THE STRUCTURAL SPECIFICITY OF α-CHYMOTRYPSIN

I. POLYPEPTIDES AS SUBSTRATES

II. N-ACYLATED PEPTIDE ESTERS AS SUBSTRATES

III. SOME REACTIVE ESTERS OF N-ACYLATED AMINO ACIDS AS SUBSTRATES

Thesis by

Gary L. Neil

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology
Pasadena, California

1966

(Submitted March 17, 1966)

Dedicated in memory of

my father

ACKNOWLEDGMENTS

I wish to express my indebtedness and deepest gratitude to Professor G. S. Hammond and to the late Professor Carl Niemann. Without their guidance, and above all, their encouragement, this work could not have been completed.

Dr. Carole Hamilton, Professor George Hein and Mr. Leonard Piszkiewicz offered many suggestions and criticisms concerning the content and organization of this thesis. I thank them all for their assistance.

I am grateful for the financial assistance given me by the California Institute of Technology in the form of scholarships and Teaching and Research Assistantships, and to the Woodrow Wilson Foundation for a 1965 Summer Fellowship.

There were many staff members, postdoctoral fellows and graduate students whose friendship and advice helped make my stay at Cal Tech both pleasant and stimulating. I thank them all most gratefully.

Throughout my entire life my parents have been a constant source of encouragement and support. I cannot hope to ever thank them enough.

Above all, I thank my wife and daughter for their constant support, encouragement and patience.

ABSTRACT

The structural specificity of α-chymotrypsin for polypeptides and denatured proteins has been examined. The primary specificity of the enzyme for these natural substrates is shown to closely correspond to that observed for model substrates. A pattern of secondary specificity is proposed.

A series of N-acetylated peptide methyl esters of varying length have been evaluated as substrates of α-chymotrypsin. The results are interpreted in terms of proposed specificity theories.

The α-chymotrypsin-catalyzed hydrolyses of a number of N-acetylated dipeptide methyl esters were studied. The results are interpreted in terms of the available specificity theories and are compared with results obtained in the study of polypeptide substrates. The importance of non-productive binding in determining the kinetic parameters of these substrates is discussed. A partial model of the locus of the active site which interacts with the $R_1'CONH-$ group of a substrate of the form $R_1'CONHCHR_2COR_3'$ is proposed.

Finally, some reactive esters of N-acylated amino acids have been evaluated as substrates of α-chymotrypsin. Their reactivity and stereochemical behavior are discussed in terms of the specificity theories available. The importance of a binding interaction between the carboxyl function of the substrate and the enzyme is suggested by the results obtained.

TABLE OF CONTENTS

Page

GENERAL INTRODUCTION 2

α-Chymotrypsin and its Substrates 2
Reaction Kinetics and Kinetic Schemes 3
Significance and Interpretation of Kinetic Parameters 5
Structural Specificity of α-Chymotrypsin 9
Present Study 16
References 18

PART I

THE STRUCTURAL SPECIFICITY OF α-CHYMOTRYPSIN: POLYPEPTIDES AS SUBSTRATES

INTRODUCTION 21

Historical Background 21
Classification of Enzyme Specificity 23

RESULTS AND DISCUSSION 26

Data 26
Primary Specificity 29
Secondary Specificity 34

REFERENCES 47

PART II

THE STRUCTURAL SPECIFICITY OF α-CHYMOTRYPSIN: N-ACYLATED PEPTIDE ESTERS AS SUBSTRATES

INTRODUCTION 51

RESULTS AND DISCUSSION 53

SECTION A: N-Acetylated Peptide Esters Containing Glycine Residues 53

Data 53
Effects of Structural Changes on the Kinetic Parameters . . 57

SECTION B: N-Acetyl-aminoacyl-L-leucine Methyl Esters 89

Data 89
Non-productive Binding 94

Page

Conclusions Regarding the Nature of Interactions at the ρ_2 and ρ_1 Loci—a Model of the ρ_1 Locus 112
Interpretation of Data in Terms of the Acyl-enzyme Kinetic Scheme 116
Comparison of k_o/K_o Values with Other Data 118
Secondary Specificity—Correlation of Experimental Results with those of the Polypeptide Study 125

EXPERIMENTAL 128

Syntheses 128
Kinetic Studies 148
Experimental Details 150

REFERENCES 154

PART III

THE STRUCTURAL SPECIFICITY OF α-CHYMOTRYPSIN: SOME REACTIVE ESTERS OF N-ACYLATED AMINO ACIDS AS SUBSTRATES

INTRODUCTION 158

RESULTS AND DISCUSSION 160

Importance of R_3-ρ_3 Binding 160
Stereospecificity 170
Effects of the Alcohol Component (R_3') of N-Acylated Amino Acid Esters on Reactivity and Stereospecificity. . . . 174
Comparison of Some Kinetic Parameters 189
Secondary Specificity 190

EXPERIMENTAL 191

Syntheses 191
Kinetic Studies 195

REFERENCES 196

PROPOSITIONS 199

STRUCTURAL SPECIFICITY OF α-CHYMOTRYPSIN:

A GENERAL INTRODUCTION

GENERAL INTRODUCTION

Enzymes may be defined as protein molecules[†] which are capable of catalyzing, or increasing the velocity of, certain chemical reactions (1,2). They are especially noted for a great degree of selectivity (specificity) in both a structural and stereochemical sense (1,3). Much attention has been focussed on this characteristic property, investigation of which has aided in the description of the chemical nature and mechanism of action of enzymes.

α-Chymotrypsin and its Substrates

Bovine pancreatic α-chymotrypsin is a simple protein composed of some 242 α-amino acid residues. It requires no coenzyme or activator. It has been shown conclusively (4,5,6,7) that the enzyme contains one active site per molecule.[‡] The enzyme is an endopeptidase, catalyzing the hydrolysis of peptide bonds inside the peptide chains of complex proteins and polypeptides. Although a primary amino acid sequence for the precursor, chymotrypsinogen, has been determined (8) and the sequence of reactions which produce the active enzyme are known, the three-dimensional (tertiary) structure has yet to be established (19).

[†] Some enzymes are simple proteins containing only covalently bonded amino acid residues; others are conjugated proteins containing, in addition, other components which are non-covalently bonded to the protein.

[‡] In a manner similar to that described by Niemann (1), the active site will be defined simply as that portion of an enzyme which is important in the overall catalysis process. The active site contains an active center, a limited region of the active site at which the actual bond-breaking or formation takes place.

α-Chymotrypsin also catalyzes the hydrolysis of a large number of synthetic model substrates of widely varying structure. Although it catalyzes the hydrolysis of substrates such as p-nitrophenyl acetate (10), dialkyl phosphofluoridates (11), derivatives of N,N'-diaryl carbamic acids (12) and esters and amides of trans-cinnamic acid (13), derivatives of acylated α-amino acids have been most extensively studied. This class of compounds serves as a reasonable model for the natural protein substrates.

Reaction Kinetics and Kinetic Schemes

The rates of α-chymotrypsin-catalyzed hydrolyses of model substrates containing a single hydrolyzable bond are in general described by the following equation:

$$-\frac{d[S]}{dt} = \frac{d[P]}{dt} = \frac{k_o[E]_o[S]}{K_o + [S]} \qquad (1)$$

where [S], [P] and $[E]_o$ represent the concentrations of substrate, product and total enzyme, respectively. The terms k_o and K_o are experimentally determined parameters characteristic for each substrate and are functions of the experimental conditions (temperature, pH, solution composition, etc.).

Equation 1 can be represented by a number of kinetic paths, three of which appear to be most pertinent.

The simplest kinetic scheme which fits equation 1 is given by:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P \qquad (2)$$

The rate equation derived from a steady-state treatment of [ES] is identical in form to equation 1, with $k_o = k_2$ and $K_o = K_s = (k_{-1} + k_2)/k_1$. ES represents an enzyme-substrate complex, formation of which has been well documented (14). In this scheme, K_o is most simply interpreted as a measure of the ability of a substrate to form an ES complex with the enzyme, and k_o as a measure of the rate of breakdown of this complex to product and free enzyme.

The second representation is given in the following equation:

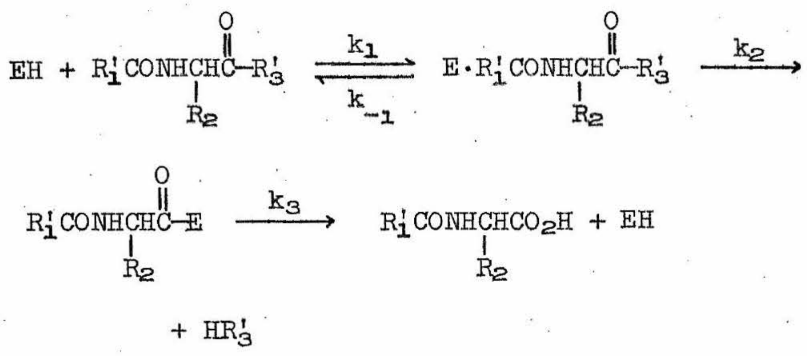
$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES_1 \xrightarrow{k_2} E + P \qquad (3)$$

$$E + S \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} ES_2 \qquad (4)$$

where $K_s = (k_{-1} + k_2)/k_1$ and $K_{sI} = k_{-4}/k_4$. Here it is assumed that the substrate may combine with the enzyme in two ways, one complex (ES_1) capable of proceeding to produce product and the other (ES_2) totally competitive with the first but non-productive. Analysis of this scheme leads to a rate equation of the form of equation 1 with $k_o = k_2K_{sI}/(K_s + K_{sI})$ and $K_o = K_sK_{sI}/(K_s + K_{sI})$. This scheme may be extended to the more general case with more binding modes, both productive and non-productive (15). Further interpretation of this scheme will be found later in this thesis.

The third representation is given by equations 5 and 6,

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} \underset{+ P_1}{ES'} \xrightarrow{k_3} E + P_2 \qquad (5)$$



$$\mathrm{EH} + \mathrm{R_1'CONH\underset{\underset{R_2}{|}}{C}H\overset{\overset{O}{\|}}{C}{-}R_3'} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{E{\cdot}R_1'CONH\underset{\underset{R_2}{|}}{C}H\overset{\overset{O}{\|}}{C}{-}R_3'} \xrightarrow{k_2} \mathrm{R_1'CONH\underset{\underset{R_2}{|}}{C}H\overset{\overset{O}{\|}}{C}{-}E} + \mathrm{HR_3'} \xrightarrow{k_3} \mathrm{R_1'CONH\underset{\underset{R_2}{|}}{C}HCO_2H} + \mathrm{EH} \qquad (6)$$

where equation 6 is a particular case of equation 5, in which ES' is the proposed acyl-enzyme intermediate (16,17) and k_2 and k_3 represent the specific rates of acylation and deacylation, respectively. The rate of the overall reaction is again given by equation 1, with $k_o = k_2k_3/(k_2 + k_3)$ and $K_o = K_sk_3/(k_2 + k_3)$ where $K_s = (k_{-1} + k_2)/k_1$. An extension of this scheme takes into account multiple modes of binding (17). Such an extension will be discussed more fully in a later part of this thesis.

Significance and Interpretation of Kinetic Parameters

The experimentally determined kinetic parameters, k_o and K_o, may bear different interpretations depending on the kinetic scheme chosen to represent the course of the reaction. These interpretations will be discussed herein in reference to the three kinetic schemes outlined above.

a) Interpretation of k_o

Consistent with the scheme indicated in equation 2, $k_o = k_2$ and is thus a direct measure of the rate of breakdown of the enzyme-substrate complex to products.

According to the multiple-binding scheme represented in equations 3 and 4, $k_o = k_2K_{sI}/(K_s + K_{sI})$. Therefore the effect of a non-productive binding mode on k_o will be to decrease k_o whenever the non-productive mode competes successfully with one which leads to products.

According to the scheme shown in equations 5 and 6, $k_o = k_2k_3/(k_2 + k_3)$ where, in the case of the acyl-enzyme intermediate mechanism, k_2 and k_3 represent the specific rates of acylation and deacylation, respectively. Thus, it is obvious that k_o can be perturbed by a change in the ratio k_3/k_2 (e.g., a change in the rate-determining step). There are two limiting cases which should be noted:

a) When acylation is the rate-determining step (i.e., $k_2 \ll k_3$) then $k_o = k_2$.

b) When deacylation is the rate-determining step (i.e., $k_2 \gg k_3$) then $k_o = k_3$.

b) Interpretation of K_o

In a manner similar to that shown for k_o, K_o can have different interpretations, depending on the choice of kinetic scheme. According to the scheme shown in equation 2, $K_o = K_s = (k_{-1} + k_2)/k_1$. It has been shown for α-chymotrypsin (18,19,20) that usually $k_2 \ll k_{-1}$, in which case $K_o = K_s \approx k_{-1}/k_1$, a simple equilibrium constant representing the dissociation of the enzyme-substrate complex to free enzyme and substrate.

As represented by equations 3 and 4, $K_o = K_s K_{sI}/(K_s + K_{sI})$. The observed K_o value will thus approximate the lower of the K_s or K_{sI} values if K_s and K_{sI} are significantly different from one another.

According to the acyl-enzyme scheme (equations 5 and 6), $K_o = K_s k_3/(k_2 + k_3)$. Again, there are two limiting cases:

a) When acylation is rate-determining (i.e., $k_2 \ll k_3$) then $K_o = K_s$.

b) When deacylation is rate-determining (i.e., $k_2 \gg k_3$) then $K_o = k_3 K_s/k_2$.

Thus it is obvious that for the acyl-enzyme mechanism, when acylation is rate-determining, this scheme is kinetically indistinguishable from the simple scheme represented by equation 2, and K_o will have the same interpretation. However, if deacylation is rate-determining, the value of K_o observed will also reflect the perturbation of k_3/k_2 on K_s.

c) Interpretation of k_o/K_o

The ratio k_o/K_o for any substrate is a composite measure of the stability of the productive complex (or complexes) and the ability of this complex to decompose to products. It is the rate constant under second order conditions (i.e., the specific rate where saturation of the enzyme is unimportant). This situation occurs when $[S] \ll K_o$. In this case equation 1 reduces to:

$$-\frac{d[S]}{dt} = \frac{d[P]}{dt} = \left(\frac{k_o}{K_o}\right)[E]_o[S] \qquad (7)$$

Further, k_o/K_o has one distinct advantage (for interpretative purposes) over the separate parameters k_o and K_o, because $k_o/K_o = k_2/K_s$ for all of the kinetic schemes described earlier. It is thus free from the ambiguities present in the interpretation of the parameters k_o and K_o.

In terms of a free energy profile, k_o/K_o is really a measure of the free energy difference between the free starting materials (enzyme and substrate) in solution, and the transition state for the reaction step immediately following formation of the productive enzyme-substrate complex (ES). In terms of the simple scheme (equation 2), this step represents formation of free product and enzyme (E+P). In terms of the acyl-enzyme scheme (equation 6), it represents formation of the acyl-enzyme (ES'). This is illustrated in Figure 1.

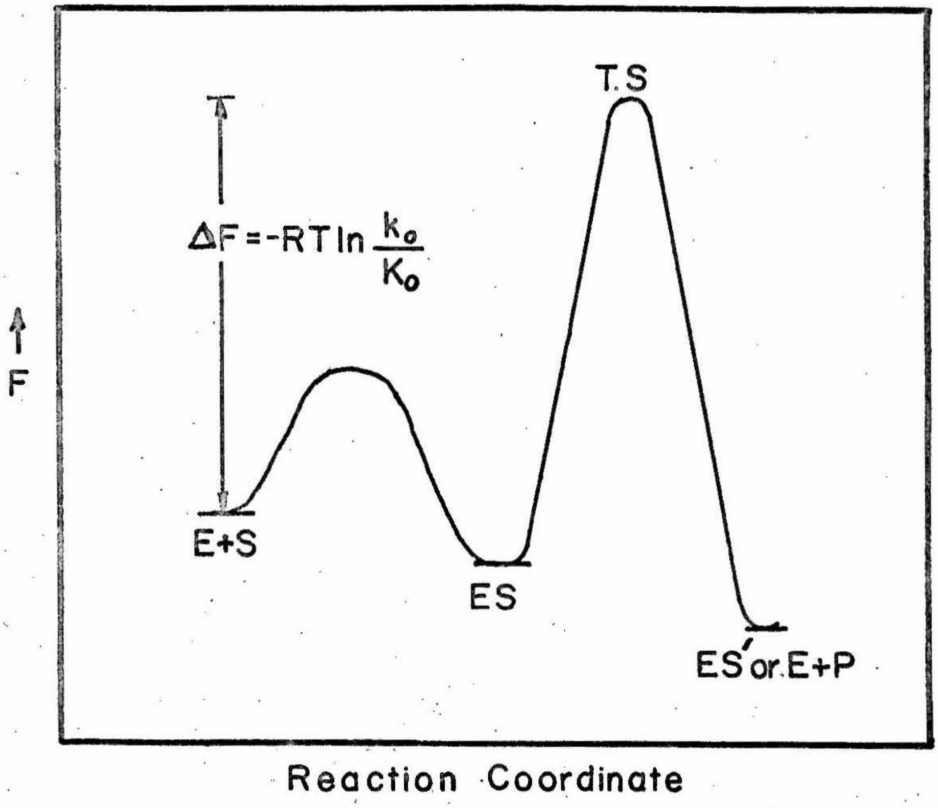


Figure 1

The kinetic data of the studies described in this thesis will be discussed, insofar as it is possible to do so, in terms of the kinetic schemes presented above.

Structural Specificity of α-Chymotrypsin

The high specificity of enzymes, that is, the strict limitation of action of a particular enzyme to a small number of closely related substances, is an important and striking characteristic. This biological phenomenon is responsible for the ordered metabolism of living material and is essential for the existence of life.

The proteinases, and α-chymotrypsin in particular, are characterized by a rather broad specificity (1,21,22). This is in contrast to some enzymes which are highly specific for only one substrate. It is true that α-chymotrypsin shows a marked preference for linkages involving aromatic amino acid residues (tryptophan, tyrosine and phenylalanine) but it is also quite active towards leucyl and methionyl bonds (see Part I of this thesis). This is true for a series of comparable acylated α-amino acid derivatives (e.g., N-acetylated α-amino acid methyl esters or amides (1,22)) as well as for polypeptides.

In addition to this side chain specificity, α-chymotrypsin shows some specificity for the N-acyl group (15,22,23,24) and considerable specificity for the carboxyl function of model substrates. These specificities are again rather broad. Stereochemical specificity has also been shown to be a relative phenomenon with model substrates (15,22).

The concept of specificity and the relation of the specificity observed in model systems to that observed with the natural substrates will be an important subject for discussion in this discourse.

a) The Hein-Niemann Theory

Results of kinetic studies on a number of asymmetric trifunctional model substrates led to Huang and Niemann's suggestion (25) that such substrates of α-chymotrypsin might combine with the active site of the enzyme in several modes, not all of which would lead to hydrolysis. A theory of structural specificity proposed by Hein and Niemann (15,22) is based on this proposition. It has been remarkably successful in explaining and correlating the kinetic data for the hydrolysis of a large number of trifunctional model substrates of the form $R_1^{\prime}CONHCHR_2COR_3^{\prime}$ (where $R_1^{\prime}CONH=R_1$ and $R_2 \neq H$). The theory has been extended to bifunctional substrates (24) with some success and its predictive value has been confirmed (26,27,28).

For a trifunctional substrate, the definition of a structure-reactivity relationship requires evaluation of the relative contributions made to both binding and orientation of the substrate at the active site by interaction of the structural components R_1, R_2, and $COR_3^{\prime}$ with this site. A model of the active site, introduced in the early stages of the study of acyl-α-amino acid derivatives (25,29), was incorporated in the theory. It was thus proposed that the active site consisted of four loci, ρ_1, ρ_2, ρ_3 and ρ_H, which are complementary to the four groups disposed about the asymmetric center of a trifunctional substrate. The locus ρ_3 is defined as having the dual purpose of binding the hydrolyzable group ($COR_3^{\prime}$) and catalyzing the hydrolytic cleavage. The locus ρ_H is postulated (primarily for the sake of completeness) as the space occupied by the α-hydrogen when $R_1-\rho_1$, $R_2-\rho_2$ and $R_3-\rho_3$ interactions occur.

The most important hypothesis of the theory is that it is not sufficient to consider only the interactions of R_1 with ρ_1, R_2 with ρ_2, and R_3 with ρ_3. Rather, in general, all possible $R_i-\rho_j$ interactions must be taken into account. Thus, several modes of binding, fully competitive with each other, are possible. Some combinations (productive complexes) proceed to products, while others (non-productive complexes) are incapable of doing so. A kinetic derivation of this scheme produces a rate equation of the familiar Michaelis-Menten form (equation 1). This has been illustrated earlier (p. 4) for the simple case of two binding modes, one productive and the other non-productive, but holds true also in the general case.

From a limited number of postulates (15), some of which are empirical in nature and others which are necessary for a successful application of the theory, it is then possible to deduce which of the possible modes are most favored. Thus, one can predict which modes and interactions will determine each of the measured parameters, k_0 and K_0.

b) The Hamilton Specificity Theory

Although the Hein-Niemann theory has been a useful tool for correlation of kinetic data, this utility has been essentially qualitative in nature. Recently, Hamilton, Niemann and Hammond (30) have presented a quantitative extension of the Hein-Niemann theory which has shown great success in correlating the observed kinetic parameters for a particular class of model substrates of α-chymotrypsin, the N-acyl-D- and L-amino acid amides.

This theory considers the overall binding of a substrate (or inhibitor) in terms of independent contributions of the various group-locus interactions. If each R-ρ interaction for each of the twelve enzyme-substrate complexes possible in the framework of the Hein-Niemann theory makes an independent contribution to the free energy of a particular complex formation, then for a particular complex, the free energy of formation should be given by the sum of the free energies associated with the four contributing interactions.[†] For example, in the case of the predominant productive complex for an L-antipode (i.e., with $R_1-\rho_1$, $R_2-\rho_2$, $R_3-\rho_3$ and $H-\rho_H$ interactions), the free energy of formation would be given by:

$$\Delta F_p = \Delta F_{11} + \Delta F_{22} + \Delta F_{33} + \Delta F_{H-H} \tag{8}$$

Equation 8 may be expressed in terms of microscopic "binding factors":

$$\overline{K}_p = \overline{K}_{11}\, \overline{K}_{22}\, \overline{K}_{33}\, \overline{K}_{HH} \tag{9}$$

where $\overline{K}_p$ is the association constant for the productive complex and the $\overline{K}_{ij}$'s are microscopic "binding factors" analogous to equilibrium constants.

Values of $\overline{K}_{ij}$ were assigned to the interactions of the seven groups in six substrate-inhibitor pairs (i.e., N-acetyl and N-nicotinyl amides of D- and L-phenylalanine, -tyrosine and -tryptophan) such that net association constants were generated which corresponded closely to the experimental $\overline{K}_o$ and $\overline{K}_I$ values. This set of $\overline{K}_{ij}$'s was then used to generate further calculated net association constants which were in excellent agreement with experimental values.

[†] It must also be assumed that the contributions of the groups to free energy of solution are also independent.

Interpretation of the parameter k_o/K_o for an L-substrate (for the multiple-binding extensions of both the simple scheme (equations 3 and 4) and the acyl-enzyme scheme (equation 6)) in terms of this theory is as follows:

$$\frac{k_o}{K_o} = k_{2p}\,\overline{K}_p = k_{2p}\,\overline{K}_{11}\,\overline{K}_{22}\,\overline{K}_{33}\,\overline{K}_{HH}$$

where k_{2p} is the rate constant for breakdown of the enzyme-substrate complex to products (simple case—equation 2) or for acylation of the enzyme (equation 6) and $\overline{K}_p$ is the association constant for the productive complex.

A test of this theory and its inherent assumptions is given in Part III of this thesis.

c) The Bender-Kézdy Specificity Theory

In a recent review on the mechanism of action of proteolytic enzymes, Bender and Kézdy (17) have presented a new semi-quantitative theory of relative specificity which applies to model substrates of α-chymotrypsin. The theory primarily employs the combination parameter k_o/K_o as its basis of comparison. According to this theory, the parameter k_o/K_o for any N-acyl-L-amino acid derivative is given with respect to that for some reference substrate by the following equation:

$$\log \frac{(k_o/K_o)_{R_1R_2X}}{(k_o/K_o)_{R_{1o}R_{2o}X_o}} = \rho^* \sigma^* + S_{R_1} + S_{R_2} \qquad (10)$$

in terms of three independent factors, one nonenzymatic reactivity term ($\rho^* \sigma^*$) and two specificity terms (S_{R_1} and S_{R_2}). These terms are defined in the following manner:

$$\rho^* \sigma^* = \log \frac{(k_o/K_o)_{R_1R_2X}}{(k_o/K_o)_{R_1R_2X_o}} \qquad (11)$$

where σ^* is the aliphatic Taft substituent constant and ρ^* is an empirical parameter dependent on the nature of the reaction and on the reaction conditions (31).

$$S_{R_1} = \log \frac{(k_o/K_o)_{R_1R_2X}}{(k_o/K_o)_{R_{1o}R_2X}} \qquad (12)$$

and

$$S_{R_2} = \log \frac{(k_o/K_o)_{R_1R_2X}}{(k_o/K_o)_{R_1R_{2o}X}} \qquad (13)$$

where R_1, R_2 and X refer to the acylamino function, the side chain and the carboxyl function, respectively, and the sub-zeros refer to groups on the reference compound.

The main difference between the Bender-Kézdy theory and that of Hein and Niemann is that in the former, the X group (R_3 or COR_3' in Hein-Niemann nomenclature) is not considered to be involved in productive binding and thus the effect on k_o/K_o of a variation of the group X solely reflects the electronic influences of X. However, in terms of the Hein-Niemann theory, an R_3-ρ_3 binding interaction is considered to be an important factor.

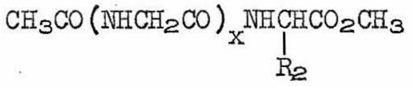
In this theory, as in the quantitative extension of the Hein-Niemann theory, the independence of the group effects is an implicit assumption.

Present Study

As stated earlier, most of the recent investigations concerning the structural specificity of α-chymotrypsin have been concerned with relatively simple synthetic model substrates (in particular, the N-acylated amino acid esters). It must be remembered, however, that α-chymotrypsin is a proteinase and as such its primary in vivo function is the catalysis of peptide bond hydrolysis in very complex protein and polypeptide molecules. The studies described in this thesis represent an attempt to find and study model substrates which more closely resemble the natural substrates of α-chymotrypsin--proteins. It is hoped that the investigations reported herein will add another span or two in the rather long bridge between studies of model substrates and the natural ones.

The first part of this thesis represents an attempt to observe and analyze the qualitative specificity pattern exhibited by α-chymotrypsin towards denatured proteins and polypeptides as indicated in the application of this enzyme in determinations of primary protein sequence. A correlation between this pattern and the quantitative data for model substrates is suggested. This study also served to suggest possible model substrates which were synthesized and used to verify the trends observed.

The first class of compounds whose kinetic parameters as substrates of α-chymotrypsin were determined consists of derivatives of N-acylated amino acid methyl esters in which the N-acyl component was lengthened by the insertion of glycine residues (Part II). This class of compounds was of the following form:



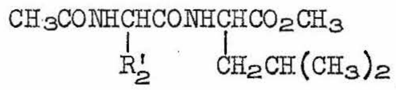
$$CH_3CO(NHCH_2CO)_xNH\underset{R_2}{\underset{|}{C}}HCO_2CH_3$$

with $R_2 = -CH_3,\ -CH(CH_3)_2,\ -CH_2\text{-}CH(CH_3)_2$

$x = 0, 1, 2, 3.$

Comparison of the kinetic parameters obtained, permitted an evaluation of the effect of lengthening the peptide chain in this manner.

The same part of this thesis (Part II) deals also with the determination of kinetic parameters for substrates of the following type:

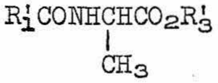


$$CH_3CONH\underset{R_2'}{\underset{|}{C}}HCONH\underset{CH_2CH(CH_3)_2}{\underset{|}{C}}HCO_2CH_3$$

R_2' = varying amino acid side chains

Part II of the thesis may be considered as an investigation of the nature of the ρ_1 locus of the Hein-Niemann active site.

Part III involves kinetic determinations on model substrates of the following type:



$R'_1\mathrm{CONH}\underset{CH_3}{\mathrm{CH}}\mathrm{CO_2}R'_3$

D and *L*

$R'_1 = -CH_3, -C_6H_5$

$R'_3 = -CH_3, -CH_2CONH_2, -CH_2CN.$

and the evaluation of the effect of variation of the ester group on reactivity and stereospecificity.

REFERENCES

1. C. Niemann, *Science*, **143**, 1287 (1964).

2. R. F. Steiner, *The Chemical Foundations of Molecular Biology*, D. Van Nostrand Company, Inc., Princeton, New Jersey (1965).

3. H. Gutfreund, in *The Enzymes*, P. D. Boyer, H. Lardy and K. Myrbäck, Eds., Academic Press, New York, N.Y., ed. 2 (1960), vol. 1, pp. 233-258.

4. E. F. Jansen, M. D. Fellows-Nutting, R. Jang and A. K. Balls, *J. Biol. Chem.*, **179**, 189, 201 (1949). E. F. Jansen and A. K. Balls, *ibid.*, **194**, 721 (1952). E. F. Jansen, R. Jang and A. K. Balls, *ibid.*, **196**, 247 (1952).

5. D. E. Koshland, Jr., D. H. Strumeyer, W. J. Ray, Jr., in *Proc. Symposium on Enzyme Models and Enzyme Systems, 1962* (Brookhaven National Laboratory, Upton, N.Y., 1962), pp. 101-119.

6. F. Sanger, *Proc. Chem. Soc.*, **1963**, 77 (1963).

7. D. H. Strumeyer, W. N. White and D. E. Koshland, Jr., *Proc. Natl. Acad. Sci. U.S.*, **50**, 931 (1963).

8. B. S. Hartley, *Nature*, **201**, 1284 (1964).

9. J. Kraut, L. C. Sieker, D. F. High and S. T. Frier, *Proc. Natl. Acad. Sci. U.S.*, **48**, 1417 (1962).

10. B. S. Hartley and B. A. Kilbey, *Biochem. J.*, **50**, 672 (1952).

11. A. K. Balls and E. F. Hansen, *Adv. Enzymol.*, **13**, 321 (1952).

12. B. F. Erlanger and W. Cohen, *J. Am. Chem. Soc.*, **85**, 348 (1963).

13. M. L. Bender, G. R. Schonbaum, and B. Zerner, *J. Am. Chem. Soc.*, **84**, 2540, 2562 (1962).

14. K. Yagi, *Adv. in Enzymol.*, **27**, 1 (1965).

15. G. E. Hein and C. Niemann, *J. Am. Chem. Soc.*, **84**, 4487, 4495 (1962).

16. H. Neurath and B. S. Hartley, *J. Cell. Comp. Physiol.*, **54** (suppl. 1), 179 (1959).

17. M. L. Bender and F. J. Kézdy, *Ann. Rev. Biochem.*, **34**, 49 (1965).

18. H. Gutfreund, *Discussions Faraday Soc.*, **15**, 167 (1955).

19. H. Gutfreund and J. M. Sturtevant, *Proc. Natl. Acad. Sci. U.S.*, **42**, 719 (1956).

20. B. Zerner and M. L. Bender, *J. Am. Chem. Soc.*, **86**, 3669 (1964).

21. H. Neurath and G. W. Schwert, *Chem. Revs.*, **46**, 69 (1950).

22. G. E. Hein and C. Niemann, *Proc. Natl. Acad. Sci. U.S.*, **47**, 1341 (1962).

23. J. B. Sumner, *J. Biol. Chem.*, **69**, 435 (1926).

24. a) J. P. Wolf, III and C. Niemann, *Biochemistry*, 2, 83 (1963).
    b) J. P. Wolf, III and C. Niemann, *ibid.*, **2**, 493 (1963).

25. H. T. Huang and C. Niemann, *J. Am. Chem. Soc.*, **73**, 3223 (1951).

26. H. R. Waite and C. Niemann, *Biochemistry*, **1**, 250 (1962).

27. J. B. Jones and C. Niemann, *Biochemistry*, **1**, 1093 (1962).

28. J. R. Rapp and C. Niemann, *J. Am. Chem. Soc.*, **85**, 1896 (1963).

29. R. B. Martin and C. Niemann, *J. Am. Chem. Soc.*, **80**, 1473 (1958).

30. C. L. Hamilton, C. Niemann and G. S. Hammond, *Proc. Natl. Acad. Sci. U.S.*, in press.

31. R. W. Taft, in *Steric Effects in Organic Chemistry*, M. S. Newman (Ed.), John Wiley and Sons, Inc., New York, N.Y., ch. 13 (1956).

I. THE STRUCTURAL SPECIFICITY OF α-CHYMOTRYPSIN: POLYPEPTIDES AS SUBSTRATES

INTRODUCTION

Historical Background

During the past decade, the structural and stereochemical specificity of proteolytic enzymes towards low molecular weight model substrates has been extensively studied. The relation of such studies to the biological specificity of these enzymes, a matter of great importance, is still an open question. In the following discussion, some of the available information concerning the specificity of α-chymotrypsin towards polypeptides and denatured proteins will be related to the considerable experimental data for an important class of model substrates, the acylated amino acid esters. A preliminary attempt in this direction has recently been reported (1) and similar but more limited data for pepsin has been analyzed (2).

More than 50 years ago, Emil Fischer suggested (3) that the specificity of proteolytic enzymes is determined by at least four important factors: the number, constitution, sequence and configuration of the amino acids in the peptide chain of a polypeptide substrate. The importance of specific amino acid residues was questioned in the early part of this century, but lack of data prevented the emergence of any clear picture.[†]

It was primarily the work of Bergmann and his collaborators in the 1930's which led to the first concrete concept of proteolytic enzyme specificity. This concept was the result of a study of the semiquantitative enzyme-catalyzed hydrolyses of a large number of small peptides of known structure and stereochemistry. Bergmann concluded (4) that the observed specificity of proteolytic enzymes depended primarily on the nature of the side chains of certain specific amino acid residues and secondarily on the nature of the residues further removed from the susceptible bond. In the next two decades, considerable attention was focussed on this primary factor, the nature of the side chain of the carboxyl residue of susceptible bonds.

[†] A brief history and pertinent references are given by M. Bergmann, Adv. Enzymol., **2**, 49 (1942).

The discovery of the esterase activity of α-chymotrypsin (5) and the subsequent development of rapid kinetic techniques based on the use of the pH-stat (6) greatly advanced the study of the specificity of α-chymotrypsin. Extensive investigations, primarily in these laboratories, have resulted in a specificity theory which correlates the kinetic data obtained for the enzyme-catalyzed hydrolysis of over 150 substrates of α-chymotrypsin (7,8).

Because of the use of α-chymotrypsin and other proteolytic enzymes in the elucidation of amino acid sequences of proteins and polypeptides, the question of the specificity of these enzymes towards larger peptides has received renewed attention in the last few years. As a result of a number of "abnormal" splittings observed (9), the concept of secondary specificity determined by amino acid residues adjacent to the susceptible residue (i.e., the residue which forms the carboxyl component of the cleavable bond) has been reintroduced (1,10,11).

Classification of Enzyme Specificity

Proteolytic enzyme specificity will be discussed here in terms of three specificity levels; primary, secondary and tertiary. Primary specificity refers to the minimum structural requirement for specificity to be manifest. For proteolytic enzymes, this is usually an amino acid residue containing a side chain, an amino group and a carboxyl function For α-chymotrypsin, the primary specificity is observed for the side chain of the amino acid residue which forms the carboxyl component of the peptide bond which is susceptible to hydrolysis. Conventionally, the primary specificity of α-chymotrypsin is towards tryptophanyl, tyrosyl, phenylalanyl and, to a lesser extent, leucyl bonds. As will be illustrated later, the primary specificity of this enzyme is even broader than this. Secondary specificity refers to the influence of amino acid residues adjacent in the peptide chain to the residue in which the primary specificity is manifest. The well-known negative influence of a prolyl residue, when it is the amino component of a peptide bond, on the susceptibility of that bond to both α-chymotrypsin- and trypsin-catalyzed hydrolyses is a common example. Tertiary specificity refers to the influence of conformation and three-dimensional structure of native, undenatured protein molecules in solution on the susceptibility of their peptide bonds towards proteolytic cleavage. One typical example is the activation of chymotrypsinogen by trypsin, in which only a few of the many bonds which should be split by trypsin (on the basis of primary specificity considerations) are cleaved (12). Another example is the unique splitting of a single peptide bond in ribonuclease by subtilisin (13).

A similarity between this proposed nomenclature for enzyme specificity and that commonly used to describe protein structure (14) may be noted. However, the present usage differs from the conventional subdivision of protein structure: the primary and secondary *specificities* of proteolytic enzymes have been defined so that both refer to the primary *structure* of a protein substrate while tertiary *specificity* refers to both the secondary and tertiary *structure* of protein substrates. Because of the lack of information concerning the effects of secondary and tertiary protein structure on susceptibility of peptide bonds towards enzymatic hydrolysis, it would be unreasonable, at the present time, to attempt a separation of these two effects.

A partial understanding of the primary specificity of α-chymotrypsin was achieved only after investigation of the kinetic behavior of a large number of model substrates. An even greater experimental effort is required in order to gain similar knowledge concerning secondary specificity. Some qualitative work in this direction has been performed by Bergmann and his collaborators and a later part of this thesis (Part II) contains a description of a more quantitative study of secondary specificity. Fortunately, results in the literature concerning the action of α-chymotrypsin towards polypeptides and denatured proteins offer some information about the secondary specificity of this enzyme. It is the purpose of the study described herein to examine some of these results from the literature and to propose as far as possible a secondary specificity pattern for α-chymotrypsin-catalyzed reactions.

The question of the tertiary specificity of proteolytic enzymes cannot yet be discussed intelligently. Relatively little information is available concerning the tertiary structure of proteins in solution and the effect of proteolytic enzymes on native proteins.

RESULTS AND DISCUSSION

Data

The specificity pattern shown by α-chymotrypsin has been examined for a limited number of peptides and proteins for which the complete amino acid sequence has been determined. Those included in the analysis are:

1. Human hemoglobin γ chain (15)
2. Human hemoglobin α chain (16)
3. Ribonuclease (17)
4. α-Corticotropin (18)
5. Glucagon (19)
6. Insulin (A and B chains) (20)
7. Horse heart cytochrome C (21)
8. Baker's yeast cytochrome C (22)
9. Tobacco mosaic virus (23,24)

Proteins whose sequences and patterns of splitting were very similar to those for the proteins mentioned above were not included. β-Corticotropin and human heart cytochrome C were omitted on this basis, because they were very similar in the above respect to α-corticotropin and horse heart cytochrome C, respectively. Both horse heart and baker's yeast cytochromes C were included because their sequences and splittings were sufficiently different to justify their inclusion.[†]

[†] Eighteen and 21 splittings, respectively, were observed for horse heart cytochrome C and baker's yeast cytochrome C. However, only 4 of these splittings occurred in sequences which were identical in the two proteins.

The peptide chains considered contain a total of 910 amino acid residues. Their amino acid compositions are shown in Table I. The size of the sample and the diverse origin of the proteins indicate that a relatively representative sample is available. The distribution of amino acids, ranging from 11 tryptophans to 78 alanine residues, is not unusual for amino acid compositions (25,26). Šorm (27) has tabulated the kinds of peptide linkages in proteins and the present sample is compatible with such a tabulation.

The data available in the literature obviously were not compiled for the purposes of the present analysis. Naturally, this raises some difficulties. The problem of minor splittings observed in enzyme-catalyzed hydrolyses is one of some importance. These minor splittings could represent bonds which are split more slowly than others or they could be a result of an impurity in the enzyme preparation employed. There is also the possibility that these splittings were present before the addition of enzyme, accounting perhaps for trace and group analyses which are occasionally reported. In the present study, all bonds which were reported cleaved by α-chymotrypsin were included. The problem of enzyme impurities seems to be a minor one. Other enzyme impurities in modern chymotrypsin preparations are exceedingly small. Furthermore, in the cases of minor splittings observed, no pattern which would implicate another enzyme becomes evident. α-Chymotrypsin causes the splitting of a lysyl-lysine bond in both α- and β-corticotropin and although in the latter case this cleavage has been attributed to a trypsin impurity in the chymotrypsin preparation (28), this was the only one of several trypsin-sensitive bonds split (29). It is obvious that either

TABLE I

Amino Acid Composition of Proteins

Amino Acid[a]	A	B	C	D	E	F	G	H	I	Total[c]
	Protein[b]									
Alanine	11	21	12	3	1	3	6	7	14	78
Arginine	3	3	4	3	2	1	2	3	11	32
Aspartic Acid	8	8	4	2	4	2	3	4	9	44
Asparagine	5	4	11	0	0	1	5	7	9	42
Cysteine	1	1	8	0	0	6	2	3	1	22
Glutamic Acid	8	4	6	4	0	6	9	7	5	49
Glutamine	4	1	6	1	3	1	3	2	11	32
Glycine	13	7	13	3	1	4	12	12	6	71
Histidine	7	10	4	1	1	2	3	4	0	32
Isoleucine	4	0	3	0	0	1	6	4	9	27
Leucine	17	18	2	1	2	6	6	8	12	72
Lysine	12	11	10	4	1	1	19	16	2	76
Methionine	2	2	4	1	1	0	2	2	0	14
Phenylalanine	8	7	3	3	2	3	4	4	8	42
Proline	4	7	4	4	0	1	4	4	8	36
Serine	11	11	15	3	4	3	0	4	16	67
Threonine	10	9	10	0	3	1	10	8	16	67
Tryptophan	3	1	0	1	1	0	1	1	3	11
Tyrosine	2	3	6	2	2	4	4	5	4	32
Valine	13	13	9	3	1	5	3	3	14	64
Total[d]	146	141	134	39	29	51	104	108	158	910

[a] Number of each amino acid found in each protein.

