_{+5}k_{+6}}{(k_{+5}+k_{-4})}}{k_{+3}k_{+6} + s\left[\frac{k_{+1}k_{+6}(k_{+2}+k_{+3})}{(k_{+2}+k_{-1})} + \frac{k_{+3}k_{+4}(k_{+5}+k_{+6})}{(k_{+5}+k_{-4})}\right]}$$

Divide by: $\left[\dfrac{k_{+1}k_{+6}(k_{+2}+k_{+3})}{(k_{+2}+k_{-1})} + \dfrac{k_{+3}k_{+4}(k_{+5}+k_{+6})}{(k_{+5}+k_{-4})}\right]$

(X)

$$v_{0,T}/e_0 = \frac{\left[\dfrac{k_{+1}k_{+2}k_{+3}k_{+6}(k_{+5}+k_{-4}) + k_{+3}k_{+4}k_{+5}k_{+6}(k_{+2}+k_{-1})}{\left[k_{+1}k_{+6}(k_{+2}+k_{+3})(k_{+5}+k_{-4}) + k_{+3}k_{+4}(k_{+5}+k_{+6})(k_{+2}+k_{-1})\right]}\right]}{\left[\dfrac{k_{+3}k_{+6}(k_{+2}+k_{-1})(k_{+5}+k_{-4})}{k_{+1}k_{+6}(k_{+2}+k_{+3})(k_{+5}+k_{-4}) + k_{+3}k_{+4}(k_{+5}+k_{+6})(k_{+2}+k_{-1})}\right] + s}$$

Therefore the total rate of substrate hydrolysis, measuring the combined rate at both ester functionalities (as by pH-stat) will follow Michaelis-Menten kinetics with $k_{cat,T}$ and $K_{M,T}$ as shown in the above equation.

From **X**:

$$k_{cat,T} = \frac{k_{+1}k_{+2}k_{+3}k_{+6}(k_{+5}+k_{-4}) + k_{+3}k_{+4}k_{+5}k_{+6}(k_{+2}+k_{-1})}{\left[k_{+1}k_{+6}(k_{+2}+k_{+3})(k_{+5}+k_{-4}) + k_{+3}k_{+4}(k_{+5}+k_{+6})(k_{+2}+k_{-1})\right]}$$

$$= \frac{\left[ k_{+1}k_{+2}k_{+3}(k_{+6}k_{+5}+k_{+6}k_{-4}) + k_{+4}k_{+5}k_{+6}(k_{+3}k_{+2}+k_{+3}k_{-1}) \right]}{\left[ (k_{+5}k_{+6}+k_{-4}k_{+6})(k_{+1}k_{+2}+k_{+1}k_{+3}) + (k_{+2}k_{+3}+k_{-1}k_{+3})(k_{+4}k_{+5}+k_{+4}k_{+6}) \right]}$$

Divide by: $(k_{+1}k_{+2}+k_{+1}k_{+3})(k_{+4}k_{+5}+k_{+4}k_{+6})$

$$k_{cat,T} = \frac{\left[ \frac{k_{+1}k_{+2}k_{+3}(k_{+6}k_{+5}+k_{+6}k_{-4}) + k_{+4}k_{+5}k_{+6}(k_{+3}k_{+2}+k_{+3}k_{-1})}{(k_{+1}k_{+2}+k_{+1}k_{+3})(k_{+4}k_{+5}+k_{+4}k_{+6})} \right]}{\left[ \frac{k_{+5}k_{+6}+k_{-4}k_{+6}}{k_{+4}k_{+5}+k_{+4}k_{+6}} + \frac{k_{+2}k_{+3}+k_{-1}k_{+3}}{k_{+1}k_{+2}+k_{+1}k_{+3}} \right]}$$

$$= \frac{\left[ \frac{k_{+1}k_{+2}k_{+3}*K_{M,C}}{k_{+1}(k_{+2}+k_{+3})} + \frac{k_{+4}k_{+5}k_{+6}*K_{M,B}}{k_{+4}(k_{+5}+k_{+6})} \right]}{K_{M,C} + K_{M,B}}$$

$$= \frac{k_{cat,B}*K_{M,C} + k_{cat,C}*K_{M,B}}{K_{M,C} + K_{M,B}}$$

$$= k_{cat,B}\left[ \frac{K_{M,C}}{K_{M,C}+ K_{M,B}} \right] + k_{cat,C}\left[ \frac{K_{M,B}}{K_{M,B}+ K_{M,C}} \right]$$

(XI)

$$\therefore\ k_{cat,T} = \frac{k_{cat,B}}{\left[1 + \frac{K_{M,B}}{K_{M,C}}\right]} + \frac{k_{cat,C}}{\left[1 + \frac{K_{M,C}}{K_{M,B}}\right]}$$

From X:

$$K_{M,T} = \frac{k_{+3}k_{+6}(k_{+2}+k_{-1})\,(k_{+5}+k_{-4})}{\left[\begin{array}{l} k_{+1}k_{+6}(k_{+2}+k_{+3})\,(k_{+5}+k_{-4}) \ + \\ k_{+3}k_{+4}(k_{+5}+k_{+6})\,(k_{+2}+k_{-1}) \end{array}\right]}$$

$$= \frac{\left[\begin{array}{l} k_{+2}k_{+3}k_{+5}k_{+6} + k_{+2}k_{+3}k_{-4}k_{+6} + \\ \quad k_{-1}k_{+3}k_{+5}k_{+6} + k_{-1}k_{+3}k_{-4}k_{+6} \end{array}\right]}{\left[\begin{array}{l} k_{+1}k_{+2}k_{+5}k_{+6} + k_{+1}k_{+2}k_{-4}k_{+6} + \\ \quad k_{+1}k_{+3}k_{+5}k_{+6} + k_{+1}\,k_{+3}k_{-4}k_{+6} + \\ \quad\quad k_{+2}k_{+3}k_{+4}k_{+5} + k_{-1}k_{+3}k_{+4}k_{+5} + \\ \quad\quad\quad k_{+2}k_{+3}k_{+4}k_{+6} + k_{-1}k_{+3}k_{+4}k_{+6} \end{array}\right]}$$

$$= \frac{\left[\begin{array}{l} k_{+2}k_{+3}k_{+5}k_{+6} + k_{+2}k_{+3}k_{-4}k_{+6} + \\ \quad k_{-1}k_{+3}k_{+5}k_{+6} + k_{-1}k_{+3}k_{-4}k_{+6} \end{array}\right]}{\left[\begin{array}{l} k_{-1}k_{+3}(k_{+4}k_{+5}+k_{+4}k_{+6}) \ + \\ \quad k_{+2}k_{+3}(k_{+4}k_{+5}+k_{+4}k_{+6}) \ + \\ \quad\quad k_{-4}k_{+6}(k_{+1}k_{+3}+k_{+1}k_{+2}) \ + \\ \quad\quad\quad k_{+5}k_{+6}(k_{+1}k_{+3}+k_{+1}k_{+2}) \end{array}\right]}$$

Divide by: $(k_{+4}k_{+5}+k_{+4}k_{+6})\,(k_{+1}k_{+3}+k_{+1}k_{+2})$

$$= \frac{\left[\dfrac{k_{-1}k_{+3}(k_{+5}k_{+6}+k_{-4}k_{+6}) \;+\; k_{+2}k_{+3}(k_{+5}k_{+6}+k_{-4}k_{+6})}{(k_{+4}k_{+6} \;+\; k_{+4}k_{+6})\,(k_{+1}k_{+2} \;+\; k_{+1}k_{+3})}\right]}{\left[\dfrac{k_{-1}k_{+3} \;+\; k_{+2}k_{+3}}{k_{+1}k_{+3} \;+\; k_{+1}k_{+2}} + \dfrac{k_{-4}k_{+6} \;+\; k_{+5}k_{+6}}{k_{+4}k_{+5} \;+\; k_{+4}k_{+6}}\right]}$$

$$= \frac{\left[\dfrac{(k_{-1}k_{+3}+k_{+2}k_{+3})\,(k_{+5}k_{+6}+k_{-4}k_{+6})}{(k_{+4}k_{+6} \;+\; k_{+4}k_{+6})\,(k_{+1}k_{+2} \;+\; k_{+1}k_{+3})}\right]}{\left[\dfrac{k_{-1}k_{+3} \;+\; k_{+2}k_{+3}}{k_{+1}k_{+3} \;+\; k_{+1}k_{+2}} + \dfrac{k_{-4}k_{+6} \;+\; k_{+5}k_{+6}}{k_{+4}k_{+5} \;+\; k_{+4}k_{+6}}\right]}$$

**(XII)** $$\therefore \; K_{M,T} = \frac{K_{M,B}{*}K_{M,C}}{K_{M,B}{+}K_{M,C}}$$

From **XI** and **XII**:

$$(k_{cat}/K_M)_T = \frac{\left[\dfrac{k_{cat,B}{*}K_{M,C} \;+\; k_{cat,C}{*}K_{M,B}}{(K_{M,B} \;+\; K_{M,C})}\right]}{\left[\dfrac{K_{M,B}{*}K_{M,C}}{K_{M,B}{+}K_{M,C}}\right]}$$

$$\therefore (k_{cat}/K_M)_T = \frac{k_{cat,B}}{K_{M,B}} + \frac{k_{cat,C}}{K_{M,C}}$$

Thus the values of $(k_{cat}/K_M)_T$, $k_{cat,T}$ and $K_{M,T}$ are same as they would be for two, independent, competitive substrates (82).

From III.1 and III.2:

$$\frac{v_{0,B}}{v_{0,C}} = \frac{k_{cat,B} * s * (K_{M,C}*(1 + s/K_{M,B}) + s)}{k_{cat,C} \quad s * (K_{M,B}*(1 + s/K_{M,C}) + s)}$$

$$= \frac{k_{cat,B} * s * K_{M,C} * (1 + s/K_{M,B} + s/K_{M,C})}{k_{cat,C} * s * K_{M,B} * (1 + s/K_{M,C} + s/K_{M,B})}$$

$$\therefore v_{0,B}/v_{0,C} = \frac{(k_{cat}/K_M)_B}{(k_{cat}/K_M)_C}$$

# 3. RESULTS

## 3.1 HPLC

### 3.1.1 Standard curves

Standard curves were determined for MocPheOCH$_2$C(O)OMe, MocPheOCH$_2$COOH and MocPheOH with solvent programs PB3 and PB5, while those for Ph(CH$_2$)$_2$C(O)OCH$_2$C(O)OMe, Ph(CH$_2$)$_2$C(O)OCH$_2$COOH and Ph(CH$_2$)$_2$COOH were determined with PB6 (Table 3.1). A linear relationship was demonstrated between the peak Area Counts (arbitrary units) and the amount of material injected (Fig. 3.1). The solvent programs produced a steeply sloping baseline; more accurate Area Counts were obtained using baseline subtraction (Fig. 3.2). The absorptivity (≡ slope) was found to be constant from day to day. The intercept was small, as expected, but somewhat variable, due to the random noise level (calculated at each Method Ready) that was in effect for a given run. Runs with high random noise gave more negative intercepts than low noise runs. The standard curve of MocPheOH with PB3 on 11 & 12 Jan '89 (noise > 300 arbitrary units) was $Area\ Counts = 4.43\times10^4 \times n - 5.85\times10^4$ ($n$ = nmol material injected) versus 05 & 06 Oct '88 (noise < 100)) where $Area\ Counts = 4.11\times10^4 \times n - 1.28\times10^4$. There are three factors contributing to this effect. The start and end of a given peak is calculated from the slope of the curve. The selected signal to noise ratio (S:N) defines the number of times above random a slope must be to be considered part of a peak. As the random noise levels rise, the steeper the slope required to be included in a peak, thus the later peaks start and sooner they end. This effect is exaggerated by the

necessity of selecting higher S:N ratios when random noise is high in order to suppress spurious peaks. This is further exaggerated by the fact that columns using very small particles, less than 10 μm diameter, tend to have long trailing edges and occasionally the appearance of having shoulders on the trailing side (86). Since high noises were only a problem with the determination of molar absorptivity with solvent program PB5 and for MocPheOCH$_2$COOH with PB3, and not for any of the papain reaction runs, the calculated intercepts were neglected. The absorptivities were calculated using a non-linear least squares regression to the linear equation:

$$A.C. = Absorptivity * n + Incpt$$

where *A.C.* is Area Counts (arbitrary units) and *n* is the amount of compound injected (nmol) and using statistical weighting (weight is proportional to $1/y_{obs}$).

**Table 3.1** HPLC standard curves.

| Compound & Solvent Program | | Slope (*Area Counts*/nmol) | Intercept (*Area Counts*) |
|---|---|---|---|
| MocPheOCH$_2$C(O)OMe | PB3 | 4.26 (±0.05) ×10$^4$ | -8.5 (±2.3) ×10$^3$ |
| | PB5 | 3.74 (±0.16) ×10$^4$ | -3.2 (±2.0) ×10$^3$ |
| MocPheOCH$_2$COOH | PB3 | 3.62 (±0.04) ×10$^4$ | 2.7 (±6.9) ×10$^3$ |
| | PB5 | 3.04 (±0.06) ×10$^4$ | -2.1 (±9.3) ×10$^3$ |
| MocPheOH | PB3 | 4.11 (±0.06) ×10$^4$ | -1.3 (±0.4) ×10$^4$ |
| | PB5 | 3.65 (±0.09) ×10$^4$ | -4.6 (±0.9) ×10$^4$ |
| Ph(CH$_2$)$_2$C(O)-OCH$_2$C(O)OMe | PB6 | 2.78 (±0.12) ×10$^4$ | -1.3 (±2.3) ×10$^4$ |
| Ph(CH$_2$)$_2$C(O)-OCH$_2$COOH | PB6 | 2.72 (±0.05) ×10$^4$ | -0.9 (±1.0) ×10$^4$ |
| Ph(CH$_2$)$_2$COOH | PB6 | 2.38 (±0.05) ×10$^4$ | 9.2 (±7.7) ×10$^3$ |

**Figure 3.1** Standard curve for MocPheOCH$_2$COOH with solvent program PB5.

The line is fitted to:
$AC = 3.04(\pm 0.06)\times 10^4 * n - 2(\pm 9)\times 10^3$, where $AC$ = Area Counts and $n$ = nmol MocPheOCH$_2$C(O)OMe injected onto the column. The *inset* is a plot of the statistical residuals; *i.e.* $(AC_{obs} - AC_{calc})/(AC_{calc})^{1/2}$ *vs.* $AC_{calc}$.