[b] Proteins considered were:

A – Human hemoglobin γ chain (15)
B – Human hemoglobin α chain (16)
C – Ribonuclease (17)
D – α-Corticotropin (18)
E – Glucagon (19)
F – Insulin (A and B chains) (20)
G – Horse heart cytochrome C (21)
H – Baker's yeast cytochrome C (22)
I – Tobacco mosaic virus (23,24)

[c] Total number of each amino acid in all the proteins considered.

[d] Total number of all amino acids in each protein considered.

inclusion or rejection of minor splittings will introduce some error into the analysis. The former course has been chosen.

From the reported experimental conditions employed in the sequence determinations considered (references 15 to 24) it is evident that conditions of concentration, temperature, pH, etc., varied from experiment to experiment. Fortunately, the conditions employed fall within a relatively narrow range. The substrate-to-enzyme ratio is usually between 300:1 and 50:1; the substrate concentration is about 1-2%; the reactions are carried out near room temperature for long times (6-24 hours); the pH is usually near the optimum for α-chymotrypsin (between 7 and 9), but no efforts have been made to control ionic strength. Recent experiments have shown that variations in pH (30,31) and ionic strength (31) can affect the relative specificity of α-chymotrypsin. These variations in experimental conditions make any analysis of the recorded results tentative. However, the enzyme-catalyzed hydrolyses were usually carried out until the rate had slowed down considerably. Therefore, it is probable that further hydrolysis, which at best might have increased the number of minor splittings, would not have significantly altered the results. Only those variations in reaction conditions which would vary the relative specificity of the enzyme might affect the conclusions.

Primary Specificity

The ideal criterion for any discussion of relative specificity is precisely determined kinetic information about the rate of cleavage of the bonds involved. Needless to say, this level of data is not currently available. One must therefore rely on a statistical analysis of relative frequency of occurrence of any particular peptide bond hydrolysis for a number of experiments performed under different conditions.

Table II lists the number of bonds split, involving any of the twenty naturally occurring amino acids as the carboxyl component of the bond. Also included is the percentage of all bonds, involving that particular amino acid as a carboxyl component, which were reported hydrolyzed by α-chymotrypsin. These amino acids range from the "classical" substrates of chymotrypsin to only four amino acids for which no bonds were split. Table II also lists experimentally determined kinetic parameters obtained for acylated amino acid methyl esters as substrates of α-chymotrypsin. The data of Table II actually represent a combination of primary and secondary specificity (tertiary specificity will be less significant because the data refers to experiments performed on denatured proteins). If the concepts of primary, secondary and tertiary specificity (as defined above) have any significance, the latter two should manifest themselves as perturbations of the primary specificity. Thus, secondary specificity effects must also be reflected in the data of Table II. In fact, it is this premise that makes possible the analysis of secondary specificity. However, because the above effect is only a perturbation, the data in Table II represent, for the most part, the primary specificity of the enzyme.

Of the twenty amino acids listed in Table II, sixteen are involved in hydrolysis at least once. Although some small fraction of these hydrolytic cleavages may be due to an impurity in an enzyme preparation, it is clear that α-chymotrypsin has a very broad primary specificity.

TABLE II

The Primary Specificity of α-Chymotrypsin for Bonds in Polypeptides

	Amino Acid	Total Residues	Number Split	%	Kinetic Data for N-Acetyl-L-Amino Acid Methyl Esters[a]			Ref.
					K_o, mM	k_o, sec.$^{-1}$	k_o/K_o, M^{-1}sec.$^{-1}$	
Group I	Tryptophan	11	11	100	.12	51.	4.2×10^5	36
	Tyrosine	32	29	91	.32	117.	3.6×10^5	37
	Phenylalanine	42	35	83	.62	67.	1.1×10^5	37
	Leucine	72	42	58	2.9	4.6	1.6×10^3	38
Group II	Methionine[b]	14	6	43	8.	18.	2.3×10^3	39
	Histidine	32	10	31				
	Glutamine	32	7	22	48.[c]	20.[c]	4.2×10^2[c]	36
	Asparagine	42	7	17	23.[d]	22.[d]	9.5×10^2[d]	40
	Threonine	67	5	7	136.	.21	1.5	41
	Lysine	76	4	5				
	Cysteine[e]	22	1	5				
	Isoleucine	27	1	4	49.	.16	3.3	42
	Valine	64	2	3	112.	.15	1.3	43
	Aspartic Acid	44	1	2				
	Glycine	71	1	1	300.	0.46	0.15	44,45
	Serine	67	1	1				
Group III	Alanine	78	0	0	739.	1.3	1.7	46
	Glutamic Acid	49	0	0				
	Proline	36	0	0				
	Arginine	32	0	0				

TABLE II--*continued*

[a] Unless otherwise noted, kinetic data were obtained in aqueous solutions at 25.0°, pH 7.9 and 0.10 *M* with respect to sodium chloride.

[b] Includes four methionine sulfone residues, two of which were split.

[c] N-acetyl-*L*-glutamic acid dimethyl ester.

[d] N-acetyl-*L*-aspartic acid dimethyl ester at *pH* 7.2.

[e] Includes 14 cysteic acid residues, one of which was split.

As indicated in Table II, the amino acids may be conveniently divided into three groups. Group I includes those amino acids which are hydrolyzed (i.e., appear as the carboxyl component of a hydrolyzed bond) a large percentage of the time they occur: the classical aromatic substrates--tryptophan, phenylalanine and tyrosine--and leucine. Methionine has also been included in this group. Of a total of 171 residues of these amino acids, 123 (72%) are hydrolyzed in chymotryptic cleavages. Group II consists of 12 amino acids which undergo hydrolysis less frequently. Arbitrarily, for purposes of further comparisons, methionine has been included in this group also. Only 46 (8%) of the 558 residues in this group are hydrolyzed. Finally, no Group III amino acids are hydrolyzed although they account for 195 residues. The amino acids have been grouped in this way so that further statistical comparisons can be made.

A comparison of the peptide hydrolysis data with the kinetic data for N-acylated-L-amino acid esters indicates that the primary specificity pattern for polypeptides is a direct reflection of the specificity observed for the model compounds. To the extent of the kinetic data available, k_o/K_o, a measure of substrate reactivity (as indicated in the General Introduction, p. 7) for the model substrates and the susceptibility of bonds of denatured proteins to hydrolysis can be correlated. This correlation is an important step forward in the evaluation of the relevance of studies on model compounds to the *in vivo* action of α-chymotrypsin.

Secondary Specificity

The classification of substrates into groups, as described above, permits a "statistical" analysis of secondary specificity to be made. For Group I amino acids, primary specificity is probably dominant: hydrolysis is the rule with only a relatively few exceptions. Examination of the amino acid sequences near those Group I residues which are not hydrolyzed may give some insight into factors which hinder hydrolysis. Conversely, for Group II amino acids, secondary specificity may be dominant: hydrolysis of these less susceptible residues may be due, to a considerable extent, to the effect of adjacent residues. Examination of the amino acid sequences surrounding those Group II residues which are hydrolyzed may reveal factors which favor hydrolysis. This is the approach which was utilized.

In order to examine the possible secondary specificity pattern, the residues found adjacent to bonds hydrolyzed were tabulated. Each such sequence was classified by the use of a tabulation which listed the amino acid residues on either side of the site of cleavage. These sequence positions were designated CO1 (the carboxyl component of the peptide bond hydrolyzed), NH1 (the amino component of the bond cleaved), CO2 (the N-acylamino acid on CO1), etc. An example of this system, which is similar to that employed by Tang (2), is shown in Figure 1. Such a system was also used to classify the sequences adjacent to Group I residues which were not cleaved.

CO3	CO2	CO1	NH1	NH2	NH3
Ser	Val	$CySO_3H$	Ser	Leu	Tyr
Ser	Leu	Tyr	Glu	Leu	Glu
		Tyr	$CySO_3H$	Asp	

← Peptide linkages hydrolyzed

Figure 1

From the number of bonds hydrolyzed (163) and the total number of sequences available, it was evident that only for the CO2 and NH1 positions was the data sufficient for even tentative conclusions regarding secondary specificity.

a) "Vacant" Positions

The effect of adjacent residues on the susceptibility of any bond to hydrolysis must be modified to account for "vacant" positions due to other splittings or terminal groups. For example, an NH1 group with a free α-carboxyl group (i.e., a "vacant" NH2 position) may exert a different effect than an NH1 residue attached to an NH2 residue. The former situation could arise in two ways: the NH1 residue could be the carboxy-terminal residue of the protein studied or it could be involved as the carboxyl component in a previous cleavage. The question of which of the two hydrolyses occurred first is an important one in the consideration of the latter of these possibilities. For example, consider the following sequence:

CO2	CO1	NH1	NH2
ala	met ↑	asg ↑	lys

sites of cleavage

in which both the methionyl-asparagine bond and the asparaginyl-lysine bond are reported cleaved. In considering the effect of the asparagine residue (NH1) on the susceptibility of the methionyl-asparagine bond it is important to know whether or not the secondary effect of the asparagine is due to the nature of the residue (i.e., its side chain) or to a free α-carboxyl group (which the asparagine residue would possess if the asparaginyl-lysine bond were hydrolyzed before the methionyl-asparagine bond). As will be shown later, the effect of a free carboxyl group on an NH1 residue (or a free amino group on a CO1 residue) on specificity is quite important.

There are two ways in which this problem of vacant positions may be taken into account:

1. In considering the effect of an NH1 or CO2 residue on a CO1-NH1 cleavage (or potential cleavage), it might be assumed that the cleavage considered is the last to occur. This would, in effect, yield the maximum number of vacant positions (CO2 or NH2).

2. The CO2 (or NH2) position might be considered "vacant" if, and only if, the primary specificity of the residue at the CO2 (or NH1) position is higher than or equal to the primary specificity of the residue in question at the CO1 position. This is a crude way of estimating which bond is first cleaved.

Use of either of these approaches will undoubtedly introduce some error in the evaluation of the effect of vacant positions on the susceptibility of peptide bonds. However, it seems that the latter approach represents a better approximation to this effect and this approach is the one employed in further discussions.

Table III lists the total number of bonds cleaved for each group of amino acids as well as the number and percentage of CO2 and NH2 positions which are "vacant" according to the above criterion. Group -I refers to sequences in which the CO1 position is occupied by a Group I residue (tryptophan, tyrosine, phenylalanine, leucine or methionine) but no hydrolysis is observed.

The most significant data in Table III concerns the vacancies occurring when Group II and Group -I residues occupy the CO1 position. Group II residues represent those which are only occasionally split. It might be expected that secondary specificity would play an important role among this set. For Group -I residues, the lack of observed hydrolysis may also be interpreted in terms of unfavorable secondary specificity effects. From Table III, it can be seen that vacant CO2 and NH2 positions are found in only 6% and 4%, respectively, of the observed hydrolyses when a Group II amino acid occupies the CO1 position. In other words, of the total of 46 splittings observed in which a Group II amino acid was the carboxyl component of the cleaved bond, only 5 of these (or 11%) occurred when either the CO2 or NH2 positions were vacant. On the other hand, vacant CO2 and NH2 positions are found in 13% and 17%, respectively, of the cases in which a Group I residue occupies a CO1 position but hydrolysis does not occur. Thus, of the 47 "non-splittings" observed when a Group I amino acids was the carboxyl component of a peptide bond, the NH2 or CO2 positions were vacant in 14 (or 30%) of the cases.

These observations are in perfect accord with the classification of α-chymotrypsin as an endopeptidase. Location at the amino-terminal

TABLE III

Effect of "Vacant" Positions on the Secondary Specificity of α-Chymotrypsin

		"Vacancies"			
		CO2 Position		NH2 Position	
CO1 Residues	Total[a] Bonds	Number	% of Total[b]	Number	% of Total[b]
Group I	123	11	9	7	6
Group -I[c]	47	6	13	8	17
Group II	46	3	6	2	4

[a] Total bonds hydrolyzed with a Group I or Group II residue at the CO1 position. For Group -I this value refers to the number of bonds in which a Group I residue occupies the CO1 position but no hydrolysis is observed.

[b] Calculated from the number of "vacant" positions and the "total bonds."

[c] No hydrolysis with a Group I residue at CO1.

end or penultimate to the carboxy-terminal end of a peptide chain appears to render a normally susceptible Group I amino acid resistant to hydrolysis. Although the data is not extensive, studies with model substrates appear to support this conclusion. For example, Fruton and Bergmann (47) found that glycyl-L-tyrosine amide was approximately 20 times more reactive than L-tyrosine amide in α-chymotrypsin-catalyzed hydrolyses.

From the data of Table III, it appears that the effect of vacant positions represents a dominant factor with regard to secondary specificity. For this reason, in the discussion of the secondary specificity effects of amino acid residues at the NH1 and CO2 positions (see below), sequences with vacancies at the CO2 and NH2 positions, respectively, were deleted from the data where these effects might influence the results.

b) NH1 Position

The effect of the nature of the amino acid residue which occurs as the amine component of a peptide bond (NH1 position) on the susceptibility of that bond towards hydrolysis is presented in Table IV. For each species of amino acid, values have been listed for the "Found" and "Expected" occurrence of that species at the NH1 position when the CO1 position is occupied by Group I, Group II and Group -I residues, respectively. When the CO1 position is occupied by a Group I or Group II residue, the "Found" value represents the number of times a particular amino acid was found at the corresponding NH1 position. Primarily for the sake of convenience, these "Found" values will include NH1 residues which are carboxy-terminal. The "Found" values will be compared to

TABLE IV

The Secondary Specificity of α-Chymotrypsin: Amino Component

Residue	Total Number in Proteins	Effective Total Number[f]	Group I NH1		Group II NH1		Group -I[a] NH1	
			Found	Expected[b]	Found	Expected[b]	Found	Expected[c]
Alanine	78	78	6	10	10	4	1	3
Arginine	32	32	4	4	1	2	0	1
Aspartic Acid	44	43	5	6	0	2	6	2
Asparagine	42	35	2	6	1	2	0	1
Cysteine	22(14)[d]	21(13)[c]	2(2)[d]	3	0	1	2(1)[d]	1
Glutamic Acid	49	49	8	7	1	2	2	2
Glutamine	32	27	2	4	2	2	0	1
Glycine	71	70	14	10	6	4	1	3
Histidine	32	22	3	4	0	2	2	1
Isoleucine	27	25	7	4	2	1	1	1
Leucine	72	39	13	10	1	4	1	2
Lysine	76	72	12	10	8	4	1	3
Methionine	14(4)[e]	9(2)[e]	1	2	0	1	1(1)[e]	0
Phenylalanine	42	8	3	6	1	2	0	0
Proline	36	36	1	5	0	2	7	2
Serine	67	66	15	9	4	3	3	3
Threonine	67	62	14	9	0	3	1	3
Tryptophan	11	0	0	1	0	1	0	0
Tyrosine	32	3	2	4	0	2	0	0
Valine	64	62	9	9	9	3	3	3
Total	910	759	123		46		32	

TABLE IV--continued

[a] No hydrolysis with Group I residue at CO1 position.

[b] Number of residues expected at this position in the absence of any specificity. Obtained by multiplying the total number of residues of any particular amino acid in the sample by the total number of residues found at that position/total number of all amino acid residues.

[c] Calculated as in b, except that the "effective" number of residues of each amino acid was used (see text).

[d] Figures in parentheses refer to cysteic acid.

[e] Figures in parentheses refer to methionine sulfone.

[f] "Effective" number of each amino acid based on the frequency of its occurrence as the carboxyl component of a hydrolyzed bond (see text).

"Expected" values which might be observed simply on the basis that any particular amino acid will appear in any position in direct proportion to the total number of amino acids of that kind available. For example, of the total 910 residues available, 78 (8.6%) are alanines. Thus, one would expect 4 (8.6% of 46) amino components (NH1 positions) associated with Group II residues at CO1 to be alanines.

In order to eliminate the dominant effect of vacant positions in hindering the hydrolysis of the highly susceptible Group I residues, NH1 "Found" values for the cases where a Group I residue occupies the CO1 position but hydrolysis is not observed (Group -I), do not include NH1 residues which were carboxy-terminal or NH1 residues which were associated with amino-terminal CO1 residues. Moreover, it was necessary to reduce the total number of each amino acid by an amount corresponding to the number of times that residue appeared as a carboxy-terminal residue (because of, for example, a "previous" cleavage). Thus, the "Expected" values for the Group -I column in Table IV are based on this "effective" number for each amino acid.

Again, the most significant information concerning secondary specificity may be obtained from the listings under Group II (the less susceptible but sometimes hydrolyzed residues) and Group -I (the "non-hydrolysis of highly susceptible residues). These data indicate that a few types of residues appear to have a favorable effect on hydrolysis of a peptide bond when they occur as the amine component (NH1) of that bond. These amino acids are found more than the expected number of times as the amine components of hydrolyzed bonds in which the carboxyl components are Group II residues and less than the expected number of times as the amine components of unhydrolyzed peptide bonds having Group I residues as the carboxyl components (Group -I). The three most favorable residues appear to be alanine, lysine and glycine. Valine, because of its high occurrence in Group II hydrolyses, might also be included.

Conversely, some amino acids appear to hinder hydrolysis when they occur as the amine component of the peptide bond. These residues are found less than the expected number of times in Group II hydrolyses and more than the expected number of times in Group -I "non-hydrolyses." These amino acids are proline, aspartic acid and perhaps histidine.

The unfavorable effect of proline as the amino component of a peptide bond on the hydrolysis of that bond has been known for some time. This is a dramatic effect which can be considered to account for all of the 3 tyrosine "non-hydrolyses" and 3 of the 7 phenylalanine "non-hydrolyses." The effect of aspartic acid may be ascribed to the free carboxylate group although this effect is not observed for glutamic acid. Confirmation of the secondary specificity of α-chymotrypsin for the other residues mentioned must await quantitative experiments on suitable model compounds.

c) CO2 Position

In Table V, data regarding the possible secondary specificity towards CO2 residues are presented. The data are presented in the same manner as in Table IV, except that the effect of vacant positions is eliminated for the Group I and Group II cases as well as for the Group -I case. "Expected" values were based on an "effective" number of each amino acid (based on consideration of the number of times that amino

TABLE V

The Secondary Specificity of α-Chymotrypsin: N-Acyl Component

Residue	Total Number in Proteins	Effective[e] Total Number	Group I CO2		Group II CO2		Group -I[a] CO2	
			Found	Expected[b]	Found	Expected[b]	Found	Expected[b]
Alanine	78	78	15	11	3	4	0	3
Arginine	32	32	4	5	2	2	0	1
Aspartic Acid	44	43	6	6	0	2	3	1
Asparagine	42	35	6	5	0	2	3	1
Cysteine	22(14)[c]	21(13)[c]	1(1)[c]	3	0	1	1	1
Glutamic Acid	49	49	8	7	0	3	2	2
Glutamine	32	26	4	4	2	1	2	1
Glycine	71	70	5	10	2	4	0	2
Histidine	32	22	4	3	1	1	5	1
Isoleucine	27	25	4	4	2	1	0	1
Leucine	72	34	7	5	3	2	2	1
Lysine	76	72	13	10	9	4	0	3
Methionine	14(4)[d]	8(2)[d]	1	1	0	0	1(1)[d]	0
Phenylalanine	42	9	2	1	0	0	0	0
Proline	36	36	3	5	5	2	1	1
Serine	67	66	8	10	2	4	4	2
Threonine	67	63	12	9	4	4	1	2
Tryptophan	11	0	0	0	0	0	0	0
Tyrosine	32	3	0	0	0	0	0	0
Valine	64	62	9	9	8	4	1	2
Total	910	754	112		43		26	

TABLE V--continued

[a] No hydrolysis with Group I residue at CO1 position.

[b] Number of residues expected at this position in the absence of any specificity. Obtained by multiplying the total number of residues of any particular amino acid by the total number of residues found at that position/total effective number of residues.

[c] Figures in parentheses refer to cysteic acid.

[d] Figures in parentheses refer to methionine sulfone.

[e] "Effective" number of each amino acid based on the frequency of its occurrence as the carboxyl component of a hydrolyzed bond.

acid appeared as an amino-terminal residue). Also, because of the demonstrated negative effect of proline at the NH1 position, CO2 residues involved in sequences with proline at the NH1 position were not included in the Group -I column.

Again, comparison of the data for Group II and Group -I occurrence suggests that some residues at the CO2 position (penultimate to the carboxyl component of the peptide bond in question) are particularly effective in promoting hydrolysis while others hinder it. Lysine, valine, proline and isoleucine, when present at the CO2 position appear to favor hydrolysis, while aspartic acid, asparagine, histidine and serine at this position appear to hinder it.

Some quantitative experiments on model compounds, described in Part II of this thesis, represent a limited but moderately successful attempt to confirm the CO2 secondary specificity pattern described above.

Perusal of Tables IV and V reveals that all conclusions regarding secondary specificity must be considered tentative at this time. The results are suggestive and the approach may be fruitful, but it is clear that a much larger sample is required before more definitive conclusions can be reached on the basis of statistical analyses. As a result of the currently accelerating rate of sequence determinations, and particularly the quantitative studies of enzyme kinetics with proteins of known sequence (32,33,34,35), the necessary data may soon be available for a definition of the secondary specificity of α-chymotrypsin and other proteolytic enzymes.

REFERENCES

1. R. L. Hill, *Adv. Protein Chem.*, **20**, 68 (1964).

2. J. Tang, *Nature*, **199**, 1094 (1963).

3. E. Fischer and P. Bergell, *Ber.*, **36**, 2592 (1903); **37**, 3103 (1904).

4. M. Bergmann and J. S. Fruton, *Adv. Enzymol.*, **1**, 63 (1941).

5. H. Neurath and G. W. Schwert, *Chem. Rev.*, **46**, 69 (1950).

6. M. Dixon and E. C. Webb, *The Enzymes*, 2nd ed., Academic Press, New York (1963), Ch. 2.

7. G. E. Hein and C. Niemann, *J. Am. Chem. Soc.*, **84**, 4895 (1962).

8. C. Niemann, *Science*, **143**, 1287 (1964).

9. Zuber, *Chymia*, **12**, 405 (1958).

10. J. Leonis, C. H. Li and D. Chung, *J. Am. Chem. Soc.*, **81**, 419 (1959).

11. C. H S. Hirs, S. Moore and W. H. Stein, *J. Biol. Chem.*, **235**, 633 (1960).

12. P. Desnuelle, in *The Enzymes*, Vol. 4, P. D. Boyer, H. Lardy and K. Myrback, Eds., Academic Press, New York (1960), Ch. 5.

13. F. M. Richards and P. J. Vithayathil, *J. Biol. Chem.*, **236**, 1380 (1961).

14. K. V. Linderstrøm-Lang, Lane Medical Lectures, Stanford U. Press, Stanford (1952), p. 58.

15. W. A. Schroeder, *et al.*, *Biochemistry*, **2**, 992 (1963).

16. R. J. Hill and W. Konigsberg, *J. Biol. Chem.*, **237**, 3151 (1962).

17. C. H W. Hirs, W. H. Stein and S. Moore, *J. Biol. Chem.*, **221**, 151 (1956); 235, 625, 633 (1956).

18. G. H. Li, et al., *Nature*, **176**, 687 (1955).

19. W. W. Bromer, *et al.*, *J. Am. Chem. Soc.*, **79**, 2798 (1957).

20. F. Sanger and E. O. P. Thompson, *Biochem. J.*, **53**, 367 (1953); F. Sanger and H. Tuppy, *Biochem. J.*, **49**, 481 (1951).

21. E. Margoliash, *J. Biol. Chem.*, **237**, 2148, 2151, 2161 (1962).

22. K. Narita, K. Titani, Y. Yaoi and H. Murakami, *Biochim. Biophys. Acta*, **77**, 688 (1963).

23. G. Funatsu, A. Tsugita and H. Fraenkel-Conrat, *Arch. Biochim. Biophys.*, **105**, 25 (1964).

24. F. A. Anderer, H. Uhlig, E. Weber and G. Schramm, *Nature*, **186**, 922 (1960).

25. G. R. Tristram and R. H. Smith, *Adv. Protein Chem.*, **18**, 227 (1963).

26. F. T. Hatch, *Nature*, **206**, 777 (1965).

27. E. Šorm, *et al.*, *Coll. Czechoslov. Chem. Communo.*, **26**, 1174 (1961); F. Šorm and B. Keil, *Adv. Protein Chem.*, **17**, 167 (1962).

28. H. Neurath, *Ann. N. Y. Acad. Sci.*, **68**, 15 (1957).

29. R. D. Shephard, *et al.*, *J. Am. Chem. Soc.*, **78**, 5067 (1956).

30. A. Platt and C. Niemann, *Proc. Natl. Acad. Sci.*, **50**, 817 (1963).

31. W. Mukatis, Ph.D. Thesis, California Institute of Technology, 1964.

32. J. A. Rupley and H. A. Scheraga, *Biochemistry*, **2**, 421 (1963).

33. M. Kaminski and C. E. Tanner, *Biochem. Biophys. Acta*, **31**, 10 (1959).

34. A. Ginsburg and H. K. Schachman, *J. Biol. Chem.*, **235**, 108 (1960).

35. J. Yon, *Bull. soc. chim. biol.*, **42**, 1263 (1960).

36. G. L. Neil, unpublished experiments.

37. R. L. Peterson, Ph.D. Thesis, California Institute of Technology, 1964.

38. G. E. Hein, J. B. Jones and C. Niemann, *Biochim. Biophys. Acta*, **65**, 353 (1962).

39. T. Kunitake, California Institute of Technology, unpublished results.

40. S. G. Cohen and J. Crossley, *J. Am. Chem. Soc.*, **86**, 4999 (1964).

41. V. K. Jones, unpublished experiments.

42. C. Hamilton, unpublished experiments.

43. H. R. Waite and C. Niemann, *Biochemistry*, **1**, 250 (1962).

44. D. W. Ingles and J. R. Knowles, Biochem. J., **96**, 71P (1965).

45. J. P. Wolf, III, R. A. Wallace, R. L. Peterson and C. Niemann, Biochemistry, **3**, 940 (1964).

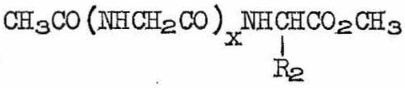
46. J. B. Jones, T. Kunitake, C. Niemann and G. Hein, J. Am. Chem. Soc., **87**, 1777 (1965).

47. J. S. Fruton and M. Bergmann, J. Biol. Chem., **145**, 253 (1942).

II. THE STRUCTURAL SPECIFICITY OF α-CHYMOTRYPSIN: N-ACYLATED PEPTIDE ESTERS AS SUBSTRATES

INTRODUCTION

As an initial attempt in the study of model substrates which closely resemble the natural protein and polypeptide substrates of α-chymotrypsin, a series of N-acylated amino acid methyl esters, in which the N-acyl group consisted of a varying number of glycine residues, was synthesized. The α-chymotrypsin-catalyzed hydrolyses of these peptide derivatives were studied. The model compounds were of the form (I):

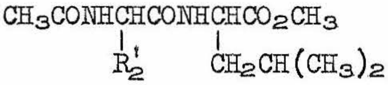


$$CH_3CO(NHCH_2CO)_x NH\underset{R_2}{\underset{|}{C}}HCO_2CH_3$$

I

where the carboxyl-terminal residue was L-alanine ($R_2 = -CH_3$), L-valine ($R_2 = -CH(CH_3)_2$) and L-leucine ($R_2 = -CH_2CH(CH_3)_2$) and the number of glycine residues in the peptide chain was varied from zero to three (i.e., $x = 0,1,2,3$).

In a further study, the α-chymotrypsin-catalyzed hydrolyses of a number of N-acetylated dipeptide esters were examined. These substrates were of the form (II):



$$CH_3CONH\underset{R_2'}{\underset{|}{C}}HCONH\underset{CH_2CH(CH_3)_2}{\underset{|}{C}}HCO_2CH_3$$

II

where the carboxy-terminal residue was L-leucine and the amino-terminal residue was varied throughout a series of amino acids.

Although both these studies represent an investigation of the ρ_1 locus (i.e., the locus at the active site which interacts with the

R_1'CONH- group of a substrate of the form $R_1'CONHCHR_2COR_3'$) they will be discussed in separate sections. The results obtained will be discussed, insofar as it is possible, in terms of the kinetic schemes and specificity theories described in the General Introduction. They will also be compared with data of other workers and with the qualitative specificity trends for polypeptide substrates described earlier in this dissertation (Part I).

RESULTS AND DISCUSSION

SECTION A:

N-Acetylated Peptide Esters Containing Glycine Residues

Data

In 1960, Braunholtz and Niemann (1) reported the kinetic parameters for the α-chymotrypsin-catalyzed hydrolysis of N-acetylglycyl-*L*-tyrosine hydrazide and compared them with the previously obtained data for N-acetyl-*L*-tyrosine hydrazide (2). These data are presented in Table I. The authors concluded that, for these two substrates, the kinetic parameters were identical (within experimental error).

In an experiment performed in these laboratories, Jones determined the kinetic parameters for the α-chymotrypsin-catalyzed hydrolysis of N-acetylglycylglycine methyl ester (3). These results, as well as the parameters obtained for N-acetylglycine methyl ester (recalculated by Ingles and Knowles (4) from data of Wolf, *et al.* (5)) are shown in Table I. Comparison of these data indicates that the N-acetylglycyl- derivative is slightly less reactive (using k_o/K_o as a criterion) than the N-acetyl-substrate. Thus, from the above results, it might have been concluded that substitution of an N-acetylglycyl- group for the N-acetyl- group in N-acetylamino acid derivatives (if it had any effect at all) resulted in a reduction of reactivity.

Results obtained by Yamashita and co-workers, however, indicated that the above conclusion might not be general. Yamashita determined and compared the kinetic parameters for the α-chymotrypsin-catalyzed hydrolyses of N-acetyl-*L*-tyrosine amide and N-acetylglycyl-*L*-tyrosine

TABLE I

Comparison of the α-Chymotrypsin-catalyzed Hydrolyses of Some N-Acetyl- and N-Acetylglycyl- Amino Acid Derivatives

Substrate	k_o sec^{-1}	K_o m*M*	k_o/K_o *M*$^{-1}$sec^{-1}	Ref.
N-Acetyl-*L*-tyrosine hydrazide[a]	0.076±0.014	29.5± 6.0	2.6 ±0.2	2
N-Acetylglycyl-*L*-tyrosine hydrazide[b]	0.076±0.014	22 ± 9	3.5 ±1.3	1
N-Acetylglycine methyl ester[c]	0.046	300	0.15	4,5
N-Acetylglycylglycine methyl ester[d]	0.21 ±0.04	2100 ±500	0.10±0.01	3
N-Acetyl-*L*-tyrosine amide[e]	1.16	153	7.6	6
N-Acetylglycyl-*L*-tyrosine amide[e]	0.095	2.77	34	6

[a] 25°, pH 7.90, buffer containing 0.02 *M* *Tris* component.

[b] 25°, pH 7.80, buffer containing 0.02 *M* *Tris* component.

[c] 25°, pH 7.9 , 1.93 *M* NaCl.

[d] 25°, pH 7.9 , 0.1 *M* NaCl.

[e] 30°, pH 7.80, buffer containing 0.1 *M* phosphate component.

amide (6) and found a substantial increase in reactivity for the latter substrate. These data are also displayed in Table I. It must be noted, however, that Yamashita's published data did not include any estimate of the error associated with the parameters determined for his substrates. As discussed later in this thesis (p. 122), the kinetic parameters for Yamashita's substrates have been recalculated using his literature data and a computer least-squares analysis which yields an estimate of the standard deviations associated with the parameters. This re-evaluation yields the following values: N-acetyl-L-tyrosine amide, $k_o/K_o = 7.5 \pm 0.2$ $\underline{M}^{-1}$ sec^{-1} and N-acetylglycyl-L-tyrosine amide, $k_o/K_o = 44 \pm 60$ $\underline{M}^{-1}$ sec^{-1}. These values may be compared with those obtained by Yamashita of 7.6 $\underline{M}^{-1}$ sec^{-1} and 34 $\underline{M}^{-1}$ sec^{-1}, respectively. Thus the apparent 4.5-fold increase in reactivity for the N-acetylglycyl- substrate (as indicated by Yamashita's k_o/K_o data) is not significant because of the large (> 100%) standard deviation in the k_o/K_o value for one of the two substrates compared.

Research of Kunitake in these laboratories (9) did, however, indicate a substantial increase in reactivity when the N-acetyl- groups of N-acetyl-L-alanine and N-acetyl-L-valine methyl esters were replaced by N-acetylglycyl- groups. The kinetic parameters for these latter substrates were redetermined as part of the present study (Table II).

The purpose of the study described herein was to determine the magnitude of the effect of lengthening the peptide chain (at least in one direction with respect to the hydrolyzable bond) on the kinetic parameters for α-chymotrypsin-catalyzed hydrolysis and to observe the extent of this effect. It was hoped that this study might yield some further information concerning the nature and size of ρ_1, the binding locus complementary to the N-acylamino- function of model substrates.

The choice of substrates for this study was dictated in part by considerations of a practical nature. Esters rather than amides were chosen for two main reasons. In the first place, the much greater reactivity of the methyl esters of N-acylated amino acids (as compared to the corresponding amides) permitted the study of substrates over a larger range of reactivity (using the convenient pH-stat technique) than could have been afforded with the less reactive amide substrates. Thus, it was possible to study the aforementioned effects on substrates ranging from the intrinsically very poor L-alanine substrates to the L-leucine substrates which are about 1000 times more reactive and which may be considered as approximate models for the "specific" substrates of α-chymotrypsin. Secondly, choice of the ester also enabled one to be confident that hydrolysis would occur only at the ester linkage and not at a peptide bond because of the much greater susceptibility of the ester linkage. This problem is not of great importance in this particular study because of the intrinsically low reactivity of the glycyl peptide linkage, but becomes increasingly important in the study of dipeptide derivatives described later (p. 89).

Glycine residues were chosen to lengthen the peptide chain primarily for synthetic reasons. This choice also minimized the possibility of non-productive binding of the type discussed in the study of dipeptide derivatives (pp. 94-112).

The experimental conditions and the kinetic parameters obtained for the α-chymotrypsin-catalyzed hydrolyses of the substrates considered in this study are shown in Table II. With the one exception noted below, all kinetic determinations were carried out with a range of substrate concentrations such that reasonably accurate separation of the parameters, k_o and K_o, could be effected.

Because of the relatively low solubility of N-acetylglycylglycylglycyl-L-alanine methyl ester and the relatively high concentrations (ca. 100 mM[†]) required to obtain accurate separation of the two kinetic parameters, k_o and K_o, it was possible to accurately determine only the value of the ratio k_o/K_o for this substrate. By carrying out the kinetic determination in a concentration range where [S], the concentration of substrate, is much less than K_o, the initial rate of the reaction is given by:

$$v_o \approx (k_o/K_o)[E]_o[S]_o \qquad (1)$$

Evaluation of the rate at $[S]_o = 1.02$ and 0.51 mM indicated that indeed the initial rate (v_o) was directly proportional to $[S]_o$ and the value of $k_o/K_o = 27 \pm 3\ M^{-1}\ sec^{-1}$ was obtained.

Effects of Structural Changes on the Kinetic Parameters

In Table III, the kinetic data of Table II is presented relative to the N-acetyl-L-amino acid methyl esters considered (i.e., x = 0). Thus, the effects on the kinetic parameters of lengthening the R_1 group in the form of an N-acetylated glycine-peptide chain may be observed and

---

[†] Examination of the K_o values of the other substrates in Table II suggested that K_o for this compound should be of the order of 100 mM.

TABLE II

The α-Chymotrypsin-catalyzed Hydrolyses of Some N-Acetylated Peptide Methyl Esters Containing a Varying Number of Glycine Residues[a]

SUBSTRATE $CH_3CO(NHCH_2CO)_xNHCHR_2CO_2CH_3$[b]	$\underline{\underline{L}}$	$[S]_o$ mM	$[E]_o$[i] 10^{-7} M	No.[c] of Expts.	k_o[i] sec^{-1}	K_o mM	k_o/K_o $M^{-1}sec^{-1}$
x	R_2						
0[d]	$-CH_3$	45 - 364	170	—	1.26 ± 0.03	739 ± 28	1.71 ± 0.06
1		24.5 - 195.6	31.7	11-0	2.1 ± 0.4	242 ± 63	9 ± 2
2		5.8 - 46.3	59.6	9-0	4.2 ± 0.3	161 ± 15	26 ± 1
3[h]		—	—	2	—	—	27 ± 3[h]
0[e]	$-CH(CH_3)_2$		—	—	0.15 ± 0.01	112 ± 12	1.34 ± 0.11
1[f]		7.10 - 56.8	32	9-0	0.36 ± 0.02	26 ± 2	13.8 ± 0.8
2		0.59 - 4.69	26	9-0	0.16 ± 0.01	3.7 ± 0.4	43 ± 4
3		2.1 - 16.6	130	11-1	0.435 ± 0.009	11.3 ± 0.4	38 ± 1
0[g]	$-CH_2CH(CH_3)_2$	0.63 - 5.04	34.6	8-0	4.6 ± 0.3	2.9 ± 0.3	1600 ± 100
1		0.29 - 2.57	0.40	13-0	7.7 ± 0.5	1.2 ± 0.2	6500 ± 800
2		0.104- 0.831	0.27	11-0	12.8 ± 1.4	0.41 ± 0.10	31000 ± 7000
3		0.173- 1.557	0.133	11-0	13.8 ± 1.0	0.73 ± 0.12	18800 ± 2800

TABLE II--continued

[a] In aqueous solutions at 25.0°, pH 7.90 ± 0.10, 0.10 M in sodium chloride.

[b] Substrates considered are the N-acetyl, N-acetylglycyl, N-acetylglycylglycyl-, and N-acetylglycylglycylglycyl- derivatives of L-alanine, L-valine and L-leucine methyl esters. The subscript x refers to the number of glycyl residues.

[c] Number of experiments performed for evaluation of k_o and K_o. The second number refers to those rejected by the statistical reiterative procedure used in this evaluation.

[d] Parameters determined by Jones, *et al.* (7).

[e] Parameters determined by Waite and Niemann (8).

[f] Parameters determined by Kunitake (9).

[g] Parameters determined by Hein, Jones and Niemann (10).

[h] The parameter k_o/K_o was determined as described in the text p. 57.

[i] Based on a molecular weight of 25,000 and a protein-nitrogen content of 16.5% for α-chymotrypsin.

TABLE III

Kinetic Parameters of Table II Expressed Relative to Those for the N-Acetyl-L-Amino Acid Methyl Esters

Substrate $CH_3CO(NHCH_2CO)_xNHCHR_2CO_2CH_3$ x	 R_2	Relative k_o[a]	Relative $\bar{K}_o$[a]	Relative k_o/K_o[a]
0	$-CH_3$	1.00 ± 0.02	1.00 ± 0.05	1.00 ± 0.05
1		1.6 ± 0.3	3.0 ± 0.8	5 ± 1
2		3.3 ± 0.3	4.6 ± 0.5	15 ± 1
3		——	——	16 ± 2
0	$-CH(CH_3)_2$	1.00 ± 0.06	1.0 ± 0.1	1.0 ± 0.1
1		2.4 ± 0.1	4.3 ± 0.7	10 ± 1
2		1.1 ± 0.1	30 ± 7	32 ± 3
3		2.9 ± 0.1	10 ± 1	28 ± 3
0	$-CH_2CH(CH_3)_2$	1.00 ± 0.06	1.0 ± 0.1	1.0 ± 0.1
1		1.7 ± 0.2	2.4 ± 0.4	4.0 ± 0.7
2		2.8 ± 0.3	7 ± 2	19 ± 5
3		3.0 ± 0.2	4.0 ± 0.8	12 ± 2

[a] From data of Table II.

comparisons may more easily be made. The magnitudes of this effect on k_o, on $\bar{K}_o$ (i.e., the reciprocal of K_o and thus a direct measure of effectiveness of binding) and on k_o/K_o are displayed.

a) Effects on k_o/K_o

The observed effects on the parameter k_o/K_o will be discussed first. As explained in the introduction, this ratio of experimentally obtained parameters has certain advantages for interpretative purposes (i.e., $k_o/K_o = k_2/K_s$ for both the simple and acyl-enzyme schemes).

According to the Bender-Kézdy specificity theory (11), variation of the R_1 group (i.e., the acylamino group) of a trifunctional model substrate with the same side chain (R_2 group) and carboxyl function (X or COR_3) should affect the k_o/K_o ratio in a manner which is independent of the nature of R_2 and X. Expressing this in a relative manner, we obtain:

$$\frac{(k_o/K_o)_{R_1R_2X}}{(k_o/K_o)_{R_{1o}R_2X}} = S^*_{R_1} \qquad (2)$$

where $R_{1o}R_2X$ refers to some reference substrate, R_{1o} is some reference R_1 group and $S^*_{R_1}$† is a constant for a particular R_1 group (relative to the R_{1o} group) which is independent of R_2 and X. Values of $S^*_{R_1}$ for the N-acetylglycyl-, N-acetylglycylglycyl- and N-acetylglycylglycylglycyl- groups (referred to the N-acetyl- group as R_{1o}) are given in Table IV (data from Tables I and II).

---

† $S^*_{R_1}$ is the antilogarithm of the Bender-Kézdy substituent constant S_{R_1} (see General Introduction, p. 14).

TABLE IV

Bender-Kézdy $S^*_{R_1}$ Values for Three N-Acylamino Groups[a]

R_1[b]	R_2	X	$S^*_{R_1}$[c]
AcglyNH–	$-H$	$-CO_2CH_3$	0.7
	$-CH_3$	$-CO_2CH_3$	5 ± 1
	$-CH(CH_3)_2$	$-CO_2CH_3$	10 ± 1
	$-CH_2CH(CH_3)_2$	$-CO_2CH_3$	4.0 ± 0.7
	$-CH_2C_6H_4OH$	$-CONH_2$	5.7 ± > 100%
	$-CH_2C_6H_4OH$	$-CONHCH_2$	1.4 ± 0.5
AcglyglyNH–	$-CH_3$	$-CO_2CH_3$	15 ± 1
	$-CH(CH_3)_2$	$-CO_2CH_3$	32 ± 3
	$-CH_2CH(CH_3)_2$	$-CO_2CH_3$	19 ± 5
AcglyglyglyNH–	$-CH_3$	$-CO_2CH_3$	16 ± 2
	$-CH(CH_3)_2$	$-CO_2CH_3$	28 ± 3
	$-CH_2CH(CH_3)_2$	$-CO_2CH_3$	12 ± 2

[a] Data from Tables I and II.

[b] R_1 groups considered are as follows: AcglyNH– ≡ $CH_3CONHCH_2CONH-$; AcglyglyNH– ≡ $CH_3CONHCH_2CONHCH_2CONH-$; AcglyglyglyNH– ≡ $CH_3CONHCH_2CONHCH_2CONHCH_2CONH-$.

[c] $S^*_{R_1}$ defined as in the text (p. 61) for $R_{10} = CH_3CONH-$.

If, for the moment, only the data for the L-alanine and L-leucine derivatives are considered, it is observed that the values of $S^*_{R_1}$ obtained from these data are essentially equal (within the calculated error) for all three R_1 groups considered. It is interesting to note that although the $S^*_{R_1}$ values obtained from the L-valine substrates are high, they are all uniformly high by a factor of about two. The value of $S^*_{R_1}$ for the N-acetylglycyl- group, obtained from the L-tyrosine amides, appears to agree quite favorably with that obtained from the L-alanine and L-leucine methyl esters but it is obvious that the very large error associated with the former value eliminates its significance. If the average $S^*_{R_1}$ value for R_1=N-acetylglycyl- obtained from the L-alanine and L-leucine methyl esters is assumed to represent the true value, then certainly the values obtained from the glycine methyl esters and the L-tyrosine hydrazides must be considered anomalous.

At first glance, it may appear that there is relatively little justification for assuming that the Bender-Kézdy relationship holds for these data. However, the success of this treatment in other instances (11) and the relatively good agreement for values obtained from the L-alanine and L-leucine methyl esters makes it worthwhile to treat the data obtained in this way and assuming the L-alanine- and L-leucine-derived values to be the true values, to discuss possible reasons for the deviations of the other values obtained.

Thus, explanations for the following observations must be presented:

1. The $S^*_{R_1}$ value for R_1 = N-acetylglycyl- obtained from the glycine methyl ester data is extremely low.

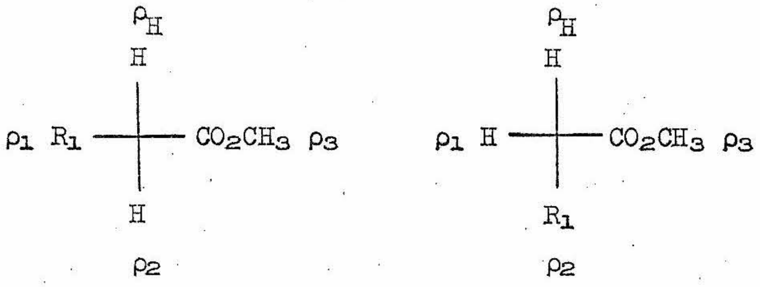
2. The $S^*_{R_1}$ values for the three R_1 groups, obtained from data for the L-valine methyl esters, are all uniformly high by the same factor.

3. The N-acetylglycyl- $S^*_{R_1}$ value obtained from data for the L-tyrosine hydrazides is low.

A reasonable explanation of the glycine derivative discrepancy involves consideration of multiple binding modes. The experimental data (Tables I and IV) indicate that the replacement of the N-acetyl-group of N-acetylglycine methyl ester by an N-acetylglycyl- group actually decreases reactivity (i.e., $S^*_{R_1} < 1$). This behavior is in direct contrast to that observed for the other derivatives studied (i.e., $S^*_{R_1} > 1$). It has been suggested (12,13) that the bifunctional substrate, N-acetylglycine methyl ester, because of its lack of a side chain (and thus a center of asymmetry) may bind to the enzyme in several modes, two of which are of major importance and both of which may be productive. This is in contrast to the trifunctional substrates which, although they may have several binding modes, have only one productive mode (14,15). The two productive modes (ES_1 and ES_2) for N-acetylglycine methyl ester are illustrated in Figure 1 and are characterized by $R_1-\rho_1$, $R_3-\rho_3$ and by $R_1-\rho_2$, $R_3-\rho_3$ interactions, respectively. Neither of these modes places a group larger than hydrogen in ρ_H and thus, from one of the postulates of the Hein-Niemann theory (15), both may be productive.

The kinetic scheme for the case of two productive modes is presented below:

$$\begin{tikzcd} E + S \arrow[r, shift left, "k_{11}"] & ES_1 \arrow[l, shift left, "k_{-11}"] \arrow[dr, "k_{21}"] & \\ & & E + P \\ E + S \arrow[r, shift left, "k_{12}"] & ES_2 \arrow[l, shift left, "k_{-12}"] \arrow[ur, "k_{22}"] & \end{tikzcd} \qquad (3)$$



ES_1 ES_2

Productive Modes for N-Acetylglycine Methyl Ester

$R_1 = CH_3CONH-$

Figure 1

A kinetic analysis of such a kinetic scheme leads to the familiar Michaelis-Menten rate equation (equation 1, p. 3). The observed parameters have the following interpretations:

$$K_o = \frac{K_{s1} K_{s2}}{K_{s1} + K_{s2}} \qquad (4a) \qquad\qquad k_o = \left(\frac{k_{21}}{K_{s1}} + \frac{k_{22}}{K_{s2}} \right) K_o \qquad (4b)$$

where $K_{s1} \approx k_{-11}/k_{11}$ and $K_{s2} \approx k_{-12}/k_{12}$

and $$k_o/K_o = k_{21}/K_{s1} + k_{22}/K_{s2} \qquad (5)$$

From equation 5, it is therefore evident that if both k_{21}/K_{s2} and k_{22}/K_{s2} are of comparable magnitude then k_o/K_o will be greater than it would be if only one productive mode were available.

For the case of N-acetylglycylglycine methyl ester , it will be assumed that one of the productive modes of Figure 1 (i.e., ES_2) is not available. This substrate contains a large R_1 group, the size of which should prevent effective R_1-ρ_2 interaction. It has been shown (16) that the ρ_2 locus is of limited size and that it contains some steric obstructions. Thus, the above assumption is not unreasonable.[†]

If N-acetylglycine methyl ester has a productive mode (ES_2) available to it which N-acetylglycylglycine methyl ester has not, it is quite reasonable to expect that the k_o/K_o value for the former substrate might be greater than that for the latter.

Using the following approximate data (from Table I) and employing some assumptions based on the kinetic behavior of the corresponding L-alanine and L-leucine derivatives (Table II), the values of k_{21}, k_{22}, K_{s1} and K_{s2} for N-acetylglycine methyl ester can be approximated.

Kinetic Data for Two Glycine Methyl Esters (from Table I)

Substrate	k_o (sec^{-1})	K_o (m*M*)	k_o/K_o (m*M*$^{-1}$$sec^{-1}$)
N-Acetylglycine methyl ester (4,5)	0.045	300	0.15×10^{-3}
N-Acetylglycylglycine methyl ester (3)	0.21	2100	0.10×10^{-3}

[†] It is possible that R_1-ρ_2 interaction (where R_1 = N-acetylglycyl-) might occur but that such a mode, because of steric or other interactions (e.g., hydrogen-bonding), might disorient R_3 at ρ_3. This situation would result in a decrease in k_{22} and an increase in K_{22}, yielding an overall decrease in the productive nature of the ES_2 complex.

Assuming that the only available (or important) productive complex for N-acetylglycylglycine methyl ester is that involving $R_1-\rho_1$ interaction (i.e., ES_1 of Figure 1 where $R_1 = CH_3CONHCH_2CONH-$) an estimate of the value of k_{21}/K_{s1} for the ES_1 productive mode of N-acetylglycine methyl ester can be obtained. The average $S^*_{R_1}$ value (for R_1 = N-acetylglycyl-) obtained from the L-alanine and L-leucine data of Table IV is 4.5. If it is assumed that the kinetic behavior of the glycine substrates (in the ES_1 productive mode) parallels that of the corresponding L-alanine and L-leucine derivatives as R_1 is varied, then:

$$\frac{k_{21}}{K_{s1}} = \frac{(k_o/K_o)_{\text{Ac-glygly-OCH}_3}{}^{\dagger}}{S^*_{R_1}} = \frac{0.10 \times 10^{-3}}{4.5} \tag{6}$$

$$= 0.022 \times 10^{-3} \text{ m}\underline{\text{M}}^{-1} \text{ sec}^{-1}$$

From equation 5 then,

$$\frac{k_{22}}{K_{s2}} = 0.15 \times 10^{-3} - 0.022 \times 10^{-3}$$

$$= 0.128 \times 10^{-3} \text{ m}\underline{\text{M}}^{-1} \text{sec}^{-1} \tag{7}$$

These values indicate that the ES_2 productive mode ($R_1-\rho_2$, $R_3-\rho_3$ interactions) contributes more to the k_o/K_o value for N-acetylglycine methyl ester than does the ES_1 mode ($R_1-\rho_1$, $R_3-\rho_3$ interactions).

From the average of the L-alanine and L-leucine data of Table III, it is observed that K_o for the N-acetyl- derivative is approximately 2.7 times that for the N-acetylglycyl- derivative. With the use of this

---

[†] Ac-glygly-OCH_3 ≡ N-acetylglycylglycine methyl ester.

value and the K_o value for N-acetylglycylglycine methyl ester, K_{s1} (the dissociation constant for the ES_1 complex of N-acetylglycine methyl ester) may be obtained:

$$\begin{aligned} K_{s1} &\approx 2.7 \times K_{o\ \text{Ac-glygly-OCH}_3} \\ &= 2.7 \times 2100\ \underline{\text{m}}\text{M}. \\ &= 5700\ \underline{\text{m}}\text{M}. \end{aligned} \tag{8}$$

Substitution of this value of K_{s1} into equation 6 yields a value of $k_{21} = 0.13\ \text{sec}^{-1}$. Substitution of $K_{s1} = 5700$ m$\underline{\text{M}}$ and $K_o = 300$ m$\underline{\text{M}}$ into equation 4a yields a value of $K_{s2} = 317$ m$\underline{\text{M}}$, and from this latter value and equation 7 a value of $k_{22} = 0.041\ \text{sec}^{-1}$ is obtained.

A summary of these calculated values is given below:

$$k_{21} = 0.13\ \text{sec}^{-1} \quad ; \quad K_{s1} = 5700\ \text{m}\underline{\text{M}} \quad ; \quad k_{21}/K_{s1} = 0.022 \times 10^{-3}\ \text{m}\underline{\text{M}}^{-1}\text{sec}^{-1}$$

$$k_{22} = 0.041\ \text{sec}^{-1} \quad ; \quad K_{s2} = 317\ \text{m}\underline{\text{M}} \quad ; \quad k_{22}/K_{s2} = 0.128 \times 10^{-3}\ \text{m}\underline{\text{M}}^{-1}\text{sec}^{-1}.$$

The value of K_{s1} (5700 m$\underline{\text{M}}$) obtained from this treatment is, of course, much higher than the observed K_o value (i.e., 300 m$\underline{\text{M}}$). The abnormally low K_o value for N-acetylglycine methyl ester has never been satisfactorily explained.† In a paper dealing with N-acetylated amino acid derivatives containing normal alkyl side chains, Jones observed that K_o increased with decreasing number of carbon atoms in the side chain (7). However, the K_o value (300 m$\underline{\text{M}}$) obtained for N-acetylglycine methyl ester (no carbon atoms in the side chain) is considerably less than the value (740 mM) obtained for N-acetyl-L-alanine methyl ester (7) (one carbon atom in the side chain). Jones found a correlation between the values of log K_0 and the number of side-chain carbon atoms for this series of substrates. Such a correlation is reproduced in Figure 2. Both the observed K_0 value and the calculated K_{S1} value for N-acetylglycine methyl ester have been included in this plot. The calculated K_{S1} value follows quite nicely the trend observed for the trifunctional homologues. Further support of the results of the above treatment is found in Rapp's study of the stereospecificity of α-chymotrypsin-catalyzed reactions (13). From comparison of the kinetic behavior of bifunctional substrates with that observed for the D-antipodes of trifunctional substrates, Rapp concluded that the ES_2 mode (i.e., R_1-ρ_2, R_3-ρ_3 interactions) was a major contributor to the reactivity of simple N-acylated glycine methyl esters.

---

† Wolf, Wallace, Peterson and Niemann (5) presented an explanation based on substrate activation via a ternary complex but recent calculation of the kinetic parameters for N-acetylglycine methyl ester (4) casts considerable doubt on this interpretation.

Because the $S^*_{R_1}$ values obtained from the L-valine derivatives (for the N-acetylglycyl-, N-acetylglycylglycyl- and N-acetylglycylglycylglycyl- groups) are uniformly about twice as high as those obtained from the corresponding L-alanine and L-leucine derivatives (Table IV), it is suggested that some factor (peculiar to the L-valine derivatives) is operable in the change from the N-acetyl- to the N-acetylglycyl- compound. A negative reactivity factor associated with N-acetyl-L-valine methyl ester (but not present in the corresponding L-alanine and L-valine substrates) which is somehow relieved in the N-acetylglycyl- compound, represents one such possibility. A β-branching effect has been invoked (8) to account for the low k_0 value (0.15 sec^{-1}) of N-acetyl-L-valine

EFFECT OF LENGTH OF SIDE CHAIN ON LOG K_o

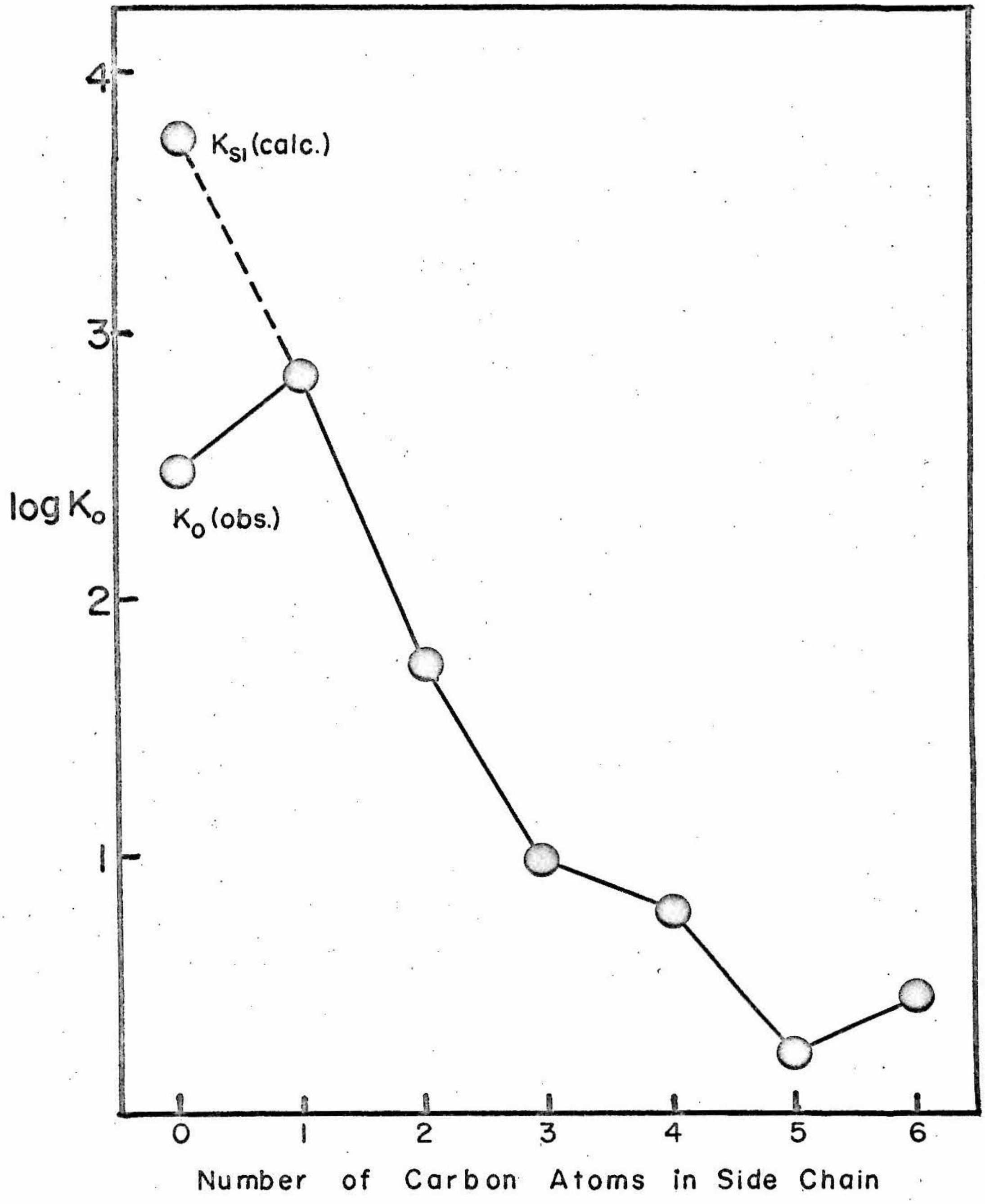
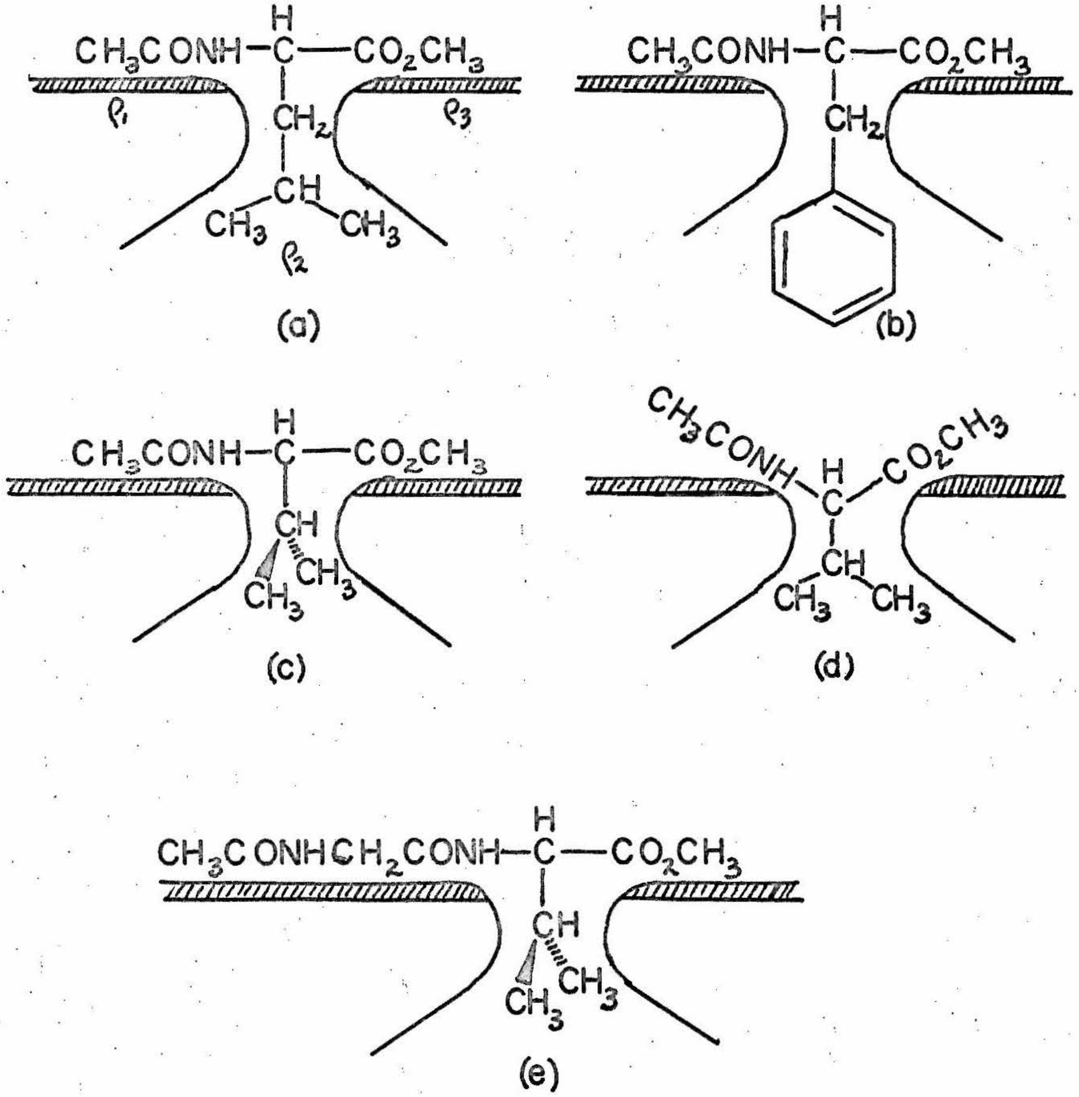


Figure 2

methyl ester and it has also been suggested that this effect produces steric hindrance to binding, thereby increasing K_o. This effect is useful in explaining the fact that N-acetyl-L-leucine methyl ester has a k_o value 30 times greater (and a K_o value 40 times lower) than the corresponding L-valine derivative (198) whereas a similar structural change (i.e., insertion of a methylene group) in the homologous unbranched side chain substrates (i.e., L-norvaline to L-norleucine) results only in a three-fold increase in k_o and a 1.5-fold decrease in K_o (7).

One of the important factors in determining the relatively low K_o value (2.9 mM (10)) of N-acetyl-L-leucine methyl ester is the similarity of its side chain to the benzyl side chain of the corresponding L-phenylalanine derivative (K_o = 0.62 mM (16)). This similarity is illustrated in schematic representations of the binding of these two substrates at the active site (Figures 3a and 3b). The steric obstructions to binding of β-branched substrate (e.g., a valine derivative) are also indicated in this set of figures. Two possible ways in which a valine derivative might be bound at the active site are illustrated in Figures 3c and 3d. In the first of these (as represented in Figure 3c) the R_1-ρ_1 and R_3-ρ_3 interactions are optimized at the expense of the R_2-ρ_2 interaction (as indicated schematically by a rotation of the β-γ carbon-carbon single bond). In the second (Figure 3d), N-acetyl-L-valine methyl ester is bound such that the R_2-ρ_2 interaction more closely resembles that of the leucine side chain with this locus. Such an interaction, however, can only be accomplished at the expense of the R_1-ρ_1 and R_3-ρ_3 interactions. It is suggested that, for N-acetyl-L-valine

SCHEMATIC REPRESENTATION OF BINDING CONFIGURATIONS


CH3CONH–C–CO2CH3
H
CH2
CH
CH3
CH3
ρ1
ρ2
ρ3
(a)
CH3CONH–C–CO2CH3
H
CH2
(b)
CH3CONH–C–CO2CH3
H
CH
CH3
CH3
(c)
CH3CONH
H
C
CO2CH3
CH
CH3
CH3
(d)
CH3CONHCH2CONH–C–CO2CH3
H
CH
CH3
CH3
(e)

a. N-acetyl-L-leucine methyl ester.
b. N-acetyl-L-phenylalanine methyl ester.
c. N-acetyl-L-valine methyl ester (energetically unfavorable).
d. N-acetyl-L-valine methyl ester (preferred).
e. N-acetylglycyl-L-valine methyl ester (preferred).

Figure 3

methyl ester, this latter combination of interactions may result in a binding mode which is energetically more favorable than that represented in Figure 3c (i.e., with optimum $R_1-\rho_1$ and $R_3-\rho_3$ interactions). This is consistent with the low k_o value and the high K_o value for N-acetyl-L-valine methyl ester. It is further suggested that because of the increased importance of the $R_1-\rho_1$ interaction, in the case of the dipeptide derivative (N-acetylglygyl-L-valine methyl ester), the preferred configuration for binding is that represented in Figure 3e. This configuration, with a diminished $R_2-\rho_2$ interaction (as compared to Figure 3d) but optimal (or close to optimal) $R_1-\rho_1$ and $R_3-\rho_3$ interactions is similar then to the less-preferred configuration (Figure 3c) for N-acetyl-L-valine methyl ester. Thus there should be an extra orientation effect on R_3 (as compared with the corresponding L-alanine and L-leucine substrates) when the N-acetyl- group of N-acetyl-L-valine methyl ester is replaced by an N-acetylglycyl- group. This should result in an increase in k_o greater than that which would be expected from the behavior of the L-alanine and L-leucine derivatives. This is what is observed. The k_o value for N-acetylglycyl-L-valine methyl ester is 2.4 times greater than that of N-acetyl-L-valine methyl ester while the increase for the corresponding L-alanine and L-leucine derivatives is only about 1.6-fold (Table III).

The direction of the effect on K_o is not quite so clear. From Table III, the increase in $\overline{K}_o$ (4.3-fold) is slightly more than would be expected on the basis of the data for the corresponding L-alanine and L-leucine derivatives (3.0- and 2.4-fold increases, respectively). This would be possible if the increased contributions to binding of the $R_1-\rho_1$ and $R_3-\rho_3$ interactions in the N-acetylglycyl- derivative more than compensated for the decreased $R_2-\rho_2$ interaction.

The low value of $S^*_{R_1}$ (for R_1 = N-acetylglycyl-) obtained from data for the L-tyrosine hydrazide substrates (Table I) is most difficult to explain. Braunholtz and Niemann (1) were certainly justified in concluding that the kinetic parameters for the two substrates studied were identical within experimental error. Although the parameters for the two substrates were determined under slightly different conditions (i.e., N-acetyl-L-tyrosine hydrazide at pH 7.9 and N-acetylglycyl-L-tyrosine hydrazide at pH 7.8) it is highly unlikely, from what is known concerning pH-effects (17,18), that this small pH difference could account for the rather large $S^*_{R_1}$ discrepancy.

An attempt has been made to treat the experimental data of this and other pertinent studies in terms of the Bender-Kézdy specificity theory. For the data available, a reasonably good correlation has been obtained and explanations for the deviations have been suggested. It is necessary, however, to realize that the above-mentioned specificity theory is only an approximate model and that its success rests on a number of assumptions (11). One of these assumptions, that is, that the binding interactions considered are completely independent of each other, although it represents a reasonably good approximation, seems to be somewhat questionable. The explanation offered for the L-valine discrepancy is, in fact, based on the proposition that one R-ρ interaction can indeed affect the others.

In conclusion, it is worthwhile mentioning that the fit obtained is very close to the quality obtained by Bender (11) for other R_1 groups.

b) Effects on k_o and K_o

1. Interpretation in terms of the simple kinetic scheme.

The data in Table III indicate that, with the exception of N-acetylglycylglycyl-L-valine methyl ester, the magnitude of k_o increases throughout the entire series studied as the length of R_1 is increased by means of insertion of glycine residues. It appears, however, from consideration of the L-leucine series, that this effect is somewhat attenuated in the tetrapeptide derivative (i.e., N-acetylglycylglycylglycyl-L-leucine methyl ester). The effectiveness of binding (as measured by $\overline{K}_o$) appears to increase up to the tripeptide derivatives but, from the data available, this effect appears to be reversed for the tetrapeptide derivatives. These effects are, of course, reflected in the previously discussed k_o/K_o ratios which indicate an increase in reactivity up to the tripeptide derivatives and an attenuation or reversal of this effect in the tetrapeptide derivatives.

The most reasonable explanation for the observed increase in both k_o and $\overline{K}_o$ is that an increase in the length of the R_1 group (in the form of an acetylated peptide chain) enhances the $R_1-\rho_1$ interaction, resulting in better binding (higher $\overline{K}_o$) and improved orientation (15) (higher k_o). This is simply a statement of the "better binding: better hydrolysis" hypothesis as expressed by Knowles (19).

Two explanations are suggested for the observed decrease in binding in the tetrapeptide derivatives (as compared with the corresponding tripeptide derivatives). The first explanation involves the postulation of a steric barrier which interferes with the N-terminal end of the peptide chain when the third glycine residue is inserted into the chain. This explanation is not very attractive for the following reasons:

1. It is well known that α-chymotrypsin catalyzes, with great efficiency, the hydrolysis of extremely long polypeptide chains (see Part I of this dissertation). In terms of the above explanation, if the steric effect were a very restrictive one, additional lengthening of the peptide chain should drastically decrease binding. Although there is a possibility that this steric barrier is a minor one and that a peptide chain of increased length would have its binding reduced no more than that observed for the tetrapeptide derivative, this explanation seems to be intrinsically unattractive.

2. From the data available, it appears that although binding does decrease (i.e., $\overline{K}_o$ decreases) k_o increases. One might expect that a steric barrier in ρ_1 which decreases binding (lowers $\overline{K}_o$) would also negatively affect orientation of R_3 in ρ_3 resulting in a lower k_o value, contrary to what is observed.

Consideration of the nature of the R_1-ρ_1 interaction may provide a clue to an alternate explanation. Neurath and Schwert (20) suggested that the R_1-ρ_1 interaction was influenced by two factors: (a) hydrogen-bonding between the peptide linkages of the R_1 group and complementary linkages in the peptide chain of the active site and (b) hydrophobic bonding between parts of the R_1 peptide chain and corresponding sub-loci on the enzyme. The results of the present study are consistent with this model and those discussed later (pp. 89-116) provide further support for it.

If indeed the two factors described above (i.e., hydrogen-bonding and hydrophobic bonding) are important in the interaction of the acylamino- function (R_1) with its complementary locus (ρ_1), then the decrease in binding observed for the tetrapeptide derivatives can be explained in terms of a "solvation" effect. It is proposed that there is a competition between a positive binding effect (due to an increased number of complementary hydrogen-bonding sites) and a negative effect (due to a relative decrease in the hydrophobic character of R_1) as R_1 is lengthened by insertion of glycine residues.

As will be more fully discussed later (p.101) there appears to be a good correlation between the free energy of transfer of an amino acid from water to a non-polar solvent (ΔG_t) and the effectiveness of binding of substrate derivatives of that amino acid to α-chymotrypsin (as measured by log K_o). The values of ΔG_t are calculated from partition coefficients (the relative solubilities of the amino acids in water and in organic solvents). It is found that binding is enhanced as the hydrophobic character of the amino acid increases (i.e., as ΔG_t decreases). As will be seen later (p.101ff.), such a treatment can be applied, with some success, to the results of a study of the α-chymotrypsin-catalyzed hydrolyses of some dipeptide derivatives.

The free energies of transfer (from water to ethanol) for glycine, glygylglycine and glygylglycylglycine have been calculated from data tabulated by Greenstein and Winitz (21) and are presented in Table V. The values of ΔG_t were calculated from the following relationship:

$$\Delta G_t = -2.303 \text{ RT} \log \frac{N_A}{N_O} \tag{9}$$

TABLE V

Free Energies of Transfer of Some Glycine Derivatives between Water and Ethanol[a]

Compound	Solubility in Water[b] log N_O	Solubility in Ethanol[b] log N_A	Relative Solubility $\log(N_A/N_O)$	ΔG_t[c] kcal.
Glycine	-1.247	-4.638	-3.391	4.64
Glycylglycine	-1.522	-5.889	-4.367	5.96
Triglycine	-2.241	-7.206	-4.965	6.78
Hydantoic acid	—	—	-0.630	0.85
Glycylglycine hydantoic acid	—	—	-1.533	2.10
Triglycine hydantoic acid	—	—	-2.253	3.08

[a] Data from reference 21.

[b] Solubilities expressed as logarithms of mole fraction solubility.

[c] Calculated as described in the test (p. 77).

where R is the universal gas constant (units of kcal.), T is the temperature (°K) and N_A and N_O represent the solubilities of the compound in ethanol and water, respectively. Solubilities were expressed in terms of mole fraction (21).

From Table V, it is apparent that the free energy of transfer from water to ethanol increases as the length of the peptide increases (indicating a decrease in hydrophobic character). A parallel behavior for uncharged derivatives is suggested by studies comparing amino acids and the amides of their corresponding α-hydroxy acids (21). In these studies it was found that the difference in ΔG_t values for two amino acids was the same as the difference in ΔG_t values for the corresponding uncharged α-hydroxy acid amides. The ΔG_t values for the hydantoic acid derivatives of glycine peptides (see Table V) also parallel those for the peptides themselves. Thus it is reasonable to assume that this type of behavior would be observed for the N-acetylated peptide methyl esters of this study.

On the basis of hydrophobic bonding alone, it would be expected that effectiveness of binding (as measured by $\overline{K}_o$) would decrease as the length of the peptide R_1 group is increased by the insertion of glycine residues, because the hydrophobic character of this group decreases. However, increasing the peptide length of the substrate also increases the opportunity for hydrogen-bonding, contributing to an increase in $\overline{K}_o$. It is suggested that there is a competition between these two opposing factors and that, although the latter factor (hydrogen-bonding) is dominant for the di- and tripeptide derivatives, for the tetrapeptide derivatives, the former factor (a decrease in the hydrophobic character of R_1) exerts its influence to such an extent that a net decrease in binding is observed. It is also possible that hydrogen-bonding becomes less important for the peptide bonds further from the carboxy-terminal end of the substrate.

The apparent constancy or slight increase in k_o for the tetrapeptide derivatives is consistent with this explanation. The decreased $\overline{K}_o$ values reflect only a shift in the equilibrium from complex to the free components in water while the increased opportunity for hydrogen-bonding permits a favorable orientation of the cleavable bond when the complex is formed. Thus, no decrease of k_o need be expected.

The anomalous behavior of the k_o and K_o (or $\overline{K}_o$) parameters for N-acetylglycylglycyl-L-valine methyl ester and its corresponding tetrapeptide derivative is extremely difficult to explain. On the basis of the results for the corresponding L-alanine and L-leucine derivatives (Table III), the k_o values for the tri- and tetrapeptide-L-valine derivatives should be about twice that for the corresponding dipeptide derivative (N-acetylglycyl-L-valine methyl ester). Experimentally (see Table II) the k_o value for the tripeptide L-valine derivative is only one-half that for the dipeptide derivative and the value for the tetrapeptide derivative is only 1.2 times that for the dipeptide derivatives. Similarly, the $\overline{K}_o$ values of the tri- and tetrapeptide L-valine derivatives should be about 2 and 1.5 times greater, respectively, than $\overline{K}_o$ for the dipeptide derivative. The $\overline{K}_o$ values are, in fact, 7 and 2.3 times greater, respectively.

As noted before, the changes in k_o/K_o for the L-valine di-, tri- and tetrapeptide derivatives closely parallel similar changes in this parameter for the corresponding L-alanine and L-leucine substrates. Any explanation for the observed "abnormal" k_o and K_o behavior of the L-valine derivatives must thus be constant with the "regular" behavior of the k_o/K_o parameter. The case of the L-valine tripeptide derivative may be taken as an example. In comparison with the corresponding L-alanine and L-leucine derivatives, the k_o and K_o values are both abnormally low. This behavior might be explained by an increased stability of the productive enzyme-substrate complex and a corresponding increase in the activation energy required in the step following complex formation. Such an explanation would require that the total free energy change required to form this transition state (for the reaction step immediately following complex formation) from the free enzyme and substrate in solution be exactly as predicted from analogy with the L-alanine and L-leucine data. There is no a priori reason why this should be so.

The observed k_o, K_o and k_o/K_o behaviors are, however, entirely consistent with increased importance of non-productive binding for the tri- and tetrapeptide L-valine derivatives. Non-productive binding is often invoked to explain results where an increase in binding (as indicated by an increase in $\overline{K}_o$) is accompanied by a decrease in k_o (14,15). That this situation might prevail in this particular case is demonstrated in a somewhat more quantitative manner by the following calculations.

The following assumptions will be made:

1. N-acetylglycyl-L-valine methyl ester and all of the L-alanine and L-leucine derivatives of this study are bound exclusively in productive modes.

2. The tri- and tetrapeptide L-valine derivatives are bound to some extent in non-productive modes.

3. Changes in the kinetic parameters of the productive modes of these latter two substrates parallel similar changes for the corresponding L-alanine and L-leucine derivatives.

Based on these assumptions, a measure of the amount of non-productive binding important in the two substrates in question can be calculated. Values of K_o can also be calculated as a check on the validity of this treatment.

The equations describing the kinetic parameters k_o and K_o for the case of two modes of binding (one productive and the other non-productive) were presented on page 4 of the General Introduction. They are as follows:

$$k_o = k_2 \frac{K_{sI}}{K_s + K_{sI}} \tag{10}$$

and

$$K_o = K_s \frac{K_{sI}}{K_s + K_{sI}} \tag{11}$$

where k_2 is the specific rate of product formation from the productive complex and K_s and K_{sI} are the dissociation constants for the productive and non-productive complexes, respectively.

In terms of the above assumptions, $k_o = k_2 = 0.36\ \text{sec}^{-1}$ and $K_o = K_s = 26$ mM for N-acetylglycyl-L-valine methyl ester (Table II). From the L-alanine and L-leucine data (Tables II and III), k_2 for the productive mode of N-acetylglycylglycyl-L-valine methyl ester should be about 1.9 times greater than that for the corresponding dipeptide derivative. Therefore, for N-acetylglycylglycyl-L-valine methyl ester, $k_2 \approx 1.9\ (0.36) = 0.68\ \text{sec}^{-1}$. Substitution of this value of k_2 and the observed k_o value (0.16 sec^{-1}) into equation 10 gives:

$$\frac{K_{sI}}{K_s + K_{sI}} = 0.24 \qquad (12a) \qquad K_{sI} \simeq 0.3\ K_s \qquad (12b)$$

By analogy with the L-alanine and L-leucine data, K_s for the tripeptide L-valine substrate should be approximately 0.5 times that of the corresponding dipeptide derivative. Therefore $K_s \approx 0.5\ (26) = 13$ mM. Substitution of this K_s value and equation 12a into equation 11 yields $K_o = (0.24)\ (13) = 3.1$ mM, a value which compares quite favorably with the observed K_o value (3.7 ± 0.4 mM).

Using the same approach for N-acetylglycylglycylglycyl-L-valine methyl ester and the data available from the corresponding L-leucine substrates (Table II) the following values are calculated: $K_{sI}/(K_s + K_{sI}) = 0.60$; $K_{sI} = 1.5\ K_s$; K_o (calculated) = 11 mM. This calculated K_o value compares very well with the observed K_o value of 11.3 ± 0.4 mM.

The observed k_o, K_o and k_o/K_o behaviors are thus entirely consistent with some kind of non-productive binding which is: (a) not important for N-acetylglycyl-L-valine methyl ester, (b) very important for the tripeptide derivative (i.e., $K_{sI} = 0.3\ K_s$) and (c) moderately important for the tetrapeptide derivative (i.e., $K_{sI} = 1.5\ K_s$). This consistency is indicated by the excellent agreement between the calculated and observed K_o values. However, no reasonable model for non-productive binding (which would not apply equally well to the L-alanine and L-leucine substrates) can be put forward on the basis of the available data. Although the explanation offered above is consistent with the data, only further experiments will permit definitive conclusions regarding the observed behavior of the kinetic parameters for these L-valine derivatives.

2. Interpretation in terms of the acyl-enzyme kinetic scheme.

It will be shown below that the data presented in this thesis and that available in the literature are not extensive enough or of the proper type to permit any meaningful interpretation of the structure-reactivity effects observed in the present study in terms of the acyl-enzyme kinetic scheme.