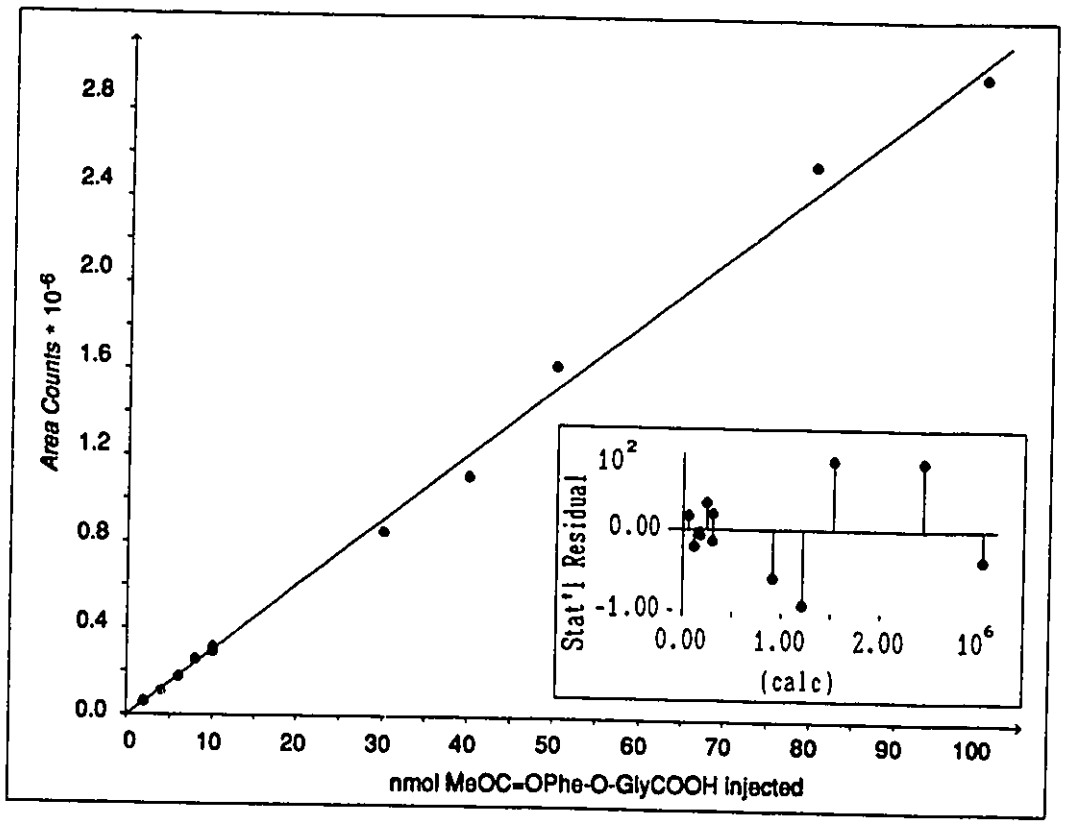

Area Counts * 10⁻⁶
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
0
10
20
30
40
50
60
70
80
90
100
nmol MeOC=OPhe-O-GlyCOOH injected
Stat'l Residual
10²
0.00
-1.00
0.00
1.00
2.00
10⁶
(calc)

**Figure 3.1**

**Figure 3.2** Errors in the standard curves for MocPheOCH$_2$C(O)OMe using solvent program PB3 with and without baseline subtraction.

**% Error** = 100 * ([S]$_{calc}$ - [S]$_{true}$)/[S]$_{true}$

The dashed (- - -) and dotted (····) lines represent linear regression to the points with (●) and without (○) baseline subtraction.

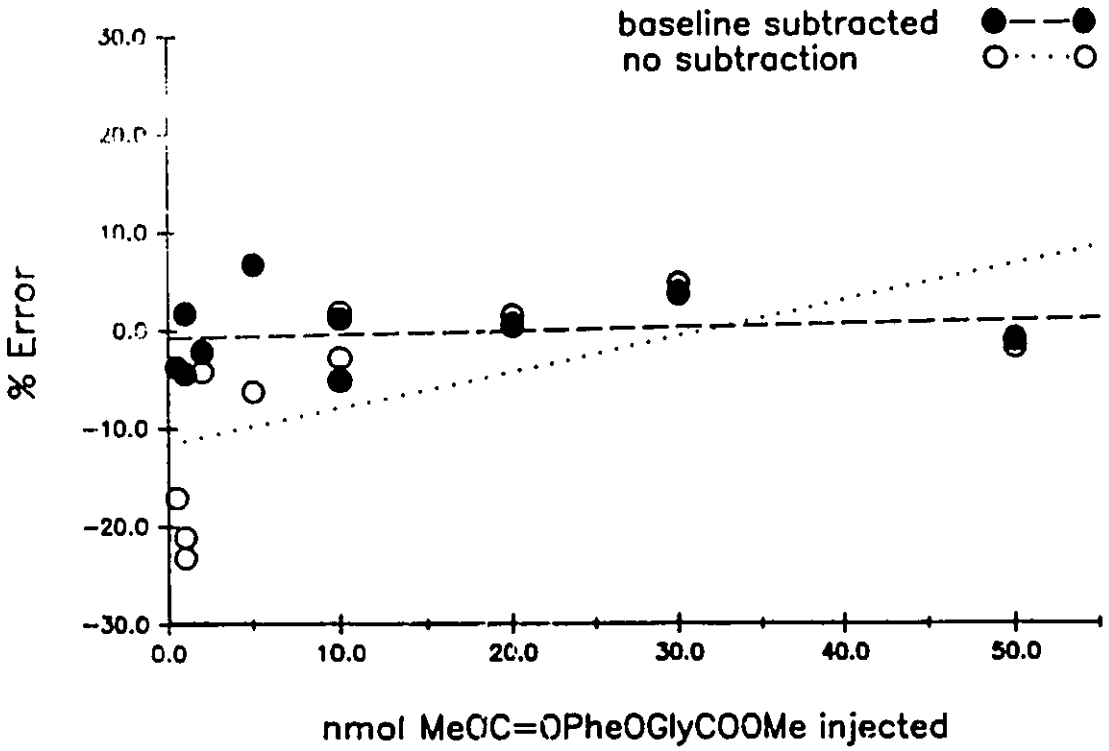

baseline subtracted
no subtraction
% Error
30.0
20.0
10.0
0.0
-10.0
-20.0
-30.0
0.0
10.0
20.0
30.0
40.0
50.0
nmol MeOC=OPheOGlyCOOMe injected

**Figure 3.2**

### 3.1.2 MocPheOCH$_2$C(O)OMe

Two different solvent programs were used with the MocPheOCH$_2$C(O)OMe reaction runs because it was found after several reaction runs that the peaks for MocPheOCH$_2$COOH and MocPheOH were eluting too close together with PB3. The lack of baseline separation between the two compounds, exacerbated by the difference in peak sizes, caused part of the area for MocPheOH peak to be included with the MocPheOCH$_2$COOH peak. This caused an increasing ratio in MocPheOCH$_2$COOH/MocPheOH area counts with reaction time, even though the height of the MocPheOH peak continued to increase. Solvent program PB5 achieved baseline separation of the two products and therefore more accurate Area Counts and a more consistent ratio of [MocPheOCH$_2$COOH]/[MocPheOH] to be observed.

In all, ten reaction runs with MocPheOCH$_2$C(O)OMe were performed. The first, using 1 μM papain and a non-silanized centrifuge tube clearly showed enzyme inactivation, as determined by Selwyn's test (Fig. 3.3) (87). The remaining runs, performed with 2 μM papain and silanized centrifuge tubes did not display apparent enzyme inactivation. Of the nine remaining runs, four gave erratic ratios of [MocPheOCH$_2$COOH]/[MocPheOH] due to the use of solvent program PB3. The remaining five runs with PB5 gave good results.

Initial velocities ($v_0/[E]_0$'s) were calculated for MocPheOCH$_2$C(O)OMe depletion and MocPheOCH$_2$COOH formation using nine runs. The MocPheOCH$_2$COOH Area Counts were not significantly increased in the four PB3 runs where part of the MocPheOH peak was included in the

**Figure 3.3** Selwyn's test for enzyme inactivation (87) in dialysis setup.

**% Reaction** = 100 * $([S]_0-[S])/[S]_0$

The solid line (——) represents the progress curve determined for $k_{cat}$ = 12.4 s-1 and Km = 25.8 mM. At 1 μM papain in unsilanized Corex tubes, there is clearly enzyme inactivation; but none is apparent at 2 μM papain in silanized tubes.

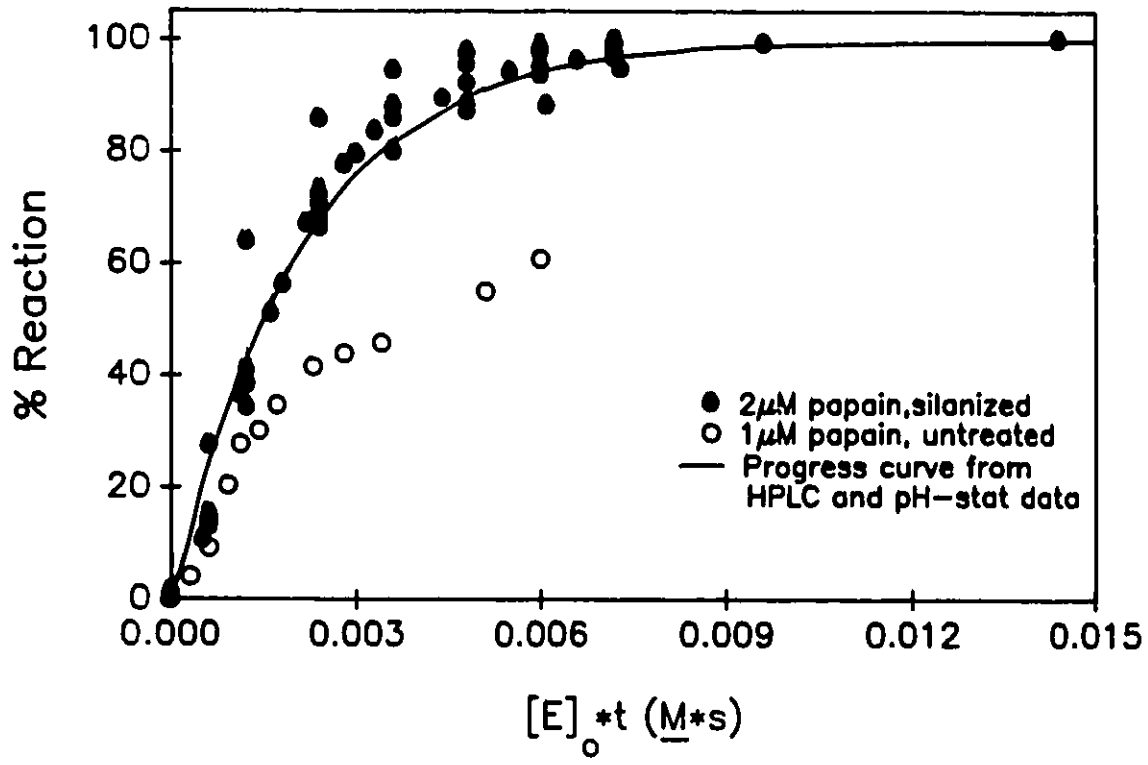

100
80
60
40
20
0
% Reaction
0.000
0.003
0.006
0.009
0.012
0.015
[E]o*t (M*s)
2μM papain,silanized
1μM papain, untreated
Progress curve from
HPLC and pH-stat data

**Figure 3.3**

area calculations and therefore did not significantly affect the calculations of $v_0/[E]_0$. For the calculation of the MocPheOH $v_0/[E]_0$'s, only the five PB5 runs were used. The depletion of $MocPheOCH_2C(O)OMe$ and formation of $MocPheOCH_2COOH$ and MocPheOH appear to follow Michaelis-Menten kinetics as expected (Section 2.3). The $v_0/[E]_0$'s calculated for $MocPheOCH_2C(O)OMe$ depletion were in excellent agreement with those determined by pH-stat (Fig. 3.4). The combined HPLC and pH-stat results are summarized in Section 3.2.

The ratio $[MocPheOCH_2COOH]/[MocPheOH]$ was determined using the five PB5 runs and found to be 8.2 ± 0.4. Plots of $[MocPheOCH_2COOH]/[MocPheOH]$ *versus* time and $[S]_0$ showed neglible slopes of 0.005 (±0.003) $min^{-1}$ and -0.07 (±0.03) $\underline{mM}^{-1}$ respectively.

### 3.1.3 $MocPheOCH_2C(O)NH_2$

$MocPheOCH_2C(O)NH_2$ was reacted with papain using the HPLC method to confirm that hydrolysis was occurring at the expected site. $MocPheOCH_2C(O)NH_2$ from the dialysis setup before addition of papain gave a single peak. After 15 min reaction of 1.9 $\underline{mM}$ $MocPheOCH_2C(O)NH_2$ with 2 $\underline{\mu M}$ papain, a gradient of 100% (0.1%TFA/$H_2O$) to 35% (0.1%TFA/$H_2O$)/65% MeCN over 6 min at 2.5 ml/min, gave peaks at 5.00 min and 5.10 min (MocPheOH). Spiking separate aliquots of the 15 min sample with pure $MocPheOCH_2C(O)NH_2$ and MocPheOH confirmed the identity of the peaks, hence confirming hydrolysis at the ester functionality ($MocPhe\hat{}OCH_2C(O)NH_2$).

**Figure 3.4** $v_0/[E]_0$ *vs.* $[S]_0$ for MocPheOCH$_2$C(O)OMe as determined by HPLC (●) and pH-stat (○).

Inset is the relative residual, i.e. $((v_0/[E]_0)_{obs} - (v_0/[E]_0)_{calc})/(v_0/[E]_0)_{calc}$ *vs.* $(v_0/[E]_0)_{calc}$.

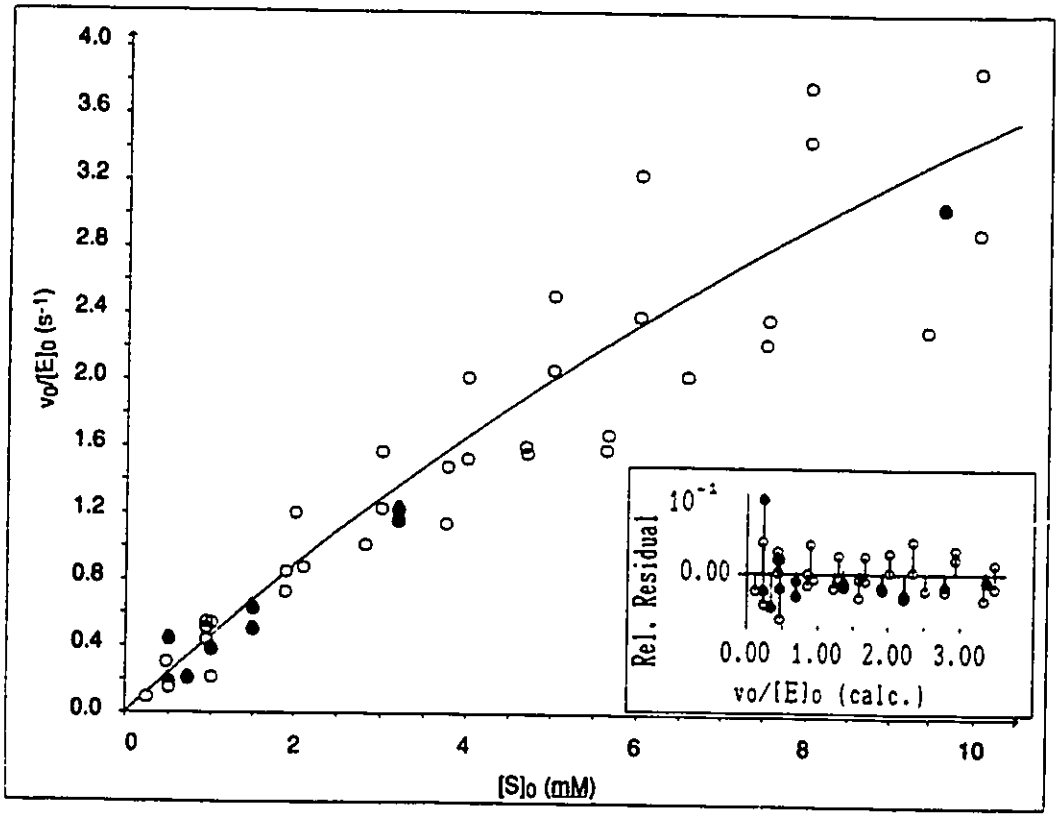

4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
vo/[E]o (s-1)
0
2
4
6
8
10
[S]o (mM)
Rel. Residual
10-1
0.00
0.00 1.00 2.00 3.00
vo/[E]o (calc.)

**Figure 3.4**

### 3.1.4 MocPheOCH$_2$COOH (as substrate)

On the basis of the HPLC runs of MocPheOCH$_2$C(O)OMe reacted with papain, particularly in light of the constant [MocPheOCH$_2$COOH]/[MocPheOH] with time, MocPheOCH$_2$COOH would be expected to be an extremely poor substrate, if at all. This was tested by reacting 3.5 mM MocPheOCH$_2$COOH with 15 μM papain for 60 min, then determining the products using solvent program PB5. HPLC analysis of the compound showed that it had partially decomposed since synthesis and contained 5.90% MocPheOH, which was taken into account when determining the amounts of products. After 60 min, 0.186 mM MocPheOH had formed, corresponding to a $v_0/[E]_0$ of $3.43 \cdot 10^{-3}$ $s^{-1}$. Using the Michaelis-Menten equation ($v_0/[E]_0 = (k_{cat}*[S]_0)/(K_M+[S]_0)$) and assuming a $k_{cat}$ of $\geq 0.01$ $s^{-1}$, 50-fold less than the lowest $k_{cat}$ measured in this study, a $K_M$ of 6.68 mM is obtained, yielding $k_{cat}/K_M \leq 1.50$ $M^{-1} \cdot s^{-1}$. Thus MocPheOCH$_2$COOH is a poor substrate, with an upper limit on the $k_{cat}/K_M$ of 1.5 $M^{-1} \cdot s^{-1}$.

### 3.1.5 Ph(CH$_2$)$_2$C(O)OCH$_2$C(O)OMe

Ph(CH$_2$)$_2$C(O)OCH$_2$C(O)OMe was examined by HPLC, as MocPheOCH$_2$C(O)OMe was, to determine the products of the reaction with papain and their ratios of formation. Three runs were done, at $[S]_0$ = 0.2, 0.5 and 2.0 mM. Ph(CH$_2$)$_2$C(O)OCH$_2$COOH was detected and quantified, but no Ph(CH$_2$)$_2$COOH was detected in any runs using solvent program PB6. After 76 min reaction time with 2 mM substrate and 6.2 μM papain, no Ph(CH$_2$)$_2$COOH peak was detected. The reaction mixture was then spiked to 0.1 mM with Ph(CH$_2$)$_2$COOH, and another sample taken 10 min later. The expected

amount of $Ph(CH_2)_2COOH$ was found, indicating that slow diffusion across the dialysis membrane was not a problem. The $Ph(CH_2)_2C(O)OCH_2COOH$ peak was 961 021 Area Counts, corresponding to 35.32 nmol. It was found using PB3 (which had a more steeply sloping baseline, making peak detection more difficult) that peaks *at least* as low as 2100 Area Counts could be detected. Conservatively assuming that a $Ph(CH_2)_2COOH$ peak of 5000 Area Counts could have been detected, there was ≤0.21 nmol present in the reaction mixture, resulting in a ratio $[Ph(CH_2)_2C(O)OCH_2COOH]/[Ph(CH_2)_2COOH] \geq 168$. Given $k_{cat}/K_M = 78.2$ for $Ph(CH_2)_2C(O)OCH_2C(O)\dagger OMe$ (*vide infra*); for $Ph(CH_2)_2COOH$, $k_{cat}/K_M \leq 0.5\ \underline{M}^{-1}\cdot s^{-1}$.

### 3.1.6 $Ph(CH_2)_2C(O)OCH_2C(O)NH_2$

As no reaction rate could be detected by pH-stat, $Ph(CH_2)_2C(O)OCH_2C(O)NH_2$ was tested as a substrate by HPLC. $Ph(CH_2)_2C(O)OCH_2C(O)NH_2$ from the dialysis setup before addition of papain gave a single peak. After reacting 5 mM $Ph(CH_2)_2C(O)OCH_2C(O)NH_2$ with 2 μM papain for 80 min, a gradient of 80% (0.1%TFA/$H_2O$)/20% MeCN to 35% (0.1%TFA/$H_2O$)/65% MeCN over 6.5 min at 2.5 ml/min, gave only the $Ph(CH_2)_2C(O)OCH_2C(O)NH_2$ peak at 3.63 min. Spiking an aliquot of the 80 min sample with $Ph(CH_2)_2COOH$ gave two peaks, with the second at 4.05 min. Using the same arguments as with $Ph(CH_2)_2C(O)\dagger OCH_2C(O)OMe$, with the $Ph(CH_2)_2C(O)OCH_2C(O)NH_2$ peak having an Area Count of 1 908 472 and assuming the same absorptivity as $Ph(CH_2)_2COOH$, the $[Ph(CH_2)_2COOH]$ was ≤0.0131 mM. This corresponds to $v_0/[E]_0 = 1.36\times10^{-3}\ s^{-1}$. Assuming again $k_{cat} \geq 0.01\ s^{-1}$, $K_M = 31.8$ mM and $k_{cat}/K_M \leq 0.31\ \underline{M}^{-1}\cdot s^{-1}$.

## 3.2 Kinetic constants - pH-stat & HPLC

With several substrates the plots of $v_0/[E]_0$ *versus* $[S]_0$ tended to have significant negative y-intercepts; that is, that the rate reached zero at significant substrate concentrations. This problem was caused by the observed tendency for the pH to drift upwards with the reaction solution. This trend is consistent and reproducible. Attempts to eliminate this drift over a ten year period using a variety of different approaches and instruments have failed to identify the source of the drift beyond it being linked to the presence of MeCN as a co-solvent or to eliminate it (A. Storer, personal communication).

Given the inability to remove the drift, it is possible to compensate for it numerically. By fitting the poorer substrates' data directly to the Michaelis-Menten equation plus offset using the program Enzfitter, better standard errors were obtained without materially changing the constants calculated. The most extreme example is $Ph(CH_2)_2C(O)OCH_2C(O)OMe$, where the observed rate decreased to zero below $[S]_0$ = 1 mM. Fitting to Michaelis-Menten equation alone gave $k_{cat}/K_M$ = 31.8 (±97.4) $M^{-1}\cdot s^{-1}$. After iteratively applying the offset until the standard error was larger than the magnitude of the calculated offset, $k_{cat}/K_M$ = 99.5 (±17.8) $M^{-1}\cdot s^{-1}$. The offset applied to $Ph(CH_2)_2C(O)OCH_2C(O)OMe$ was the largest applied to any of the substrates studied. Iterative applications of offset resulted in a smooth approach to zero offset in typically one or two, occasionally three iterations. The offsetted $v_0/[E]_0$'s were used with Cornish-Bowden's algorithm to calculate the kinetic constants.

When the Michaelis-Menten equation plus offset was applied to many of the better substrates, positive offsets were calculated. Applying these offsets resulted in the standard errors in $k_{cat}$ and $K_M$ increasing. Since negative drifts were not observed with any of the pH-stat experiments performed, this positive offset appears to be an artifact and fitting was to the Michaelis-Menten equation only.

The results of the pH-stat work are summarized in Table 3.2a. Previously unpublished pH-stat results which are necessary to the Discussion are listed in Table 3.2b. These kinetics determinations were performed by L. Sans Cartier and T. Hirama. The pH-stat and HPLC results for MocPheOCH$_2$C(O)OMe are summarized in Table 3.3.

**Table 3.2** Kinetic constants as determined by pH-stat.

**3.2a** Results from this study.
**3.2b** Previously determined and unpublished results.

| Substrate | $k_{cat}/K_M$ ($\underline{M}^{-1}\cdot s^{-1}$) | $k_{cat}$ ($s^{-1}$) | $K_M$ ($\underline{mM}$) | No. pt. |
|---|---|---|---|---|
| **Table 3.2a** | | | | |
| *N*-AcPheGly↑OMe | 120 000 ± 5000 | 9.4 ± 0.2 | 0.078 ± 0.004 | 37 |
| *O*-Acphenyllactyl-Gly↑OMe | 1100 ± 100 | 2.1 ± 0.2 | 1.9 ± 0.3 | 32 |
| MocPheGly↑OMe | 39 000 ± 4000 | 6.3 ± 0.2 | 0.16 ± 0.02 | 34 |
| Moc Phe↑$OCH_2C{=}ONH_2$ | 2200 ± 100 | 4.7 ± 0.2 | 2.2 ± 0.2 | 48 |
| Moc Phe↑$OCH_2C{=}OCH_3$ | 39 ± 3 | 1.6 ± 1.0 | 40 ± 28 | 29 |
| Moc Phe↑ $OCH_2C{=}OCH_2CH_3$ | 96 ± 5 | 15 ± 28 | 160 ± 300 | 27 |
| $Ph(CH_2)_2C(O)$-$OCH_2C(O)$↑OMe | 79 ± 4 | 0.67 ± 0.07 | 8.4 ± 1.2 | 26 |
| $HOCH_2C(O)$↑OMe* | ≈0 | - | - | |
| **Table 3.2b** | | | | |
| *N*-benzoylGly↑OMe | 140 ± 6 | 2.8 ± 0.2 | 20 ± 2 | 33 |
| *O*-benzoyl$OCH_2C(O)$↑OMe | 17 ± 2 | 0.23 ± 0.04 | 14 ± 4 | 49 |
| MocGlyGly↑OMe | 40 ± 2 | 1.8 ± 0.2 | 45 ± 6 | 22 |
| *N*-AcGly↑OEt** | 2.4 ± 0.2 | 1.4 ± 0.1 | 590 ± 100 | 16 |
| MocPhe↑OMe | 7.12 ± 0.04 | - | - | 9 |
| $Ph(CH_2)_2C(O)$Gly↑OMe*** | 710 ± 30 | 4.7 ± 0.2 | 5.9 | - |

* No detectable rate at 2 $\underline{mM}$ and [papain] = 2 $\underline{\mu M}$.
** $k_{cat}/K_M$ for *N*-AcGly↑OMe = 2.25 $\underline{M}^{-1}\cdot s^{-1}$ (59).
*** From (38).

**Table 3.3** Kinetic constants for MocPheOCH$_2$C(O)OMe as determined by pH-stat and HPLC.

**Direct fit** refers to kinetic constants determined from $v_0/[E]_0$'s and $[S]_0$'s by the method of Cornish-Bowden (82).

**From ratios** refers to kinetic constants determined from the relationships shown in Section 2.3 and the ratio of $k_{cat}/K_M$(MocPheOCH$_2$C(O)↑OMe) : $k_{cat}/K_M$(MocPhe↑OCH$_2$C(O)OMe) = 8.2 ± 0.4. The constants for the overall reaction were calculated from the results of the direct fits of HPLC results for MocPheOCH$_2$C(O)↑OMe and MocPhe↑OCH$_2$C(O)OMe data.

The 'best' kinetic constants, which are used in the remainder of this work are shown in **bold type**.

| Substrate | $k_{cat}/K_M$ ($M^{-1}.s^{-1}$) | $k_{cat}$ ($s^{-1}$) | $K_M$ (mM) | No. pt. |
|---|---|---|---|---|
| **DIRECT FIT:** | | | | |
| MocPheOCH$_2$C(O)OMe (OVERALL RATE) | | | | |
| pH-stat + HPLC | **480 ± 30** | **12 ± 5** | **26 ± 11** | 44 |
| HPLC only | 380 ± 20 | 23 ± 21 | 61 ± 58 | 9 |
| pH-stat only | 520 ± 40 | 10 ± 3 | 20 ± 8 | 35 |
| MocPheOCH$_2$C(O)↑OMe HPLC only | 320 ± 20 | 34 ± 71 | 105 ± 230 | 8 |
| MocPhe↑OCH$_2$C(O)OMe HPLC only | 52 ± 18 | 0.43 ± 0.34 | 8.3 ± 8.7 | 5 |
| **FROM RATIOS:** | | | | |
| MocPheOCH$_2$C(O)OMe (OVERALL) | 370 ± 30 | 8 ± 15 | 2.9 ± 8.2 | - |
| MocPheOCH$_2$C(O)↑OMe | **430 ± 35** | - | - | - |
| MocPhe↑OCH$_2$C(O)OMe | **52 ± 4** | - | - | - |

# 4. DISCUSSION

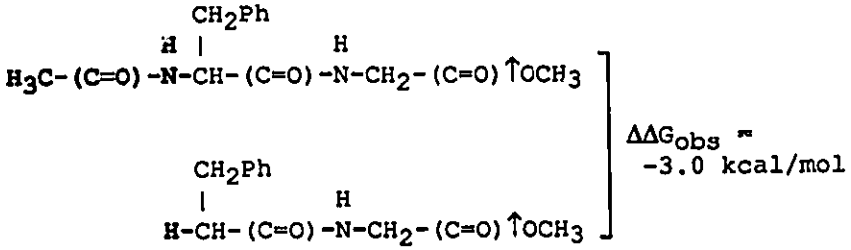
## 4.1 Interactions in the $S_2$ and $S_1$ subsites

By systematically varying substituents of substrates to fulfil specificity requirements for papain, we have been able to map out interactions in the $S_2$, $S_1$, $S_1'$ and $S_2'$ subsites. The incremental binding energies of the hydrogen bonding at the $P_2$ NH, $P_1$ NH, $P_1'$ C=O and the $P_2'$ NH, as well as of hydrophobic interactions of the $P_2$ Phe side chain were estimated. In contrast to similar recently published studies (57), we do not find it necessary to invoke any 'special' effects or interactions to explain the observed kinetic constants, only a consideration of standard binding energetics and entropy phenomena (62).