In terms of the simple kinetic scheme, the two kinetic parameters k_o and K_o fairly well define the kinetic course of an α-chymotrypsin-catalyzed hydrolysis (see p. 3). On the other hand, the acyl-enzyme kinetic scheme (see p. 5) results in considerable ambiguity with regard to the interpretation of k_o and K_o and renders such interpretation much more difficult. In general, variation of k_o depends on two difficult-to-determine parameters k_2 and k_3, while K_o is a function of k_2, k_3 and K_s (see p. 6). For example, an observed increase in k_o and decrease in K_o can be interpreted, in terms of the simple scheme, as the result of enhanced binding and concomitant improvement in the orientation of the hydrolyzable function. In terms of the acyl-enzyme scheme, however, this observed behavior can arise in several ways, as indicated by the following:

1. K_s constant, k_3 invariant and k_2 increasing.
2. K_s constant, both k_2 and k_3 increasing with k_2 increasing more than k_3.

3. K_s decreasing with:
   (a) k_2 or k_3 decreasing while the other increases,
   (b) k_2 and k_3 increasing,
   or (c) k_2 or k_3 invariant while the other increases.

Thus, it is obvious that, in addition to the evaluation of k_o and K_o, some knowledge is required of the variation of the parameters k_2, k_3 and K_s with structural changes for an adequate treatment of a structure-reactivity study such as that presented in this dissertation. Such knowledge is possible only in certain cases. In the case of limit-type substrates, in which either acylation or deacylation is definitely the rate-determining step, some of the ambiguities mentioned above become unimportant. Also, under some conditions, certain esters (e.g., the p-nitrophenyl ester) permit the direct determination of k_2 and k_3. Experiments of this type are, however, fraught with experimental difficulties.

The proponents of the acyl-enzyme hypothesis suggest (22) that the above limit cases are represented by two classes of substrates. On the one hand, a primary amide of an acylated amino acid will have acylation as its rate-determining step. This case is kinetically indistinguishable from the simple scheme. On the other hand, the methyl esters of acylated aromatic amino acids (tryptophan, tyrosine and phenylalanine) will have rates determined by the deacylation step. Even in this limit case, in which $k_o = k_3$ and $K_o = (k_3 K_s)/k_2$, a knowledge of the magnitude of either k_2 or K_s is necessary for any complete description of a structure-reactivity relationship. Moreover, deacylation is not completely rate-determining for methyl esters of all N-acylated amino acids. In the case of N-acetylglycine methyl ester, calculations have been performed using the experimentally obtained values of k_2 and k_3 for the corresponding p-nitrophenyl ester (22). The results of these calculations indicate that, for N-acetylglycine methyl ester, k_2 is considerably less than k_3 and thus acylation is the rate-determining step. As the methyl esters of N-acetyl-L-alanine, N-acetyl-L-valine and N-acetyl-L-leucine are intermediate in reactivity between the corresponding glycine and aromatic amino acid derivatives, it is suggested that, for the former derivatives, neither step will be completely rate-determining and k_2 and k_3 will probably be of the same order of magnitude.

A very rough estimate of the relative magnitudes of k_2 and k_3 for some of these substrates can be obtained from the kinetic parameters for two esters of the same acylated amino acid. One must assume, however, that changes in K_s for a series of different esters of the same acylated amino acid are insignificant (i.e., that R_3 does not contribute to binding).[†]

Consider the methyl and β-chloroethyl esters of N-acetyl-L-norvaline. The kinetic data for these two substrates are given below:

Substrate	k_o (sec^{-1})	K_o (mM)
N-Acetyl-L-norvaline methyl ester (23)	2.70	10.2
N-Acetyl-L-norvaline β-chloroethyl ester (23)	3.72	0.79

[†] This assumption (an important one in the Bender-Kézdy specificity treatment (11)) is a tenuous one. In a later part of this dissertation, evidence is put forward which tends to discredit it. Therefore the results obtained on the basis of this assumption must be viewed with caution.

Now,

$$K_o = \frac{k_3}{k_2 + k_3} K_s \qquad (13a) \qquad\qquad K'_o = \frac{k'_3}{k'_2 + k'_3} K'_s \qquad (13b)$$

where the unprimed symbols refer to the methyl ester and the primed ones to the β-chloroethyl ester. From the above-mentioned assumption, $K_s = K'_s$. Also, because the acyl-enzyme intermediate is the same in both cases, $k_3 = k'_3$. From the above data and equations 13a and 13b, the following relationship is obtained:

$$0.8\ k'_2 - 10.2\ k_2 = 9.4\ k_3 \qquad (14)$$

Because k_2, k'_2 and k_3 are all positive numbers, it is obvious (from equation 14) that $k'_2 > 12\ k_3$. This indicates that deacylation is certainly the rate-determining step for N-acetyl-L-norvaline β-chloroethyl ester. Thus, we can assume that $k_3 \approx k'_o = 3.72\ \text{sec}^{-1}$. Now, for the methyl ester,

$$k_o = \frac{k_2\ k_3}{k_2 + k_3} \qquad (15)$$

and substituting $k_o = 2.70\ \text{sec}^{-1}$ and $k_3 = 3.72\ \text{sec}^{-1}$ into equation 15, one obtains $k_2 = 10\ \text{sec}^{-1}$. Thus this treatment indicates that $k_2 \approx 2.6\ k_3$ for N-acetyl-L-norvaline methyl ester, assuring that, for this substrate, neither step can be considered completely rate-determining.

Similar data for esters of N-acetyl-L-leucine are not available. However, N-acetyl-L-leucine methyl ester ($k_o = 4.6\ \text{sec}^{-1}$, $K_o = 2.9$ mM (10)) is of the same order of reactivity as N-acetyl-L-norvaline methyl ester and it is not unreasonable to assume that, for the former compound, k_2 and k_3 will also be of the same order of magnitude.

A similar treatment of data for the methyl and β-chloroethyl esters of N-acetyl-L-valine (8) yields the following values for N-acetyl-L-valine methyl ester: $k_2 = 0.45$ sec^{-1}, $k_3 = 0.23$ sec^{-1} and thus $k_2 \simeq 2\ k_3$.

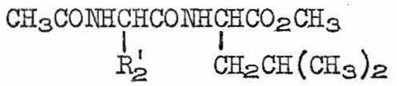
Similarly, from data for the methyl and glycolamide esters of N-acetyl-L-alanine (see Part III of this dissertation) the following values of the parameters for N-acetyl-L-alanine methyl ester are obtained: $k_2 = 1.95$ sec^{-1}, $k_3 = 3.55$ sec^{-1} and thus $k_2 \approx 0.5\ k_3$.

Therefore, if the assumption on which the above calculations are based is correct, it is obvious that for the N-acetylated methyl esters of L-alanine, L-valine and probably L-leucine, neither acylation nor deacylation may be considered rate-determining. Thus, a knowledge of the variation of both k_2 and k_3 with modification of structure (changing R_1) would be required for any significant interpretation of the results of this study in terms of the acyl-enzyme kinetic scheme. These data is not available. No explanation of the anomalous k_o and K_o values for the tri- and tetrapeptide L-valine derivatives in terms of suitable variation of k_2 and k_3 is any more reasonable than the explanation offered on the basis of the simple kinetic scheme (p. 81). A decrease in k_3 from the dipeptide to the tripeptide case (accompanying an expected k_2 increase) would account for the observed values. However, as will be discussed later (p. 118), there is no reason to expect that a given structural change should affect acylation and deacylation in opposite ways. On the contrary, a parallel behavior is expected. Even then, it is not clear why this situation should apply only to the L-valine derivatives and not to the L-alanine- and L-leucine- derivatives.

SECTION B:

N-Acetyl-aminoacyl-L-leucine Methyl Esters

In the preceding section, the α-chymotrypsin-catalyzed hydrolyses of some N-acetylated peptide methyl esters were discussed. In those substrates, peptide length was varied by means of insertion of glycine residues into the peptide chain. Attention is now focussed on a series of N-acetylated dipeptide esters of the form:



$$CH_3CONH\underset{|\atop R_2'}{CH}CONH\underset{|\atop CH_2CH(CH_3)_2}{CH}CO_2CH_3$$

II

where the carboxy-terminal residue is L-leucine and the amino-terminal residue is varied throughout a series of amino acids. The side chain of this amino-terminal residue is represented as $-R_2'$. The α-chymotrypsin-catalyzed hydrolyses of these substrates were studied and the effect of the amino-terminal residue (i.e., the R_2' side chain) on the kinetic parameters obtained will now be discussed.

Data

The N-acetyl-aminoacyl-L-leucine methyl esters were chosen as the model series for a number of reasons. The methyl ester was chosen for reasons discussed in the previous section (p. 56) in order that undesired hydrolysis at the peptide bond would be negligible with respect to cleavage at the ester linkage. Derivatives of L-leucine were chosen because they represent a relatively good approximation (in re-

TABLE VI

The α-Chymotrypsin-catalyzed Hydrolyses of Several N-Acetyl-*L*-aminoacyl-*L*-leucine Methyl Esters[a]

Aminoacyl Component[b]	$[S]_o$ m*M*	$[E]_o$[c] 10^{-8} *M*	No. of Expts.[d]	Max. % Reaction[e]	k_o[c] sec^{-1}	K_o m*M*	k_o/K_o m*M*$^{-1}$ sec^{-1}
glycine	0.286 - 2.571	4.0	11-0	4	7.7 ± 0.5	1.2 ± 0.2	6.5 ± 0.8
L-alanine	0.147 - 1.324	4.18	12-0	5.6	6.7 ± 0.5	0.81 ± 0.12	7.8 ± 0.9
β-alanine	0.265 - 2.381	8.44	10-0	3.6	4.8 ± 0.5	2.2 ± 0.4	2.2 ± 0.3
L-α-aminobutyric acid	0.095 - 0.758	4.12	10-1	14	6.0 ± 0.4	0.31 ± 0.04	19 ± 2
L-valine	0.157 - 1.411	3.84	14-0	7	3.4 ± 0.2	0.26 ± 0.05	13 ± 3
L-norvaline	0.068 - 0.542	3.67	10-1	15	4.8 ± 0.1	0.11 ± 0.02	44 ± 8
L-leucine	0.059 - 0.474	5.64	11-0	14	3.2 ± 0.2	0.18 ± 0.04	18 ± 4
L-norleucine	0.079 - 0.634	4.10	13-0	20	6.6 ± 0.2	0.17 ± 0.02	39 ± 6
L-isoleucine	0.050 - 0.450	9.36	11-0	19	3.4 ± 0.3	0.30 ± 0.04	11 ± 1
L-phenylalanine	0.020 - 0.163	9.90	12-0	26	1.8 ± 0.1	0.078 ± 0.010	23 ± 2
L-proline	0.101 - 0.807	3.92	11-1	12	8.2 ± 0.6	0.46 ± 0.06	18 ± 2
L-aspartic acid	1.317 -10.456	224	10-0	2.6	1.1 ± 0.1	20 ± 3	0.055 ± 0.006
L-phenylalanine (amide)[f]	0.067 - 0.263	400	4	15	——	——	0.11 ± 0.01

TABLE VI--continued

[a] In aqueous solutions at 25.0°, pH 7.90 ± 0.10, 0.10 M in sodium chloride.

[b] The aminoacyl component refers to the amino-terminal residue of the dipeptide derivative. The carboxy-terminal residue is, in all cases, L-leucine.

[c] Based on an enzyme molecular weight of 25,000 and a protein-nitrogen content of 16.5%.

[d] First number refers to number of kinetic runs performed; second number refers to the number of points rejected by a statistical reiterative procedure.

[e] The extent of reaction (i.e., the percentage of substrate hydrolyzed) over a 4 minute period for the lowest initial substrate concentration.

[f] Parameters and conditions for N-acetyl-L-phenylalanyl-L-leucine amide (see text).

activity) to the "specific" substrates of α-chymotrypsin (i.e., derivatives of L-tryptophan, L-tyrosine and L-phenylalanine).†

The kinetic parameters and pertinent experimental conditions for the α-chymotrypsin-catalyzed hydrolyses of 13 N-acetyl-aminoacyl-L-leucine derivatives are presented in Table VI.

It is well known that alkyl esters of N-acylated amino acids are more reactive than the corresponding amides in α-chymotrypsin-catalyzed hydrolyses. For example, the observed k_o/K_o values for N-acetyl-L-phenylalanine methyl ester and N-acetyl-L-phenylalanine amide are approximately 3.8×10^4 $\underline{M}^{-1}\mathrm{sec}^{-1}$ and 1.0 $\underline{M}^{-1}\mathrm{sec}^{-1}$, respectively (24). Thus, for the class of substrates considered in this study, it was expected that hydrolysis would take place at the ester linkage at such a rate that, during the time this reaction was followed, hydrolysis of the peptide bond would be negligible. This was confirmed by the experiment described below.

Because α-chymotrypsin has its highest primary specificity for peptide bonds involving aromatic amino acids as the carboxyl component (see Part I of this dissertation) it is reasonable to assume that if hydrolysis of peptide bonds in the N-acylated dipeptide esters were to be significant, it should be observed to the greatest extent in N-acetyl-L-phenylalanyl-L-leucine methyl ester. As a test of this, the corresponding amide substrate was synthesized and the following kinetic parameter for N-acetyl-L-phenylalanyl-L-leucine amide was determined: k_o/K_o =

---

† It would have been impossible to accurately determine kinetic parameters for the corresponding dipeptide ester derivatives of the aromatic amino acids, as their reactivity would have exceeded the limits of the experimental method employed (the pH-stat).

0.11 ± 0.01 $\underline{\text{mM}}^{-1}\text{sec}^{-1}$ (Table VI). The initial substrate concentrations permitted by the limited solubility of this substrate were such that the kinetics over the range of substrate concentrations studied were very close to being first order in substrate concentration and thus an accurate determination of the individual parameters, k_o and K_o, was impossible. A very rough estimate of approximately 1 m$\underline{\text{M}}$ for the value of K_o was obtained, but this is certainly no more than an order-of-magnitude approximation. However, it was possible to obtain a relatively accurate value of k_o/K_o, as indicated above. Thus it is evident that N-acetyl-$\underline{\text{L}}$-phenylalanyl-$\underline{\text{L}}$-leucine amide (k_o/K_o = 0.11 $\underline{\text{mM}}^{-1}\text{sec}^{-1}$) is about 200 times less reactive than N-acetyl-$\underline{\text{L}}$-phenylalanyl-$\underline{\text{L}}$-leucine methyl ester (k_o/K_o= 23 $\underline{\text{mM}}^{-1}\text{sec}^{-1}$). From this result it may therefore be concluded that hydrolysis of peptide bonds in the dipeptide ester derivatives was indeed negligible.

The possibility of hydrolysis at the primary amide bond of N-acetyl-$\underline{\text{L}}$-phenylalanyl-$\underline{\text{L}}$-leucine amide has not been excluded. Indeed, the analysis of the kinetic data was based on the assumption that only hydrolysis at the phenylalanylleucine peptide bond occurred. Because of the higher primary specificity of phenylalanyl linkages as compared to leucyl linkages (see Part I of this thesis) this is a reasonable assumption. In any case, the added possibility of primary amide hydrolysis would be included in the kinetic determination. The kinetic method employed (the pH-stat) measures only the total reaction (i.e., the sum of both hydrolyses, if both occur) and therefore, if primary amide hydrolysis were to occur, the implication would be that the peptide bond hydrolysis was even slower than the measured parameter indicates.

Non-productive Binding

The most striking trend observed for the data presented in Table VI is the apparent decrease in k_o as K_o decreases. This is illustrated in Figure 4, a plot of k_o vs. K_o for the model substrates studied. This behavior at least superficially indicates that as the binding in the enzyme-substrate complex is improved, the ability of this complex to proceed to products is impaired. This relationship is contrary to the "better-binding: better reaction" hypothesis (as expressed by Knowles (19)) which holds in many cases, especially for a similar series of N-acetyl-L-amino acid methyl esters (19).

The above phenomenon (both k_o and K_o decreasing) has been observed before (15) with other structural variations but the results of the present study represent one of the most extensive presentations of such data. A most obvious interpretation of the present data involves consideration of non-productive binding, a phenomenon which can at least qualitatively produce the observed effects.

A "rotational" type of non-productive binding has been invoked to explain the reduction in both k_o and K_o when, for example, an N-benzoyl group is substituted for the N-acetyl- group in N-acetyl-L-alanine methyl ester (see Part III of this dissertation). "Rotational" non-productive binding is the type proposed originally in the Hein-Niemann specificity theory (see p. 10). It is, however, not the sort of non-productive binding which appears important for the substrates of Table VI.

The most likely candidate for a "rotational" non-productive mode would be one involving an R_1-ρ_2 interaction. However, as shown in the preceding section, replacement of the N-acetyl- group (R_1)

k_o as a Function of K_o for the α-Chymotrypsin-catalyzed Hydrolyses of Some N-Acetyl-aminoacyl-L-leucine Methyl Esters

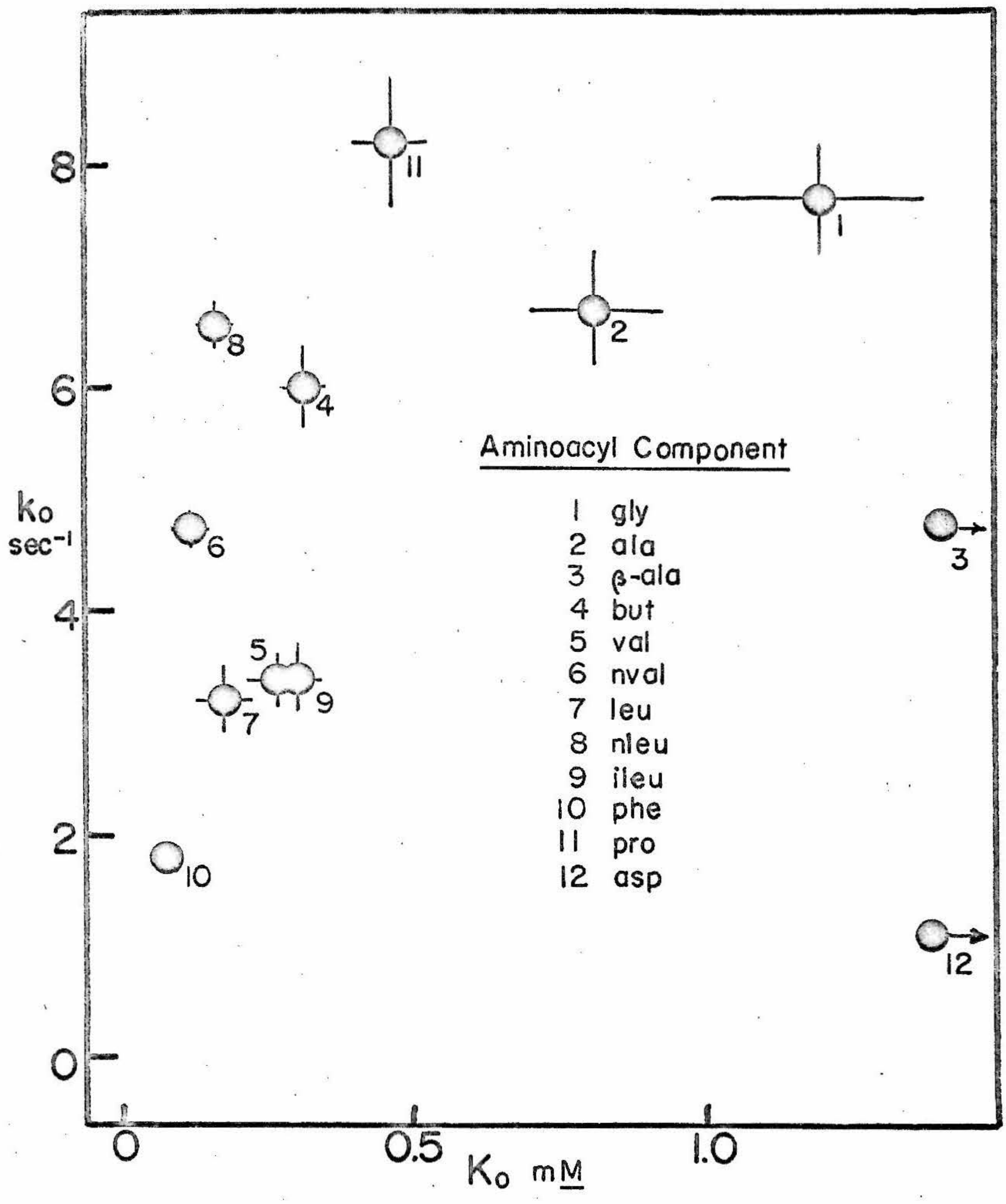
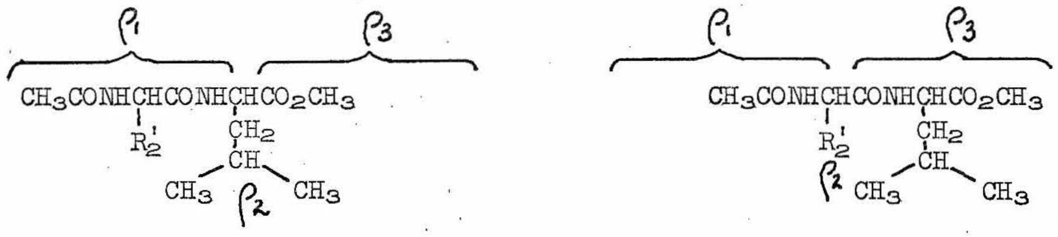


Figure 4

of N-acetyl-L-amino acid methyl esters by an N-acetylglycyl- group results in a definite increase in k_o with a decrease in K_o. If the N-acetylglycyl- group had an affinity for the ρ_2 locus, such that interaction of the R_1 group at this locus resulted in a favorable non-productive binding mode, the k_o value would have decreased, contrary to what is observed. Other considerations (see p. 66) indicate that $R_1-\rho_2$ interaction, where $R_1 = CH_3CONHCHR_2'CONH-$, is not particularly favorable. Therefore a different type of non-productive binding (hereafter referred to as "translational" non-productive binding) will be proposed for the substrates of the present study. In these terms, for dipeptide ester derivatives, the productive mode will involve interaction of the side chain of the carboxy-terminal residue (L-leucine) at ρ_2, whereas the non-productive mode will involve "translation" of the substrate molecule in the active site so that the side chain of the amino-terminal residue (R_2') interacts with this locus. Thus, the productive mode will be characterized by $CH_3CONHCHR_2'CONH-\rho_1$, $CH_2CH(CH_3)_2-\rho_2$ and $CO_2CH_3-\rho_3$ interactions and the non-productive mode by $CH_3CONH-\rho_1$, $R_2'-\rho_2$ and $\mathrm{CONH\underset{\overset{|}{CH_2CH(CH_3)_2}}{C}HCO_2CH_3}-\rho_3$ interactions. These two binding modes are schematically represented in Figure 5:



Productive Mode

Non-productive Mode

Figure 5

Because of the extremely low reactivity of the peptide bond (as compared to the ester linkage) in these substrates, a mode in which a peptide linkage interacts at the active center of ρ_3 may indeed be considered non-productive.

Although non-productive binding may not be the only factor involved in the observed behavior of k_o and K_o, as will be discussed later, it is suggested that it is a very important factor, and the results will be discussed in terms of this hypothesis. With the introduction of certain assumptions, outlined below, non-productive binding may be discussed in more quantitative terms.

As described in the General Introduction of this dissertation, the observed kinetic parameters (based on the simple kinetic scheme) for a case in which one productive mode and one non-productive mode are considered, are given by the following equations:

$$k_o = k_2 \frac{K_{sI}}{K_s + K_{sI}} \tag{10}$$

$$K_o = K_s \frac{K_{sI}}{K_s + K_{sI}} \tag{11}$$

where K_s and K_{sI} represent the dissociation constants of the productive and non-productive complexes, respectively, and k_2 represents the specific rate with which the productive complex proceeds to products.

The relative effectiveness of productive and non-productive binding can be represented by x in the following equation:

$$x = \frac{K_{sI}}{K_s} \tag{16}$$

Then, from equations 10 and 11:

$$k_o = \frac{x}{1 + x} k_2 \tag{17}$$

and

$$K_o = \frac{x}{1 + x} K_s \tag{18}$$

Therefore, if k_2 is known in addition to the experimentally determined parameters (k_o and K_o) then x may be determined from equation 17. Then, K_s and K_{sI} may be obtained from x and equations 18 and 16, respectively.

a) Estimation of k_2

It is quite obvious that experimental values of k_2 for the substrates in question are not available. Therefore, some estimation of k_2 values is required.

The first assumption that will be made is that k_2 is essentially invariant throughout the series of substrates considered. Yamashita (25) has determined the kinetic parameters for some glycyl-aminoacyl-*L*-tyrosine amides. For such substrates, because of inherently large binding ability of the tyrosine side chain, it is reasonable to assume that non-productive binding is absent (or at least relatively unimportant) and that the k_o values represent true k_2 values.[†] It is observed that in 4 of the 5 substrates studied, k_o is essentially invariant: k_o = 9.3, 8.3, 11.1 and 11.7 sec^{-1} for the cases where the aminoacyl component is *L*-alanine, *L*-valine, *L*-leucine and *L*-norleucine, respectively (25).

[†] It should also be noted that even in terms of the acyl-enzyme kinetic scheme, acylation is undoubtedly rate-determining (i.e., $k_2 \ll k_3$) for these amide substrates. Therefore, in the absence of non-productive binding, k_o truly equals k_2.

The results of the non-productive binding treatment will show that K_s (the dissociation constant for the productive binding mode) does not vary greatly throughout the series and thus, if the "better-binding: better reaction" hypothesis (19) holds for the productive mode, little variation of k_2 would be expected. The above assumption regarding k_2 is admittedly not very vigorously supported but it seems to be a good enough approximation for the rather rough conclusions that will be based on it.

The highest value of k_o obtained for the dipeptide series (Table VI) was $k_o = 8.2 \pm 0.6$ sec^{-1} for N-acetyl-L-prolyl-L-leucine methyl ester. Also, $k_o = 7.7 \pm 0.5$ sec^{-1} for N-acetylglycyl-L-leucine methyl ester. A reasonable choice of k_2 would be the highest value of k_o obtained (indicating the absence of non-productive binding). Because of the low affinity of the glycine "side chain" (i.e., $R_2' = H-$) for ρ_2, it seems reasonable to assume that non-productive binding will be relatively unimportant for N-acetylglycyl-L-leucine methyl ester. Therefore, from the above data, a value of $k_2 = 8.0$ sec^{-1} was chosen.

b) Calculation of K_s and K_{sI} Values

Using a constant value of k_2 (i.e., 8.0 sec^{-1}) K_s and K_{sI} values for several substrates were calculated from the data of Table VI and equations 16, 17 and 18, so that appropriate comparisons could be made. These values are presented in Table VII. The corresponding errors in the calculated K_s and K_{sI} values were obtained from the following equations (based on the method of calculation) (26):

TABLE VII

Dissociation Constants for the Productive and Non-productive Modes of Some N-Acetyl-aminoacyl-*L*-leucine Methyl Esters[a]

Substrate	K_s[b,c]	log K_s	K_{sI}[b,d]	log K_{sI}
Aminoacyl Component				
glycine	1.3 ± 0.2	0.11	32 ± 62	1.51
L-alanine	1.0 ± 0.15	0.00	5.0 ± 2.8	0.70
β-alanine	3.7 ± 0.8	0.57	5.5 ± 2.3	0.74
L-α-aminobutyric acid	0.41 ± 0.06	-0.39	1.25 ± 0.42	0.10
L-valine	0.68 ± 0.11	-0.17	0.68 ± 0.14	-0.17
L-norvaline	0.19 ± 0.03	-0.72	0.29 ± 0.06	-0.54
L-leucine	0.45 ± 0.10	-0.35	0.30 ± 0.09	-0.52
L-norleucine	0.21 ± 0.03	-0.68	0.98 ± 0.28	-0.01
L-isoleucine	0.70 ± 0.11	-0.15	0.52 ± 0.14	-0.28
L-phenylalanine	0.36 ± 0.05	-0.44	0.10 ± 0.02	-1.00
L-aspartic acid	145 ± 25	2.18	23 ± 6	1.36
L-proline	~ 0.46 ± 0.06[e]	-0.34	very large[e]	

[a] Data calculated from data of Table VI in the manner described in the text.

[b] Errors estimated as described in the text.

[c] Dissociation constant of productive mode in units of *mM*.

[d] Dissociation constant of non-productive mode in units of *mM*.

[e] Calculated on the basis of no non-productive binding.

$$K_s \pm \Delta K_s = \frac{K_o \pm \Delta K_o}{A \pm \Delta A} = \frac{K_o}{A} \pm \frac{1}{A}\left[(\Delta K_o)^2 + (K_s)^2(\Delta A)^2\right]^{\frac{1}{2}} \qquad (19)$$

and $$K_{sI} \pm \Delta K_{sI} = (x \pm \Delta x)(K_s \pm \Delta K_s) = x K_s \pm (K_s \Delta x + x \Delta K_s) \qquad (20)$$

where x, Δx, A and ΔA are given by:

$$x \pm \Delta x = \frac{k_o \pm \Delta k_o}{(k_2 - k_o) \pm \Delta k_o} = \frac{k_o}{k_2 - k_o} \pm \frac{\Delta k_o}{k_2 - k_o}(1 + x^2)^{\frac{1}{2}} \qquad (21)$$

and $$A \pm \Delta A = \frac{x + \Delta x}{(1 + x) \pm \Delta x} = \frac{k_o}{k_2} \pm \frac{\Delta k_o}{k_2} \qquad (22)$$

These errors represent a minimum error based on the standard deviations for k_o and K_o and do not take into account errors involved in the estimation of $k_2 = 8.0 \text{ sec}^{-1}$.

c) Discussion of K_s and K_{sI} Values

The experimental and calculated data of the study described herein may be conveniently and fruitfully treated in terms of linear free energy relationships. Such an approach will permit an evaluation of the types of binding involved in the interactions of structural components of the substrates at the two binding loci, ρ_1 and ρ_2.

With such a treatment, Knowles (19) has succeeded in correlating the free energy of enzyme-substrate binding (as measured by log K_o) with the hydrophobic character of the amino acid side chains for a series of N-acetylated L-amino acid methyl esters. From this correlation, he concluded that the predominant and most important interaction in the binding of such substrates to α-chymotrypsin is of a hydrophobic nature.

A plot employed by Knowles is reproduced in Figure 6. As a measure of the hydrophobic character of an amino acid side chain, Knowles employed the parameter ΔG_t, the free energy of transfer of an amino acid from water to a non-polar solvent (see p. 77). The ΔG_t values (for three different solvents) were calculated from relative solubility data[†] with the use of equation 9 (p. 77). The K_o data is that summarized in reference 19 with one addition: the calculated value of log K_{sI} obtained for N-acetylglycine methyl ester (p. 68) is also included and, at least in the case of the ethanol plot, this value corresponds more closely to the line obtained than does that obtained from the observed K_o.

In the linear free energy correlations employed in this dissertation (and indeed in those employed by Knowles (19)), there are two implicit assumptions:

1. The overall free energy of transfer (ΔG_t) of an amino acid (or amino acid derivative) is considered to be the sum of independent contributions (to this free energy) by the structural components of that molecule. This assumption is supported by the results of relative solubility studies with various derivatives of the amino acids (e.g., hydantoic acids, hydantions, α-hydroxy acid amides, etc. (21)). In these studies, it was found that differences in ΔG_t values for two amino acids or two amino acid derivatives (in which the only structural

---

[†] See references 19 and 21 for original sources. The ΔG_t (ethanol) value for proline, employed in later comparisons, was calculated by this author from the solubilities of L-proline in water (21) and ethanol (27), respectively. A value of ΔG_t (ethanol) = 2.1 kcal. was obtained.

Correlation of the Hydrophobic Character of the Side Chains of Amino Acids and K_o Values for the α-Chymotrypsin-catalyzed Hydrolyses of the Corresponding N-Acetyl-L-Amino Acid Methyl Esters

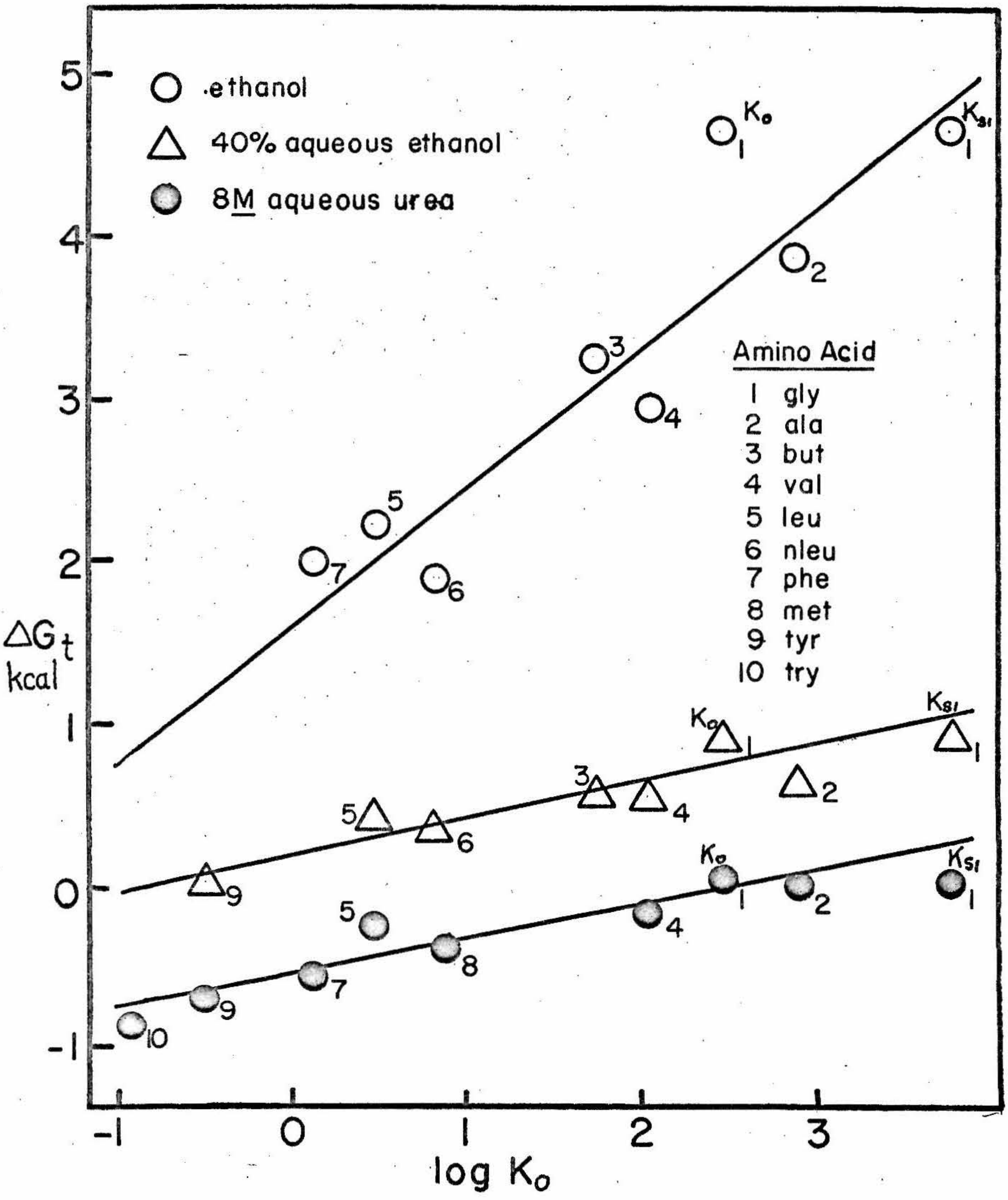


Figure 6

variation was in the amino acid side chain) were independent of the nature of the derivative employed and were a function only of the nature of the side chain. Because of this, ΔG_t values for free amino acids can be employed in correlations with other parameters (e.g., log K_o) for some derivative of those amino acids.

2. The total free energy of binding in an enzyme-substrate complex is assumed to be represented as the sum of independent contributions by the free energies associated with the R-ρ interactions present in this complex. The validity of this assumption is supported by the results of the Hamilton treatment of specificity (28).

Thus, a linear free energy correlation between ΔG_t values (for free amino acids) and log K_o values (for a derivative of those amino acids) compares the hydrophobic character of the amino acid side chains with the strength of the interaction of those side chains with a complementary locus at the active site. Figure 6 is an example of such a correlation. The slope of the line obtained in such a correlation is a measure of the partitioning of the side chain in the systems "solvent-water" and "enzyme locus-water" and is thus a measure of the hydrophobic character of the side chain - locus interaction. For any such line, if the slope is greater than unity, the enzyme locus is less selective (or less hydrophobic) than the solvent corresponding to the line. If the slope is less than unity, then the locus is more selective (or more hydrophobic) than that solvent.

From Figure 6 it appears that the enzyme locus which interacts with the side chain of N-acetylated amino acid methyl esters (i.e., ρ_2) is only slightly less polar than ethanol (slope $\approx$ 0.9) but much less polar than 40% ethanol (slope = 0.15) or 8 M aqueous urea (slope = 0.17). For all further comparisons, the values of ΔG_t employed will be those involved in transfer from water to ethanol.

One apparent trend in the results of the present study is that K_o for the N-acetylated aminoacyl-L-leucine methyl esters appears to decrease as the hydrophobic character of the side chain of the N-terminal amino acid becomes more pronounced, a behavior similar to that observed for the N-acetylated L-amino acid methyl esters (Figure 6). A plot of ΔG_t (ethanol) vs. log K_o for the N-acetylated dipeptide methyl esters is shown in Figure 7. A reasonably good correlation is obtained for a straight line of slope approximately 2.6. One notable deviation from this line is N-acetyl-L-aspartyl-L-leucine methyl ester which has a K_o much higher than would be expected from the hydrophobic character of its side chain. The implications of this deviation will be discussed later. If translational non-productive binding is important in this class of substrates, the slope of this line will not tell much about the hydrophobic character of the interaction involved because the locus with which the N-terminal amino acid side chain (R_2') interacts is not determined. However, consideration of the calculated K_s and K_{sI} values which reflect the strength of the interaction of R_2' with ρ_1 (or more properly a ρ_1 sub-locus) and ρ_2, respectively, will shed more light on this question.

In Figure 8, values of ΔG_t (ethanol) are plotted against the logarithms of the calculated K_{sI} values (Table VII) for the dipeptide derivatives. A reasonably good correlation is obtained for a line with slope of approximately 1.3. Notable deviations from this line are the

Correlation of the Hydrophobic Character of the Side Chains of Amino Acids and K_o Values for the α-Chymotrypsin-catalyzed Hydrolyses of the Corresponding N-Acetyl-L-aminoacyl-L-leucine Methyl Esters

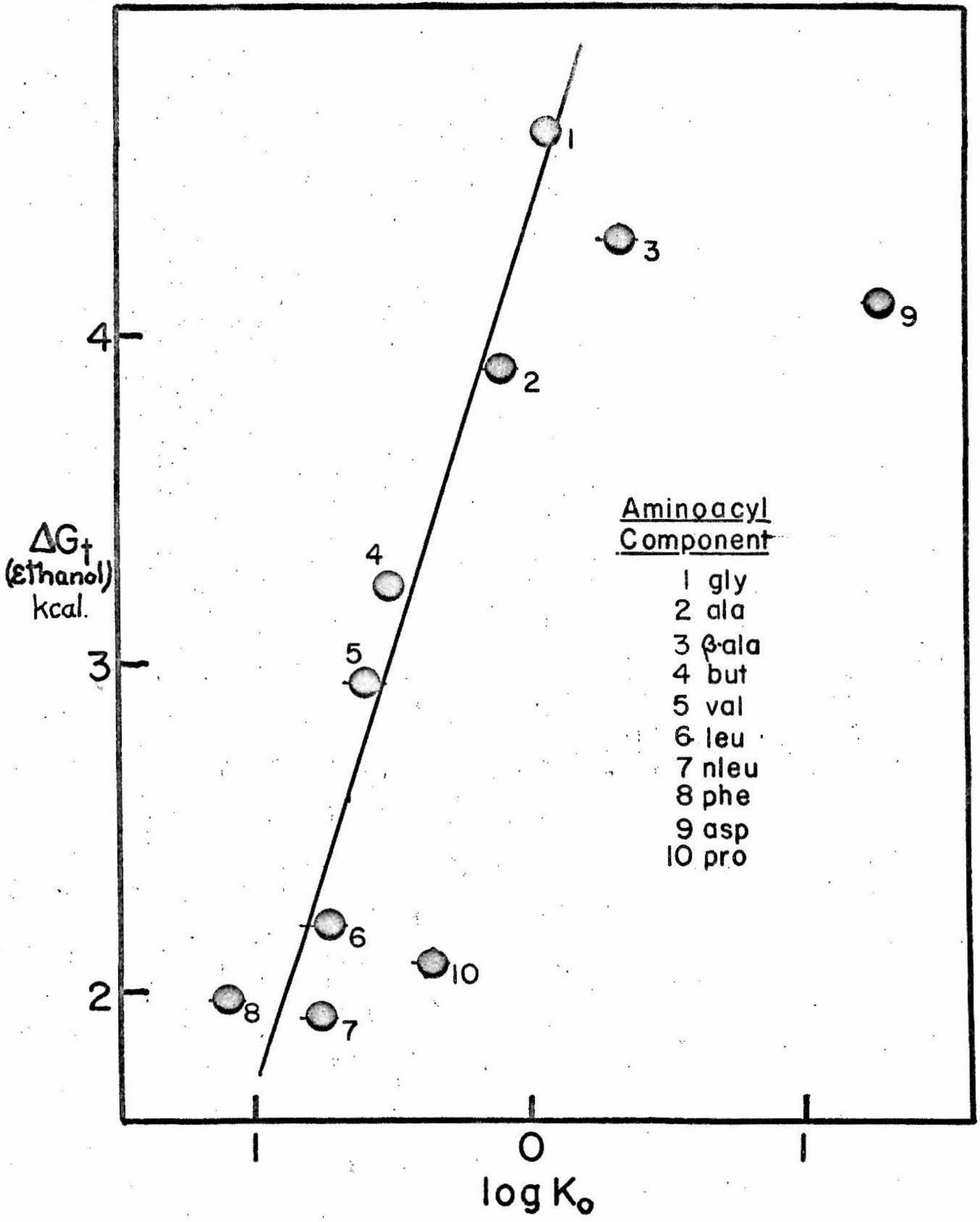


Figure 7

Hydrophobic Character of the $R_2'-\rho_2$ Interaction
in the Non-productive Binding Mode for Some
N-Acetyl-_L_-aminoacyl-_L_-leucine Methyl Esters

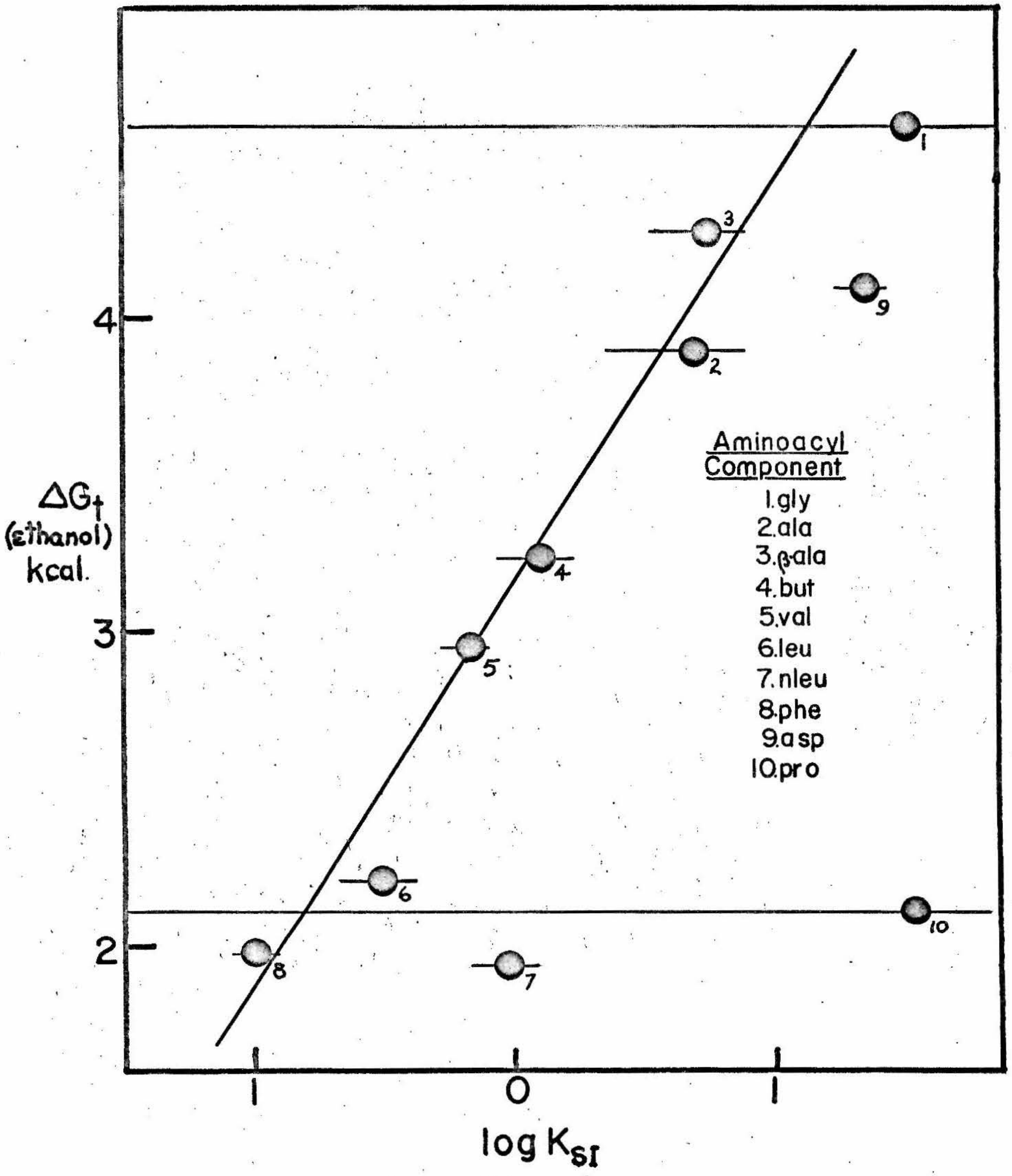


Figure 8

points for the L-norleucyl, L-prolyl and L-aspartyl derivatives. For these substrates, the K_{sI} values are higher than would be expected. Note that the errors associated with the K_{sI} values for the glycyl and prolyl derivatives are so large (> 100%) that these values are not particularly significant.

Figure 8 represents a correlation of the hydrophobic character of the R_2' side chain with the effectiveness of the R_2'-ρ_2 interaction present in the non-productive mode. Considering the rather crude approximations on which the calculation of K_{sI} values are based, it is pleasing that the slope of the line obtained (i.e., approximately 1.3) is reasonably close to that obtained for the N-acetylated-L-amino acid methyl esters (0.9). Both of these correlations represent the same type of interaction (i.e., the interaction of an amino acid side chain with the ρ_2 locus) and the slopes would be expected to be the same. That this is approximately the case argues in support of translational non-productive binding as proposed.

In Figure 9, values of ΔG_t (ethanol) are plotted against the logarithms of the calculated K_s values (Table VII) for the dipeptide derivatives. Once again a reasonable correlation is obtained. Now, however, a line with slope approximately 3.8 is obtained. This slope is much greater than that obtained for the corresponding K_{sI} values (i.e., 1.3). Here, the most notable deviations are the points for the β-alanyl- and L-aspartyl derivatives. The high slope obtained in this correlation indicates that the interaction of R_2' with its complementary ρ_1 sub-locus has less hydrophobic character than the corresponding interaction with ρ_2 (see Figure 8). This agrees quite favorably with the picture de-

Hydrophobic Character of the R'_2-ρ_1 Sub-locus Interaction in the Productive Binding Mode for Some N-Acetyl-L-aminoacyl-L-leucine Methyl Esters.

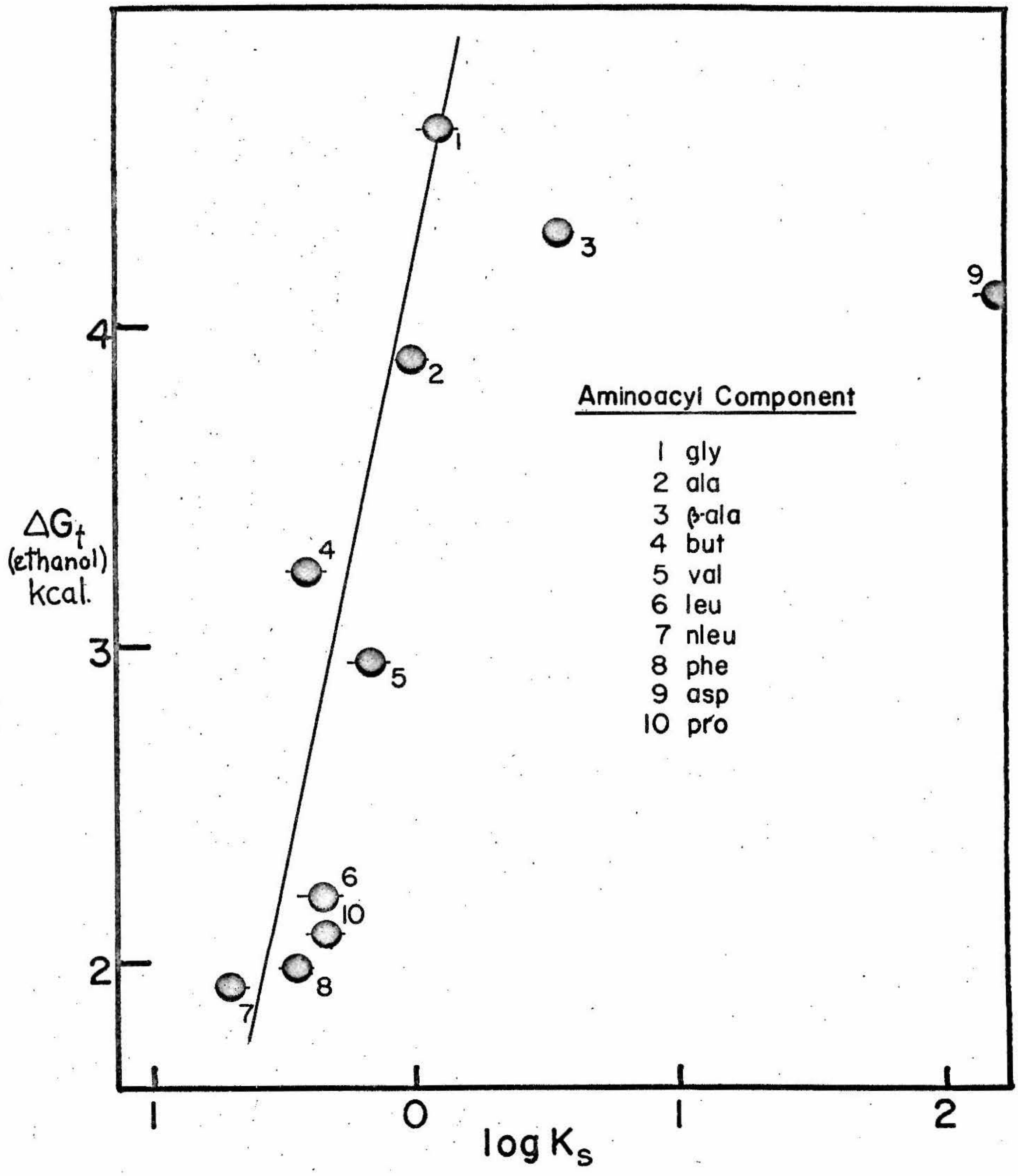
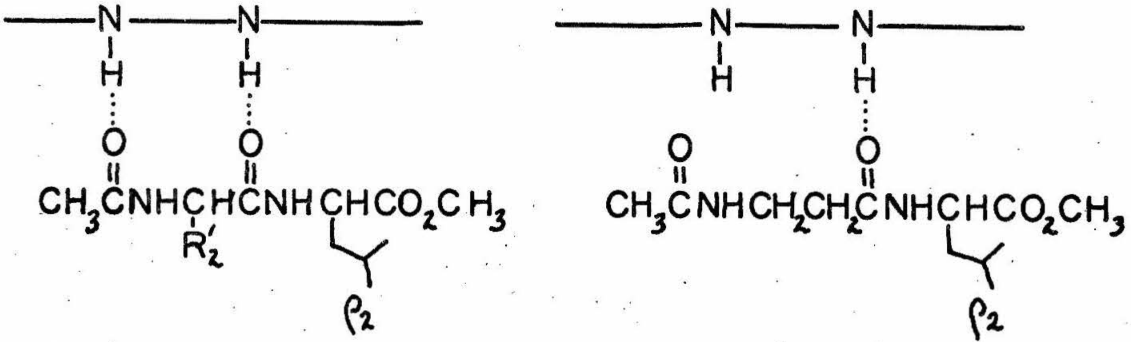


Figure 9

scribed in the study of N-acetylated glycine-containing peptide esters (p. 76). In that study, it was suggested that although hydrophobic interactions played a part in the overall R_1-ρ_1 interaction and were perhaps important in the observed increase in K_o for the tetrapeptide derivatives, other interactions of the R_1 moiety with the ρ_1 locus (e.g., hydrogen-bonding) were the predominant features of this overall interaction.

The deviations noted above in the ΔG_t-log K_s correlation (Figure 9) tend to support the above representation of the ρ_1 locus and also, it will be suggested, add another dimension to it. The large K_s value of N-acetyl-β-alanyl-L-leucine methyl ester can be interpreted in terms of diminished opportunity for hydrogen-bonding between the acetamido-group of the substrate and a corresponding sub-locus in ρ_1 (possibly a peptide linkage in the protein chain). The additional methylene group in this β-amino acid derivative can, because of a steric restriction, serve to eliminate the possibility of this particular hydrogen bond. This is illustrated in the following schematic figure:


N
H
O
CH$_3$CNHCHCNHCHCO$_2$CH$_3$
R$'_2$
ρ_2
CH$_3$CNHCH$_2$CH$_2$CNHCHCO$_2$CH$_3$

α-amino acid derivative

β-alanine derivative

Figure 10

It may be noted that indeed the K_s value for the β-alanine derivative (3.7 ± 0.8 mM) is even greater than those for the glycine or L-alanine derivatives (1.2 ± 0.2 mM and 1.0 ± 0.15 mM, respectively).

The abnormally high value of K_s calculated for the L-aspartic acid derivative (145 ± 25 mM) and the high value of K_o obtained (20 ± 3 mM) strongly suggest a repulsive interaction between the carboxylate[†] side chain of this substrate and the locus with which it interacts. In view of the magnitude of this effect, it is quite possible that a negatively charged species (e.g., an aspartic or glutamic acid residue in the protein chain) is present in this ρ_1 sub-locus. Such a species would exert a repulsive interaction on the aspartyl side chain and prevent effective binding. Results of the polypeptide study (Part I of this dissertation) support this view. In that study, it was observed that an aspartic acid residue penultimate to the residue which forms the carboxyl component of a cleavable peptide bond (i.e., the CO2 residue as defined in Part I) appears to hinder the hydrolysis of that bond.

The calculated K_{sI} value for N-acetyl-L-aspartyl-L-leucine methyl ester is also higher than that expected on the basis of the values for the other substrates studied (Fig. 8). However, this deviation is much less than that observed for the K_s value (Fig. 9) and this deviation might possibly reflect a long-range interaction of the carboxylate side chain with the proposed negative charge in the ρ_1 sub-locus.

[†] It is reasonable to assume that the β-carboxyl group ($pK_a \approx 3.6$) of N-acetyl-L-aspartyl-L-leucine methyl ester will be essentially completely ionized at pH 7.90, the pH at which the kinetic determination was carried out.

It would have been interesting, and undoubtedly very informative, to have studied the kinetics of the α-chymotrypsin-catalyzed hydrolysis of N-acetyl-*L*-lysyl-*L*-leucine methyl ester in order to test the hypothesis of the presence of a negative charge at the ρ_1 sub-locus. Results of the polypeptide study (Part I) indicate that α-chymotrypsin has a high secondary specificity for lysine at the CO2 position (i.e., lysine residues at this position appear to favorably influence the susceptibility of a peptide bond). If indeed a negative charge is present in the ρ_1 sub-locus, then it would be expected that K_o and K_s for the *L*-lysyl dipeptide derivative would be very low. The substrate N-acetyl-*L*-lysyl-*L*-leucine methyl ester was indeed synthesized. However, because of the extreme hygroscopicity of both this compound and its hydrochloride derivative, all attempts to purify this substrate were unsuccessful. Determination of the kinetic parameters of this substrate would serve to support or refute the hypothesis of a negatively charged moiety at the ρ_1 sub-locus.

There appears to be no reasonable explanation for the abnormally high K_{sI} value obtained for N-acetyl-*L*-norleucyl-*L*-leucine methyl ester and this must be considered an anomalous case.

Conclusions Regarding the Nature of Interactions at the ρ_2 and ρ_1 Loci--a Model of the ρ_1 Locus

The results of this study support the hypothesis (19,20,14) that the forces involved in R_2-ρ_2 interactions are predominantly hydrophobic in nature. Although Peterson (16) has shown that other forces (steric, hydrogen-bonding etc.) must also be considered, there seems to be no doubt that hydrophobic interactions are of primary importance.

On the basis of the results reported in this dissertation and in other discussions (20), a partial model of the ρ_1 locus may be proposed. This model will be discussed in terms of two sub-loci; one which interacts with the backbone of a substrate peptide chain, and one which interacts with the side chain of the amino acid residue in the peptide chain which is penultimate to that which is directly involved in the hydrolytic cleavage.

Results of the previously described study of glycine-containing peptide derivatives (p. 50ff.) indicated that some factor other than hydrophobic bonding was of major importance in determining the interaction of the R_1 moieties of these substrates at the ρ_1 locus. It was suggested that hydrogen-bonding between the peptide R_1 group of the substrate and the enzyme might be an important contributor to such an interaction. This could occur in one of the following ways:

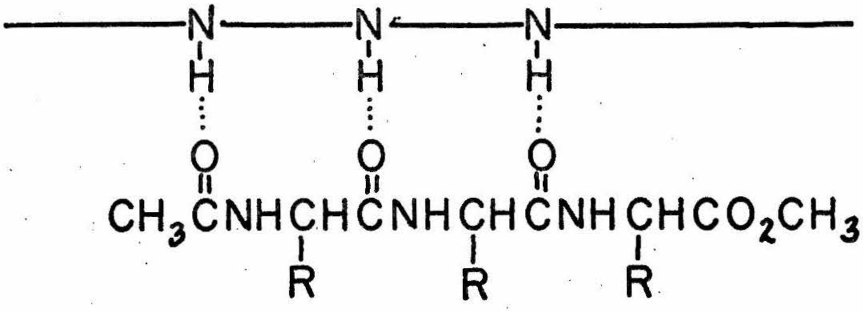
1. Hydrogen-bonding between the carbonyl oxygens of the substrate peptide linkages and a hydrogen donor on the enzyme (e.g., the hydrogens on the peptide nitrogens of the polypeptide backbone of the enzyme).

2. The converse of situation 1 in which the substrate acts as the hydrogen donor and the enzyme as the acceptor.

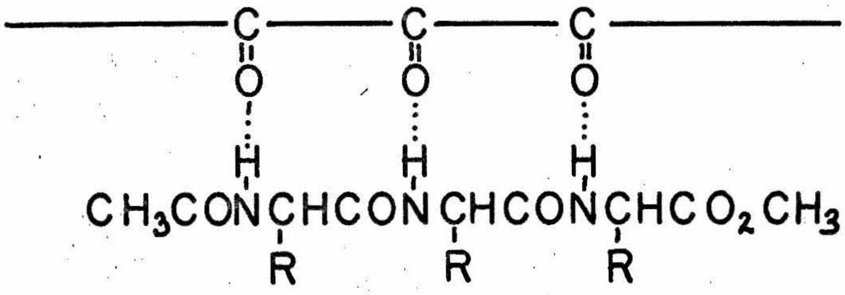
3. A situation similar to a "pleated-sheet" (29) arrangement wherein both types of interactions (1 and 2) could occur at the same time.

These three possibilities are represented schematically in Figure 11.

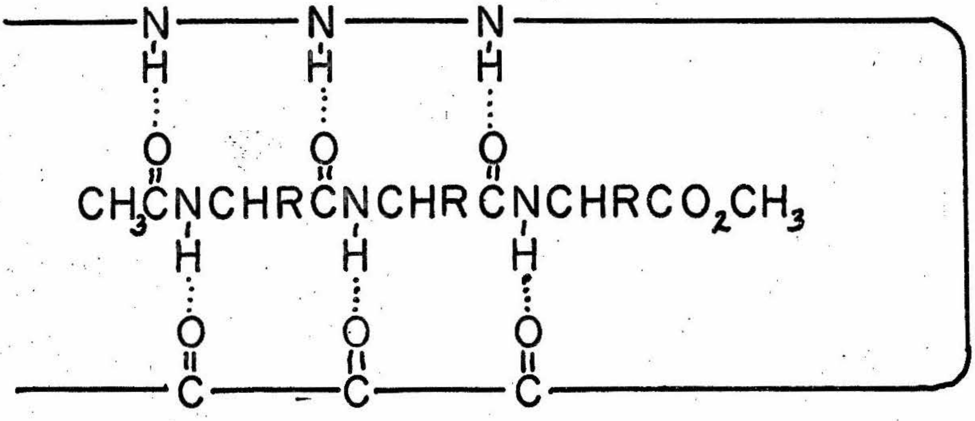
ENZYME-SUBSTRATE HYDROGEN-BONDING



Situation 1



Situation 2



Situation 3

Figure 11

These same results indicate that hydrogen-bonding is of more importance than hydrophobic bonding in the R_1-ρ_1 interaction for this type of substrate. Binding appears to increase (at least up to the tripeptide derivatives) although the hydrophobic character of the R_1 moiety decreases.

The rather high K_o and K_s values obtained for the β-alanyl dipeptide derivative also support this picture. The apparent agreement of the K_{sI} value of this substrate with the correlation obtained (Figure 8) might tend to dispute this. However, if hydrogen bonding is important both in binding and in orientation of the COR_3 group, then one might expect that k_o for this substrate might be lowered both by non-productive binding and by a decrease in k_2 itself because of poor orientation. Such a change would result in a higher calculated K_{sI} value and a K_s value which, although reduced, would still be higher than that expected from the rest of the data.

Two sets of results suggest that hydrophobic bonding is of some importance in the overall R_1-ρ_1 interaction. The decrease in binding observed for the glycine-containing tetrapeptide derivatives (Table III) is consistent with the proposed hydrophobic character of this locus. Also, the correlation between free energy of productive binding (log K_s) and the hydrophobic character of the side chain of the amino-terminal residue (in the dipeptide derivatives) suggests that there is some hydrophobic character to the R_1-ρ_1 interaction.

The data obtained from the kinetic study of the L-aspartyl dipeptide derivative and from the study of secondary specificity towards polypeptides (Part I) suggest the presence of a negatively charged species (perhaps a glutamate or aspartate anion) located in the ρ_1 sub-locus at which the R_2' side chain of dipeptide derivative interacts. The presence of such a species is also consistent with the lower hydrophobic character of this sub-locus (as compared with the ρ_2 locus). It is interesting and perhaps significant to note that in the sequence of α-chymotrypsin (30), aspartic acid is the residue penultimate to the serine residue implicated at the active center.

In conclusion, the results of these studies suggest that the ρ_1 locus can be partially characterized by two sub-loci, one of which makes possible hydrogen-bonding interactions and another which has some hydrophobic character but which appears also to contain an anionic species.

Interpretation of Data in Terms of the Acyl-enzyme Kinetic Scheme

The observed behavior of k_o and K_o for the dipeptide derivatives (Table VI) can of course be rationalized in terms of the acyl-enzyme kinetic scheme without invoking the non-productive binding hypothesis. In fact, almost any behavior of k_o and K_o can be explained by assuming suitable variation of the parameters k_2, k_3 and K_s when these latter parameters have not been determined. An example of such a possibility (i.e., for the case where k_o increases while K_o decreases) was illustrated in the previous section (p. 84). However, because $k_o = k_2k_3/(k_2+k_3)$ in terms of the acyl-enzyme scheme, it is evident that a decrease in k_o must involve a decrease in one or both of k_2 and k_3.

It seems worthwhile to discuss one of the many possibilities in terms of some other available data. Yamashita (25) determined the kinetic parameters for the α-chymotrypsin catalyzed hydrolyses of some glycyl-aminoacyl-L-tyrosine amides. He found, for example (see p. 98), that $k_o = k_2 = 9.3\ \mathrm{sec}^{-1}$ and $11.1\ \mathrm{sec}^{-1}$ for these substrates when the aminoacyl component was L-alanine and L-leucine, respectively.[†] The k_o values for the corresponding N-acetyl-aminoacyl-L-leucine methyl esters are $6.7\ \mathrm{sec}^{-1}$ and $3.2\ \mathrm{sec}^{-1}$, respectively (Table VI). The K_o values for these substrates are 0.81 mM and 0.18 mM, respectively (Table VI). Therefore, if we reasonably assume that variations in k_2 for the dipeptide ester derivatives should parallel variations in k_2 for the tyrosine amide derivatives, then the only way the observed decrease in k_o between the L-alanyl and L-leucine dipeptide ester derivatives can be explained is by a suitable decrease in k_3.

However, the postulates of the Bender specificity theory concerning k_3 (11,31) appear to argue against a decrease in k_3. This theory involves a rigid enzyme and an unspecified (binding-type) interaction which reduces the rotational freedom of the substrate within the active site. Bender suggests that the kinetic specificity of deacylation is determined solely by differences in entropy of activation and uses this argument to explain the higher deacylation rate of the acyl-enzyme of a specific substrate (e.g., N-acetyl-L-tyrosyl-chymotrypsin) as compared to that for a non-specific substrate (e.g., acetyl-chymotrypsin). It seems then that a structural modification which results in an increase in binding (as indicated by a decrease in K_o or K_s) should serve to make the acyl-enzyme ground state more rigid and thus reduce the entropy of activation to the transition state for deacylation. This then should result in an increased k_3, contrary to the behavior necessary to explain the results of the present study.

[†] It should be noted that these amide substrates represent the limit case in which acylation is the rate-determining step (i.e., $k_2 \ll k_3$) and thus $k_o = k_2$.

Non-productive binding may also be taken into account in the acyl-enzyme kinetic scheme (11) and such a treatment is employed in Part III of this thesis. However, without some knowledge of the magnitudes of the variations of k_2 and k_3 with structural modification, detailed discussion of the results of the present study in these terms would not be fruitful. It can be argued that, because of the mechanistic similarity of acylation and deacylation, k_2 and k_3 might vary in a similar manner (in both magnitude and direction) for the structural changes considered in this study. If this were true, the conclusions reached would not qualitatively differ from those presented on the basis of the simple scheme. Once again, lack of appropriate data precludes conclusive discussion in these terms.

Comparison of k_o/K_o Values with Other Data

Values of the parameter k_o/K_o obtained by Yamashita and co-workers for several aminoacyl-L-tyrosine amides (32,33) and five glycyl-aminoacyl-L-tyrosine amides (25) are presented in Tables VIII, IX and X. The ratio k_o/K_o, as described previously, is a measure of the free energy change required to go from free enzyme and free substrate in solution to the transition state for the kinetic step immediately following formation of the enzyme-substrate complex. The k_o/K_o values for the above-mentioned substrates will be compared with corresponding data for the N-acetyl-aminoacyl-L-leucine methyl esters (Table VI) in terms of the Bender-Kézdy specificity theory.

TABLE VIII

Comparison of k_o/K_o Values for Some Glycyl-aminoacyl-L-tyrosine Amides[a] and Some Aminoacyl-L-tyrosine Amides[b]

Aminoacyl Component	$(k_o/K_o)_I$[a] $M^{-1}sec^{-1}$	$(k_o/K_o)_{II}$[b] $M^{-1}sec^{-1}$	$(k_o/K_o)_I/(k_o/K_o)_{II}$[c]
glycine	25.2 ± 0.1	3.6 ± 0.1	7.0 ± 0.2
L-alanine	98.9 ± 0.3	12.8 ± 0.2	7.7 ± 0.1
L-valine	390 ± 3	54 ± 1	7.2 ± 0.1
L-leucine	520 ± 9	100 ± 4	5.2 ± 0.2
L-norleucine	560 ± 5	81 ± 2	6.9 ± 0.2

[a] Data for glycyl-aminoacyl-L-tyrosine amides from reference 25. Kinetics determined in aqueous solutions at 30°, pH 7.7 (0.1 M phosphate buffer).

[b] Data for aminoacyl-L-tyrosine amides from references 32 and 33. Kinetics determined in aqueous solutions at 30°, pH 8.0 (0.1 M phosphate buffer).

[c] Reactivity ratio based on $(k_o/K_o)_I$ and $(k_o/K_o)_{II}$.

TABLE IX

Comparison of k_o/K_o Values for Some N-Acetyl-aminoacyl-L-leucine Methyl Esters[a] and Some Aminoacyl-L-tyrosine Amides[b]

Aminoacyl Component	$(k_o/K_o)_I$[a] $M^{-1}sec^{-1}$	$(k_o/K_o)_{II}$[b] $M^{-1}sec^{-1}$	$(k_o/K_o)_I/(k_o/K_o)_{II}$[c] $\times 10^{-2}$
glycine	6.4 ± 0.9	3.6 ± 0.1	18 ± 3
L-alanine	7.8 ± 0.9	12.8 ± 0.2	6.1 ± 0.7
β-alanine	2.2 ± 0.3	4.0 ± 0.3	5.5 ± 0.9
L-α-aminobutyric acid	19 ± 2	52 ± 9	3.7 ± 0.7
L-valine	13 ± 3	54 ± 1	2.4 ± 0.6
L-norvaline	44 ± 8	84 ± 1	5.2 ± 0.9
L-leucine	18 ± 4	100 ± 4	1.8 ± 0.4
L-norleucine	39 ± 6	81 ± 2	4.8 ± 0.8
L-phenylalanine	23 ± 2	50 ± 20	5 ± 2
L-proline	18 ± 2	44 ± 66	4.1 ± > 100%

[a] Data for N-acetyl-aminoacyl-L-leucine methyl esters from Table VI. Kinetics determined in aqueous solutions at 25.0°, pH 7.90, 0.10 M in sodium chloride.

[b] Data for aminoacyl-L-tyrosine amides from references 32 and 33. Kinetics determined in aqueous solutions at 30°, pH 8.0 (0.1 M phosphate buffer).

[c] Reactivity ratio based on $(k_o/K_o)_I$ and $(k_o/K_o)_{II}$.

TABLE X

Comparison of k_o/K_o Values for Some N-Acetyl-aminoacyl-L-leucine Methyl Esters[a] and Some Glycyl-aminoacyl-L-tyrosine Amides[b]

Aminoacyl Component	$(k_o/K_o)_I$ [a] $mM^{-1} sec^{-1}$	$(k_o/K_o)_{II}$ [b] $M^{-1} sec^{-1}$	$(k_o/K_o)_I/(k_o/K_o)_{II}$ [c]
glycine	6.4 ± 0.9	25.2 ± 0.1	250 ± 40
L-alanine	7.8 ± 0.9	98.9 ± 0.3	78 ± 9
L-valine	13 ± 3	390 ± 3	34 ± 8
L-leucine	18 ± 4	520 ± 9	35 ± 8
L-norleucine	39 ± 6	560 ± 5	70 ± 10

[a] Data for N-acetyl-aminoacyl-L-leucine methyl esters from Table VI. Kinetics determined in aqueous solutions at 25.0°, pH 7.90, 0.10 M in sodium chloride.

[b] Data for glycyl-L-aminoacyl-L-tyrosine amides from reference 25. Kinetics determined in aqueous solutions at 30°, pH 7.7 (0.1 M phosphate buffer).

[c] Reactivity ratio based on $(k_o/K_o)_I$ and $(k_o/K_o)_{II}$.

Yamashita made no estimate of the standard deviations associated with the kinetic parameters for the substrates which he studied. The error values in Tables VIII, IX and X were computed by this author (with the use of a computer least-squares analysis) from the primary data of references 25, 32 and 33. The accuracy and validity of the data presented in these tables might well be questioned. The kinetic technique employed was the Conway microdiffusion technique which, although it is adequate for the substrates studied, is certainly no more accurate than the pH-stat technique employed in the present studies. For the case of the aminoacyl-L-tyrosine amides, the kinetic parameters were determined from plots of [S]/v *vs.* [S] containing only three points (i.e., three initial substrate concentrations over a four- to five-fold range). The straight lines fitting these plots were calculated by means of a least-squares analysis! Such calculations are certainly of questionable significance.† The standard deviations for the individual parameters (k_o and K_o) were often quite large but, because conditions were chosen such that in most cases the reaction rates were approximately first order in substrate concentration, the calculated standard deviations in the parameter k_o/K_o were usually small. However, the "incomplete" nature of the kinetic analysis casts some suspicion even on these latter values.

In terms of the Bender specificity theory, if one makes the additional assumption (as Bender implies (11)) that the specificity for the parts of a complex R_1 moiety are independent, then the S_{R_1}

---

† Indeed the standard deviations calculated from a least-squares treatment on three points (Table VIII) represent only a minimum estimation of error.

relationship given in the General Introduction of this dissertation (p. 14) may be modified:

i.e.,
$$\log \frac{(k_o/K_o)_{r_a r_b R_2 X}}{(k_o/K_o)_{r_{ao} r_{bo} R_2 X}} = S_{R_1} = S_{r_a} + S_{r_b} \tag{23}$$

where r_a, r_b and r_{ao}, r_{bo} are components of R_1 and R_{1o} respectively.

Therefore, the following relationship should hold:

$$\log \frac{(k_o/K_o)_{r_a r_b R_2 X}}{(k_o/K_o)_{r_{ao} r_b R_{2o} X_o}} = \rho_x^* \sigma_x^* + S_{R_2} + S_{r_a}$$
$$= \log \frac{(k_o/K_o)_{r_a r_b' R_2 X}}{(k_o/K_o)_{r_{ao} r_b' R_{2o} X_o}} \tag{24}$$

Thus, $\frac{(k_o/K_o)_{r_a r_b R_2 X}}{(k_o/K_o)_{r_{ao} r_b R_{2o} X_o}}$ should be a constant for given r_a, r_{ao}, R_2, X and X_o groups as r_b is varied.

An illustration of this type of treatment is given by the data of Table VIII where $(k_o/K_o)_I \equiv (k_o/K_o)_{r_a r_b R_2 X}$, $(k_o/K_o)_{II} \equiv (k_o/K_o)_{r_{ao} r_b R_{2o} X_o}$, $r_a r_b R_2 X$ represents a glycyl-aminoacyl-L-tyrosine amide and $r_{ao} r_b R_{2o} X_o$ represents the corresponding aminoacyl-L-tyrosine amide (note that $R_2 = R_{2o}$ and $X = X_o$ in these comparisons). Thus, by the argument presented above, the ratio $(k_o/K_o)_I/(k_o/K_o)_{II}$ should be a constant independent of the aminoacyl component (r_b). It is evident from Table VIII that this is true and that the ratio in question is approximately six to seven.[†]

By the same arguments, the ratio of k_o/K_o values for the N-acetyl-aminoacyl-L-leucine methyl esters to those for the corresponding aminoacyl-L-tyrosine amides should also be a constant independent of the aminoacyl component. These data are presented in Table IX. With respect to the results of this table it may be noted that, with the exception of the cases where the aminoacyl component is glycine or L-leucine, there is less than a factor of three separating the highest and lowest values of the ratio $(k_o/K_o)_I/(k_o/K_o)_{II}$.[‡] The average value of this ratio, with the above-mentioned deviations omitted, is approximately 450.

A similar constancy of the k_o/K_o ratios should be expected in a comparison of the k_o/K_o parameters for the N-acetyl-aminoacyl-L-leucine methyl esters with those for the glycyl-aminoacyl-L-tyrosine amides. These data are presented in Table X. Again, with the exception of the case where the aminoacyl component is glycine, the values agree roughly within a factor of two, with an average ratio of about 54.

The L-leucine deviation in the results of Table IX may not be unexpected. Yamashita reported experimental difficulties in the kinetic study of this tyrosine amide derivative (32). The glycine deviations observed in both correlations (Tables IX and X) may arise because of some property associated with the lack of a center of asymmetry in the glycine residue penultimate to the L-tyrosine residue. One possibility is that the terminal free amino group of both of these substrates might be able to interact with the proposed anionic species in the ρ_1 sub-locus (see p. 116) in such a manner as to lower the reactivity of these substrates. The masking of the free amino group in the N-acetylated dipeptide ester derivatives would serve to eliminate the possibility of this interaction, yielding an increase in reactivity.

[†] Actually, it was this correlation which Bender used as a test of the hypothesis of independence of specificity for the parts of an R_1 group.

[‡] Bender states (11) that reproducibility within a factor of two between different laboratories must be accepted and therefore the above correlation is a reasonable one.

In conclusion, although these latter two comparisons are not nearly as good as the correlation between the two sets of Yamashita's data, it must be remembered that, whereas in the data of Table VIII, the difference in reactivity considered is only about 7-fold, in the data of Tables IX and X, the differences in reactivity are roughly 500-fold and 50-fold, respectively. Therefore, in the light of what has been said regarding the accuracy of Yamashita's data and in view of the limitations of the theory employed, the correlations obtained may be considered adequate.

Secondary Specificity--Correlation of Experimental Results with those of the Polypeptide Study

The results of the study of the α-chymotrypsin-catalyzed hydrolyses of N-acetylated dipeptide esters illustrate graphically the importance of secondary specificity, at least for this class of substrates. Using values of k_0/K_0 as criteria of reactivity (Table VI), a 700-fold secondary specificity effect is observed between the unreactive L-aspartyl derivative and the relatively reactive L-norleucyl derivative.

One of the major objectives of the present study was to correlate the kinetic results with the secondary specificity trends observed in the α-chymotrypsin-catalyzed hydrolysis of polypeptides (Part I). It appears that the experimental results do correspond roughly to the qualitative effects observed for CO2 residues in polypeptides (i.e., the residues penultimate in sequence to those which contain the carboxyl component of the cleavable bond). In addition, the observation that the secondary specificity effects seen for the dipeptide ester derivatives roughly parallel those observed for the aminoacyl- and glycyl-aminoacyl-L-tyrosine amides lend credence to the suggestion that the specificity pattern observed for the ester derivatives reflects that found for polypeptides.

The amino acid residues which may be considered in this correlation are (in order of their reactivity-enhancing properties in the dipeptide esters): L-phenylalanine > L-proline ≈ L-leucine, > L-valine ≈ L-isoleucine > L-alanine ≈ glycine >> L-aspartic acid.

With regard to the L-phenylalanine residue, it will be noted (Part I, Table II, p. 31) that because of its large frequency of occurrence at the CO1 position in polypeptides, the sample available for discussion of its secondary specificity effect was so small that no meaningful conclusions could be made. L-Proline, however, was noted as one of the residues which clearly appeared to assist hydrolysis and therefore its beneficial effect in model substrates was to be expected. The polypeptide data for L-leucine indicated no marked effect one way or the other and thus the beneficial effect of this residue noted in the model study may be regarded as slightly anomalous. Both L-valine and

L-isoleucine (although to a much smaller extent in the latter case) were observed to have beneficial effects in the polypeptide study. This result is in fair agreement with results of the model study. The residues L-alanine and glycine may be regarded as having approximately neutral effects in both studies. Finally, in both studies, the L-aspartyl residue was observed to markedly hinder hydrolysis.

Therefore, in view of the very approximate and qualitative treatment of secondary specificity of α-chymotrypsin towards polypeptides, it appears that results of the model study agree quite well with this specificity pattern. These preliminary results should encourage further work in this rather important area of secondary specificity.

EXPERIMENTAL

Syntheses

All melting points were corrected with the use of a calibration curve. Analyses were performed by Spang Microanalytical Laboratories, Ann Arbor, Michigan, and by Elek Microanalytical Laboratories, Torrance, California. Infrared and NMR spectra were taken of most of the compounds. These spectra were consistent with the structures proposed.

i. L-Alanine methyl ester hydrochloride

Esterification of L-alanine with methanol and thionyl chloride by the method of Brenner and Huber (34,35) and two recrystallizations from methanol-ethyl ether gave L-alanine methyl ester hydrochloride (75%).

ii. L-Valine methyl ester hydrochloride

Esterification of L-valine by the method of Brenner and Huber (34,35) and two recrystallizations from methanol-ethyl ether gave L-valine methyl ester hydrochloride (80%), m.p. 150.0-153.0°, with decomposition (lit. m.p. 150.0-154.5°, with decomposition) (35).

iii. L-Leucine methyl ester hydrochloride

This compound was prepared as described by Hein, Jones and Niemann (10) in 93% yield. The crude product was recrystallized from methanol-ethyl ether and used directly in further syntheses.

iv. N-Acetylglycyl-L-alanine methyl ester and
v. N-acetylglycyl-L-valine methyl ester

These two substrates were synthesized in these laboratories by

Dr. Kunitake (9). After re-drying in vacuo they were used in kinetic studies.

vi. Glycylglycine

This compound was prepared from glycine ethyl ester hydrochloride by way of 2:5 diketopiperazine and glycylglycine hydrochloride monohydrate by the method of Greenstein and Winitz (36) in ca. 50% overall yield, m.p. 215-220°, with decomposition.

vii. N-Acetylglycylglycine

Glycylglycine (0.130 mole) was suspended in 330 ml. glacial acetic acid. Acetic anhydride (0.33 mole) was added. The suspension was heated to boiling with subsequent solution of most of the solid. The solution was quickly cooled to room temperature. The solvent was removed under vacuum, dry toluene being added to entrain the remaining acetic acid when the distillation became sluggish. The resulting solid was dried for one hour at 40°. The crude product was twice recrystallized from ethanol with Norit decolorization yielding N-acetylglycylglycine (90%), m.p. 178.0-179.5° (lit. m.p. 179°) (37).

viii. N-Acetylglycylglycyl-L-alanine methyl ester

N-Acetylglycylglycine (0.03 mole) and triethylamine (0.03 mole) were added to 100 ml. chloroform (Spectrograde) in a 500 ml., three-neck, round-bottom flask fitted with a drying tube, dropping funnel, thermometer and magnetic stirring bar. Most of the solid dissolved. The mixture was cooled to -10° in a methyl cellusolve-dry ice bath and isobutyl chloroformate (0.032 mole) was added slowly while the temperature was maintained between -15° and -5°. After a further 10 minutes, a precooled solution of L-alanine methyl ester hydrochloride (0.03 mole) and triethylamine (0.03 mole) in 100 ml. chloroform was added. The mixture was allowed to warm to room temperature and was stirred overnight. The resulting heavy white suspension was filtered (with difficulty). The crude white crystalline product was washed with 20 ml. water, dried and recrystallized four times from isopropanol (twice with Norit decolorization) yielding powdery white crystals of N-acetylglycylglycyl-L-alanine methyl ester (15%), m.p. 208.5-209.5° and $[\alpha]_D^{25}$ -(22.1 ± 1)° (c, 2% in methanol).