### 4.1.1 $P_2$ NH hydrogen bonding

The improvement to $k_{cat}/K_M$ by -3.0 kcal/mol upon the addition of an acetamide group to $Ph(CH_2)_2C(O)Gly\uparrow OMe$ (Scheme 4.1)[4] indicates that the

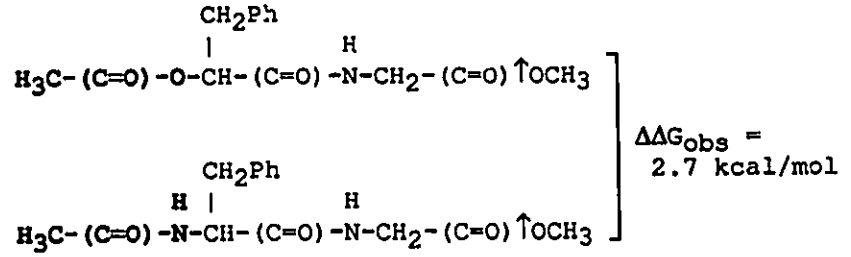
[4] In all the schemes in the Discussion, $\Delta\Delta G_{obs} = -RT*\ln((k_{cat}/K_M)_{first}/(k_{cat}/K_M)_{second})$. A positive $\Delta\Delta G_{obs}$ means that the second substrate listed has a higher $k_{cat}/K_M$, a negative $\Delta\Delta G_{obs}$ means that the first substrate is better.




*N*-AcPheGly↑OMe *vs.* Ph(CH$_2$)$_2$C(O)Gly↑OMe

**Scheme 4.1**

substrate's $P_3$-$P_2$ amide hydrogen bonds with the enzyme. Exchanging an ester for the amide group decreases specificity by $\Delta\Delta G_{obs}$ = 2.7 kcal/mol (Scheme 4.2), suggesting that it is the amide NH that is primarily

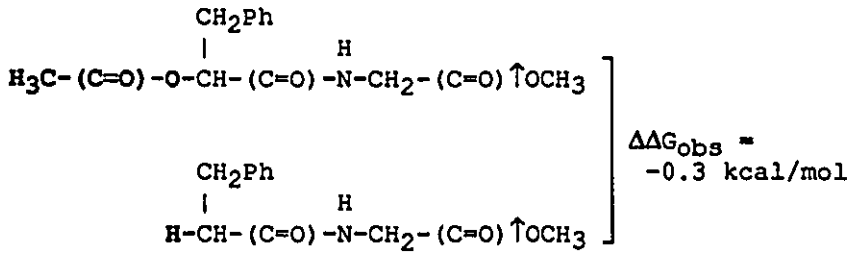



*O*-Ac-phenyllactlyGly↑OMe *vs.* *N*-AcPheGly↑OMe

**Scheme 4.2**

responsible for the observed binding energy.

If the $P_3$ C=O were to hydrogen bond to the enzyme, then a more negative $\Delta\Delta G_{obs}$ between O-Ac-phenyllactylGly↑OMe and Ph(CH$_2$)$_2$C(O)Gly↑OMe would be expected than the weakly favourable $\Delta\Delta G_{obs}$ = -0.3 kcal/mol (Scheme 4.3). Part of the 0.3 kcal/mol may be accounted for by differences in




*O*-Ac-phenyllactlyGly↑OMe *vs.* $Ph(CH_2)_2C(O)Gly\uparrow OMe$

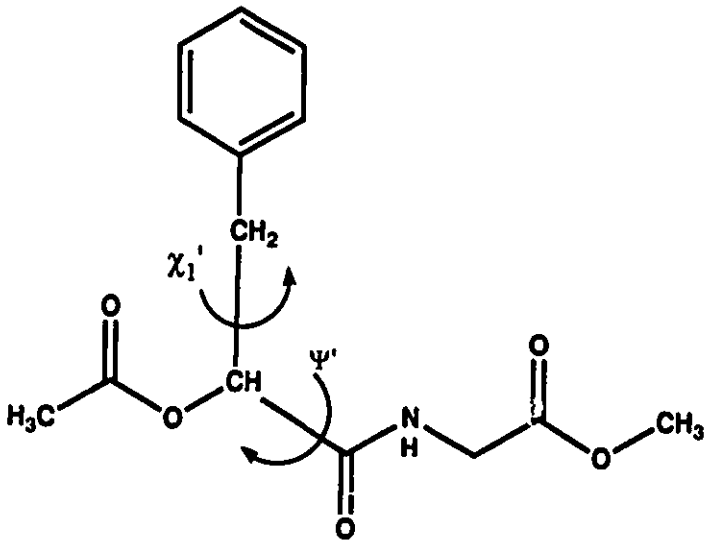
**Scheme 4.3**

conformational entropy. The presence of the acetyl group of *O*-Ac-phenyllactlyGly↑OMe will hinder internal rotations of the $\chi_1'$ and $\Psi'$ bonds of this Phe analog (Fig. 4.1), meaning that less entropy must be lost on forming the enzyme-substrate complex. Although the oxygen to nitrogen exchange between O-Ac-phenyllactlyGly↑OMe and N-Ac-PheGly↑OMe is a measure of incremental specificity energy rather than incremental binding energy, it is argued below (Section 4.4) that the energetic penalty of inserting a substrate oxygen adjacent to an enzyme's hydrogen bond acceptor is small.

Two other lines of evidence support the assumption that the $P_3$ C=O does not hydrogen bond the enzyme. In the x-ray crystallographic structures of α-chloroketone inhibited papains where the inhibitor extends as far

**Figure 4.1** *O*-AcphenyllactylGlyOMe

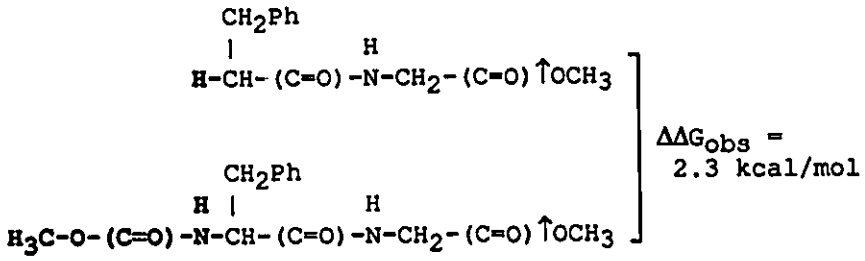
The arrows represent the two dihedral bond rotations that become significantly more restricted upon the addition of and acetyl group to $Ph(CH_2)_2C(O)GlyOMe$. The $\Psi'$ and $X_3$ angles are so named by analogy to the convention of naming the $\Psi$ and $X$ angles of peptides (97).



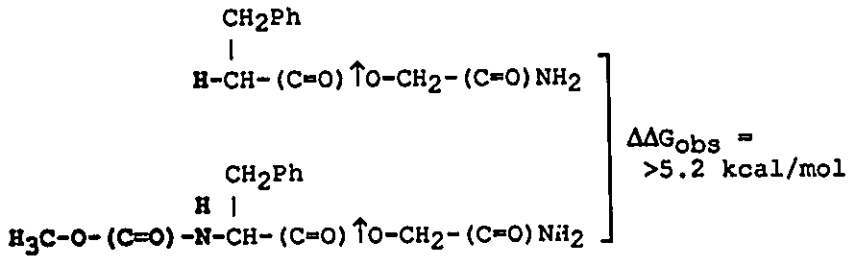

*O*-Ac-phenyllactylGlyCOOMe

**Figure 4.1**

as the $P_5$ subsite, the $P_3$ C=O is not in contact with the enzyme. Rather, it is solvent exposed (5). Secondly, the $\Delta\Delta G_{obs}$ between $Ph(CH_2)_2C(O)Gly\uparrow OMe$ and $MocPheGly\uparrow OMe$ (2.3 kcal/mol) (addition of $P_3$-$P_2$ carbamate, Scheme 4.4) is much smaller than between




$Ph(CH_2)_2C(O)Gly\uparrow OMe$ *vs.* $MocPheGly\uparrow OMe$



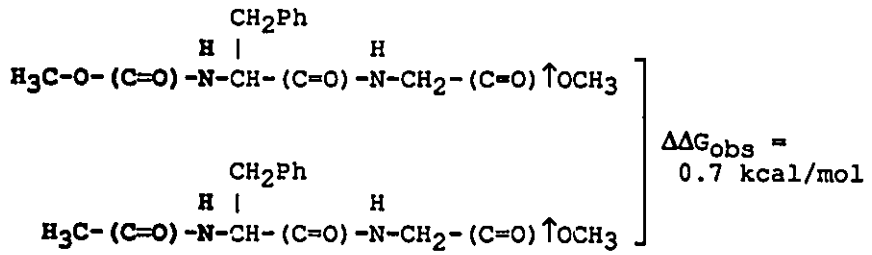

$Ph(CH_2)_2C(O)\uparrow OCH_2C(O)NH_2$ *vs.* $MocPhe\uparrow OCH_2C(O)NH_2$

**Scheme 4.4**

$Ph(CH_2)_2C(O)\uparrow OCH_2C(O)NH_2$ and $MocPhe\uparrow OCH_2C(O)NH_2$ (>5.2 kcal/mol) (addition of $P_2$-$P_1$ carbamate). This indicates that the full hydrogen bonding capabilities of the $P_3$-$P_2$ carbamate are not being utilized as compared to the $P_2$-$P_1$ carbamate. In the latter case it is known from the x-ray structures (5) that both the $P_2$ C=O and the $P_1$ NH can hydrogen bond to the enzyme, leading to a more negative $\Delta\Delta G_{obs}$.

*4.1.1.1 Carbamate versus Amide*

The use of a carbamate group in the place of an amide in $MocPheGly\uparrow OMe$ leads to a loss of specificity by $\Delta\Delta G_{obs}$ = 0.7 kcal/mol (Scheme 4.5).



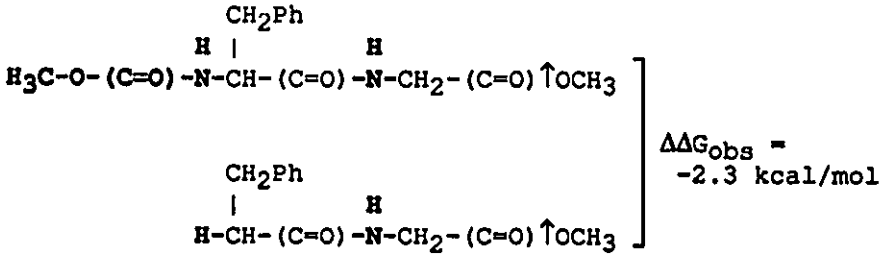

MocPheGly↑OMe *vs.* *N*-AcPheGly↑OMe

**Scheme 4.5**

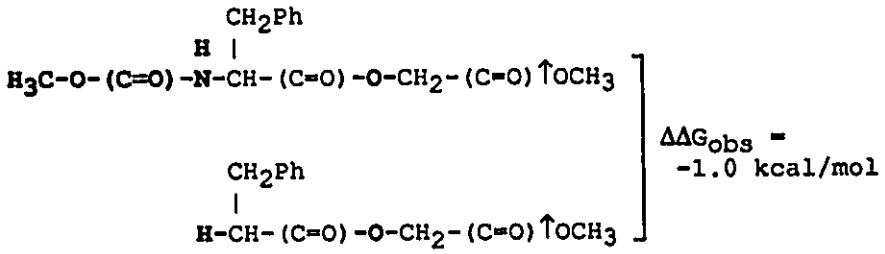
This effect is attributable to the added oxygen of the carbamate changing the electron density distribution about the -C(O)-NH- functionality. The effect of the oxygen cannot easily be predicted; however, if it causes increased electron density on the NH, causing both a shorter N-H bondlength and a decreased positive charge on the proton, a weaker hydrogen bond donor will result. This qualitative description is currently being investigated quantitatively using distributed multipole analysis (88) to calculate molecular charge distributions.

*4.1.1.2 Hydrogen bond strength*

The $\Delta\Delta G_{obs}$'s on addition of a $P_2$ NH range from -3.0 to -1.0 kcal/mol and clearly depend on the presence of other enzyme-substrate interactions. Scheme 4.6 illustrates that the $\Delta\Delta G_{obs}$ of the addition of a methyl



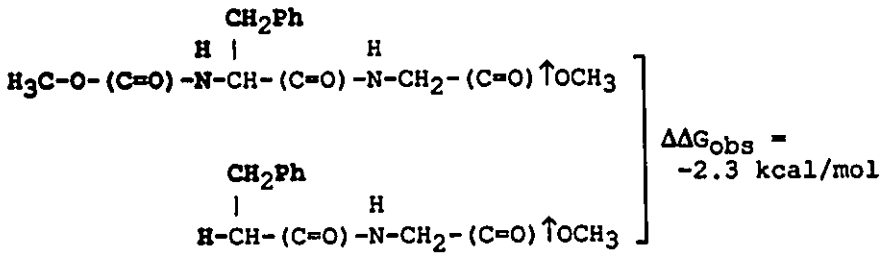

MocPheGly↑OMe *vs.* Ph(CH$_2$)$_2$C(O)Gly↑OMe



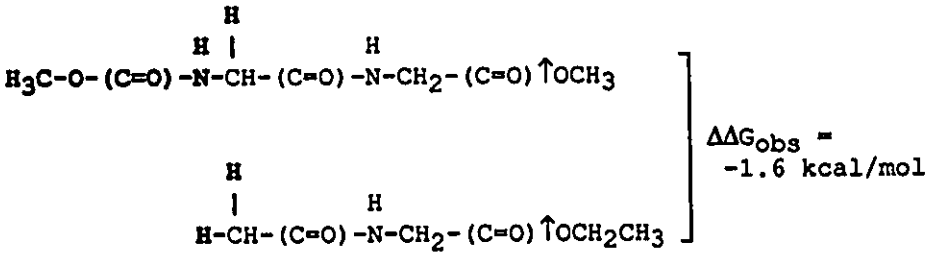

MocPheOCH$_2$C(O)↑OMe *vs.* Ph(CH$_2$)$_2$C(O)OCH$_2$C(O)↑OMe

**Scheme 4.6**

carbamate group depends on the presence of other enzyme-substrate contacts. Specifically, the presence of the $P_1$ NH makes the added $P_2$ NH more favourable than when there is an ester at $P_2$-$P_1$. Scheme 4.7 illustrates that the same effect is seen in the presence or absence of the $P_2$ Phe side chain.




MocPheGly↑OMe *vs.* Ph(CH$_2$)$_2$C(O)Gly↑OMe



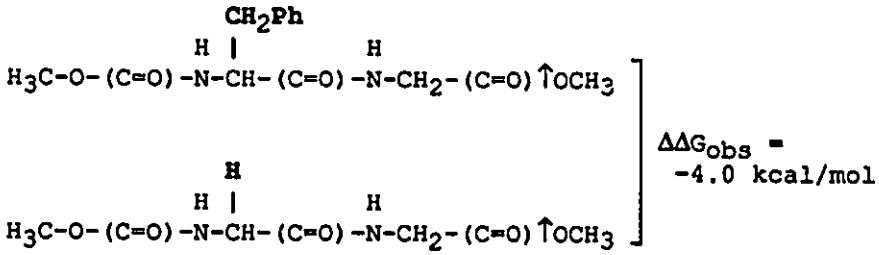

MocGlyGly↑OMe *vs.* *N*-AcGly↑OEt

**Scheme 4.7**

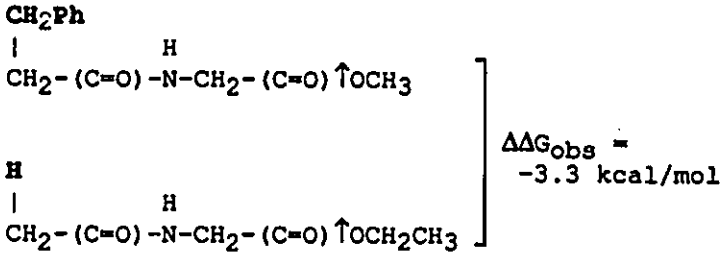
## 4.1.2 $P_2$ side chain interactions

A number of different side chains have been investigated in the $S_2$ subsite (unpublished data). As the $S_2$ binding subsite is progressively

occupied with hydrophobic interactions, $k_{cat}/K_M$ increases. On adding a benzyl group to MocGlyGly↑OMe to give MocPheGly↑OMe, $k_{cat}/K_M$ improves by -4.0 kcal/mol (Scheme 4.8).




MocPheGly↑OMe *vs.* MocGlyGly↑OMe




$Ph(CH_2)_2C(O)Gly\uparrow OMe$ *vs.* *N*-AcGly↑OEt

**Scheme 4.8**

Computationally it is possible to partly dissect out the contribution of the conformational entropy of the Phe *versus* Gly to the binding energy. The calculations of Brant *et al.* (89) indicated that the conformational entropy of a Gly residue about the Φ and Ψ angles is only slightly higher than that of an amino acid with a non-β-branched side chain. The upper limit of this difference has been estimated to be on the order of 1.5 cal/K/mol, corresponding to a maximum effect of 0.4 kcal/mol. If the total decrease in conformational entropy is directly converted into binding energy, $k_{cat}/K_M$ would increase by a maximum factor of 2.1 with respect to a substrate with a $P_2$ Gly. This establishes the absolute upper limit of the contribution of the decrease in dihedral bond

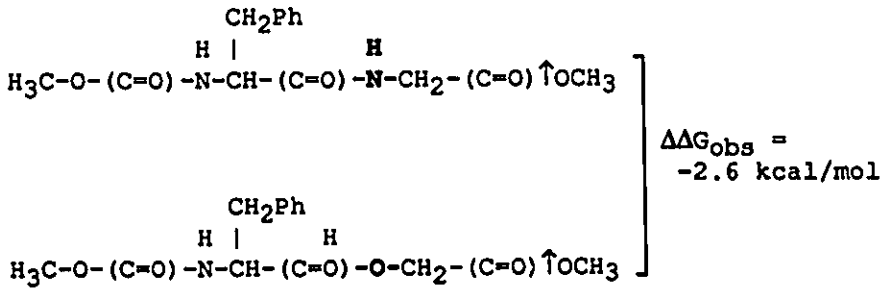
rotational entropy to the $\Delta\Delta G_{obs}$ on adding a Phe side chain; in fact, the effect would be expected to be smaller than this.

As is the case with the hydrogen bonding interactions, the effect of side chain binding in $S_2$ is interdependent with other substrate-enzyme interactions. The added side chain has the greatest effect when the $P_2$ NH and the $P_2$-$P_1$ amide are present as witnessed by the difference between N-AcGlyŶOEt[5] *vs.* $Ph(CH_2)_2C(O)GlyŶOMe$ and MocGlyGlyŶOMe *vs.* MocPheGlyŶOMe (Scheme 4.8).

## 4.1.3 $P_1$ hydrogen bonding

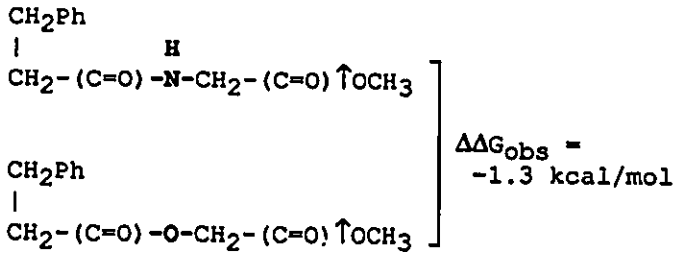
The presence of the hydrogen bond to the $P_1$ NH was demonstrated and its strength evaluated as for the $P_2$ NH.

Substitution of a $P_2$-$P_1$ ester for the amide demonstrates that the $P_1$ NH is hydrogen bonded to the enzyme (Scheme 4.9) with $\Delta\Delta G_{obs}$ = -1.2 to -2.6 kcal/mol. The x-ray structures of α-chloroketone inhibited papains (5) indicate that the Asp 158 backbone C=O is the probable hydrogen bond acceptor.

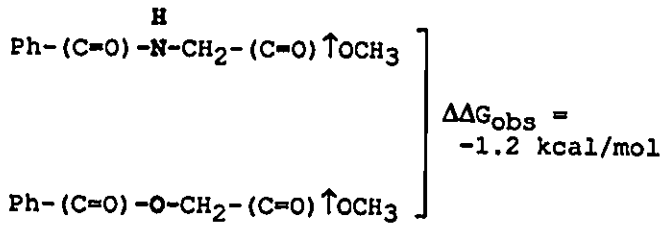



MocPheGlyŶOMe *vs.* $MocPheOCH_2C(O)ŶOMe$

[5] Williams *et al.* (59) found $k_{cat}/K_M$ = 2.25 $\underline{M}^{-1}\cdot s^{-1}$ for *N*-AcGlyCOŶOMe as opposed to $k_{cat}/K_M$ = 2.4 $\underline{M}^{-1}\cdot s^{-1}$ for *N*-AcGlyCOŶOEt reported here, showing a negligible contribution from the ethyl *versus* the methyl leaving group and indicating that the comparison with $PhCH_2CH_2C(O)GlyŶOMe$ is valid.




$Ph(CH_2)_2C(O)Gly\uparrow OMe$ *vs.* $Ph(CH_2)_2C(O)OCH_2C(O)\uparrow OMe$




*N*-benzoylGly↑OMe *vs.* O-benzoylglycolic acid methyl ester

**Scheme 4.9**

The x-ray structures also indicate that the $P_2$ C=O should be hydrogen bonded, as is evidenced by the more favourable $\Delta\Delta G_{obs}$ between $MocPhe\uparrow OCH_2C(O)NH_2$ *versus* $Ph(CH_2)_2C(O)(\uparrow)OCH_2C(O)NH_2$ as opposed to $MocPheGly\uparrow OMe$ *versus* $MocPheOCH_2C(O)\uparrow OMe$ (Scheme 4.4). It is also in accord with the results indicating that substitution of the $P_2$ C=O with C=S decreases $k_{cat}/K_M$ for the different cysteine proteases examined (58,59).

Once again, the $\Delta\Delta G_{obs}$ of the $P_1$ NH is dependent on other enzyme-substrate interactions, with the presence of the $P_2$ NH and the $P_2$ Phe side chain cooperating to increase the effect of the $P_1$ NH.

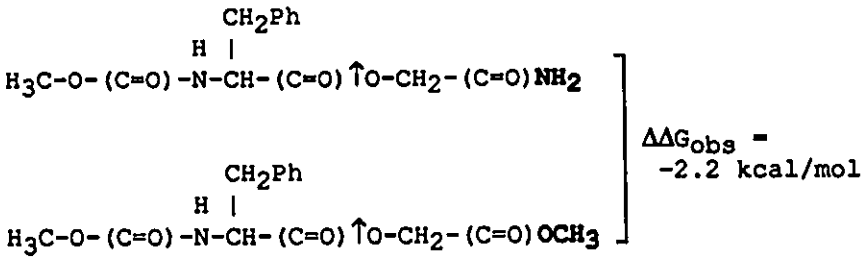
## 4.2 $S_1'$ hydrogen bond donor and $S_2'$ hydrogen bond acceptor

### 4.2.1 MocPheO$CH_2C(O)NH_2$ based substrates

#### *4.2.1.1 Presence of hydrogen bonds*

The kinetic constants of MocPhe↑$OCH_2C(O)NH_2$ and related substrates demonstrate that there are hydrogen bonding partners for the $P_1'$-$P_2'$ amide group of peptide substrates.

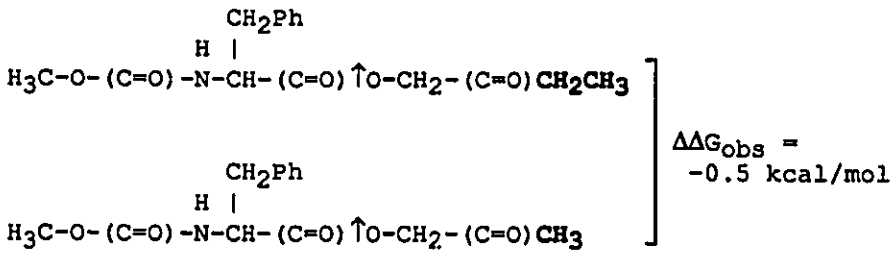
The specificity difference between MocPhe↑$OCH_2C(O)NH_2$ and MocPhe↑$OCH_2C(O)OMe$ (Scheme 4.10) can be attributed to hydrogen bonding




MocPhe↑$OCH_2C(O)NH_2$ *vs.* MocPhe↑$OCH_2C(O)OMe$

**Scheme 4.10**

effects or to unfavourable steric interactions between the ester's methyl group of MocPhe↑$OCH_2C(O)OMe$ and the enzyme. Isosteric analogs of each compound that were not hydrogen bonding in the region under consideration were synthesized (Scheme 4.11). The difference between

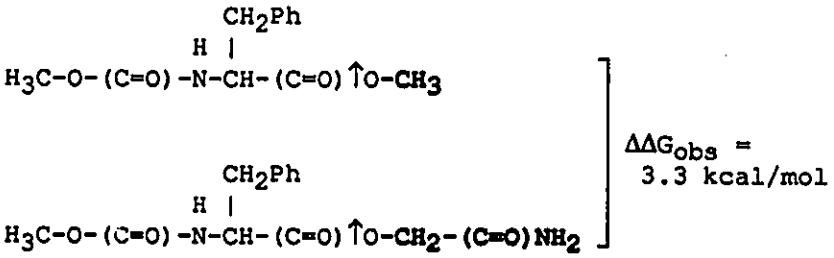



MocPhe↑OCH$_2$C(O)CH$_2$CH$_3$ *vs.* MocPhe↑OCH$_2$C(O)CH$_3$

**Scheme 4.11**

MocPhe↑OCH$_2$C(O)CH$_3$ and MocPhe↑OCH$_2$C(O)CH$_2$CH$_3$ should reflect steric effects only. The $\Delta\Delta G_{obs}$ = -0.5 kcal/mol indicates that there is not a problem sterically with the size of the MocPhe↑OCH$_2$C(O)OMe leaving group and therefore that the difference between MocPhe↑OCH$_2$C(O)NH$_2$ and MocPhe↑OCH$_2$C(O)OMe is an hydrogen bonding effect.

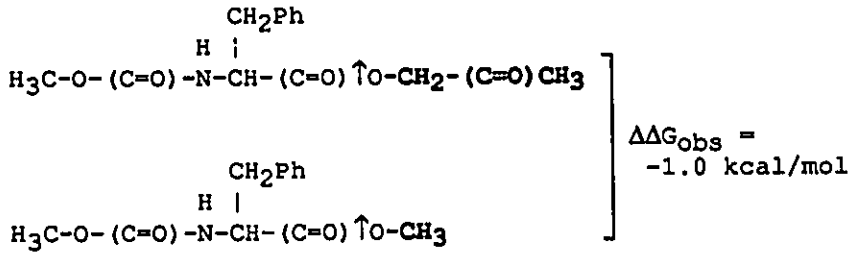
Comparison of MocPhe↑OMe and MocPhe↑OCH$_2$C(O)NH$_2$ (Scheme 4.12) gives $\Delta\Delta G_{obs}$ = 3.3 kcal/mol. It is possible to separate out the C=O and -NH$_2$ binding effects.