Analysis	Calculated:	C: 46.32	H: 6.61	N: 16.28
$C_{10}H_{17}N_3O_5$ (259.3)	Found:	C: 46.09 46.15	H: 6.43 6.55	N: 16.21 14.95

ix. N-Acetylglycylglycyl-L-valine methyl ester

This compound was prepared in a manner analogous to that for the preceding compound (viii). After stirring overnight, a small amount of suspended solid was filtered off. The resulting filtrate was washed with 15 ml. portions of N aqueous hydrochloric acid, 10% aqueous sodium carbonate and water. The organic phase was dried over anhydrous sodium sulfate, then stripped of solvent under vacuum, leaving a yellow oil which crystallized on addition of ethyl acetate. The crude product was recrystallized three times from ethyl acetate (twice with Norit decolorization) yielding white powdery crystals of N-acetylglycylglycyl-L-valine methyl ester (20%), m.p. 178.0-179.0°, $[\alpha]_D^{25}$ -(20.0 ± 0.3)° (c, 2% in methanol).

Analysis				
$C_{12}H_{21}N_3O_5$ (287.3)	Calculated:	C: 50.16	H: 7.37	N: 14.63
	Found:	C: 50.24	H: 7.06	N: 14.78

x. N-Acetylglycylglycyl-L-leucine methyl ester

This compound was prepared in the same manner as the preceding compound (ix). The crude product was recrystallized three times from ethyl acetate (once with Norit decolorization) yielding flat white crystals of N-acetylglycylglycyl-L-leucine methyl ester (20%), m.p. 175.5-177.0°, $[\alpha]_D^{25}$ -(27.7 ± 0.6)° (c, 2% in methanol).

Analysis				
$C_{13}H_{23}N_3O_5$ (301.3)	Calculated:	C: 51.81	H: 7.69	N: 13.95
	Found:	C: 51.65 51.73	H: 7.74 7.86	N: 13.92

xi. N-Chloroacetylglycylglycine

This compound was prepared by the method of Fischer (38) as modified by Sluyterman and Veenendaal (39) from 2:5 diketopiperazine and chloroacetyl chloride, yielding N-chloroacetylglycylglycine (70%). The crude material was twice recrystallized from water to give white needle-like crystals, m.p. 174.0-175.0° (lit. m.p. 175-176°) (39).

xii. Glycylglycylglycine

N-Chloroacetylglycylglycine (0.050 mole) in 37 ml. water at 40° was neutralized with 3.3 ml. of 15 N aqueous ammonia. A further 60 ml. of 15 N ammonia (warmed to 40°) was then added; the flask was stoppered and allowed to stand at 40° for 2½ hours. The solvent was then removed under vacuum with the aid of absolute ethanol. A large part of the ammonium chloride present was removed by washing the solid residue with warm methanol. Purification was effected by dissolving the crude material in a minimum amount of warm water and precipitating it with ethanol. This procedure was repeated twice. The yield of pure product was 80%, m.p. 244-246°, with decomposition (lit. m.p. 246°, with decomposition) (40).

xiii. N-Acetylglycylglycylglycine

Acetylation of glycylglycylglycine (0.03 mole) using the method described for the synthesis of N-acetylglycylglycine (vii) yielded yellow-white crystals of N-acetylglycylglycylglycine (97°). The crude product was dried over silica gel and KOH and was twice recrystallized from ethanol yielding fine white crystals, m.p. 212.0-214.0° (lit. m.p. 214°) (41).

Equivalent Weight   Calculated: 231.2
Found: 237.0

xiv. N-Acetylglycylglycylglycyl-L-alanine methyl ester

This compound was prepared with the use of a mixed-anhydride technique similar to that described by Vaughan and Osato (42). A suspension of N-acetylglygylglycylglycine (0.014 mole) and triethylamine (0.014 mole) in 60 ml. chloroform and 30 ml. toluene in a 300 ml., round-bottom, three-neck flask (fitted with a dropping funnel, drying tube, thermometer and magnetic stirring bar) was cooled to -5° in a methyl cellusolve-dry ice bath. Isobutyl chloroformate (0.015 mole) was added slowly while the temperature was maintained between -10 and -5°. After a further half hour at this temperature, a precooled solution of L-alanine methyl ester hydrochloride (0.014 mole) and triethylamine (0.014 mole) in 25 ml. chloroform was added. The cooling bath was removed and the mixture was stirred at room temperature overnight. The mixture was then filtered (with considerable difficulty). When most of the solvent had been filtered off, the resulting gel was triturated with hexane leaving a white solid which was filtered and washed with two 10 ml. portions of water. The crude product was dried under vacuum over phosphorous pentoxide (yield 60%), m.p. 245-247°. The crude product was purified by dissolving it in water and running the solution through an ion exchange resin (Amberlite MB-3 cation-anion exchange) with subsequent removal of the solvent from the eluent under vacuum. Recrystallization from methanol yielded the pure product, m.p. 251.0-252.0°.

Analysis

$C_{12}H_{20}N_4O_6$ (316.3)	Calculated:	C: 45.56	H: 6.37	N: 17.71
	Found:	C: 45.52	H: 6.28	N: 17.64
Saponification Equivalent	Calculated:	316.3		
	Found:	314.0		

xv. N-Acetylglycylglycylglycyl-L-valine methyl ester

This compound was prepared in the same manner as the preceding compound (xiv). The crude product, m.p. 203-205°, with decomposition, was purified by ion exchange (Amberlite MB-3 cation-anion exchange resin). Evaporation of the solvent from the eluent yielded white crystals of the desired product, m.p. 226-228°. Two recrystallizations from absolute ethanol gave the pure product, m.p. 227.5-228.0°, $[\alpha]_D^{25}$ -(29.8 ± 0.6)° (c, 1.4% in water).

Analysis	Calculated:	C: 48.83	H: 7.02	N: 16.27
$C_{14}H_{24}N_4O_6$ (344.4)	Found:	C: 48.62	H: 6.88	N: 16.07
Saponification Equivalent	Calculated:	344.4		
	Found:	347.0		

xvi. N-Acetylglycylglycylglycyl-L-leucine methyl ester

This compound was prepared in the same manner as the preceding compound (xv). The crude product, m.p. 208-209°, with decomposition, was recrystallized from absolute ethanol yielding a partially purified product, m.p. 213.0-214.0°, with decomposition. The product was further purified by ion exchange (Amberlite MB-3 cation-anion exchange resin). After evaporation of the eluent solvent (under vacuum) fine white crystals were obtained, m.p. 217.0-218.0°. Recrystallization from absolute ethanol did not change the melting point.

Analysis	Calculated:	C: 50.27	H: 7.31	N: 15.63
$C_{15}H_{26}N_4O_6$ (358.4)	Found:	C: 50.44	H: 7.36	N: 15.68

xvii. N-Acetylglycyl-L-leucine methyl ester

I. This compound was prepared using a mixed-anhydride technique similar to that described by Vaughan and Osato (42). To a solution of N-acetylglycine (0.04 mole) and triethylamine (0.04 mole) in a mixture of 60 ml. toluene and 60 ml. chloroform (cooled to -5° in a methyl cellusolve-dry ice bath) was added isobutylchloroformate (0.04 mole) over a period of 15 minutes while the temperature was maintained between -15° and -5° C. The reaction mixture set to a semi-solid mass. The mixture was stirred mechanically at the above temperature for a further 25 minutes. Then a precooled solution of L-leucine methyl ester hydrochloride (0.04 mole) and triethylamine (0.04 mole) in 80 ml. chloroform was added. The solid subsequently dissolved. The mixture was allowed to warm to room temperature and was stirred overnight. The organic solution was washed with 10 ml. portions of water and aqueous sodium bicarbonate (5%). The organic phase was dried over anhydrous sodium sulfate and stripped of solvent under vacuum, leaving a straw-yellow oil (60%). Attempts at crystallization from conventional solvents at room temperature (or greater) were unsuccessful. In order to purify the compound, a benzene solution (5%) was chromatographed on neutral alumina (activity one). The compound was eluted with benzene-methanol. The main fraction was eluted with 1% methanol in benzene. The recovery was 94%. The solvent was evaporated under vacuum leaving a viscous, colorless oil. Crystallization from absolute ether at -30° was effected. It was found that the product (fine white crystals) was very hygroscopic. All manipulations during recrystallization were carried out in a "dry-bag" under nitrogen, and the compound was dried and stored in a desiccator over phosphorous pentoxide. After two recrystallizations from absolute ether at -30° the pure white crystalline product was obtained, m.p. 70.0-71.0° (sealed tube).

Analysis	Calculated:	C: 54.08	H: 8.25	N: 11.47
$C_{11}H_{20}N_2O_4$ (244.3)	Found:	C: 54.20 53.97	H: 8.38 8.37	N: 11.36

II. This compound (xvii) was also prepared from glycyl-L-leucine methyl ester hydrochloride in the following manner. Glycyl-L-leucine $[\alpha]_D$ -34.6° (c = 4, in water) (0.027 mole) was esterified with methanol and thionyl chloride in the usual manner as described by Brenner and

Huber (34). After the solvent was evaporated under vacuum, the remaining white powder was washed with hexane and dried. The product was obtained in 98% yield. One recrystallization from ethanol-ethyl ether gave glycyl-L-leucine methyl ester hydrochloride, m.p. 170.0-171.0° (m.p. 170.5-171.5° as reported by Kunitake (9)).

The ester hydrochloride (0.02 mole) was added to a mixture of sodium acetate (0.04 mole) in 20 ml. acetic anhydride which had been cooled in an ice bath. The mixture, protected from atmospheric moisture with a drying tube containing calcium chloride, was stirred at ca. 0° for 30 minutes. The cooling bath was then removed and the mixture was stirred at room temperature overnight. Most of the acetic anhydride was removed under vacuum (ca. 200 μ Hg) at room temperature. Ten ml. water was added to the remaining residue and the mixture was heated on a steam bath for 10 minutes to destroy the rest of the acetic anhydride. After cooling, the mixture was made slightly alkaline with solid sodium bicarbonate. Twenty ml. of water was added to make the mixture more convenient to handle. The mixture was filtered. The filtrate and the solid residue were both extracted with five 10 ml. portions of chloroform. The two chloroform extracts were combined and washed with 10 ml. portions each of aqueous hydrochloric acid (5%), aqueous sodium bicarbonate (5%) and water. The organic phase was dried one hour over anhydrous magnesium sulfate and most of the solvent was removed at room temperature under vacuum. The remaining solvent was removed under vacuum at 30° and ca. 30 μ Hg, yielding a colorless, viscous oil (90% overall yield). The crude product was purified in the usual manner by alumina chromatography, crystallization and recrystallization (as indicated above (xvii--procedure I)). Two recrystallizations from ethyl ether at -30° gave the pure product, m.p. 70.0-71.0° (sealed tube).

xviii. N-Acetyl-L-alanyl-L-leucine methyl ester

L-Alanyl-L-leucine, $[\alpha]_D^{24}$ -17° (c, 5% in water) (0.002 mole) was esterified in the usual Brenner-Huber fashion (34) with thionyl chloride and methanol. When the solvent was removed (under vacuum) from the reaction mixture, it gave the product, L-alanyl-L-leucine methyl ester hydrochloride (85%), a white fluffy solid. The product was very hygroscopic and attempts to recrystallize it from ethanol-ethyl ether resulted in an oil. The mother liquor was decanted and the remaining solvent was removed under vacuum (200 μ Hg at room temperature) yielding a white foamy solid. This solid was dried over phosphorous pentoxide (4 hours at 200 μ Hg). The product was used directly in the acetylation. The acetylation was carried out as described in the preparation of N-acetylglycyl-L-leucine methyl ester (xvii--procedure II) except that an "inverse-addition" procedure was used. The ester hydrochloride (under vacuum at 200 μ Hg) was cooled in an ice bath. A previously cooled mixture of sodium acetate (0.005 mole) in 6 ml. acetic anhydride was then added to the cold ester hydrochloride. A drying tube was attached and the mixture was stirred at ca. 0° for 30 minutes. The cooling bath was then removed and the mixture was stirred at room temperature overnight. The work-up was the same as that described for the preparation of N-acetylglycyl-L-leucine methyl ester (xvii--procedure II) and the product obtained, N-acetyl-L-alanyl-L-leucine methyl ester, was a white crystalline solid, m.p. 116.5-117.5°. Two recrystallizations, from isopropyl ether gave white fibrous crystals, m.p. 118.5-119.2°.

Analysis	Calculated:	C: 55.79	H: 8.59	N: 10.85
$C_{12}H_{22}N_2O_4$ (258.3)	Found:	C: 55.63	H: 8.58	N: 10.89

xix. N-Acetyl-L-valyl-L-leucine methyl ester

L-Valyl-L-leucine (0.002 mole) was esterified in the usual Brenner-Huber fashion (34). Removal of the solvent under vacuum and recrystallization from ethanol-ethyl ether gave white crystals of L-valyl-L-leucine methyl ester hydrochloride (85%), m.p. 178.5-179.5° (lit. m.p. 174-176°) (43), $[\alpha]_D^{26}$ -(12.3 ± 0.6)° (c, 0.9% in water), (lit. $[\alpha]_D$ -12.6°) (c, 1% in water) (43).

The ester hydrochloride was acetylated in the usual manner (see xvii--procedure II) with acetic anhydride and sodium acetate. Recrystallization of the crude product from ethyl acetate-hexane gave white fibrous crystals of N-acetyl-L-valyl-L-leucine methyl ester, m.p. 137.0-139.0° (slow heating); m.p. 119-120° (fast heating). Apparently there is a phase change which occurs at about 119-120°. $[\alpha]_D^{27.5}$ -(56.8 ± 0.4)° (c, 2% in methanol).

Analysis	Calculated:	C: 58.71	H: 9.15	N: 9.78
$C_{14}H_{26}N_2O_4$ (286.4)	Found:	C: 58.68	H: 9.14	N: 9.85

xx. N-Acetyl-L-phenylalanyl-L-leucine methyl ester

L-phenylalanyl-L-leucine monohydrate (0.0017 mole) $[\alpha]_D^{24}$ -20° (c, 1% in aqueous sodium bicarbonate) was esterified by the method of Brenner and Huber (34). Removal of solvent under vacuum gave white crystals of L-phenylalanyl-L-leucine methyl ester hydrochloride. Because the product was quite hygroscopic no further purification (other than drying over phosphorous pentoxide at 300 μ Hg) was attempted.

The ester hydrochloride was acetylated with acetic anhydride and sodium acetate using the "inverse addition" procedure as described in the preparation of N-acetyl-L-alanyl-L-leucine methyl ester (xviii). The crude product was obtained in 92% overall yield. Two recrystallizations from isopropyl ether gave fine white crystals of N-acetyl-L-phenylalanyl-L-leucine methyl ester, m.p. 124.5-125.5°; $[\alpha]_D^{27.5}$ -(16.8 ± 0.4)° (c, 2% in methanol).

Analysis				
$C_{18}H_{26}N_2O_4$ (334.4)	Calculated:	C: 64.64	H: 7.84	N: 8.38
	Found:	C: 64.66	H: 7.84	N: 8.42

xxi. N-Acetyl-L-prolyl-L-leucine methyl ester

L-Prolyl-L-leucine hydrate (0.002 mole) $[\alpha]_D^{24}$ -75° (c, 2% in N aqueous hydrochloric acid) was esterified by the method of Brenner and Huber (34). Removal of the solvent under vacuum (400 μ Hg) gave a yellow-white solid which was very hygroscopic. The crude product was dissolved in methanol, and decolorized with Norit. When ethyl ether was added, the product oiled out. The mother liquor was decanted and the remaining solvent was removed under vacuum (400 μ Hg) yielding a slightly yellowish solid. The product was dried over phosphorous pentoxide at 400 μ Hg and was used directly in the acetylation.

The ester hydrochloride was acetylated (using acetic anhydride and sodium acetate) with the "inverse addition" technique (xviii). The crude product was a straw-white viscous oil (90% overall yield). This oil resisted numerous crystallization attempts.

The product was purified by column chromatography. A 2% solution of the crude product in benzene was chromatographed on neutral alumina (activity one). Elution was effected with methanol-benzene. The main fraction was eluted with 1% methanol in benzene (recovery 90%). The compound was further purified by molecular distillation using a small sublimation apparatus (30-50 μ Hg, 70-90°). The distilled N-acetyl-L-prolyl-L-leucine methyl ester was transferred to a sealed flask under dry nitrogen.

Analysis				
$C_{14}H_{24}N_2O_4$ (284.4)	Calculated:	C: 59.13	H: 8.51	N: 9.85
	Found:	C: 58.81	H: 8.65	

xxii. N-Acetyl-β-alanyl-L-leucine methyl ester

β-Alanyl-L-leucine (0.0012 mole) was esterified by the method of Brenner and Huber (34). The methyl ester hydrochloride obtained was a colorless oil which was dried over phosphorous pentoxide at 300 μ Hg and used directly in the acetylation.

Acetylation of the ester hydrochloride in the usual manner, using the "inverse addition" technique (xviii) gave, after two recrystallizations, fine white crystals of N-acetyl-β-alanyl-L-leucine methyl ester (78% overall), m.p. 100.0-101.0°, $[\alpha]_D^{27.5}$ -(25.7 ± 0.3)° (c, 2% in methanol).

Analysis				
$C_{12}H_{22}N_2O_4$ (258.3)	Calculated:	C: 55.79	H: 8.59	N: 10.85
	Found:	C: 55.76	H: 8.52	N: 11.05

xxiii. N-Acetyl-L-α-aminobutyryl-L-leucine methyl ester

L-α-Aminobutyryl-L-leucine (0.0012 mole) was esterified by the method of Brenner and Huber (34) giving white crystals of L-α-aminobutyryl-L-leucine methyl ester hydrochloride. The crude product was dried under vacuum and used directly in the acetylation.

Acetylation of the ester hydrochloride in the usual manner using the "inverse addition" technique (xviii) gave the crude product, a white solid (80%). Two recrystallizations from ethylacetate-hexane gave fine white crystals of N-acetyl-L-α-aminobutyryl-L-leucine methyl ester, m.p. 102.0-103.2°, $[\alpha]_D^{27.5}$ -(59.4 ± 0.4)° (c, 2% in methanol).

Analysis				
$C_{13}H_{24}N_2O_4$ (272.3)	Calculated:	C: 57.33	H: 8.88	N: 10.29
	Found:	C: 57.44	H: 9.04	N: 10.07

xxiii. N-Acetyl-L-leucyl-L-leucine methyl ester

L-Leucyl-L-leucine (0.0005 mole) was esterified by the method of Brenner and Huber (34). Crystallization of the oil obtained (from ethanol-ethyl ether) gave the methyl ester hydrochloride, fine white crystals (93%), m.p. 177-178°. One recrystallization from ethanol-ethyl ether gave L-leucyl-L-leucine methyl ester hydrochloride, m.p. 178.0-178.5°.

The ester hydrochloride was acetylated in the usual manner (xvii--procedure II) with acetic anhydride and sodium acetate. The product obtained was an oil which crystallized on standing (71% overall), m.p. 102-105°. Two recrystallizations from ethyl acetate-hexane gave thin white needles of N-acetyl-L-leucyl-L-leucine methyl ester, m.p. 108.5-110.5°, $[\alpha]_D^{27.5}$ -(66.0 ± 0.5)° (c, 1.5% in methanol).

Analysis				
$C_{15}H_{28}N_2O_4$ (300.4)	Calculated:	C: 59.97	H: 9.39	N: 9.33
	Found:	C: 59.83	H: 9.22	N: 9.35

xxiv. N-Acetyl-L-norvalyl-L-leucine methyl ester

Three ml. of a saturated solution of dry hydrogen bromide in glacial acetic acid (ca. 25% weight/weight) was added to N-carbobenzoxy-L-norvalyl-L-leucine methyl ester (0.0013 mole). The mixture, protected from atmospheric moisture with a drying tube containing calcium chloride, was stirred at room temperature for one hour. Evolution of carbon dioxide was noted immediately. The mixture was then treated with 50 ml. anhydrous ethyl ether. The ester hydrobromide which separated was filtered off, washed with ether and dried over phosphorous pentoxide. Recrystallization from methanol-ether gave L-norvalyl-L-leucine methyl ester hydrobromide, a white hygroscopic solid which was dried and stored over phosphorous pentoxide.

The ester hydrobromide was acetylated in the usual manner, using the "inverse addition" technique (xviii). The white crystalline product was obtained in 75% overall yield. Two recrystallizations from ethyl acetate-hexane gave N-acetyl-L-norvalyl-L-leucine methyl ester, m.p. 123.5-124.5°, $[\alpha]_D^{27.5}$ -(60.2 ± 0.4)° (c, 2% in methanol).

Analysis				
$C_{14}H_{26}N_2O_4$ (286.4)	Calculated:	C: 58.71	H: 9.15	N: 9.78
	Found:	C: 58.68	H: 8.95	N: 9.74

xxv. N-Acetyl-L-norleucyl-L-leucine methyl ester

N-Carbobenzoxy-L-norleucyl-L-leucine methyl ester (0.0013 mole) was decarbobenzoxylated with hydrogen bromide-glacial acetic acid in the same manner as that described in the preparation of N-acetyl-L-norvalyl-

L-leucine methyl ester (xxiv) to give L-norleucyl-L-leucine methyl ester hydrobromide (98%).

The ester hydrobromide was acetylated in the usual manner with acetic anhydride and sodium acetate. The white crystalline product was obtained in 86% overall yield, m.p. 123.0-124.0°. Two recrystallizations from ethyl acetate-hexane gave white needles of N-acetyl-L-norleucyl-L-leucine methyl ester, m.p. 125.5-126.5°, $[\alpha]_D^{27.5}$ -(53.2 ± 0.4)° (c, 2% in methanol).

Analysis				
	Calculated:	C: 59.97	H: 9.30	N: 9.33
$C_{15}H_{28}N_2O_4$ (300.4)	Found:	C: 59.90	H: 9.28	N: 9.11

xxvi. N-Acetyl-L-aspartyl(β-benzyl ester)-L-leucine methyl ester

N-t-butyloxycarbonyl-L-aspartic acid β-benzyl ester (0.0073 mole) and L-leucine methyl ester hydrochloride (0.0073 mole) were condensed (in chloroform solution) by the usual mixed-anhydride technique (42) using isobutylchloroformate (0.0073 mole). The conventional work-up procedure was employed and the product obtained (N-t-butyloxycarbonyl-L-aspartyl(β-benzyl ester)-L-leucine methyl ester) was a light yellow oil which crystallized on standing at 4° for two days. The yield of white crystals was 81%. The crude product was used in the following synthesis without any purification.

N-t-butyloxycarbonyl-L-aspartyl(β-benzyl ester)-L-leucine methyl ester (0.002 mole) was treated with 10 ml. saturated dry hydrogen chloride in glacial acetic acid (ca. 5%). The mixture (protected from atmospheric moisture by a drying tube filled with calcium chloride) was stirred for one hour at room temperature. The solvent was then removed under vacuum (1.5 mm. Hg at 35°). The last traces of solvent were removed at 300 μ Hg and room temperature, yielding a white hygroscopic glass, L-aspartyl(β-benzyl ester)-L-leucine methyl ester hydrochloride (96%). This product was used directly in the subsequent aceylation.

The ester hydrochloride was acetylated in the usual manner with acetic anhydride and sodium acetate with the "inverse-addition" technique (xviii), yielding a colorless oil which crystallized on standing at 4° (yield 90%, 86% overall). After one recrystallization, fine white crystals of N-acetyl-L-aspartyl(β-benzyl ester)-L-leucine methyl ester, m.p. 77.5-78.5°, were obtained.

xxvii. N-Acetyl-L-aspartyl-L-leucine methyl ester

A solution of N-acetyl-L-aspartyl(β-benzyl ester)-L-leucine methyl ester (xxvi) (0.001 mole) in 10 ml. methanol was added to a mixture of 150 mg. palladium black catalyst in 20 ml. methanol which had been stirring under hydrogen for one hour in an atmospheric-hydrogenation apparatus. The mixture was hydrogenated for 4 hours (one atmosphere, room temperature). The catalyst was then filtered off, washed with methanol, and the filtrate was evaporated to dryness under vacuum leaving a glass which was dried at 300 μ Hg over P_2O_5 (yield 97%). The product was hygroscopic and further manipulations were carried out (where possible) in a dry bag under dry nitrogen. Two recrystallizations from ethyl acetate-hexane gave fine white crystals of N-acetyl-L-aspartyl-L-leucine methyl ester, m.p. 108.5-109.5°. The compound was stored in a desiccator over phosphorous pentoxide.

Analysis	Calculated:	C: 51.65	H: 7.34	N: 9.27
$C_{13}H_{22}N_2O_6$ (302.3)	Found:	C: 51.81	H: 7.33	N: 9.18

xxviii. ε-N-*t*-butyloxycarbonyl-α-N-carbobenzoxy-*L*-lysyl-*L*-leucine methyl ester

ε-N-*t*-butyloxycarbonyl-α-N-carbobenzoxy-*L*-lysine (0.0066 mole) and *L*-leucine methyl ester hydrochloride (0.0066 mole) were condensed (in chloroform solution) by the usual mixed-anhydride technique (42) using isobutylchloroformate (0.0066 mole). The usual work-up procedure was employed, and the product obtained (89%) was a light yellow oil which crystallized on standing at room temperature under vacuum (300 μ Hg), m.p. 96.0-98.0°. This product was used in a subsequent synthesis without further purification.

xxix. α-N-Acetyl-ε-N-*t*-butyloxycarbonyl-*L*-lysyl-*L*-leucine methyl ester

A solution of ε-N-*t*-butyloxycarbonyl-α-N-carbobenzoxy-*L*-lysyl-*L*-leucine methyl ester (xxviii) (0.002 mole) in 10 ml. methanol containing glacial acetic acid (0.0024 mole) was added to a mixture of 300 mg. palladium black catalyst in 20 ml. methanol which had been stirring under hydrogen in an atmospheric-hydrogenation apparatus for one hour. The mixture was hydrogenated for 4 hours (one atmosphere, room temperature). The catalyst was filtered off, washed with methanol and the solvent was evaporated from the combined filtrate and wash under vacuum, leaving a colorless oil. Dry ethyl ether was added and then evaporated off, leaving a glass (95%). The crude ε-N-*t*-butyloxycarbonyl-*L*-lysyl-*L*-leucine methyl ester acetate was used directly in the acetylation after drying at 200 μ Hg over phosphorous pentoxide.

The ester acetate was acetylated in the usual manner with acetic anhydride and sodium acetate with the "inverse-addition" technique (xviii) giving the product in 75% yield (71% overall). One recrystallization from ethyl acetate-hexane gave α-N-acetyl-ε-N-t-butyloxycarbonyl-L-lysyl-L-leucine methyl ester, m.p. 99.0-101.0°.

xxx. α-N-acetyl-L-lysyl-L-leucine methyl ester hydrochloride

To α-N-acetyl-ε-N-t-butyloxycarbonyl-L-lysyl-L-leucine methyl ester (0.0013 mole) in a 10 ml. flask (fitted with a drying tube containing calcium chloride) was added 6 ml. glacial acetic acid saturated with dry hydrogen chloride (ca. 5%). Evolution of carbon dioxide was noted immediately. The solution was stirred at room temperature for one-half hour. Fifty ml. absolute ether was added and a gummy precipitate formed. After standing overnight at 10° (protected from moisture) the mixture was filtered under a dry nitrogen atmosphere. The white solid obtained was dried over phosphorous pentoxide. The sample was recrystallized three times from methanol-ethyl ether yielding a very hygroscopic white amorphous solid. Because of the extreme hygroscopicity, it was impossible to obtain an accurate melting point.

Analysis				
$C_{15}H_{30}N_3O_4Cl$ (351.9)	Calculated:	C: 51.20	H: 8.59	N: 11.94
	Found:	C: 48.37	H: 8.34	N: 12.09

In an attempt to purify this compound, the free amine was prepared from the hydrochloride in the manner of Paul and Anderson (44). Methylene chloride (15 ml.) was added to the hydrochloride (0.0005 mole). Dry ammonia was bubbled through the mixture. The suspension appeared to dissolve with subsequent formation of another precipitate. This precipitate (NH_4Br) was filtered off with the aid of sodium sulfate. Addition of ligroin (30-60) to the filtrate resulted in a fine white precipitate. The product was filtered off and was found to be extremely hygroscopic. All attempts to recrystallize this product resulted in a viscous oil. The NMR spectrum was consistent with the structure of α-N-acetyl-L-lysyl-L-leucine methyl ester, but all further attempts at purification (recrystallization, chromatography) were abortive.

xxxi. L-Leucine amide hydrochloride

This compound was prepared by the method of Greenstein and Winitz (45) yielding, after two recrystallizations from methanol-ether, thin plates, m.p. 235.0-236.0° (lit. m.p. 236.0-237.0°) (45), $[\alpha]_D^{26}$ +(9.3 ± 0.3)° (c, 5% in water) (lit. $[\alpha]_D^{25}$ + 9.5°) (c, 5% in water) (45). The pK_a of this amine hydrochloride was determined from a titration curve by the procedure of Almond, Kerr and Niemann (46). The value obtained was $pK_a(NH_3^+) = 7.95 \pm 0.03$.

N-Acetyl-L-phenylalanyl-L-leucine amide

N-Acetyl-L-phenylalanyl-L-leucine methyl ester (0.0005 mole) was allowed to stand at room temperature for 4 days in a sealed pressure bottle containing 5 ml. methanol saturated with anhydrous ammonia. The mixture was then concentrated to near-dryness under vacuum (25 mm. Hg and 35°). This evaporation was repeated several times after the addition, each time, of about 10 ml. methanol. The residual crystals were filtered with the aid of hexane. Two recrystallizations from ethanol-water gave white needles of N-acetyl-L-phenylalanyl-L-leucine amide, m.p. 253-254°.

Kinetic Studies

The kinetic parameters for the α-chymotrypsin-catalyzed hydrolyses of the substrates in this study were determined by a procedure described earlier (47), with modifications described below. The thermostatted reaction cell referred to in reference 47 has been modified to hold a 20 ml. beaker. All kinetic runs were conducted in aqueous solutions, 0.10 M in sodium chloride, at 25° and pH 7.90 ± 0.10.

The α-chymotrypsin used in these studies was a bovine, salt-free preparation obtained from Worthington Biochemicals (lot number C.D.I. 6066-67). The protein-nitrogen content was determined by this author and J. R. Rapp. The method employed was the micro-Kjeldahl procedure of Redeman (48). A value of 14.91 ± 0.06% was obtained (13). Values for enzyme concentrations employed in the determination of kinetic parameters were based on an enzyme molecular weight of 25,000 and a protein-nitrogen content of 16.5%.

The initial rates of all α-chymotrypsin-catalyzed hydrolyses considered in this study are described by equation:

$$v_o = \frac{k_o[E]_o[S]_o}{K_o + [S]_o} \qquad (25)$$

where v_o is the initial rate of the reaction, $[E]_o$ and $[S]_o$ are the initial concentrations of enzyme and substrate, respectively, and K_o and k_o are two experimentally obtained parameters.

Many of the substrates considered in this study are "fast" substrates which have relatively high values of k_o and low values of K_o. In order to conveniently and accurately follow the course of the reaction of these "fast" substrates with the pH-stat, it was necessary to care-fully choose initial concentrations of both substrate and enzyme.

1. Substrate Concentration

The initial substrate concentrations, $[S]_o$, which are varied over an eight- to nine-fold range should be chosen such that one of the intermediate values of $[S]_o$ approximates the value of K_o for the substrate. When this requirement is satisfied, the most accurate values of k_o and K_o are obtained.

2. Enzyme Concentrations

Two important factors must be considered in the choice of $[E]_o$, the initial substrate concentration:

a. Extent of Reaction

Integration of equation 25 yields the desired initial enzyme concentration as a function of initial substrate concentration, extent of reaction and the ratio K_o/k_o. This is shown in equation 26:

$$[E]_o = \frac{1}{\Delta t}\left\{\frac{K_o}{k_o}\left(\ln\frac{[S]_o}{[S]_t} + \frac{[S]_o - [S]_t}{K_o}\right)\right\} \tag{26}$$

where Δt is the length of time during which the reaction is followed, $[S]_o$ is the initial substrate concentration and $[S]_t$ is the substrate concentration at time t.

In order to avoid such complications as inhibition by products and excessive curvature of velocity traces, it is desirable to limit the extent of reaction as much as possible. It may be seen from equation 26 that a "fast" substrate (i.e., low K_o and high k_o) will require the use of extremely low enzyme concentrations if a reasonably small extent (i.e., < 20%) of the reaction is to be followed.

b. Adsorption of Enzyme on Glass Surfaces

A lower operational limit of about 10^{-8} M for initial enzyme concentration is established because α-chymotrypsin, like many proteins, is adsorbed on glass surfaces. This adsorption may result in atypical kinetic behavior when $[E]_o$ is lower than this limit (16,49).

In addition to the difficulties described above, there are other experimental problems associated with the determination of kinetic parameters for "fast" substrates such as those employed in the present study. The concentration of sodium hydroxide used to follow the reaction must be low (ca. 5×10^{-3} to 1×10^{-2} M). This results in significant blank reactions and difficulties in base standardization because of carbon dioxide absorption. In addition, because of the relatively high extents of reaction necessary and the correspondingly curved reaction velocity traces obtained, great care must be exercised to follow the initial part of the reaction in order to obtain accurate initial velocities.

Experimental Details

1. It was found that it was necessary to follow a relatively large extent of reaction (up to 25%) for the enzyme-catalyzed hydrolyses of many of the substrates studied. Therefore, extreme care was taken to "catch" the reaction as soon as possible (i.e., a few seconds) after enzyme was added to the reaction mixture.

2. Initial enzyme concentrations were often very low (ca. 10^{-7} to 10^{-8} M). In order to minimize adsorption, the stock enzyme solution was prepared so that its concentration was about 10^3 times as concentrated as the desired $[E]_o$. The enzyme stock solution was injected

into the 10 ml. reaction solution by means of a syringe delivering 0.996×10^{-2} ml. The syringe was rinsed 10-20 times with the stock solution prior to each delivery. At the end of each run, the electrodes and stirrer were washed with copious amounts of distilled water. The above procedure was employed whenever the initial enzyme concentration ($[E]_o$) was less than 1×10^{-6} M. Otherwise, the more conventional technique (i.e., injection of 1 ml. enzyme stock solution) (47) was used.

In all cases, the pH of the enzyme stock solution was adjusted to pH 7.9 immediately before injection into the reaction solution. Contrary to another report (16), this procedure was found to be necessary even for the more dilute enzyme solutions so that an accurate trace of the initial portion of the reaction could be obtained.

3. Although a stream of nitrogen was introduced into the reaction vessel above the solutions, carbon dioxide was still absorbed slowly by the basic reaction mixture. Because of the low concentrations of sodium hydroxide employed, this blank reaction was often significant. Because of its poor reproducibility, in most cases this blank reaction was determined on each run (i.e., at each substrate concentration) prior to addition of enzyme. This blank reaction included: the carbon dioxide blank, the substrate blank (i.e., the hydroxide-ion-catalyzed hydrolysis of the substrate) and the very small amount of reaction catalyzed by any enzyme left adsorbed on the electrodes or stirrer. Correction for this blank reaction was made on each run.

4. In most cases, the enzyme-catalyzed reaction was followed for only 4 minutes, rather than the conventional 8 minutes (47) in order to reduce the extent of reaction and to minimize adsorption of the

enzyme (see Appendix I of reference 16).

5. All stock solutions were thermostatted at 25.0° and were transferred to the reaction vessel as quickly as possible.

The primary kinetic data consisted of automatically recorded traces of the rates of production of acid from the α-chymotrypsin-catalyzed hydrolyses at several concentrations of the substrates. The data were corrected for the blank reaction as described above and analyzed with a Datatron 220 digital computer programmed as described previously (50).

The kinetic parameter k_o/K_o determined for N-acetyl-L-phenylalanyl-L-leucine amide was obtained in the manner described by Kurtz and Niemann (52) from the following equation:

$$\frac{k_o}{K_o} = \left(\frac{k_o}{K_o}\right)' \left\{1 + \frac{[BH_2^+]}{[BH_2]}\right\} \qquad (27)$$

where $(k_o/K_o)'$ is the observed parameter and $[BH_2^+]/[BH_2]$ is the ratio of the concentrations of the ionized and unionized forms of one of the products (L-leucine amide). This ratio may easily be obtained from the pKa(NH_3^+) value for L-leucine amide (see Syntheses) and the pH of the reaction solution (pH 7.90). This correction factor (i.e., $\{1 + [BH_2^+]/[BH]\}$) corrects for the effect of the buffering ability of the product on the parameter k_o (53).

The standard deviations for k_o/K_o parameters were determined from the individual parameters, k_o and K_o, and their standard deviations. The equation used is based on the method by which the individual parameters were determined (50,51) and is given by the following:

$$\Delta\left(\frac{k_o}{K_o}\right) = \frac{k_o}{K_o}\left\{\left(\frac{\Delta K_o}{K_o}\right)^2 - \left(\frac{\Delta k_o}{k_o}\right)^2\right\}^{\frac{1}{2}} \tag{28}$$

where $\Delta\ (k_o/K_o)$ is the standard deviation for the parameter k_o/K_o and ΔK_o and Δk_o are the standard deviations for K_o and k_o, respectively.

The errors in the ratios $(k_o/K_o)I/(k_o/K_o)II$ etc. (Tables VIII, IX and X) were calculated from the standard relationships (36):

$$\frac{A + a}{B + b} = C \pm \rho \quad \text{and} \quad \rho = \frac{1}{B}\left(a^2 + b^2\,\frac{A^2}{B^2}\right)^{\frac{1}{2}} \tag{29}$$

REFERENCES

1. J. T. Braunholtz and C. Niemann, Biochim. Biophys. Acta, **45**, 375 (1960).
2. J. T. Braunholtz, R. J. Kerr and C. Niemann, J. Am. Chem. Soc., **81**, 2852 (1959).
3. J. B. Jones, unpublished experiments in these laboratories.
4. D. W. Ingles and J. R. Knowles, Biochem. J., **96**, 71P (1965).
5. J. P. Wolf, III, R. A. Wallace, R. L. Peterson and C. Niemann, Biochemistry, **3**, 940 (1964).
6. T. Yamashita, J. Biochem. (Tokyo), **48**, 651 (1960).
7. J. B. Jones, T. Kunitake, C. Niemann and G. E. Hein, J. Am. Chem. Soc., **87**, 1777 (1965).
8. H. R. Waite and C. Niemann, Biochemistry, **1**, 250 (1962).
9. T. Kunitake, unpublished experiments in these laboratories.
10. G. E. Hein, J. B. Jones and C. Niemann, Biochim. Biophys. Acta, **65**, 353 (1962).
11. M. L. Bender and F. Kézdy, Ann. Rev. Biochem., **34**, 49 (1965).
12. J. P. Wolf, III and C. Niemann, Biochemistry, **2**, 82 (1963).
13. J. R. Rapp, Ph.D. Thesis, California Institute of Technology, 1964.
14. G. E. Hein and C. Niemann, Proc. Natl. Acad. Sci. U.S., **47**, 1341 (1962).
15. G. E. Hein and C. Niemann, J. Am. Chem. Soc., **84**, 4487, 4495 (1962).
16. R. L. Peterson, Ph.D. Thesis, California Institute of Technology, 1965.
17. A. Platt and C. Niemann, Proc. Natl. Acad. Sci. U.S., **50**, 817 (1963).
18. W. A. Mukatis, Ph.D. Thesis, California Institute of Technology, 1965.
19. J. R. Knowles, J. Theoret. Biol., **9**, 213 (1965).
20. H. Neurath and G. W. Schwert, Chem. Rev., **46**, 69 (1950).

21. J. P. Greenstein and M. Winitz, *Chemistry of the Amino Acids*, John Wiley and Sons, Inc., New York, N.Y., vol. 1, pp. 545-558 and references therein.

22. B. Zerner and M. L. Bender, *J. Am. Chem. Soc.*, **86**, 3669 (1964).

23. J. B. Jones and C. Niemann, *Biochemistry*, **1**, 1093 (1962).

24. B. Zerner, R. P. M. Bond and M. L. Bender, *J. Am. Chem. Soc.*, **86**, 3674 (1964).

25. T. Yamashita, *J. Biochem. (Tokyo)*, **48**, 846 (1960).

26. F. Daniels, J. H. Mathews, J. W. Williams, P. Bender and R. A. Alberty, *Experimental Physical Chemistry*, McGraw-Hill Book Co., Inc., 6th ed., 1962, p. 402.

27. N. A. Lange, *Handbook of Chemistry*, McGraw-Hill Book Co., Inc., 10th ed., 1961, p. 661.

28. C. L. Hamilton, C. Niemann and G. S. Hammond, *Proc. Natl. Acad. Sci. U.S.*, in press.

29. L. Pauling and R. B. Corey, *Proc. Natl. Acad. Sci. U.S.*, **37**, 251, 729 (1951).

30. B. S. Hartley, *Nature*, **201**, 1284 (1964).

31. M. L. Bender, F. J. Kézdy and C. R. Gunter, *J. Am. Chem. Soc.*, **86**, 3714 (1964).

32. N. Izumiya and T. Yamashita, *J. Biochem. (Tokyo)*, **46**, 19 (1959).

33. T. Yamashita and N. Izumiya, *J. Biochem. (Tokyo)*, **46**, 991 (1959).

34. M. Brenner and W. Huber, *Helv. Chim. Acta*, **36**, 1109 (1953).

35. T. H. Applewhite, H. Waite and C. Niemann, *J. Am. Chem. Soc.*, **80**, 1462 (1958).

36. See reference 21, vol. 2, p. 803.

37. R. E. Steiger, *Helv. Chim. Acta*, **17**, 563 (1934).

38. E. Fischer, *Ber.*, **39**, 2893 (1906).

39. L. A. Æ Sluyterman and H. J. Veenendaal, *Rec. Trav. Chim.*, **71**, 137 (1952).

40. See reference 27, p. 504.

41. J. Chouteau, *Compt. rend.*, **232**, 2314 (1951).

42. J. R. Vaughan and R. L. Osato, *J. Am. Chem. Soc.*, **74**, 676 (1952).

43. K. Luebke and E. Schroeder, *Ann.*, **665**, 205 (1963).

44. R. Paul and G. W. Anderson, *J. Org. Chem.*, **27**, 2094 (1962).

45. See reference 21, vol. 2, p. 1188.

46. H. R. Almond, R. J. Kerr and C. Niemann, *J. Am. Chem. Soc.*, **81**, 2856 (1959).

47. T. H. Applewhite, R. B. Martin and C. Niemann, *J. Am. Chem. Soc.*, **80**, 1457 (1958).

48. C. E. Redeman, *Ind. Eng. Chem.*, **11**, 635 (1939).

49. R. L. Bixler and C. Niemann, *J. Am. Chem. Soc.*, **81**, 1412 (1959).

50. H. I. Abrash, A. N. Kurtz and C. Niemann, *Biochim. Biophys. Acta*, **45**, 378 (1960).

51. K. A. Booman and C. Niemann, *J. Am. Chem. Soc.*, **78**, 3642 (1956).

52. A. N. Kurtz and C. Niemann, *Biochemistry*, **1**, 238 (1962).

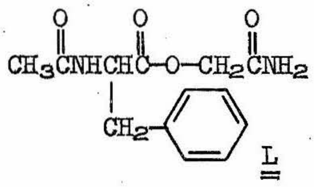
53. A. N. Kurtz, Ph.D. Thesis, California Institute of Technology, 1960.

III. THE STRUCTURAL SPECIFICITY OF α-CHYMOTRYPSIN: SOME REACTIVE ESTERS OF N-ACYLATED AMINO ACIDS AS SUBSTRATES

INTRODUCTION

In this chapter, attention will be focussed on the α-chymotrypsin-catalyzed hydrolyses of some reactive esters of N-acylated amino acids. The effects of the R_3' component of substrates of the form $R_1'CONHCHR_2COR_3'$ on reactivity and stereospecificity will be discussed in terms of the kinetic schemes and specificity theories outlined in the General Introduction. The implications of the possible contribution of $R_3-\rho_3$ interaction to binding and to reactivity of the enzyme-substrate complex will form the basis of much of the discussion.

Rather early in the study of the esterase activity of α-chymotrypsin, a search for a class of neutral, water soluble acylated α-amino acid esters capable of functioning as specific substrates for α-chymotrypsin led to the preparation, by Kerr and Niemann (1), of N-acetyl-L-phenylalanine glycolamide ester (I).

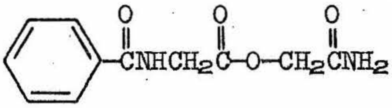


I

This compound indeed proved to be an excellent substrate for α-chymotrypsin (2). In fact, its reactivity was so great that, under reasonable experimental conditions, accurate determination of the kinetic parameters was impossible.

In a later study of the α-chymotrypsin-catalyzed hydrolyses of a series of analogues of N-acetylglycine methyl ester, Wolf and Niemann

(3) were unable to determine the kinetic parameters for N-benzoylglycine glycolamide ester(II). Two factors, a rapid hydroxide-ion-catalyzed



II

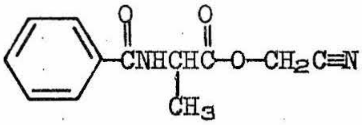
hydrolysis and an exceptionally fast enzyme-catalyzed reaction (for a glycine derivative) appeared to prevent this evaluation.

The glycolamide esters of N-acylated amino acids are really depsipeptide derivatives. They are analogous to N-acylated aminoacylglycine amides with an oxygen atom replacing the NH moiety of the peptide bond. Thus, investigation of these compounds appeared to present another approach to the study of the secondary specificity of α-chymotrypsin, more fully discussed in the preceding parts of this dissertation. Therefore, in part, the study discussed in this chapter represents an introductory, but very limited, attempt in the model study of the secondary specificity of α-chymotrypsin for residues which correspond to the amine component of the peptide bond in the α-chymotrypsin-catalyzed hydrolyses of polypeptides (see Part I).

RESULTS AND DISCUSSION

A new and successful attempt was made to evaluate the kinetic parameters for N-benzoylglycine glycolamide ester(II). The modifications of technique described in the Experimental section of Part II were employed. The results are displayed in Table I.

The glycolamide esters of three trifunctional substrates (N-benzoyl-L-alanine, N-benzoyl-D-alanine and N-acetyl-L-alanine) were also synthesized and studied as substrates of α-chymotrypsin. Because of the rather peculiar stereochemical behavior observed for the N-benzoylalanine glycolamide esters, the α-chymotrypsin-catalyzed hydrolyses of the cyanomethyl esters of the two antipodes of N-benzoylalanine (III) (the synthetic precursors of the corresponding glycolamide esters) were also studied.



D and L

III

The kinetic parameters for the above-mentioned substrates, as well as those for the corresponding methyl esters are reported in Table I.

Importance of R_3-ρ_3 Binding

There has been, and still is, a great deal of controversy concerning the importance of a binding interaction between the carboxyl component of a substrate of α-chymotrypsin (i.e., $R_3 = COR_3'$) and the

TABLE I

Experimental Conditions and Kinetic Parameters for the α-Chymotrypsin-catalyzed Hydrolyses of Various Esters of some N-Acylated Amino Acids[a]

Substrate	$[S]_o$ mM	$[E]_o$[b] 10^{-7} M	Max.[c] % Reaction	Number[d] of Expts.	k_o[b] sec^{-1}	K_o mM	k_o/K_o $M^{-1} sec^{-1}$
Benzoylglycine methyl ester[e]	—	—	—	—	0.225 ±0.006	5.3 ±0.4	42 ± 3
Benzoylglycine glycolamide ester	0.15 - 1.18	4.0	26	12-0	2.37 ±0.07	0.12 ±0.02	$(2.0 \pm 0.3) \times 10^4$
Benzoyl-L-alanine methyl ester[f]	1.9 -15.4	340	—	10-0	0.23 ±0.01	9.6 ±0.6	24.1 ± 0.3
Benzoyl-L-alanine glycolamide ester	0.097 -0.875	1.6	5	9-1	0.36 ±0.02	0.34 ±0.03	1050 ± 90
Benzoyl-L-alanine cyanomethyl ester	0.0356-0.321	1.32	16	9-1	0.428 ±0.007	0.065 ±0.004	6620 ± 380
Benzoyl-D-alanine methyl ester[f]	1.1 - 8.8	540	—	10-1	0.0071±0.0002	2.2 ±0.2	3.1 ± 0.2
Benzoyl-D-alanine glycolamide ester	0.043 -0.345	11.0	14	9-1	0.085 ±0.001	0.130 ±0.005	660 ± 20
Benzoyl-D-alanine cyanomethyl ester	0.0186-0.167	3.72	25	9-1	0.0934±0.0013	0.0133±0.0015	7020 ± 790
Acetyl-L-alanine methyl ester[g]	45 - 364	170	—	—	1.26 ±0.03	739 ± 28	1.71 ± 0.06
Acetyl-L-alanine glycolamide ester	7.27 -58.2	37	8	9-0	3.55 ±0.14	37.1±2.6	96 ± 5

TABLE I--*continued*

[a] In aqueous solutions at 25.0°, pH 7.90 ± 0.10 and 0.10 *M* in sodium chloride.

[b] Based on an enzyme molecular weight of 25,000 and a protein-nitrogen content of 16.5%.

[c] The extent of reaction during the time the reaction was followed (for the lowest initial substrate concentration).

[d] First number refers to the number of kinetic runs performed; second number refers to the number of points rejected by a statistical reiterative procedure.

[e] Parameters determined by J. R. Rapp (29).

[f] Parameters determined by W. A. Mukatis (31).

[g] Parameters determined by Jones, *et al.* (30).

active site of the enzyme. Indeed, this is perhaps the only major point of difference between the Bender specificity theory (4) and the Hein-Niemann (5) or Hamilton (6) treatments. Hein and Niemann (5) proposed that the rather dramatic decrease in K_o observed between an amide and an ester substrate was due primarily to enhanced R_3-ρ_3 interaction for the ester (as compared with the amide). On the other hand, Bender (7) attributes this decrease in K_o to the effects of a change in the rate-determining step (see p. 7). He states (7) that "the real K_s's of the L-ester and L-amide differ only by a small factor† and that the values of K_m(app)‡ of the ester and its corresponding amide are different because of the differing ratios $(k_3/(k_2 + k_3))$ in the two reactions." Furthermore, he considers that there are no differences in the K_s values for various esters of the same acylated amino acid (7,8). In the Bender-Kézdy treatment of specificity, the authors conclude (4) that "the effect of a variation of the group X [the carboxyl function] on k_{cat}/K_m(app),‡ appears in general, not to involve productive binding of X to the enzyme, but rather to reflect electronic influences of X." This statement is the basis for equation 1 which contains one of the terms important in the overall Bender-Kézdy description of specificity (4):

$$\log \frac{(k_o/K_o)_{R_1R_2X}}{(k_o/K_o)_{R_1R_2X_o}} = \sigma^*\rho^* \qquad (1)$$

---

†This factor K_s(amide)/K_s(ester) is, according to Bender (7) about 2.9 for N-acetyl-L-tryptophan derivatives and about 5 for N-acetyl-L-phenylalanine derivatives.

‡K_m(app) and k_{cat} in Bender's terminology are equivalent to K_o and k_o in the terminology of this dissertation.

where the terms are those explained in the General Introduction (p. 14). It might be noted that if binding of X (i.e., R_3) to the active site is important, equation 1 would have to be modified to include some binding factor which is dependent on the nature of X(R_3). There is some evidence which suggests that R_3-ρ_3 binding is a factor in many cases. This evidence will be presented below.

The most extensive set of data to which equation 1 applies is that for several esters of hippuric acid (i.e., N-benzoylglycine). Pertinent data for several of these substrates are presented in Table II. The k_o/K_o data are those of Nelson and co-workers (9), Epand and Wilson (10) and Zerner (11) (as tabulated by Zerner and Bender (7)). The k_o/K_o value for the glycolamide ester was obtained from results of the present study (Table I). The Taft substituent constants (σ^*) are those tabulated by Leffler and Grunwald (12) except for the values for $-C_6H_4NO_2$, $-CH_2CONH_2$ and $-CH_2CH_2N^+(CH_3)_3$ which were calculated as described below.

The rate of the hydroxide-ion-catalyzed hydrolysis (H_2O, 25°) of p-nitrophenyl acetate is about 57 times greater than that for methyl acetate (4). The ρ^* value for this reaction is +1.47 (13). The σ^* value for $-CH_3$ is zero (12) and thus:

$$\sigma^*_{p\text{-}NO_2\phi-} \approx \frac{\log 57}{1.47} = +\,1.20 \qquad (2)$$

The ρ^* value for the ionization of carboxylic acids (H_2O, 25°) is + 1.72 (13). The pK_a values for acetic and carboxamidoacetic acids (i.e., CH_3CO_2H and $NH_2COCH_2CO_2H$) are 4.756 (14,15,16) and 3.64 (16,17),

TABLE II

Kinetic Parameters for the α-Chymotrypsin-catalyzed Hydrolyses of some Esters of Hippuric Acid[a]

Substrate $\phi CONHCH_2CO_2R$ R	k_o/K_o $M^{-1}sec^{-1}$	log k_o/K_o	$\sigma_R{}^{*b}$
$-CH_3$	60 39	1.78 1.59	0.00
$-C_2H_5$	42	1.62	-0.10
$-n-C_3H_7$	67	1.83	-0.115
$-i-C_3H_7$	22 23	1.34 1.36	-0.190
$-n-C_4H_9$	156	2.19	-0.130
$-i-C_4H_9$	69	1.84	-0.125
$-(CH_2)_2\overset{+}{N}(CH_3)_3$	890	2.95	+0.71
$-C_6H_4NO_2(p)$	~5700	3.75	+1.20
$-CH_2CONH_2$	2×10^4	4.30	+0.65

[a] In aqueous solutions at pH 7.90 and 25.0°. Data from references 9, 10 and 11 (as tabulated in reference 7).

[b] Taft substituted constants for the alcohol component of the esters. See text for references.

respectively. Again the σ^* value for $-CH_3$ is zero and therefore

$$\sigma^*_{-CH_2CONH_2} \approx \frac{4.756 - 3.64}{1.72} = + 0.65 \qquad (3)$$

From the rate constants for the hydroxide-ion- and hydrogen-ion-catalyzed hydrolyses (H_2O, 25°) of acetyl choline (18) and ethyl acetate (18,19), $\sigma^*_{-CH_2CH_3} = -0.10$ and $\rho^* = +1.47$ (13) and the Taft equation for this reaction (13):

$$\log \left(\frac{k}{k_o}\right)_{OH^-} - \log \left(\frac{k}{k_o}\right)_{H^+} = + 1.47\ \sigma^* \qquad (4)$$

a σ^* value of + 0.85 for $R = -CH_2CH_2\overset{+}{N}(CH_3)_3$ is obtained.

A Taft linear free energy plot of $\log (k_o/K_o)$ for the substrates of Table II against the Taft substituent constants for the alcohol components of the esters ($\sigma_R{}^*$) is displayed in Figure 1. Although the scatter in the data for the simple alkyl esters is quite large, if the glycolamide ester point is considered anomalous, a line of correlation may be drawn, yielding a ρ^* value of about + 1.6. This value, it will be noted, is reasonably close to that obtained for the hydroxide-ion-catalyzed hydrolysis of acetates (i.e., +1.47 (13)). As reported by Bender (4), a similar treatment for esters and amides of N-acetyl-L-tryptophan and N-acetyl-L-phenylalanine yields similar results, with ρ^* values close to that for the hydroxide-ion-catalyzed hydrolysis of acetates. The most striking deviation from the line drawn (Figure 1), is the extremely high k_o/K_o value of the glycolamide ester. This derivative is about 20 times more reactive (using k_o/K_o as a criterion) than

Free Energy Plot of log (k_o/K_o) for the α-Chymotrypsin-catalyzed Hydrolyses of a Series of Esters of N-Benzoylglycine vs. the Taft Substituent Constants (σ^*_R) for the Alcohol Components of the Esters

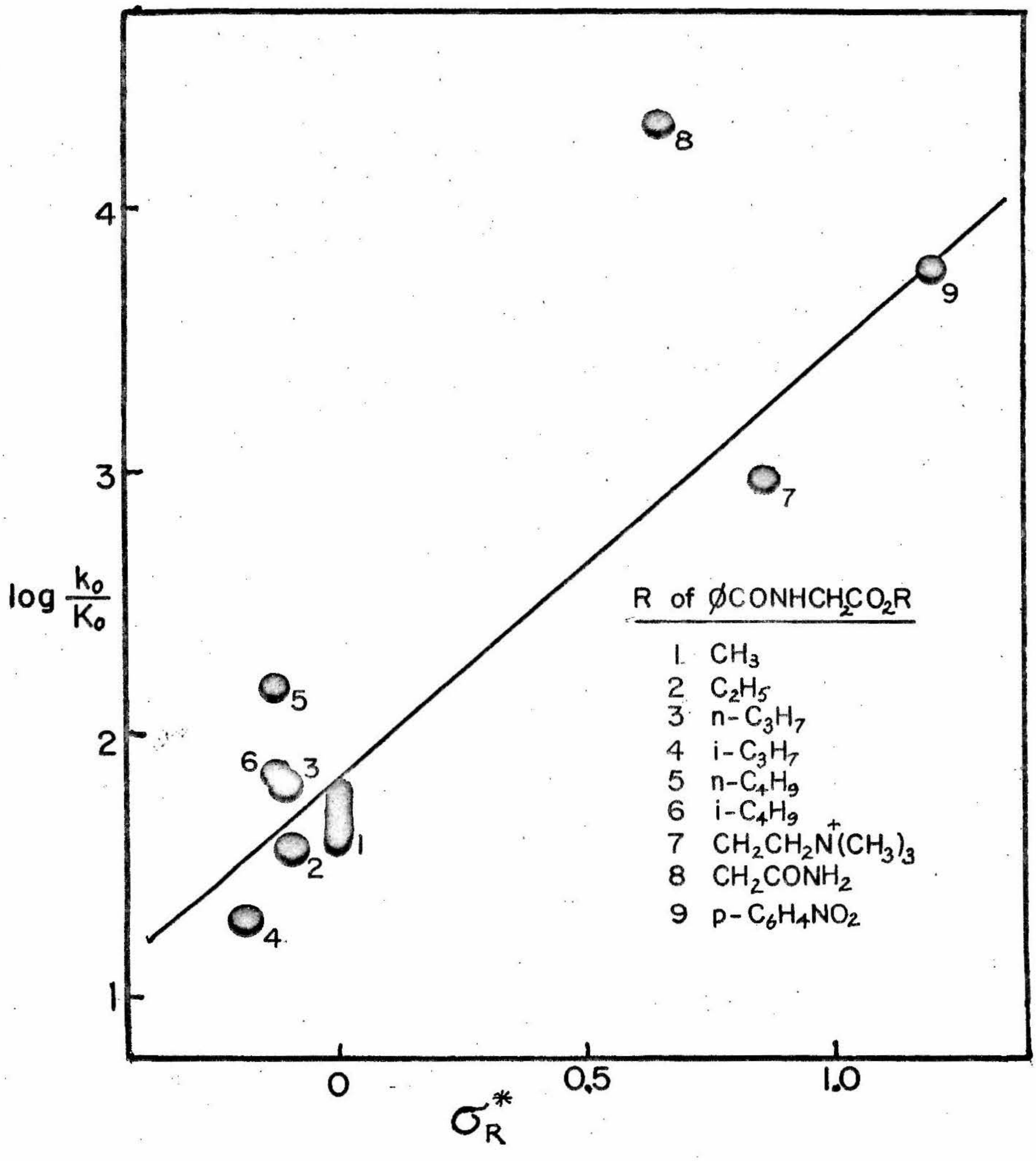


Figure 1

the choline bromide ester although the σ^* value for the latter compound is slightly greater (+0.85) than that for the glycolamide ester (+0.65). Indeed, the glycolamide ester is more than three times as reactive as the p-nitrophenyl ester which has a σ^* value almost twice that for the glycolamide ester. The only reasonable explanation for this phenomenon is some sort of binding interaction between the glycolamide ester moiety (R_3) and the enzyme. This enhanced R_3-ρ_3 binding might increase reactivity in one or both of the following ways: (a) it might decrease the parameter K_o because of the beneficial effect of the R_3-ρ_3 interaction on K_s or (b) it might increase k_o because of an improved orientation of the hydrolyzable function (COR_3') at the active center (i.e., an increase in k_2). Both of these cases imply some kind of R_3-ρ_3 interaction.

With regard to the alkyl esters, Bender (7) has calculated, with the aid of certain approximations, the "true" K_s values for these substrates. Although the variations in the K_s values obtained are not large, Knowles (20) has been able to correlate these values (i.e., log K_s) with the hydrophobic character of the alkyl substituent. He found that K_s decreases (i.e., binding increases) as the non-polar character of the alcohol moiety of the ester increases and he suggests that R_3-ρ_3 interaction with some hydrophobic character does contribute to binding. This is contrary to the view of Zerner and Bender (7) that K_s is essentially invariant with change of R_3.

The importance of an R_3-ρ_3 interaction is also suggested by some kinetic data for the α-chymotrypsin-catalyzed hydrolyses of two esters of N-acetyl-L-norvaline. The k_o values for the methyl and isopropyl esters of this amino acid derivative are 2.70 ± 0.17 sec^{-1} and 2.92 ± 0.04 sec^{-1}, respectively (21). In terms of the acyl-enzyme scheme (see p. 6):

$$k_o = \frac{k_2\ k_3}{k_2 + k_3} \qquad (5)$$

Thus,

$$2.70 = \frac{k_2\ k_3}{k_2 + k_3} \qquad (6) \qquad\qquad 2.90 = \frac{k_2'\ k_3}{k_2' + k_3} \qquad (7)$$

where k_2 and k_2' are the specific rates of acylation for the methyl and isopropyl esters, respectively, and k_3 is the specific rate of deacylation for the common N-acetyl-L-norvalyl-chymotrypsin. It is apparent from equations 6 and 7 that $k_2' > k_2$. Because the σ_R^* value for R = $-CH(CH_3)_2$ equals -0.19 (12) and that for R = $-CH_3$ is zero (12), the implication is that ρ^* for acylation must be negative. The ρ^* values for $k_o/K_o = k_2/K_s$ for substrates of α-chymotrypsin are usually positive (4) (i.e., $\rho^* \approx +1$ to +3). In terms of the Bender-Kézdy theory which assumes a constant K_s value, a positive ρ^* value for acylation would be expected. The results for these two esters of N-acetyl-L-norvaline are entirely contrary to this. This data suggests that an R_3-ρ_3 interaction may indeed affect the rate of acylation (k_2).

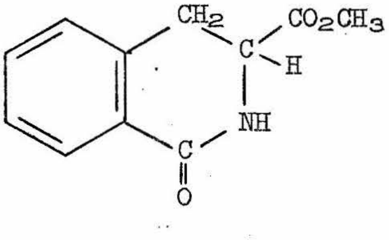
Additional evidence supporting the importance of the R_3-ρ_3 interaction is also available from the observation by Isaacs and Niemann (22) that acetyl-α-chymotrypsin stereospecifically acetylates one enantiomer of butan-2-ol and from observations of Balls and co-workers (23,24,25) on the effect of alcohols on the hydrolysis of p-nitrophenyl acetate by α-chymotrypsin.

Thus, although the magnitude of its effect is uncertain, the importance of an R_3-ρ_3 interaction to the observed kinetic behavior of substrates of α-chymotrypsin cannot be denied.

Stereospecificity

Undoubtedly, the usual antipodal specificity of α-chymotrypsin is for the L-antipode of amino acid derivatives. However, the view of an obligatory absolute stereospecificity in favor of the L-antipodes began to lose ground when it became evident that stereochemical preference was relative rather than absolute (5a), and indeed it was negated when, in several instances, the D-antipodes were found to be more reactive (5a,26,27,28).

In 1960, an inversion of the usual antipodal stereospecificity of α-chymotrypsin was demonstrated (28) when it was observed that the D-antipode of 3-carbomethoxydihydroisocarbostyril(IV) was hydrolyzed in an α-chymotrypsin-catalyzed reaction at a rate much greater than the corresponding L-antipode. A theory (5) developed to explain the above



IV

observations also predicted, in general terms, the existence of other examples of diminished stereochemical specificity and even inversion for substrates of the type $R_1'CONHCHR_2COR_3'$.

In a recent communication, Rapp and Niemann (27) describe an inversion of the usual antipodal specificity in the α-chymotrypsin-catalyzed hydrolysis of N-picolinylalanine methyl ester. This result has been at least partially explained in terms of the above theory (29).

The kinetic behavior of some N-acylated alanine methyl esters illustrates the effect of the R_1' group (in $R_1'CONHCHR_2COR_3'$) on stereospecificity. The kinetic parameters for N-acetyl- and N-benzoyl-D- and L-alanine methyl esters are given in Table III. The ratio of k_o/K_o values for the L- and D-antipodes of an N-acylated amino acid derivative is an excellent measure of stereospecificity. The ratio $(k_o/K_o)_{\underline{L}}/(k_o/K_o)_{\underline{D}}$ is a measure of the preference for the L-antipode over the D-enantiomer in an α-chymotrypsin-catalyzed reaction. This ratio is greater than 570 for the N-acetylalanine methyl esters but is only about 8 for the N-benzoyl-derivative (Table III). This indicates a substantial decrease in stereospecificity for the N-benzoyl-derivatives. The diminished stereochemical specificity for the N-benzoyl-derivatives (as compared to the N-acetyl-derivatives) has been attributed to the effects of non-productive binding. The results have been explained in terms of the increased importance of a non-productive binding mode in the N-benzoyl-L-substrate.

In terms of the Hein-Niemann theory (5), the productive mode for the L-antipode of a trifunctional substrate has been proposed to be that involving R_1-ρ_1, R_2-ρ_2, R_3-ρ_3 and H-ρ_H interactions (Figure 2a). The productive mode for the D-antipode is that involving R_1-ρ_2, R_2-ρ_1, R_3-ρ_3 and H-ρ_H interactions (Figure 2b). One possible mode of binding for the L antipode (its importance depending on the nature of R_1) might

TABLE III

Kinetic Parameters for the α-Chymotrypsin-catalyzed Hydrolyses of N-Acetyl- and N-Benzoyl-D- and L-Alanine Methyl Esters[a]

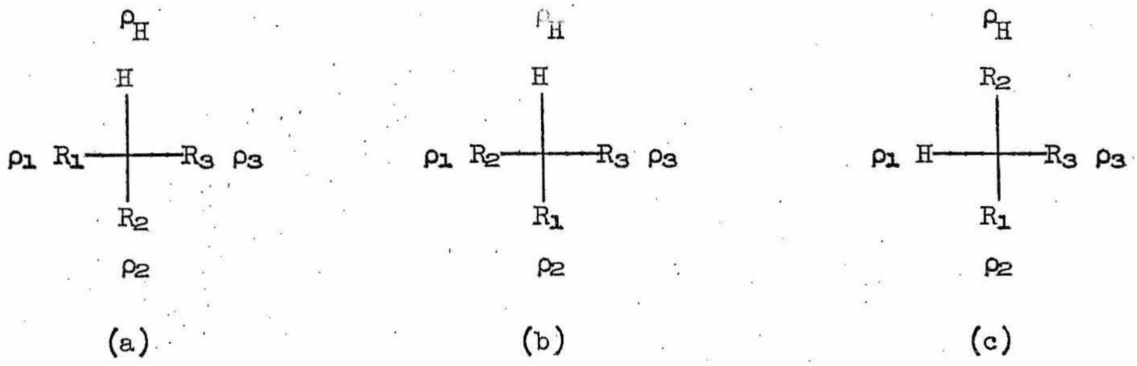
Substrate	k_o sec^{-1}	K_o mM	k_o/K_o $M^{-1}sec^{-1}$	$\frac{(k_o/K_o)_L}{(k_o/K_o)_D}$[c]	Ref.
N-Acetyl-L-alanine methyl ester	1.26	739	1.71	> 570	30
N-Acetyl-D-alanine methyl ester[b]	< 0.001	~ 300	< 0.003		29
N-Benzoyl-L-alanine methyl ester	0.23	9.6	24	7.8	31
N-Benzoyl-D-alanine methyl ester	0.0071	2.2	3.1		31

[a] In aqueous solutions at 25.0°, pH 7.90, 0.10 M in sodium chloride. Data from Table I and reference 29.

[b] Estimate of parameters from reference 29.

[c] Relative reactivity of a D-L antipodal pair.

be that with $R_1-\rho_2$, $R_2-\rho_H$, $R_3-\rho_3$ and $H-\rho_1$ interactions (Fig. 2c). This mode, although it has an $R_3-\rho_3$ interaction, will, for alanine derivatives have a group larger than hydrogen (i.e., $R_2 = -CH_3$) in ρ_H and thus from one of the experimentally supported postulates of the above theory (5a), will be non-productive.



L-antipode
Productive Mode

D-antipode
Productive Mode

L-antipode
Non-productive Mode

Figure 2

Thus, a modification of R_1 (i.e., $R_1 = R_1'CONH-$) such that its affinity for ρ_2 is enhanced, should serve to increase the reactivity of the D-antipode (i.e., lower K_o and raise k_o). However, for the L-antipode (because of the increased importance of non-productive binding) a decrease in both K_o and k_o should occur (see p. 6). This is indeed what is observed when the N-acetamido- group of the N-acetylalanine methyl esters is replaced by an R_1 group (i.e., the benzamido- group) which, because of its increased hydrophobic character, would be expected to interact strongly with the ρ_2 locus (see Part II, p. 112).

The interpretation of the results obtained for esters of N-benzoyl-D- and L-alanine (Table I) will be made in terms of the binding modes represented in Figure 2.

Effects of the Alcohol Component (R_3') of N-Acylated Amino Acid Esters on Reactivity and Stereospecificity

In Table IV, the kinetic parameters for the substrates of Table I are expressed relative to the values for the methyl esters. In Table V, the $(k_o/K_o)_{\underline{L}}/(k_o/K_o)_{\underline{D}}$ ratios for the methyl, glycolamide and cyanomethyl esters of the N-benzoylalanines are presented.

The ratio $(k_o/K_o)_{\underline{L}}/(k_o/K_o)_{\underline{D}}$ may be regarded as a measure of stereoselectivity. From Table V, it is observed that whereas N-benzoyl-L-alanine methyl ester is about 8 times as reactive as its corresponding D-antipode, the L-glycolamide ester is only 1.6 times as reactive as its D-antipode, and indeed, for the cyanomethyl esters, the D-substrate appears to be slightly more reactive than the L-antipode (although when experimental error is considered, the reactivities of the latter two substrates may be considered to be equal). Thus, as the reactivity of both the antipodes increases (Table IV), a decrease in stereospecificity is observed to the point where, in the case of the cyanomethyl esters, the D-antipode is just as reactive, if not more so, than the L-antipode.

Accommodation of this behavior (i.e., a change in stereospecificity as a function of the alcohol portion of an N-acylated amino acid ester) in terms of the specificity theories outlined in the General Introduction, presents some difficulties.

In terms of the Bender-Kézdy specificity theory (4);

$$\log \frac{(k_o/K_o)_{\underline{L}-X}}{(k_o/K_o)_{\underline{L}-X_A}} = \rho_{\underline{L}}^* \ (\sigma_X^* - \sigma_{X_A}^*) \qquad (8)$$

and

$$\log \frac{(k_o/K_o)_{\underline{D}-X}}{(k_o/K_o)_{\underline{D}-X_A}} = \rho_{\underline{D}}^* \ (\sigma_X^* - \sigma_{X_A}^*) \qquad (9)$$

TABLE IV

Kinetic Parameters of Table I Expressed Relative to the Respective Methyl Esters

Relative k_o Values

Acyl Component[a]	Alcohol Component[b] $-OCH_3$	$-OCH_2CONH_2$	$-OCH_2CN$
N-Benzoylglycyl-	1.0	10.5	—
N-Benzoyl-L-alanyl-	1.0	1.6	1.9
N-Benzoyl-D-alanyl-	1.0	12.0	13
N-Acetyl-L-alanyl-	1.0	2.8	—

Relative K_o Values

Acyl Component[a]	Alcohol Component[b] $-OCH_3$	$-OCH_2CONH_2$	$-OCH_2CN$
N-Benzoylglycyl-	1.0	0.023	—
N-Benzoyl-L-alanyl-	1.0	0.035	0.0068
N-Benzoyl-D-alanyl-	1.0	0.059	0.0060
N-Acetyl-L-alanyl-	1.0	0.050	—

Relative k_o/K_o Values

Acyl Component[a]	Alcohol Component[b] $-OCH_3$	$-OCH_2CONH_2$	$-OCH_2CN$
N-Benzoylglycyl-	1.0	470	—
N-Benzoyl-L-alanyl-	1.0	44	275
N-Benzoyl-D-alanyl-	1.0	210	2260
N-Acetyl-L-alanyl-	1.0	56	—

---

[a] Acyl component refers to $R_1'CONHCH(R_2)CO-$ in the substrate $R_1'CONHCH(R_2)COR_3'$.

[b] Alcohol component refers to R_3' in the substrate $R_1'CONHCH(R_2)COR_3'$.

TABLE V

The Stereospecificity of the α-Chymotrypsin-catalyzed Hydrolyses of Esters of N-Benzoyl-D- and L-Alanine

Alcohol Component (R_3') of $\phi CONHCH(CH_3)CO_2R_3'$	$(k_o/K_o)_L/(k_o/K_o)_D$[a]
$-CH_3$	7.8 ± 0.5
$-CH_2CONH_2$	1.60 ± 0.15
$-CH_2C{\equiv}N$	0.94 ± 0.12

[a] Ratio of the k_o/K_o values (Table I) for the L- and D-antipodes considered.

where the subscripts L-X and $\underline{L}$-X_A refer to two different esters of N-benzoyl-$\underline{L}$-alanine, the subscripts $\underline{D}$-X and $\underline{D}$-X_A refer to the corresponding $\underline{D}$-antipodes, $\rho_L{}^*$ and $\rho_D{}^*$ are the Taft reaction parameters for the $\underline{L}$- and $\underline{D}$-antipodes, respectively and $\sigma_X{}^*$ and $\sigma_{X_A}{}^*$ are the Taft substituent constants for the alcohol portions of the two esters considered.

From equations 8 and 9, the following equation may be obtained:

$$\log \frac{(k_o/K_o)_{\underline{L}-X}}{(k_o/K_o)_{\underline{D}-X}} = (\sigma_X{}^* - \sigma_{X_A}{}^*)(\rho_{\underline{L}}{}^* - \rho_{\underline{D}}{}^*) + \log \frac{(k_o/K_o)_{\underline{L}-X_A}}{(k_o/K_o)_{\underline{D}-X_A}} \qquad (10)$$

Therefore, a plot of the term in the left hand side of equation 10 vs. ($\sigma_X{}^* - \sigma_{X_A}{}^*$) when X_A refers to the methyl esters should yield a straight line of slope ($\rho_{\underline{L}}{}^* - \rho_{\underline{D}}{}^*$). Such a plot for the three esters of the N-benzoylalanines is shown in Figure 3 (data from Table V). From this figure (in which errors in the k_o/K_o ratios are considered) it is observed that the three points available do not correlate particularly well with a straight line. A straight line giving the best fit yields a value of $\rho_L{}^* - \rho_D{}^*$ of about -0.7. Whether or not a straight line should indeed be drawn, is a moot point, and later discussion will suggest that it should not. The assumption that Bender makes regarding constancy of K_s for various esters has already been questioned (pp. 160-169) and later discussion will suggest that orientation of COR_3' (X) at the active site is an important factor. Thus, changes in k_2 with COR_3' will reflect this latter factor as well as the electronic effects of the alcohol component (R_3').

Stereospecificity of the α-Chymotrypsin-catalyzed Hydrolyses of Three Esters of N-Benzoylalanine as a Function of the Taft Substituent Constant for the Alcohol Component of the Ester

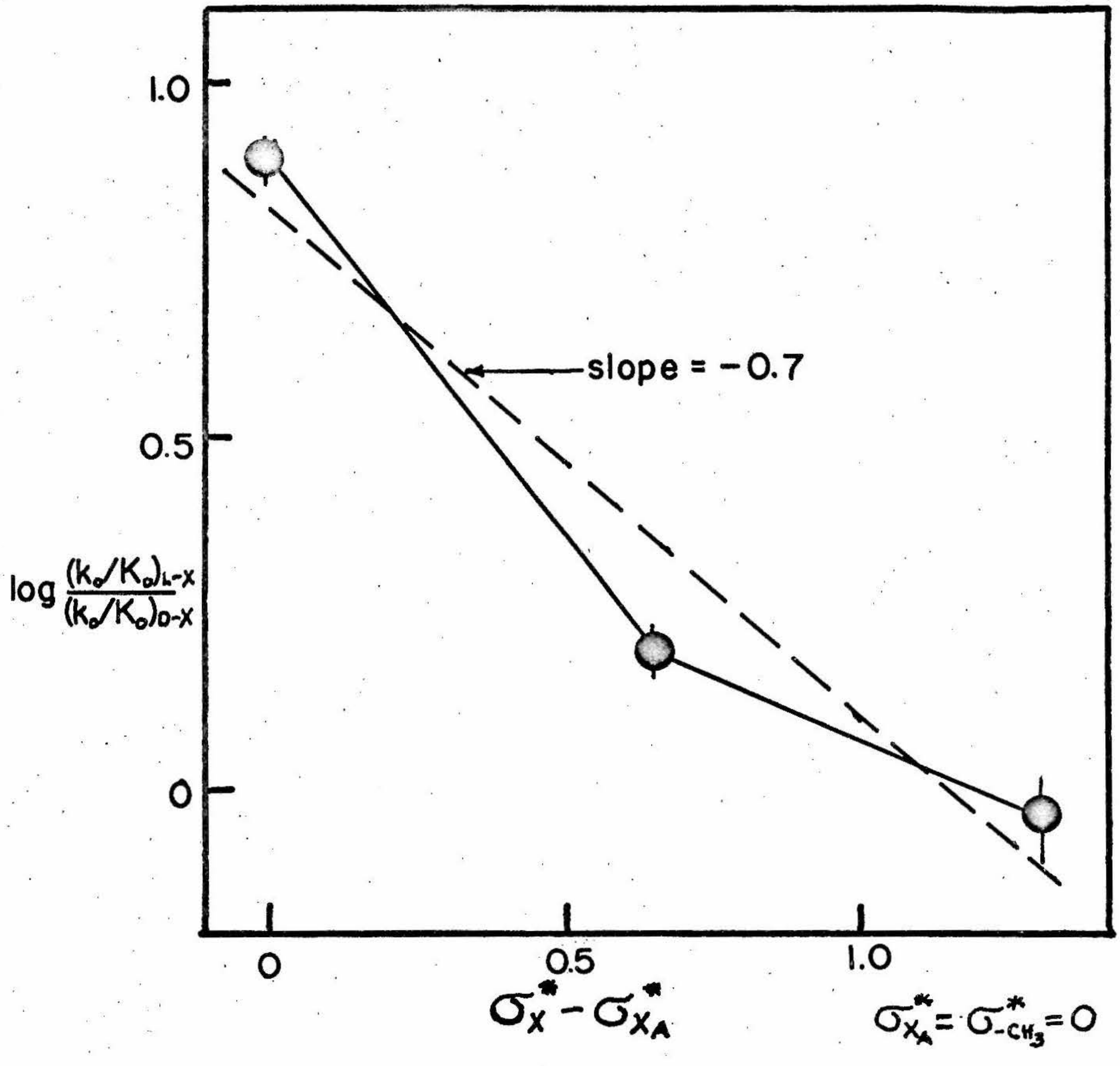


Figure 3

In terms of the Hamilton extension of the Hein-Niemann theory (6), relative stereospecificity may be expressed in terms of $R_{\underline{L}/\underline{D}}$ where:

$$R_{\underline{L}/\underline{D}} = \frac{(k_o/K_o)_{\underline{L}\text{-}R_3}}{(k_o/K_o)_{\underline{D}\text{-}R_3}} \Bigg/ \frac{(k_o/K_o)_{\underline{L}\text{-}R_{30}}}{(k_o/K_o)_{\underline{D}\text{-}R_{30}}} \tag{11}$$

where $\underline{L}$-R_3 and $\underline{D}$-R_3 refer to some ester of N-benzoyl-$\underline{L}$- and $\underline{D}$-alanine, respectively, and $\underline{L}$-R_{30} and $\underline{D}$-R_{30} refer to reference esters of the two antipodes of N-benzoylalanine.

When the productive modes for the two antipodes are represented as in Figure 2, $R_{\underline{L}/\underline{D}}$ is given by equation 12 (see p. 13).

$$R_{\underline{L}/\underline{D}} = \frac{k_{2\underline{L}\text{-}R_3}}{k_{2\underline{D}\text{-}R_3}} \cdot \frac{k_{2\underline{D}\text{-}R_{30}}}{k_{2\underline{L}\text{-}R_{30}}} \cdot \frac{\bar{K}_{11}\,\bar{K}_{22}\,\bar{K}_{33}\,\bar{K}_{HH}}{\bar{K}_{12}\,\bar{K}_{21}\,\bar{K}_{33}\,\bar{K}_{HH}} \cdot \frac{\bar{K}_{12}'\,\bar{K}_{21}'\,\bar{K}_{33}'\,\bar{K}_{HH}'}{\bar{K}_{11}'\,\bar{K}_{22}'\,\bar{K}_{33}'\,\bar{K}_{HH}'} \tag{12}$$

where the $\bar{K}_{ij}$'s represent the "microscopic binding factors" for productive modes of the $\underline{L}$-R_3 and $\underline{D}$-R_3 substrates and the $\bar{K}_{ij}'$'s represent the corresponding binding factors for the $\underline{L}$-R_{30} and $\underline{D}$-R_{30} reference substrates.