MocPhe↑OCH$_3$ *vs.* MocPhe↑OCH$_2$C(O)NH$_2$

**Scheme 4.12**

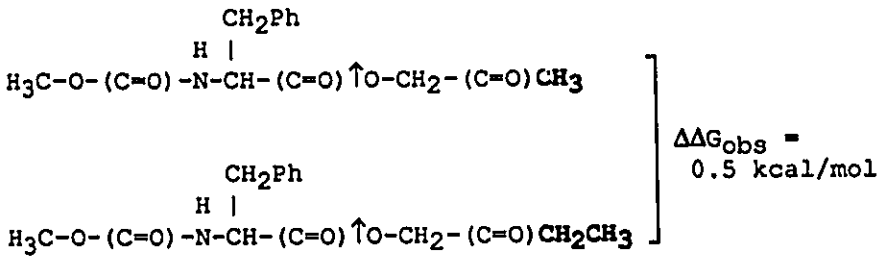
The hydrogen bonding contribution of the carbonyl portion may be estimated thus. The observed $\Delta\Delta G_{obs}$ between MocPhe↑OCH$_2$C(O)CH$_3$ and MocPhe↑OMe is -1.0 kcal/mol (Scheme 4.13). Lowe & Williams (90)



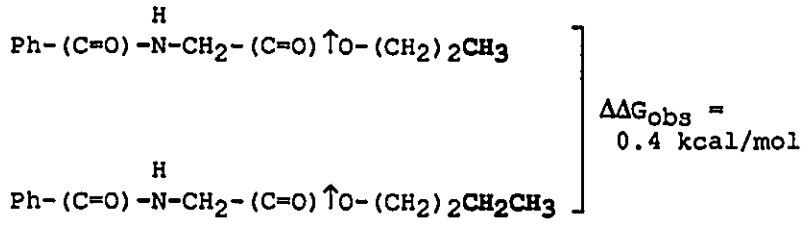

MocPhe↑OCH$_3$ *vs.* MocPhe↑OCH$_2$C(O)CH$_3$

**Scheme 4.13**

found that the $k_{cat}/K_M$'s of N-benzoylGly↑OMe (133 $\underline{M}^{-1}\cdot s^{-1}$), ↑-OEt ($k_{cat}/K_M$ = 133 $\underline{M}^{-1}\cdot s^{-1}$) and ↑-O*n*Pr (118 $\underline{M}^{-1}\cdot s^{-1}$, $\Delta\Delta G_{obs}$ < 0.1 kcal/mol) were the same. Thus, the favourable $\Delta\Delta G_{obs}$ between MocPhe↑OMe and MocPhe↑OCH$_2$C(O)CH$_3$ may be attributed to the binding of the carbonyl, with no apparent contribution from the methyl group (↑-OCH$_2$C(O)**CH$_3$**). For the comparison of N-benzoylGly↑O*n*Pr with MocPhe↑OCH$_2$C(O)CH$_3$ to be valid, it is necessary that the leaving group's carbons interact with the enzyme in the same conformation in each substrate. This assumption is validated by noting that on adding a methylene to either leaving group, the $k_{cat}/K_M$ is improved by the same amount (Scheme 4.14).




MocPhe↑OCH$_2$C(O)CH$_3$ *vs.* MocPhe↑OCH$_2$C(O)CH$_2$CH$_3$

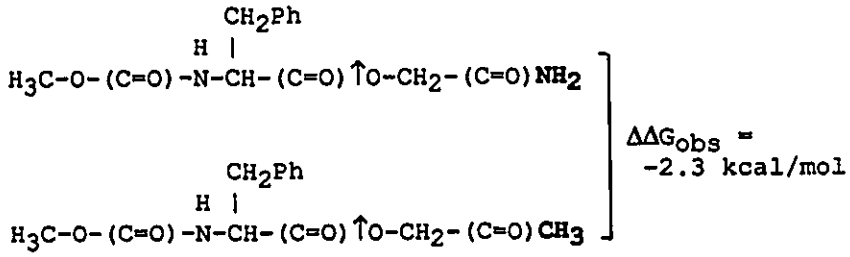



*N*-benzoylGly↑O-*n*Pr *vs.* *N*-benzoylGly↑O-*n*Bu
(from (90))

**Scheme 4.14**

That there is no difference between the methyl and ethyl esters of N-benzoylglycine as substrates (90) suggests that the $P_1'$ carbonyl carbon of MocPhe↑$OCH_2C(O)NH_2$ would not be expected to contribute favourably to the apparent binding energy through van der Waals' contacts. The observed $\Delta\Delta G_{obs}$ = -1.0 kcal/mol for the $P_1'$ carbonyl is therefore due primarily to effects of the carbonyl oxygen, forming an hydrogen bond with a donor on the enzyme. The observed $\Delta\Delta G_{obs}$ does not arise from very favourable dispersion energy of the oxygen interacting with a non-polar part of the enzyme since one would then expect an equivalent effect on going from an ethyl to an *n*-propyl leaving group as the terminal methyl of the *n*-propyl is isosteric with the carbonyl oxygen.

Given the contribution of -1.0 kcal/mol by the carbonyl group, the remaining -2.3 kcal/mol of the $\Delta\Delta G_{obs}$ between MocPhe↑OMe and MocPhe↑$OCH_2C(O)NH_2$ is accounted for by hydrogen bonding of the $-NH_2$. This is the same as the binding energy observed comparing MocPhe↑$OCH_2C(O)NH_2$ with MocPhe↑$OCH_2C(O)CH_3$ (Scheme 4.15). The $-NH_2$ to



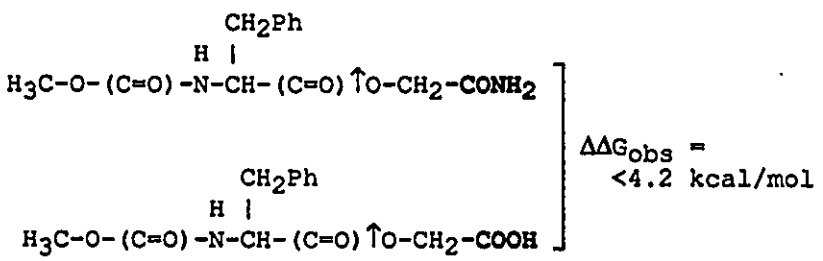

MocPhe↑OCH$_2$C(O)NH$_2$ *vs.* MocPhe↑OCH$_2$C(O)CH$_3$

**Scheme 4.15**

$-CH_3$ substitution is not generally acceptable for determining incremental binding energies, however in this case the methyl group does not appear to affect the kinetics. The lack of effect of this methyl group appears to be a function of three factors. Firstly, on the basis of the kinetics of N-benzoylGly↑OMe, ↑-OEt and ↑-O*n*Pr, this methyl group does not appear to interact with the enzyme (90). Secondly, any interactions of the methyl with the hydrogen bond acceptor in $S_2'$ will be dominated by dispersion forces, which will be favourable, leading to only a slight reduction of $\Delta\Delta G_{obs}$. Thirdly, there is the possibility that the methyl group would have reduced the accessibility of bulk water to the hydrogen bond acceptor at $S_2'$. If this were the case, empirical evidence indicates that if it is an uncharged hydrogen bond acceptor (e.g. C=O), then the various effects will tend to cancel out (68). On the other hand, leaving an unpaired charged hydrogen bond acceptor (e.g.-$COO^-$) will lead to $\Delta\Delta G_{obs} > \Delta\Delta G_{bind}$ (68). Since $\Delta\Delta G_{obs}$ is very similar to those observed for the $P_2$ and $P_1$ NH hydrogen bonds, which are known to involve dipole-dipole interactions, it is probable that the $P_2'$ NH hydrogen bond is also dipole-dipole. It is possible that a weak charge-dipole interaction is being observed, but the lack of obvious candidates as the charged side chain in the x-ray structure of papain argues against this. The incremental binding energy of the

MocPhe↑$OCH_2C(O)NH_2$···:B-Enzyme hydrogen bond is therefore concluded to be on the order of -2.3 kcal/mol.

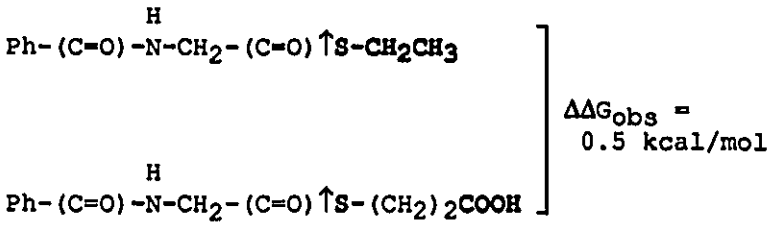
Further evidence for an hydrogen bond acceptor is the very poor $k_{cat}/K_M$ < 1.5 $\underline{M}^{-1}\cdot s^{-1}$ for MocPhe↑$OCH_2COOH$ (Scheme 4.16). Removal of the $-NH_2$




MocPhe↑$OCH_2C(O)NH_2$ *vs.* MocPhe↑$OCH_2COOH$

**Scheme 4.16**

accounts for 2.3 kcal/mol of this with the other > 1.9 kcal/mol being accounted for by the unfavourable carboxylate-dipole interaction. Finally, a carboxylate shifted by one methylene further away from the catalytic site interacts favourably with the enzyme, indicating a strong location dependence of the binding energy penalty of the $-COO^-$ (90) (Scheme 4.17).




*N*-benzoylGly↑S-Et *vs.* *N*-benzoylGly↑S$(CH_2)_2$COOH
(from (90))

**Scheme 4.17**

*4.2.1.2 Added nucleophiles*

The evidence for the presence of hydrogen bonding partners for the $P_1'$ C=O and the $P_2'$ NH is in direct contrast with the observations of

Brubacher and Bender (39), who found that the presence of this amide functionality did not increase the second order rate constants for added nucleophiles in the deacylation of *trans*-cinnamoyl-papain. The work of Alecio *et al.* (47), however, supports the role of the $P_1'$-$P_2'$ amide in the hydrolysis of amides.

Brubacher and Bender (39) found that there was no decrease in the second order rate constants using L-tryptophanamide *versus* L-tryptamine or methyl L-tryptophanate. Similarly, the similarity of specificity of glycinamide ($k_{4,II}$ = 0.53 $\underline{M}^{-1}\cdot s^{-1}$) and aminoacetonitrile ($k_{4,II}$ = 0.48 $\underline{M}^{-1}\cdot s^{-1}$) implies that the $-NH_2$ of glycinamide does not form an hydrogen bond.

There are likely two different principles in effect with the tryptophan-based nucleophiles and the glycine-based nucleophiles. The large difference between the $k_{4,II}$'s for L-tryptophanamide and glycinamide in spite of their very similar inherent nucleophilicities and $pK_a$'s indicates that there is binding of the tryptophan side chain in a quasi-equilibrium step before nucleophilic attack by the amine group on the acyl enzyme. The lack of effect of the presence of the $P_1'$-$P_2'$ amide indicates that the binding of the tryptophan side chain dominates the association of nucleophile with enzyme and that the amide (or methyl ester) is in a different orientation than the amide of $MocPhe\uparrow OCH_2C(O)NH_2$. The difference in position need not be large given the strong length-dependence and directionality of hydrogen bonds (particularly hydrogen bond donors (91)) and given the fact that a weak hydrogen bond is worse than none at all (92).

In the case of glycinamide and aminoacetonitrile, it is unlikely, due to entropic effects, that an adsorptive complex is formed between the enzyme and the nucleophiles in the $S_1'$ subsite. The loss of entropy entailed in binding a small molecule to an enzyme can be estimated variously to be worth 10 to 13 kcal/mol (Section 4.5.2.1). Clearly, the two hydrogen bonds through $-C(O)NH_2$, even in the presence of strong dispersion energies acting on the $C_\alpha$ of glycinamide (for which there is no evidence for or against) will not lead to appreciable binding of the nucleophile to the enzyme in a quasi-equilibrium step before nucleophilic attack. The observed $k_{4,II}$'s represent then, the attack of the free nucleophile onto the acyl-enzyme.

### 4.2.2 $Ph(CH_2)_2C(O)OCH_2C(O)NH_2$ based substrates

$Ph(CH_2)_2C(O)OCH_2C(O)NH_2$ was found not to be a substrate on the basis of pH-stat and HPLC results. By HPLC, an upper limit of $k_{cat}/K_M \leq 0.31\ \underline{M}^{-1}\cdot s^{-1}$ was established. As discussed previously, the difference between $Ph(CH_2)_2C(O)-$ and MocPheŤ$OCH_2NH_2$ is a reflection of the hydrogen bonding energies of the $P_2$-$P_1$ amide group.

Asboth and Polgar (93) reported a $k_{cat}/K_M = 2\ \underline{M}^{-1}\cdot s^{-1}$ for N-benzoylO$CH_2$Ť$NH_2$. Given that β-phenylpropionyl- containing substrates have higher $k_{cat}/K_M$'s than N-benzoyl- containing substrates (38), it appears that this observed $k_{cat}/K_M$ is due to either contamination of the substrate or an insufficiently accurate rate assay. Rates were determined in that study by detecting with a ninhydrin assay the ammonia liberated during the hydrolysis of the amide. This was done by taking an aliquot of the reaction solution, adding a ninhydrin solution,

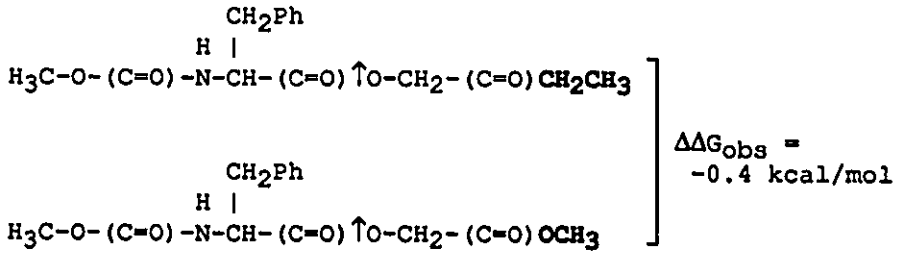
heating at 100°C for 15 min, cooling on ice for 3 min, then measuring $A_{570}$, centrifuging if necessary to remove turbidity. This method of rate determination is more prone to errors than the HPLC method of this study. The [papain] = 0.4 to 0.6 μM, was lower than the 2 μM used in this study. The $Ph(CH_2)_2C(O)OCH_2C(O)NH_2$ used here was shown to be pure by HPLC and it was demonstrated using a higher enzyme concentration that there was no product formation of any kind.

## 4.3 Ester for amide substitution

The $\Delta\Delta G_{obs}$'s for the $P_2$ and $P_1$ NH's have been estimated, at least in part, on the substitution of amide functionalities with esters at the site of the NH in question and therefore represent estimates of incremental specificity energies. On the basis of the ester *versus* amide comparisons alone, it is possible to argue that the incremental binding energies of the hydrogen bonded NH's are considerably less than the -2.6 to -2.7 kcal/mol observed incremental specificity energies and that the unfavourable interaction of the oxygen with the enzyme's hydrogen bond acceptor contributes most of the $\Delta\Delta G_{obs}$'s. However, our results indicate that the energetic penalty on placing an oxygen adjacent to the enzyme's hydrogen bond acceptor is small; on the order of 0.4 kcal/mol.

For the $P_2'$ NH, the binding energy of -2.3 kcal/mol was determined independently of the ester and represents an incremental binding energy. The penalty to $k_{cat}/K_M$ on placing an ester's alcoholic oxygen adjacent to the enzyme's hydrogen bond acceptor in the $S_2'$ subsite is

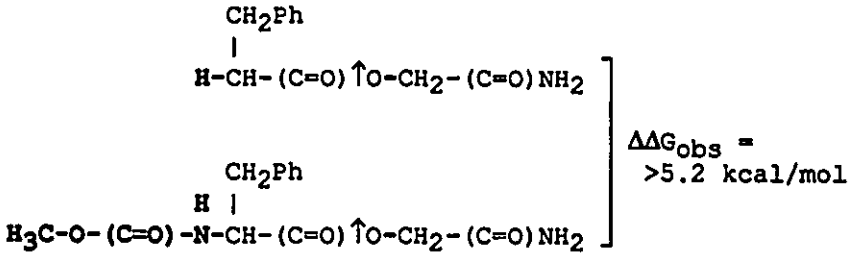
approximately 0.4 kcal/mol as determined by the comparison of MocPhe↑$OCH_2C(O)OMe$ with MocPhe↑$OCH_2C(O)CH_2CH_3$ (Scheme 4.18).




MocPhe↑$OCH_2C(O)CH_2CH_3$ *vs.* MocPhe↑$OCH_2C(O)OMe$

**Scheme 4.18**

Similarly, the energetic penalty of placing an ester at $P_2$-$P_1$ is not large; the $\Delta\Delta G_{obs}$ on substituting an ester for an amide at $P_2$-$P_1$ is due primarily to the loss of favourable binding interactions of the $P_1$ NH rather than addition of unfavourable interactions of the $P_1$ -O-. This is illustrated by considering the addition of a methyl carbamate to $Ph(CH_2)_2C(O)OCH_2C(O)NH_2$, increasing specificity by $\Delta\Delta G_{obs} < -5.2$ kcal/mol (Scheme 4.19). This energy must arise primarily from the hydrogen




$Ph(CH_2)_2C(O)$↑$OCH_2C(O)NH_2$ *vs.* MocPhe↑$OCH_2C(O)NH_2$

**Scheme 4.19**

bonding of the carbonyl group and of the NH. There is no hydrogen bond donor in the position that the $sp^3$ oxygen would be expected to occupy and the terminal methyl group is positioned adjacent to Gly 66 NH. A binding energy of better than -5.2 kcal/mol is too large to have been contributed by the C=O alone; even assuming 1 kcal/mol in favourable dispersion energy being contributed by the $H_3C$-O-, a contribution on the

order of 2 kcal/mol from the NH is required. This compares favourably with the -2.6 kcal/mol observed between the ester and amide functionalities at this site and therefore represents primarily the favourable binding energy of the NH rather than the unfavourable interaction of the oxygen. Using the figure of 0.4 kcal/mol for unfavourable binding of an ester in place of an amide, the incremental binding energies may be estimated at -2.3 kcal/mol for the $P_2$ NH and -2.2 kcal/mol for the $P_1$ NH.

Thus, the main contributor to the incremental specificity energies between ester and amide containing substrates is the favourable hydrogen bonding energy of NH (-2.2 to -2.3 kcal/mol) and only a minor contribution from the unfavourable oxygen-hydrogen bond acceptor contacts (≈ 0.4 kcal/mol). Papain expresses a specificity *for* NH's at $P_2$, $P_1$ and $P_2'$, rather than a specificity *against* oxygens.

## 4.4 Interdependence of interaction energies

### 4.4.1 Combined $P_2$ NH, $P_2$ side chain and $P_1$ NH effects

The observed interdependence of effects at the $P_2$ NH, the $P_2$ side chain and the $P_2$-$P_1$ amide are typical of enzyme-substrate interactions. When a favourable interaction such as an hydrogen bond is added to a enzyme-substrate complex, there is a loss of internal entropy of the complex. This includes vibrational and overall rotational motions of the substrate and enzyme relative to each other, as well as bond rotations that become more restricted. Other factors such as increased solvent exclusion (which may or may not be entropy driven) may also contribute

to the energy change. This loss of entropy causes, then, an increase in the free energy of the enzyme-substrate complex (E·S) and the transition state ($ES^{\ddagger}$). The magnitude of the energy increase depends on the tightness of binding of the substrate before the addition of the hydrogen bond. The incremental specificity energy of the hydrogen bond is offset to a certain extent by the higher energy of the enzyme-substrate complexes (*i.e.* E·S and $ES^{\ddagger}$). A 'loose' complex where there is considerable freedom of motion of the substrate in the active site will undergo a greater entropy loss than a 'tight' complex. The observable effect of this condition is that on the addition of an hydrogen bond of a given incremental specificity energy to a substrate, the observed change in free energy ($\Delta\Delta G_{obs}$) depends on the presence of other interactions and will be greater when it is added to a good *versus* a poor substrate.

This effect is illustrated by comparing the addition of a $P_1$ NH to $MocPheOCH_2C(O)\uparrow OMe$ *versus* adding the same $P_1$ NH to *O*-benzoylglycolic acid methyl ester or $Ph(CH_2)_2C(O)OCH_2C(O)\uparrow OMe$ (Scheme 4.11). The former case yields $\Delta\Delta G_{obs}$ = -2.6 kcal/mol whilst for the latter cases, $\Delta\Delta G_{obs}$ = -1.2 and -1.3 kcal/mol respectively. The presence of the extra $P_2$ NH (and $P_2$ Phe side chain compared with *O*-benzoylglycolic acid methyl ester) in $MocPheOCH_2C(O)\uparrow OMe$ reduces the motions of complexed substrate and enzyme molecules relative to each other, making the entropic penalty upon addition of the $P_1$ NH less than with the other two substrates.

Alternately phrased: because the entropy of the enzyme-substrate complex is higher for poor substrates, the restriction of motion necessary to form the hydrogen bond with the $P_1$ NH raises the energy of

the enzyme-substrate complex (E·S) and the transition state ($ES^{\ddagger}$) *more* for poor substrates than for good. In each case, the NH group interacts equally strongly with the hydrogen bond acceptor, but for poor substrates a larger portion of this energy compensates for the higher free energy of the complex and therefore appears as a poorer $\Delta\Delta G_{obs}$.

Precisely the same effect is observed with the $P_2$ NH and the $P_2$ Phe side chain as with the $P_1$ NH, as discussed Sections 4.1.1 and 4.1.2 and in Schemes 4.6, 4.7 and 4.8.

It is also in accord with the results of Asboth *et al.* (58), as compared with those of Williams *et al.* (59). Substitution of the $P_2$ C=O of Z-PheGly$\hat{}$OEt ($k_{cat}/K_M$ = 120 000 $\underline{M}^{-1}\cdot s^{-1}$ for papain) with C=S decreased specificity by 2.7 to 4.9 kcal/mol for several cysteine proteases, including 4.4 kcal/mol for papain (58). These $\Delta\Delta G_{obs}$'s represent the loss of favourable C=O hydrogen bonded energy plus the unfavourable energy of fitting the C=S into a site designed for the smaller C=O. For *N*-benzoylGly$\hat{}$OMe ($k_{cat}/K_M$ = 246 $\underline{M}^{-1}\cdot s^{-1}$), $\Delta\Delta G_{obs}$ = 1.6 kcal/mol on substituting the $P_2$ carbonyl oxygen with a sulfur atom (59). As with the substrates reported in this study, the effect of oxygen to sulfur substitution at $P_2$ C=O depends on other enzyme-substrate interactions. The better the substrates, the more unfavourable the resulting effect.

The 'signalling' and 'non-signalling' interactions of Brocklehurst *et al.* (57), when cast in the context of these results, become clear. The description of signalling, that is, the requirement for the $P_2$ C=O, plus one of the $P_2$ Phe or the $P_1$ NH to 'signal' the catalytic site and enhance catalysis is simply a case of the co-operativity of binding

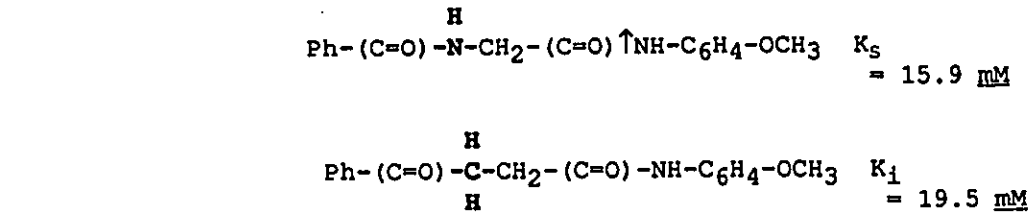
effects, with more of the incremental binding energy being expressed in binding as the specificity of the inhibitor increases.

*4.4.1.1 Conformational distortion*

Whilst there is no direct evidence from these kinetics results, other studies have shown that conformational distortions are likely to be important in papain's catalytic mechanism. Conformational distortions by their definition effect catalytic enhancement, whereas simply tighter binding, as reflected in $K_S$ is only binding in the ground state and does not necessarily lead to a lowering of the activation energy.

There are several lines of evidence for conformational distortion occurring in papain-substrate complexes in those substrates which possess the three important contacts about the $S_2$ binding subsite.

In a number of studies where $K_S$'s have been determined and where similar compounds are being compared, it has been found that large changes in $k_{cat}/K_M$ are often accompanied by only small changes in $K_S$ (or $K_i$). For example, Lowe and Yuthavong (32) compared the $K_S$ of N-benzoylGly-*p*-anisidide (acylation is rate determining, and $\therefore K_M = K_S$) with the $K_i$ of 3-benzoylpropionyl-*p*-anisidide (Scheme 4.20) which is not a




*N*-benzoylGlyC(O)-*p*-anisidide *vs.* 3-benzoylpropionylC(O)-*p*-anisidide

**Scheme 4.20**

substrate, but is a competitive inhibitor. They found that the dissociation constants for the two compounds were the same, meaning that changes in $k_{cat}/K_M$ may be attributed to increased $k_{+2}$. Since the substrate does not bind tighter in forming the Michaelis complex, the catalytic enhancement must arise from conformational distortion, with the $P_1$ NH providing the additional binding energy. Mattis *et al.* (26) found with the peptides mansyl-Gly-Gly-Val-Glu-Leu-Gly where the Leu residue was either the L- or D- enantiomer, that $\Delta\Delta G_{obs}$ = -2.7 kcal/mol (favouring L-Leu) whereas the difference in the free energy of the $K_S$'s was only -0.5 kcal/mol.

The binding of the competitive inhibitors α-*N*-benzoyl-D-ArgOEt or *N*-benzoyl β-aminoethanol increases papain's reactivity to the four alkylating reagents chloro- and iodo-acetamide and -acetate (94). This is explainable by either a more nucleophilic sulfur or a conformational change making the active site cysteine more accessible to the alkylating agents. In either case, binding of an inhibitor remote from the catalytic site (necessarily remote otherwise alkylation would be prevented, as it is by *N*-benzoylGlyOEt at [S]»$K_M$ (94)) causes change in the reactivity of papain.

In the resonance Raman studies of dithioacylpapain of Storer, Carey and co-workers, all thiono ester substrates containing Gly at $P_1$ have been

observed to form acylenzyme intermediates in which the Gly residue is in a relaxed conformation, referred to as the B-conformer (38). A shift of the ≈1090 $cm^{-1}$ peak to ≈1065 $cm^{-1}$ has been observed in substrates containing the $P_2$ NH, the $P_2$ Phe side chain as well as the $P_2$-$P_1$ amide bond. All other substrates not containing all three groups do not exhibit this shift. This peak is a complex vibrational mode and isotopic substitution with $^2H$ or $^{15}N$ has demonstrated a contribution from the $P_1$ NH (95,96). The shift in peak position is diagnostic of a change in the environment about the $P_1$ NH. Other spectral evidence shows that the $P_1$ Gly residue remains as a relaxed B-conformer. Thus, the kinetic effects of these three interactions also seem to cause subtle structural changes in the acylenzyme intermediate structure that are not seen with poorer substrates (A. Storer, P. Carey, P. Tonge, personal communication).

While the resonance Raman results indicate that there are changes in the dithioacylenzyme structure, more directly geometrical evidence is provided by the x-ray crystallographic structures of α-chloroketone (5) and E-64 (6) inhibited papains. The binding of these inhibitors is accompanied by a widening of the active site cleft by approximately 1 Å relative to the structure of native papain (3).

The movement of the walls of the active site cleft includes residues in the $S_2$ binding subsite. Of the residues involved in binding the substrate molecule at $S_2$, those from the right domain include Val 157, Asp 158, and Ala 160. These residues sequentially bracket His 159, the active site histidine residue. Movement of these residues *en bloc* away

from the left domain will necessarily also cause movement of the His 159 away from the active site Cys 25.

Thus, the kinetic evidence for conformational distortion in the papain-substrate complex, i.e. the enhanced $k_{+2}$ with small or no changes in $K_S$ is confirmed by the observation of structural changes spectroscopically and x-ray crystallographically. It is possible that conformational distortion effects may contribute to the observed interdependence of $\Delta\Delta G_{obs}$'s. For example, if by more tightly in the active site cleft a substrate more thoroughly excludes solvent, this will lead to a locally more hydrophobic environment and thus a strengthening of the enzyme-substrate hydrogen bonds. In the absence of any evidence that any putative conformational distortions will have this sort of effect, we propose that the main cause of the interdependence of $\Delta\Delta G_{obs}$'s is entropy compensation with the possibility of other, unspecified contributions.

*4.4.1.2 Summary*

The interdependence of interactions at the $P_2$ NH, the $P_2$ Phe side chain and the $P_2$-$P_1$ amide have been demonstrated. The reason for the stronger binding at each of these sites in the presence of the other two is that as the number of enzyme-substrate interactions increases, the entropic cost per interaction decreases. The incremental binding energies of the $P_2$ NH ($\Delta\Delta G_{obs}$ = -2.3 kcal/mol) and the $P_1$ NH ($\Delta\Delta G_{obs}$ = -2.2 kcal/mol) were estimated by subtracting a penalty of 0.4 kcal/mol for the unfavourable energy of placing an oxygen in those positions from the observed incremental specificity energies. The $\Delta\Delta G_{obs}$ = -4.0 kcal/mol

for the comparison of MocPheGly↑OMe *versus* MocGlyGly↑OMe (Scheme 4.8) represents an estimate of the incremental binding energy of the Phe side chain in the $S_2$ subsite.

There is evidence, on the basis of previous studies, that the binding energy supplied by the added interactions induces conformational changes in the enzyme that increase catalytic enhancement. Once the energetic cost of this distortion is paid for, additional binding energy will be manifested as tighter binding, that is, as a lower $K_S$. If, for example, a widening of the inter-domain cleft of papain causes movement of the loop Val 157 to Ala 160, causing a change in the catalytic residues that enhances catalysis, then once an optimum widening has been induced, additional binding energy will cause tighter papain-substrate binding.

### 4.4.2 Strengths of $P_1'$ C=O and $P_2'$ NH hydrogen bonds

The relative strengths of the two observed hydrogen bonds ($P_1'$ C=O, $\Delta\Delta G_{obs}$ = -1.0 kcal/mol; $P_2'$ NH, $\Delta\Delta G_{obs}$ = -2.3 kcal/mol) is probably attributable to the same effects as interactions on the acyl group side of the catalytic site. The progressive addition of favourable binding interactions in the substrate results in the entropic cost of forming the enzyme-substrate complex being spread over more interactions. This is likely to be particularly applicable in the case of the $P_1'$-$P_2'$ amide. Upon forming the hydrogen bond to the $P_1'$ C=O, the movement of the $P_2'$ NH becomes very restricted due to the rigid, planar geometry of the amide; making the entropic cost of forming the $P_2'$ NH hydrogen bond very small and therefore making $\Delta\Delta G_{obs}$ = -2.3 kcal/mol a good estimate of the incremental binding energy. The incremental binding energy of

the $P_1'$ C=O is probably considerably more favourable than the $\Delta\Delta G_{obs}$ = -1.0 kcal/mol observed here.

## 5. CONCLUSIONS

We have shown that the active site cleft of papain has hydrogen bonding partners for the peptide backbone of substrates extending at least from the $S_2$ to the $S_2'$ subsites. The incremental binding energies of these hydrogen bonds and of the $P_2$ Phe side chain have been estimated (Table 5.1). The observed interdependence of binding energies at each site on the acyl side of the catalytic site is characteristic of enzyme-substrate interactions and is largely due to a decreased entropic penalty with successive additions of favourable interactions.

Application of the knowledge of sites of potential favourable interactions will permit the more intelligent design of cysteine protease inhibitors.

**Table 5.1** Summary of results.

| Site | $\Delta\Delta G_{bind}$ (kcal/mol) | $\Delta\Delta G_{obs}$ (kcal/mol) |
|---|---|---|
| $P_2$ NH | -2.3 | |
| $P_2$ Phe side chain | -4.0 | |
| $P_1$ NH | -2.2 | |
| $P_1$' C=O | | -1.0 |
| $P_2$' NH | -2.3 | |

## 6. REFERENCES


1. Brocklehurst, K., Willenbrock, F. and Salih, E. (1987) in *Hydrolytic enzymes* (Neuberger, A. and Brocklehurst, K.eds) pp. 39-158, Elsevier Biomedical Press, Amsterdam

2. Drenth, J., Jansonius, J.N., Koekoek, R., Swen, H.M. and Wolthers, B.G. (1968) Nature 218, 929-932

3. Kamphuis, I.G., Kalk, K.H., Swarte, M.B.A. and Drenth, J. (1984) *J. Mol. Biol.* **179**, 233-256

4. Priestle, J.P., Ford, G.C., Glor, M., Mehler, E.L., Smit, J.D.G., Thaller, C. and Jansonius, J.N. (1984) *Acta Crystallogr. Sect. A.* **40**, 17

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

6. Varughese, K.I., Ahmed, F.R., Carey, P.R., Hasnain, S., Huber, C.P. and Storer, A.C. (1989) *Biochemistry* **28**, 1330-1332

7. Baker, E.N. (1977) *J. Mol. Biol.* **115**, 263-277

8. Baker, E.N. (1980) *J. Mol. Biol.* **141**, 441-484

9. Heinemann, U., Pal, G.P., Hilgenfeld, R. and Saenger, W. (1982) *J. Mol. Biol.* **161**, 591-606

10. Hilgenfeld, R. (1987) *Hochauflösende Proteinkristallographie: Aminosauresequenz und dreidimensionale Struktur der Cysteinproteinase Calotropin D1*, Ph. D. thesis, Freien Universitat Berlin,

11. Kamphuis, I.G., Drenth, J. and Baker, E.N. (1985) *J. Mol. Biol.* **182**, 317-329

12. Richardson, J.S. (1981) *Adv. Protein Chem.* **34**, 167-339

13. Murzin, A. and Finklestein, A.V. (1988) *J. Mol. Biol.* **204**, 749-769

14. Barbas, C.F., III and Wong, C.-H. (1988) *Tetrahedron Lett.* **29**, 2907-2910

15. Lee, H., Takahashi, K., Kodera, Y., Ohwada, K., Tsuzuki, T., Matsushima, A. and Inada, Y. (1988) *Biotech. Lett.* **10**, 403-407

16. Cantacuzene, D., Pascal, F. and Guerreiro, C. (1987) *Tetrahedron* **43**, 1823-1823

17. Lowe, G. (1976) *Tetrahedron* **32**, 291-302

18. Polgar, L. and Halasz, P. (1982) *Biochem. J.* **207**, 1-10

19. Brocklehurst, K. (1987) in *Enzyme mechanisms* (Page, M.I. and Williams, A.eds) pp. 140-158, The Royal Society of Chemistry, London

20. Polgar, L. (1974) *FEBS Lett.* **47**, 15-18

21. Lewis, S.D., Johnson, F.A. and Shafer, J.A. (1976) *Biochemistry* **15**, 5009-5017

22. Lewis, S.D., Johnson, F.A. and Shafer, J.A. (1981) *Biochemistry* **20**, 48-51

23. Johnson, F.A., Lewis, S.D. and Shafer, J.A. (1981) *Biochemistry* **20**, 44-48

24. Dixon, H.B.F. (1976) *Biochem. J.* **153**, 627-629

25. Sluyterman, L.A.Æ. and Wijdenes, J. (1976) *Eur. J. Biochem.* **71**, 383-391

26. Mattis, J.A., Henes, J.B. and Fruton, J.S. (1977) *J. Biol. Chem. 252, 6776-6782*

27. Lavery, R., Pullman, A. and Wen, Y.K. (1983) *Int. J. Quantum Chem.* **24**, 353-371

28. Rullman, J.A.C., Bellido, M.N. and van Duijnen, P.Th. (1989) *J. Mol. Biol.* **206**, 101-118

29. Thole, B.T. and van Duijnen, P.Th. (1983) *Biophys. Chem.* **18**, 53-59

30. Jencks, W.P. (1969) *Catalysis in chemistry and enzymology*, McGraw-Hill, New York

31. Fastrez, J. (1983) *Eur. J. Biochem.* **135**, 339-341

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

33. Lowe, G. and Yuthavong, Y. (1971) *Biochem. J.* **124**, 117-122

34. O'Leary, M.H., Urberg, M. and Young, A.P. (1974) *Biochemistry* **13**, 2077-2081

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

36. Asboth, B. and Polgar, L. (1983) *Biochemistry* **22**, 117-122

37. Storer, A.C. and Carey, P.R. (1985) *Biochemistry* **24**, 6808-6818

38. Storer, A.C., Angus, R.H. and Carey, P.R. (1988) *Biochemistry* **37**, 264-268

39. Brubacher, L.J. and Bender, M.L. (1966) *J. Amer. Chem. Soc.* **88**, 5871-5879

40. Lowe, G. and Williams, A. (1965) *Biochem. J.* **96**, 189-193

41. Ozaki, Y., Pliura, D.H., Carey, P.R. and Storer, A.C. (1982) *Biochemistry* **21**, 3102-3108

42. Malthouse, J.P.G., Gamcsik, M.P., Boyd, A.S.F., Mackenzie, N.E. and Scott, A.I. (1982) *J. Amer. Chem. Soc.* **104**, 6811

43. Desmazeaud, M.J. (1972) *Biochimie* **54**, 1109-1114

44. Berger, A. and Schechter, I. (1970) *Phil. Trans. R. Soc. Lond. B* **257**, 249-264

45. Schechter, I. and Berger, A. (1968) *Biochem. Biophys. Res. Commun.* **32**, 898-902

46. Schechter, I. and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27**, 157-162

47. Alecio, M.R., Dann, M.L. and Lowe, G. (1974) *Biochem. J.* **141**, 495-501

48. Polgar, L. (1987) in *Hydrolytic enzymes* (Neuberger, A. and Brocklehurst, K.eds) pp. 159-200, Elsevier, Amsterdam

49. Brocklehurst, K., Kowlessur, D., O'Driscoll, M., Patel, G., Quenby, S., Salih, E., Templeton, W., Thomas, E.W. and Willenbrock, F. (1987) *Biochem. J.* **244**, 173-181

50. Brocklehurst, K., Brocklehurst, S.M., Kowlessur, D., O'Driscoll, M., Patel, G., Salih, E., Templeton, W., Thomas, E.W., Topham, C.M. and Willenbrock, F. (1988) *Biochem. J.* **256**, 543-558

51. Willenbrock, F. and Brocklehurst, K. (1985) *Biochem. J.* **227**, 521-528

52. Willaims, A. (1987) in *Enzyme mechanisms* (Page, M.I. and Williams, A.eds) pp. 123-139, The Royal Society of Chemistry, London

53. Chen, M.M.L. and Hoffmann, R. (1976) *J. Amer. Chem. Soc.* **98**, 1647-1653

54. Kice, J.L. (1980) *Adv. Phys. Org. Chem.* **17**, 65-181

55. Kowlessur, D., O'Driscoll, M., Topham, C.M., Templeton, W., Thomas, E.W. and Brocklehurst, K. (1989) *Biochem. J.* **259**, 443-452

56. Jencks, W.P. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4046-4050

57. Kowlessur, D., Topham, C.M., Thomas, E.W., O'Driscoll, M., Templeton, W. and Brocklehurst, K. (1989) *Biochem. J.* **258**, 755-764

58. Asboth, B., Majer, Z. and Polgar, L. (1988) *FEBS Lett.* **233 (2)**, 339-341

59. Williams, A., Lucas, E.C., Rimmer, A.R. and Hawkins, H.C. (1972) *J. Chem. Soc. Perkin Trans. I* 627-633

60. Thornton, E.K. and Thornton, E.R. (1978) in *Transition states in biochemical processes* (Gandour, R.D. and Schowen, R.L.eds) pp. 3-76, Plenum Press, New York

61. Schowen, R.L. (1978) in *Transition states of biochemical processes* (Gandour, R.D. and Schowen, R.L.eds) pp. 77-114, Plenum Press, New York

62. Jencks, W.P. (1975) *Adv. Enzymol.* **43**, 219-410

63. Kraut, J. (1988) *Science* **242**, 533-539

64. Illuminati, G. and Mandolini, L. (1981) *Acc. Chem. Res.* **14**, 95-102

65. Page, M.I. and Jencks, W.P. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 1678-1683

66. Page, M.I. (1984) in *The chemistry of enzyme action* (Page, M.I.ed) pp. 1-54, Elsevier Science Publishers B.V., Amsterdam

67. Fersht, A.R. (1985) *Enzyme structure and mechanism*, 2nd Ed. W.H. Freeman and Company, New York

68. Fersht, A.R. (1988) *Biochemistry* **27**, 1577-1580

69. Fersht, A.R. (1987) *Trends Biochem. Sci.* **12**, 301-304

70. Rich, D.H. (1986) in *Proteinase inhibitors* (Barrett, A.J. and Salvesen, G.eds) pp. 153-178, Elsevier, Amsterdam

71. Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brain, J., Kos, J. and Turk, V. (1988) *EMBO J.* **7**, 2593-2599

72. Barrett, A.J., Rawlings, N.D., Davies, M.E., Machleidt, W., Salvesen, G. and Turk, V. (1986) in *Proteinase Inhbitors* (Barrett, A.J. and Salvesen, G.eds) pp. 5150-569, Elsevier Science Publishers BV, Amsterdam

73. Bjorck, L., Akesson, P., Bohus, M., Trojnar, J., Abrahamson, M., Olafsson, I. and Grubb, A. (1989) *Nature* **337**, 385-386

74. Sluyterman, L.A.Æ. and Wijdenes, J. (1970) *Biochim. Biophys. Acta* **200**, 593-595

75. Carey, P.R., Angus, R.H., Lee, H. and Storer, A.C. (1984) *J. Biol. Chem.* **259**, 14357-14360

76. Nefkens, G.H.L. and Nivard, R.J.F. (1965) *Rec. Trav. Chim. Pays Bas* **84**, 1315-1318

77. Cohen, S.G. and Weinstein, S.Y. (1964) *J. Amer. Chem. Soc.* **86**, 5326-5330

78. Ingles, D.W. and Knowles, J.R. (1968) *Biochem. J.* **108**, 561-569

79. Brocklehurst, K., Carlsson, J., Kierstan, M.P.J. and Crook, E.M. (1973) *Biochem. J.* **133**, 573-584

80. Blumberg, S., Schecter, I. and Berger, A. (1970) *Eur. J. Biochem.* **15**, 97-102

81. Cornish-Bowden, A. (1975) *Biochem. J.* **149**, 305-312

82. Cornish-Bowden, A. (1976) *Principles of enzyme kinetics*, Butterworths, London

83. Johansen, G. and Lumry, R. (1961) *Compt. Rend. Trav. Lab. Carlsberg* **32**, 185-214

84. Whitaker, J.R. and Bender, M.L. (1965) *J. Amer. Chem. Soc.* **87**, 2728-2737

85. Huang, C.Y. (1979) *Methods Enzymol.* **63**, 54-84

86. Johnson, E.L. and Stevenson, R. (1978) *Basic liquid chromatography*, Varian Associates, Inc., Palo Alto

87. Selwyn, M.J. (1965) *Biochim. Biophys. Acta* **105**, 193-195

88. Stone, A.J. (1981) *Chem. Phys. Lett.* **83**, 233-239

89. Brant, D.A., Miller, W.G. and Flory, P.J. (1967) *J. Mol. Biol.* **23**, 47-65

90. Lowe, G. and Williams, A. (1965) *Biochem. J.* **96**, 199-204

91. Baker, E.N. and Hubbard, R.E. (1984) *Prog. Biophys. Molec. Biol.* **44**, 97-179

92. Wilkinson, A.J., Fersht, A.R., Blow, D.M. and Winter, G. (1983) *Biochemistry* **22**, 3581-3586

93. Asboth, B. and Polgar, L. (1977) *Acta Biochim. Biophys. Acad. Sci Hung.* **12**, 223-230

94. Whitaker, J.R. (1969) *Biochemistry* **8**, 4591-4597

95. Lee, H., Storer, A.C. and Carey, P.R. (1983) *Biochemistry* **22**, 4781-4789

96. Storer, A.C., Lee, H. and Carey, P.R. (1983) *Biochemistry* **22**, 4789-4796

97. IUPAC-IUB Joint Commission on Biochemical Nomenclature (1985) *J. Biol. Chem.* **260**, 14-42

98. Atkins, P.W. (1986) *Physical chemistry*, 3rd Ed. W.H. Freeman and Co., New York