If the $\bar{K}_{ij}$'s and $\bar{K}_{ij}'$'s are independent of each other (i.e., if the contributions of each R-ρ interaction to the overall binding of the substrate are independent) then $\bar{K}_{11} = \bar{K}_{11}'$, $\bar{K}_{22} = \bar{K}_{22}'$, $\bar{K}_{12} = \bar{K}_{12}'$, $\bar{K}_{21} = \bar{K}_{21}'$ and equation 12 reduces to:

$$R_{\underline{L}/\underline{D}} = \frac{k_{2\underline{L}\text{-}R_3}}{k_{2\underline{D}\text{-}R_3}} \cdot \frac{k_{2\underline{D}\text{-}R_{30}}}{k_{2\underline{L}\text{-}R_{30}}} \tag{13}$$

This is a result which is very similar to that obtained earlier (p.177) from the Bender-Kézdy treatment. Thus, if $k_{2\underline{L}\text{-}R_3}/k_{2\underline{D}\text{-}R_3}$ is a constant independent of R_3, then $R_{\underline{L}/\underline{D}}$ should also be a constant independent of R_3. From Table V, it is obvious that $R_{\underline{L}/\underline{D}}$ is not a constant, and thus the implication is that $k_{2\underline{L}-R_3}/k_{2\underline{D}-R_3}$ is indeed dependent on the nature of R_3.

The explanation which will be proposed to account for the observed results indeed is concerned with the suggestion that $k_{2\underline{L}-R_3}/k_{2\underline{D}-R_3}$ is a function of R_3. Moreover, it will be suggested that this change in $k_{2\underline{L}-R_3}/k_{2\underline{D}-R_3}$ with changing R_3 is a result of the effect of the R_3-ρ_3 interaction on the orientation of the ester function at the active site.

If the data are regarded such that the cyanomethyl esters represent the "true" stereospecificity of the enzyme-catalyzed hydrolysis and the methyl esters represent an anomalous increase in stereospecificity, the following interpretation of this increase may be considered. The productive binding modes proposed for the two antipodes of the esters of N-benzoyalanine are represented by the following interactions: L-antipode--ϕCONH-ρ_1, CH_3-ρ_2, R_3-ρ_3 and H-ρ_H; D-antipode--CH_3-ρ_1, ϕCONH-ρ_2, R_3-ρ_3 and H-ρ_H (Fig. 2).

It is proposed that, although the productive mode for the D-antipode represents a favorable binding mode and is the only productive mode available to this antipode, for the methyl ester it is not a particularly good productive mode because the dominant R_1-ρ_2 interaction results in a poor orientation of R_3 at the active center. However, in the case of the more reactive glycolamide and cyanomethyl esters, it is proposed that effective R_3-ρ_3 interaction alleviates this situation of poor orientation, resulting in greatly increased reactivity of the productive mode for the D-antipode.

The benzamido- (ϕCONH-) group resembles the benzyl side chain (ϕCH_2-) of the very well-bound and reactive phenylalanine substrates (32), and indeed, it is because of this similarity that non-productive binding of N-benzoyl-L-alanine derivatives is of such great importance. However, it must be noted that the phenyl ring of the benzamido- group is one atom further from the α-carbon of the amino acid than it is in the benzyl side chain. Because the L-phenylalanine derivatives (which bind to α-chymotrypsin with a benzyl-ρ_2 interaction) have very high k_0 (and indeed k_2) values (8,32) it is reasonable to assume that a ϕCH_2-ρ_2 interaction provides an optimum R_3 orientation and that the R_3 orientation in a substrate with a $\phi CONH$-ρ_2 interaction represents some deviation from this optimum. It is worthwhile noting that the methyl ester of β-phenylpropionic acid ($\phi CH_2CH_2CO_2CH_3$) is a substrate for α-chymotrypsin ($k_0 \simeq 0.018$ sec^{-1}, $K_0 = 3.9$ mM) (33). Indeed, this substrate has a higher k_0 value than N-benzoyl-D-alanine ($k_0 = 0.0071$) sec^{-1}; $K_0 =$ 2.2 mM (31)). Because this substrate lacks the carboxamido function of acylated amino acid derivatives, ϕCH_2-ρ_1 interaction would be unlikely, and thus the productive mode for this substrate would undoubtedly involve a ϕCH_2-ρ_2 interaction. In spite of the lack of an R-ρ_1 interaction (considered by Hein and Niemann (5) to be important for orientation of R_3 at the active center) this substrate still has a higher k_0 value than that for N-benzoyl-D-alanine methyl ester. This suggests strongly that the productive binding mode for the latter substrate does not provide particularly good orientation of R_3 at the active center.

It is therefore proposed that an increased R_3-ρ_3 interaction (e.g., in the glycolamide and cyanomethyl esters) undoubtedly at some expense of the $\phi CONH$-ρ_2 interaction, provides a more nearly optimum orientation of the ester carbonyl group of the D-antipode at the active center of the enzyme, resulting in an increase in reactivity greater than that which would be expected solely on the basis of electronic effects.

For the corresponding L-antipodes, this R_3-ρ_3 orientation effect will not be as important as for the D-antipodes. The ϕCONH-ρ_1 interaction (present in the productive mode of the L-antipodes) may well be considered to provide optimum orientation of the carboxyl function at the active center, even in the case of the methyl ester. This R_1 group contains the necessary functionality (i.e., the -CONH-moiety) which is present in the natural polypeptide substrates of α-chymotrypsin. Thus, the increase in reactivity of the L-antipode productive complex (as measured by k_2) with a change of R_3, may merely reflect the electronic influence of R_3' (the alcohol component of the ester).

An increase in substrate reactivity, as measured by $k_o/K_o = k_2/K_s$, for the more reactive esters of the D-antipode may be a function of three different effects: the electronic effect of the alcohol component of the ester (σ_R^*) on k_2; a possible binding effect (i.e., change of K_s due to the R_3-ρ_3 interaction); and improved orientation of the hydrolyzable function (R_3) at the active center because of an enhanced R_3-ρ_3 interaction (leading to an increase in k_2). For the L-antipodes, however, this latter effect will not be operable. On this basis, the greater increase in the reactivity of the D-antipodes, as compared to the L-antipodes, can be and have been explained.

Some calculations based on the data of Table I supporting the above argument will now be presented. The D-antipodes will be considered first. The small difference in the k_o values for the glycolamide and cyanomethyl esters of N-benzoyl-D-alanine (0.085 sec^{-1} and 0.0934 sec^{-1}, respectively) suggest that, at least in the case of the cyanomethyl ester, deacylation is the rate determining step. Therefore, $k_o \simeq k_3$ for this substrate. Employing a value of $k_3 = 0.0934$ sec^{-1}, values of k_2 and K_s for the corresponding methyl and glycolamide esters may be calculated from the data of Table I and the relationships

$$k_o = \frac{k_2\ k_3}{k_2 + k_3} \qquad (14)$$

and

$$K_o = \frac{k_3}{k_2 + k_3} K_s\ . \qquad (15)$$

The values thus obtained for the methyl ester are $k_2 = 0.0077$ sec^{-1}, $k_3 = 0.0934$ sec^{-1}, and $K_s = 2.4$ m$\underline{M}$. For the glycolamide ester the corresponding values are $k_2 = 0.94$ sec^{-1}, $k_3 = 0.0934$ sec^{-1} and $K_s = 1.4$ m$\underline{M}$. It is observed that k_2 for the glycolamide ester is about 120 times greater than k_2 for the methyl ester. If this increase were due only to electronic effects of the alcohol component, a ρ^* value for acylation of about +3.2 would be obtained. Using this ρ^* value and the σ^* value of Table II for the cyanomethyl ester (+ 1.30), a value of $k_2 \approx 110$ sec^{-1} is obtained for the cyanomethyl ester. From this value of k_2, and $k_3 = 0.0934$ sec^{-1}, a value of $K_s = 15.7$ m$\underline{M}$ for N-benzoyl-$\underline{D}$-alanine cyanomethyl ester is obtained. This value of K_s seems unreasonably high when it is noted that the K_s value for the methyl ester is only 2.4 m$\underline{M}$.

If, however, one considers that much of the difference in the k_2 values for the methyl and glycolamide esters is due to the above-mentioned orientation effect and assumes, as an approximation, that the ρ^* value for acylation in the absence of this effect would be similar to that for the hydroxide-ion-catalyzed hydrolysis of acetates (i.e., about

+ 1.5), then values of $k_2 = 9.4$ sec^{-1} and $K_s = 1.34$ mM for the cyanomethyl ester of the D-antipode are obtained. This latter value is a more reasonable one for the dissociation constant of this enzyme-substrate complex. However, because a small but significant decrease in stereospecificity (1.6-fold, from Table V) is observed between the glycolamide and cyanomethyl esters, the orientation effect proposed to explain an abnormal increase in k_2 between the methyl and glycolamide esters should also be operable for the former pair. This effect should be considerably less for this case and thus the actual parameters for N-benzoyl-D-alanine cyanomethyl ester may be represented as 110 sec^{-1} $\gg k_2 > 9.4$ sec^{-1} and 15.7 mM $\gg K_s > 1.34$ mM. The values on the left of the above inequalities are based on the assumption that the effects on k_2 (orientation and electronic) between the glycolamide and cyanomethyl esters are the same as those between the methyl and glycolamide esters. The values on the right are based on the absence of an orientation effect between the glycolamide and cyanomethyl esters. Therefore, the actual values should lie between the two extremes, and because the orientation effect is small for this case, k_2 and K_s should both be better approximated by the lower values than by the higher ones.

Now, if the esters of N-benzoyl-L-alanine are considered, the kinetic parameters (k_o and K_o) may be represented by equations 16 and 17 (4):

$$k_o = \frac{k_2}{\frac{k_2}{k_3} + \alpha} \tag{16}$$

$$K_o = \frac{K_s}{\frac{k_2}{k_3} + \alpha} \tag{17}$$

where $\alpha = 1 + K_s/K_{sI}$ and K_s and K_{sI} are the dissociation constants for the productive and non-productive complexes, respectively.

Because both the productive and non-productive complexes involve R_3-ρ_3 interactions, one reasonable approximation which may be made is that α is essentially constant (independent of R_3) for the substrates considered. If now the value of ρ^* for acylation is considered to approximate that for the hydroxide-ion-catalyzed hydrolysis of acetates ($\rho^* \sim +1.5$) so that $k_{2_{-OCH_2CN}} \approx 10k_{2_{-OCH_2CONH_2}} \approx 100k_{2_{-OCH_3}}$, then from equations 16 and 17, the approximate relative magnitudes of the K_s values for the three esters may be calculated. The relative K_s values thus obtained for the methyl, glycolamide and cyanomethyl esters of benzoyl-L-alanine are in the ratios of 1:0.22:0.39, indicating some increase of binding for the more reactive esters. These ratios are in reasonable agreement with those obtained from the above calculations for the D-antipodes ($1{:}0.58{:}0.56 < X^{\dagger} \ll 6.5$) when it is taken into account that, for the D-antipodes, an increased R_3-ρ_3 interaction will occur only at some expense of the R_1-ρ_2 interaction, resulting in a decrease in K_s which is not as large as would be expected simply from consideration of the increased R_3-ρ_3 interaction. This would account for a smaller decrease in K_s for the glycolamide and cyanomethyl esters of the D-antipode as compared to that calculated for the corresponding L-antipodes.

[†] "X" represents the relative K_s value of N-benzoyl-D-alanine cyanomethyl ester (as compared with that for the methyl ester). From the discussion above, X can only be estimated by the limits shown.

From the ratio of the K_s values for the glycolamide and cyanomethyl esters of the L-antipodes (0.22:0.39) a very rough approximation of K_s for the D-cyanomethyl ester can be made. As an approximation, it will be assumed that the K_s values for the glycolamide and cyanomethyl esters of the D-antipode will be roughly proportional to those for the L-antipodes. Therefore, from the above ratio and the K_s value for the D-glycolamide ester (i.e., K_s = 1.4 mM) a value of $K_s \approx$ 2.5 mM for N-benzoyl-D-alanine cyanomethyl ester is obtained. This value of K_s yields a value of $k_2 \approx 17$ sec^{-1} for the same substrate. The above value of K_s satisfies the limits (i.e., 15.7 mM $\gg K_s >$ 1.34 mM) discussed above.

The calculated and approximated values of k_2, k_3 and K_s for the esters of N-benzoyl-D-alanine are summarized below:

Calculated Kinetic Parameters for Esters
of N-Benzoyl-D-Alanine

Ester	k_2 sec^{-1}	k_3 sec^{-1}	K_s mM
Methyl	0.0077	0.0934	2.4
Glycolamide	0.94	0.0934	1.4
Cyanomethyl	~ 17	0.0934	~ 2.5

From the above data for the D-antipodes, the data of Table I and the calculated K_s ratios for the L-antipodes, an approximation to the free energy profiles for these substrates can be made. These profiles are illustrated in Figure 4 and represent the free energy changes involved in going from free enzyme and substrate in solution to the transition state for formation of the acyl-enzyme intermediate (see General

Introduction, p. 8). The dotted lines are included only for illustrative purposes and represent only hypothetical parts of the profile. The K_s values for the L-antipodes cannot be calculated; only their approximate ratios are available. These values must, however, be greater than the corresponding K_s values for the D-antipodes (Table I and p. 186). Therefore, the free energy minima for the L-antipode ES complexes are hypothetical and are included only to illustrate the K_s ratios and the fact that these K_s values will be larger than those for the corresponding D-substrates. The three profiles (Figure 4a,b,c) have been normalized so that in all cases the free energy of the free substrate and enzyme in solution is set equal to zero. The free energy difference between E+S and ES is based on a standard state of 1 M concentration for one of the reactants (i.e., E or S). These figures provide an illustration (in free energy terms) of the effects proposed to explain the stereospecificity behavior of esters of N-benzoylalanine.

a) Conclusion

The effect of variation of the alcohol component (R_3') on the stereospecificity of the α-chymotrypsin-catalyzed hydrolyses of esters of N-benzoyl-D- and L-alanine is graphically illustrated by the data of Table V. Because the only structural change in these antipodal pairs is a variation in the R_3' group, the observed variation in stereospecificity must be a function of the nature of this group. It has been suggested that the R_3' group can affect reactivity in one or both of the following general ways: a) by an electronic effect or b) by a binding effect. Although electronic effects can account for a variation in reactivity, they alone do not adequately explain the observed variation in the rela-

Free Energy Profiles for the α-Chymotrypsin-catalyzed Hydrolyses of Three Esters of Benzoyl-D- and L-Alanine

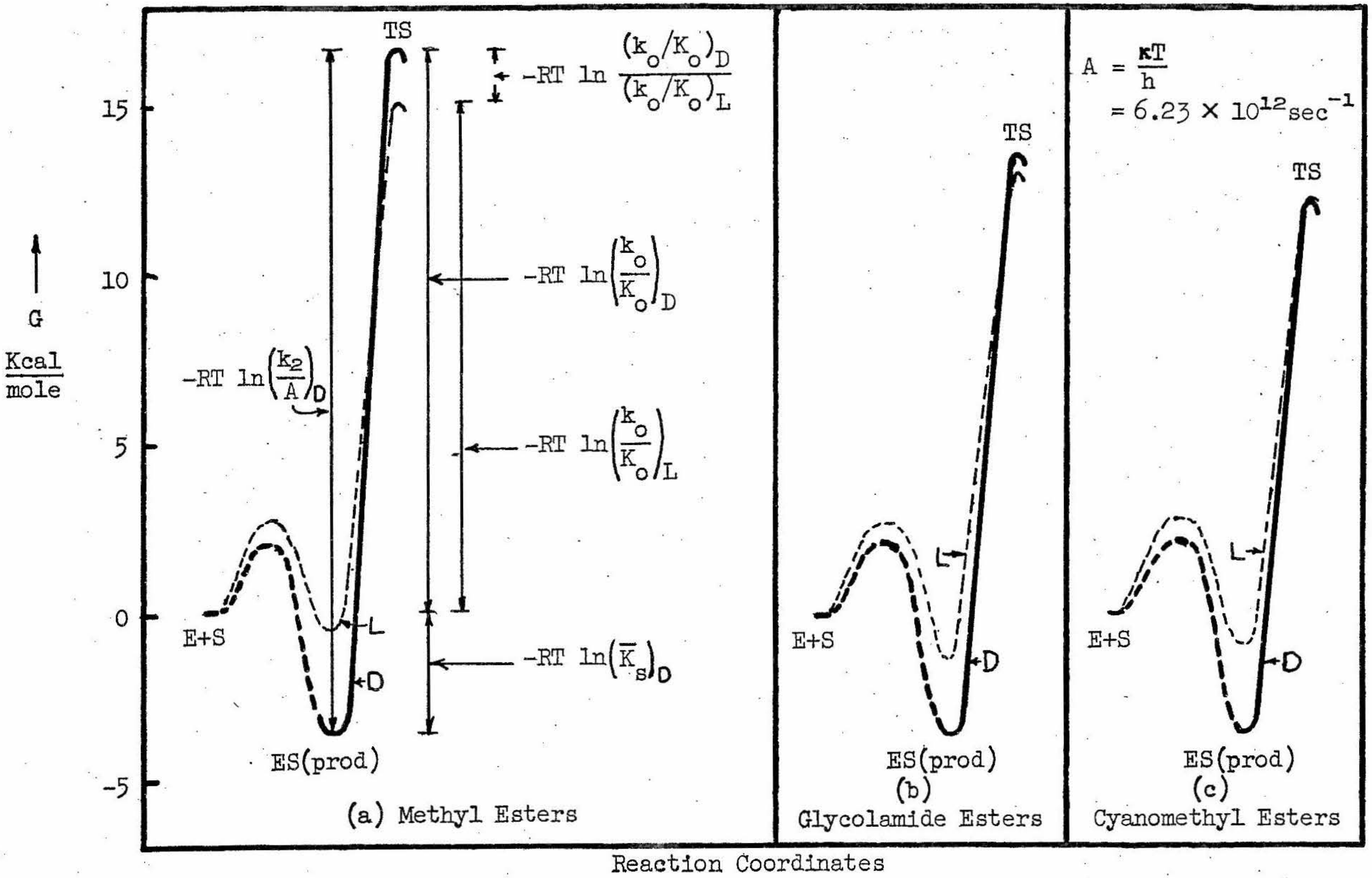


Figure 4

tive reactivity (stereospecificity) of the D-L pairs. Therefore, an explanation for this stereochemical behavior has been proposed based on the relative importance of an R_3-ρ_3 binding interaction which results in different orientation effects for the hydrolyzable functions of the two antipodal series. It was found that the results of this study were consistent with a ρ^* value (for acylation) for both antipodal series of the order of that observed for hydroxide-ion-catalyzed ester-hydrolysis when the effect of an improved R_3-ρ_3 interaction on the orientation of the carboxyl function in the reactive esters of the D-antipodes is taken into consideration.

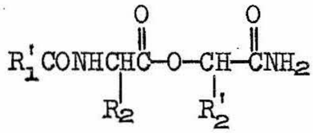
Comparison of Some Kinetic Parameters

One gratifying result of this study is the similarity of the increase in reactivity observed for the glycolamide ester of N-acetyl-L-alanine (as compared to the methyl ester) with that observed for the corresponding N-benzoyl-L-alanine derivatives. From Table IV, it is evident that the k_o/K_o values for the glycolamide esters of these two L-alanine derivatives are 56 and 44 times greater, respectively, than the k_o/K_o values for the corresponding methyl esters.

On the other hand, the increase in reactivity (210-fold) observed for the corresponding N-benzoyl-D-alanine substrates more closely resembles the 470-fold increase for the N-benzoylglycine derivatives than it does the reactivity increase for the L-alanine substrates (Table IV). These results support the premise (see Part II, p. 69) that the productive mode of the D-alanine derivatives resembles the predominant productive mode of the bifunctional substrates (a mode characterized by an R_1-ρ_2 interaction). The extra increase observed for the glycine derivative may, however, reflect some increased importance of the other possible productive mode for bifunctional substrates (i.e., one involving an R_1-ρ_1 interaction). Thus, the similarity of the productive complexes of bifunctional substrates and of D-antipodes of amino acid derivatives is supported by this data.

Secondary Specificity

The present study has shown that glycolamide esters of N-acylated amino acids are substantially more reactive than the usual methyl ester substrates (see Tables I and IV). It is suggested that further study of the α-chymotrypsin-catalyzed hydrolyses of derivatives of N-acylated amino acid glycolamide esters (IV),



$$R_1'CONH\underset{R_2}{\underset{|}{CH}}\overset{O}{\overset{\|}{C}}\text{-}O\text{-}\underset{R_2'}{\underset{|}{CH}}\text{-}\overset{O}{\overset{\|}{C}}NH_2$$

IV

in which R_2' is varied in the form of an amino acid side chain, might lead to a greater insight into the nature of the secondary specificity of α-chymotrypsin for amino acid residues which represent the amine component of peptide bonds in the natural polypeptide substrates.

It is hoped that the study described herein will stimulate interest in such a study.

EXPERIMENTAL

Syntheses

i. N-Benzoylglycine glycolamide ester

Hippuric acid (0.055 mole) was converted by the method of Schwyzer (34) to N-benzoylglycine cyanomethyl ester (55%), m.p. 102.4-103.4° (lit. m.p. 102.6-103.6°) (34).

The cyanomethyl ester (0.029 mole) was converted by the method of Wolf and Niemann (3) to N-benzoylglycine glycolamide ester (49%), m.p. 124.0-125.0° (lit. m.p. 124.1-125.0°) (3). A mixture (ca. 50-50) of this compound and a sample prepared by Wolf had a m.p. 123.8-125.0°.

ii. N-Benzoyl-L-alanine

L-Alanine was benzoylated via a Schotten-Baumann type reaction with benzoylchloride and aqueous sodium hydroxide (35). The crude product (86%) was recrystallized from hot water, dried at 60°, 200 μ Hg, over phosphorous pentoxide for 5 hours, yielding shiny plates of N-benzoyl-L-alanine, m.p. 147.0-147.5° (lit. m.p. 136-138° (36), 144-145° (37), 150-151° (38)); $[\alpha]_D^{25}$ +(37.4 ± 0.3)° (c, 18% in N aqueous sodium hydroxide) (lit. for N-benzoyl-D-alanine, $[\alpha]_D^{25}$ -36.0° (c, 19% in N aqueous sodium hydroxide) (39)).

Equivalent Weight		
	Calculated:	193.2
	Found:	193.3

iii. N-Benzoyl-D-alanine

This compound was prepared as described for the L-antipode (ii). Two recrystallizations from water and drying at 60°, 200 μ Hg, over phosphorous pentoxide yielded shiny plates of N-benzoyl-D-alanine, m.p. 146.5-147.5°, $[\alpha]_D^{25}$ -(36.0 ± 0.3)° (c, 18% in N aqueous sodium hydroxide) (lit. $[\alpha]_D^{25}$ -36.0° (c, 19% in N aqueous sodium hydroxide) (39)).

iv. N-Acetyl-L-alanine

This compound was prepared from L-alanine by the method described by Karrer (40) using acetic anhydride and sodium acetate. The crude product was recrystallized twice from ethyl acetate yielding large transparent crystals of N-acetyl-L-alanine, m.p. 123.5-124.5°, $[\alpha]_D^{25}$ -(65.0 ± 0.5)° (c, 2% in water) (lit. $[\alpha]_D^{20}$ -66.2° (c, 2% in water) (41)).

v. N-Benzoyl-L-alanine cyanomethyl ester

This compound was prepared from N-benzoyl-L-alanine (ii) and chloroacetonitrile by the method of Schwyzer (34) (67%). Two recrystallizations from ethanol-isopropyl ether gave white needles of N-benzoyl-L-alanine cyanomethyl ester, m.p. 108.5-109.0°, $[\alpha]_D^{24}$ -(32.9 ± 0.8)° (c, 3.2% in methanol).

Analysis				
	Calculated:	C: 62.06	H: 5.21	N: 12.07
$C_{12}H_{12}N_2O_3$ (232.3)	Found:	C: 62.03	H: 5.20	N: 12.11

vi. N-Benzoyl-D-alanine cyanomethyl ester

This compound was prepared in the same manner as that described for the L-antipode (v) (73%). Two recrystallizations from ethanol-isopropyl ether gave white needles of N-benzoyl-D-alanine cyanomethyl ester, m.p. 108.5-109.0°, $[\alpha]_D^{24}$ +(33.2 ± 0.8)° (c, 3.1% in methanol).

Analysis	Calculated:	C: 62.06	H: 5.21	N: 12.07
$C_{12}H_{12}N_2O_3$	Found:	C: 62.25	H: 5.22	N: 12.10

vii. N-Benzoyl-L-alanine glycolamide ester

This compound was prepared *via* the imido-ester hydrochloride (42) in a manner similar to that used for N-benzoylglycine glycolamide ester (i) and for N-acetyl-L-phenylalanine glycolamide ester (l). N-Benzoyl-L-alanine cyanomethyl ester (v) (0.02 mole) was added to 200 ml. benzene and warmed, with subsequent solution of most of the solid. The mixture was cooled to room temperature and 8.0 ml. of methanol in benzene (2.48 M) was added. Dry hydrogen chloride was bubbled through the mixture for 10 minutes. A moderate amount of a gummy substance appeared on the walls of the flask. The flask was then stoppered and allowed to stand at room temperature for one hour. The mixture was refluxed for 2 hours and then the solvent was removed by distillation at atmospheric pressure, leaving a white amorphous solid. Hexane was added; the mixture was heated and triturated on a steam bath for one-half hour, then cooled and filtered, leaving a white solid (92%). The crude product was dissolved in ethyl acetate; the organic solution was washed with small portions of aqueous sodium bicarbonate (10%) and water; the organic phase was dried over anhydrous magnesium sulfate and the solvent was removed under vacuum. Two recrystallizations from ethanol-hexane gave fine white crystals of N-benzoyl-L-alanine glycolamide ester, m.p. 132.5-133.0°, $[\alpha]_D^{23}$ -(11.3 ± 0.4)° (c, 3.2% in methanol).

Analysis	Calculated:	C: 57.59	H: 5.64	N: 11.20
$C_{12}H_{14}N_2O_4$ (250.3)	Found:	C: 57.64	H: 5.55	N: 11.02

viii. Benzoyl-D-alanine glycolamide ester

This compound was prepared in the same manner as the L-antipode (vii). Two recrystallizations from ethanol-hexane gave white crystals of N-benzoyl-D-alanine glycolamide ester, m.p. 132.5-133.0°, $[\alpha]_D^{23}$ $+(11.1 \pm 0.4)°$ (c, 3.4% in methanol).

Analysis				
$C_{12}H_{14}N_2O_4$ (250.3)	Calculated:	C: 57.59	H: 5.64	N: 11.20
	Found:	C: 57.64	H: 5.58	N: 11.09

ix. N-Acetyl-L-alanine cyanomethyl ester

This product was prepared from N-acetyl-L-alanine in the same manner as that described for the preparation of N-benzoyl-L-alanine cyanomethyl ester (v). Two recrystallizations from ethanol-isopropyl ether gave shiny plates of N-acetyl-L-alanine cyanomethyl ester (50%), m.p. 99.0-100.0°.

x. N-Acetyl-L-alanine glycolamide ester

N-acetyl-L-alanine cyanomethyl ester (ix) (0.02 mole) was added to 200 ml. benzene (Spectrograde). Most of the solid dissolved. Eight ml. of methanol in benzene (2.48 M) was added and dry hydrogen chloride was bubbled through the stirred mixture for about 5 minutes. A water-white gum formed on the sides of the flask. The flask was stoppered and left standing at room temperature for one hour. The benzene was distilled off at atmospheric pressure, leaving a white amorphous residue. The residue was dissolved in ethyl acetate and heated at ca. 40° for an hour. Hexane was added and crystallization occurred after several days at 4°. The product was twice recrystallized from ethyl acetate-isopropyl ether yielding powdery white crystals of N-acetyl-L-alanine glycolamide ester, m.p. 172.0-173.5°, $[\alpha]_D^{24}$ $-(14.2 \pm 0.9)°$ (c, 3% in methanol).

Analysis	Calculated:	C: 44.67	H: 6.43	N: 14.89
$C_7H_{12}N_2O_4$ (188.2)	Found:	C: 44.50	H: 6.50	N: 14.70

Kinetic Studies

The procedures employed in these studies are described in the Experimental section of Part II.

REFERENCES

1. R. J. Kerr and C. Niemann, *J. Org. Chem.*, **23**, 304 (1958).
2. T. H. Applewhite, H. Waite and C. Niemann, *J. Am. Chem. Soc.*, **80**, 1465 (1958).
3. J. P. Wolf, III and C. Niemann, *Biochemistry*, **2**, 493 (1963).
4. M. L. Bender and F. J. Kézdy, *Ann. Rev. Biochem.*, **34**, 49 (1965).
5. (a) G. E. Hein and C. Niemann, *J. Am. Chem. Soc.*, 84, 4487, 4495 (1962); (b) *Proc. Natl. Acad. Sci. U.S.*, **47**, 1341 (1961).
6. C. L. Hamilton, C. Niemann and G. S. Hammond, *Proc. Natl. Acad. Sci. U.S.*, in press.
7. B. Zerner and M. L. Bender, *J. Am. Chem. Soc.*, **86**, 3669 (1964).
8. B. Zerner, R. P. M. Bond and M. L. Bender, *J. Am. Chem. Soc.*, **86**, 3674 (1964).
9. G. H. Nelson, J. L. Miles and W. J. Canady, *Arch. Biochem. Biophys.*, **96**, 545 (1962).
10. R. M. Epand and I. B. Wilson, *J. Biol. Chem.*, **238**, 1718 (1963).
11. Observations of Dr. B. Zerner as published in reference (7).
12. J. E. Leffler and E. Grunwald, *Rates and Equilibria of Organic Reactions*, John Wiley and Sons, Inc., New York, N.Y. (1963), p. 222.
13. R. W. Taft, Jr., *Steric Effects in Organic Chemistry*, John Wiley and Sons, Inc., New York, N.Y. (1956), ed. 1, p. 556.
14. H. S. Harned and R. W. Ehlers, *J. Am. Chem. Soc.*, **54**, 1350 (1932).
15. D. A. MacInnes and T. Shedlovsky, *J. Am. Chem. Soc.*, **54**, 1429 (1932).
16. A. Albert and E. P. Sergeant, *Ionization Constants of Acids and Bases--a Lab Manual*, John Wiley and Sons, Inc., New York, N.Y. (1962), p. 124.
17. G. H. Jeffery and A. I. Vogel, *J. Chem. Soc.*, 1101 (1934).
18. J. Butterworth, D. D. Eley and G. S. Stone, *Biochem. J.*, **53**, 30 (1953).
19. J. E. Potts and E. S. Amis, *J. Am. Chem. Soc.*, **71**, 2112 (1949).

20. J. R. Knowles, *J. Theoret. Biol*. **9**, 213 (1965).

21. J. B. Jones and C. Niemann, *Biochemistry*, **1**, 1093 (1962).

22. N. S. Isaacs and C. Niemann, *Biochim. Biophys. Acta*, **44**, 196 (1960).

23. A. K. Balls and F. L. Aldrich, *Proc. Natl. Acad. Sci. U.S.*, **41**, 190 (1955).

24. A. K. Balls and H. N. Wood, *J. Biol. Chem*., **219**, 245 (1956).

25. C. E. McDonald and A. K. Balls, *J. Biol. Chem*., **221**, 993 (1956).

26. S. G. Cohen, J. Crossley, E. Khedouri and R. Zand, *J. Am. Chem. Soc*., **84**, 4163 (1962).

27. J. R. Rapp and C. Niemann, *J. Am. Chem. Soc*., **85**, 1896 (1963).

28. G. E. Hein, R. B. McGriff and C. Niemann, *J. Am. Chem. Soc*., **82**, 1830 (1960).

29. J. R. Rapp, Ph.D. Thesis, California Institute of Technology, 1963.

30. J. B. Jones, T. Kunitake, C. Niemann and G. E. Hein, *J. Am. Chem. Soc*., **87**, 1777 (1965).

31. W. A. Mukatis, Ph.D. Thesis, California Institute of Technology, 1965.

32. R. L. Peterson, Ph.D. Thesis, California Institute of Technology, 1965.

33. M. L. Bender, *J. Am. Chem. Soc*., **84**, 2582 (1962).

34. R. Schwyzer, M. Feurer, B. Iselin and H. Kagi, *Helv. Chim. Acta*, **38**, 80 (1955).

35. J. P. Greenstein and M. Winitz, *Chemistry of the Amino Acids*, John Wiley and Sons, Inc., New York, N.Y., vol. 2 (1961), pp. 1267-1269.

36. M. S. Dunn, M. P. Stoddard, L. R. Rubin and R. C. Bovie, *J. Biol. Chem*., **151**, 241 (1943).

37. R. Mozingo, D. E. Wolf, S. A. Harris and K. Folkers, *J. Am. Chem. Soc*., **65**, 1013 (1943).

38. *Biochemisches Handlexicon*, Berlin, **4**, 449.

39. J. O'Brien and C. Niemann, *J. Am. Chem. Soc*., **79**, 80 (1957).

40. P. Karrer, K. Escher and P. Widmer, Helv. Chim. Acta, 9, 301 (1926).

41. S. M. Birnbaum, L. Levintow, R. B. Kingsley and J. P. Greenstein, J. Biol. Chem., 194, 455 (1952).

42. S. Grudzinski, J. Strumillo and A. Kotelko, Roczniki Chemii, 35 729 (1961).

PROPOSITIONS

PROPOSITION I

The validity of the results of a study of the inhibition of α-chymotrypsin-catalyzed hydrolysis by α-keto analogues of amino acids is questioned. It is proposed that the K_I's of these inhibitors be re-evaluated and that further investigation of the behavior of these compounds be carried out.

* * * * * * * * *

Recently, J. D. Geratz (1) has reported a study of α-keto analogues of amino acids as inhibitors of α-chymotrypsin (and other enzymes). The results (at least those with α-chymotrypsin) are quite startling and are summarized below:

α-Keto analogues of tryptophan, phenylalanine and tyrosine (indole-3-pyruvic acid (IPVA), phenylpyruvic acid (PPVA) and p-hydroxyphenylpyruvic acid (HPPVA)) were found to be strong inhibitors of the esterase and proteolytic activity of α-chymotrypsin. The first two inhibitors appear to be much more effective than their non-keto analogues (indole-3-propionic acid (IPA) and phenylpropionic acid (PPA)) at least with respect to inhibition of esterase activity when L-tyrosine ethyl ester (TEE) is the substrate. This disparity is not as great, however, when casein is the substrate (inhibition of proteolytic activity).

The extremely low K_I values obtained for the pyruvic acid derivatives (TEE as a substrate) are particularly amazing. The K_I value for IPVA was 0.07 mM. This is almost as low as the value (0.063 mM) reported for benzo[f]quinoline, the most effective small-molecule competitive inhibitor of α-chymotrypsin yet described (2). The K_I's obtained for PPVA and HPPVA were 0.86 mM and 0.28 mM, respectively. Those obtained for PPA and IPA were 23.5 mM and 3.90 mM, respectively.

These results are extremely difficult to explain on the basis of our knowledge of the specificity of α-chymotrypsin. Geratz gives no real explanation. At least superficially, it would seem unreasonable to expect HPPVA (K_I = 0.28 mM) to bind to α-chymotrypsin 300 times more efficiently than the substrate analogue, N-acetyl-L-tyrosine (K_I = 80 mM (3)).

It is proposed that the nature of the experiments performed in the evaluation of the inhibitory power of the pyruvic acid derivatives was such that the reliability of the data obtained is extremely questionable. Pertinent experimental details are outlined below:

1. Esterase inhibition: Assay mixtures contained ca. 10 μg/ml. α-chymotrypsin, [TEE] varying from 5 to 40 mM in 0.1 M imidazole buffer (pH 7.0). Concentrations of inhibitor employed were, [IPVA] = 5×10^{-5} M, [HPPVA] = 2×10^{-4} M and [PPVA] = 5×10^{-4} M. Incubation was carried out at 37°. Inhibition was competitive.

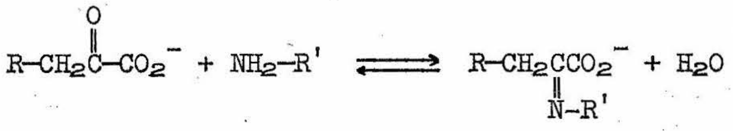
2. Proteolytic inhibition: Assay mixtures contained ca. 3 μg/ml. α-chymotrypsin, 0.5 g. % casein, 5% ethanol in 0.1 M Tris buffer (pH 8.0). Per cent inhibition was determined after 16 min. at 37° for concentrations of inhibitor ca. 10^{-3} M.

These experiments were unfortunate for the following reasons:

1. In the esterase experiments, the buffer (imidazole) is a reasonably effective inhibitor for the enzyme (i.e., K_I = 45 mM vs. Ac-L-valOCH$_3$ (2)). Thus, a considerable amount of the reaction will be

inhibited by this compound. However, because the studies with IPA and PPA were also carried out in this buffer, this buffer-inhibition cannot account for the disparity between the keto- and non-keto analogues.

2. It is known that derivatives of pyruvic acid react readily with amines (including imidazole derivatives). P. Zuman (4,5) has studied the reaction



$$R\text{–}CH_2\overset{O}{\overset{\|}{C}}\text{–}CO_2^- + NH_2\text{–}R' \rightleftharpoons R\text{–}CH_2\underset{N\text{–}R'}{\underset{\|}{C}}CO_2^- + H_2O$$

I II III

and some equilibrium constants ($K = [III]/[I][II]$) are given below:

R	R'	K	Ref.
ϕ	–H	0.29	4
ϕ	$-CH_2CO_2H$	2.50	4
CH_3^-	$R'NH_2$ = histidine	24.5 (imidazole N)	5
		1.99 (primary N)	5

It has also been reported that pyruvic acid reacts (at physiological pH's) with Tris (6).

Thus, there appears to be a distinct possibility that the actual inhibitor in the experiments described is some adduct of the pyruvic acid derivative. In the esterase study, this could be an adduct with imidazole or the substrate (TEE). In the proteolytic experiment, a Tris-adduct could be the inhibitor. Although there is no precedent for such compounds as inhibitors for α-chymotrypsin, because they possess more functionality than the pyruvic acids themselves, it is not inconceivable that they might function as very efficient inhibitors. The extremely interesting possibility of the formation of an inhibiting adduct with the active site of the enzyme itself cannot be ruled out. There are two histidine residues and an N-terminal isoleucine residue implicated at the active site of α-chymotrypsin (7). The experiments proposed below should shed more light on these possibilities.

The inhibition of esterase activity by pyruvic acid derivatives should be evaluated against an acylated amino acid ester with the use of a pH-stat (i.e., in the absence of a buffer). High values of K_I would support the suggestion that buffer- or substrate-adducts were the inhibiting species in Geratz' experiment. If, however, very low K_I values are again obtained, the possibility of reversible pyruvic acid-enzyme adduct formation should be further investigated. Difference-spectra and O^{18}-labelling experiments should be useful here.

REFERENCES

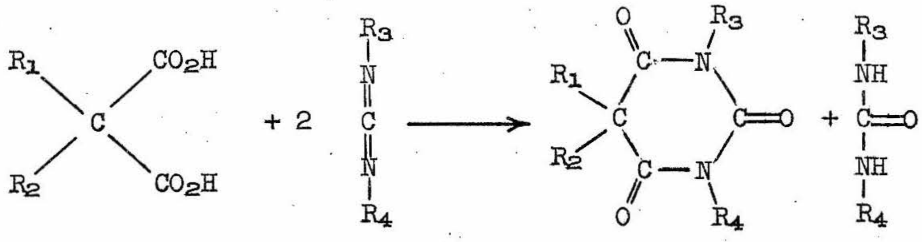
1. J. D. Geratz, Arch. Biochem. Biophys. , **111**, 134 (1965).
2. R. A. Wallace, A. N. Kurtz and C. Niemann, Biochemistry, **2**, 824 (1963).
3. R. J. Foster, H. J. Shine and C. Niemann, J. Am. Chem. Soc., **77**, 2378 (1955).
4. P. Zuman, Coll. Czechoslov. Chem. Comm., **15**, 839 (1951).
5. P. Zuman, Chem. Listy, **45**, 40 (1951); CA, **45**, 9031.
6. P. F. Duggan, D. M. X. Donnely and D. P. Melody, Irish J. Med. Sci., № 460, 163 (1964); CA, **61**, 5990b.
7. M. L. Bender and F. J. Kézdy, Ann. Rev. Biochem., **34**, 49 (1965).

PROPOSITION II

The difference between aromatic and aliphatic carbodiimides in the reaction with malonic acids is pronounced. It is proposed that this difference may be explained more readily by electronic effects than by steric effects as suggested by Bose and Garratt.

* * * * * * * *

Bose and Garratt (1,2) have recently reported a novel synthesis of substituted barbituric acids (III) which involves the reaction between malonic acids (I) and carbodiimides (II).



I II III

Table I shows the results of syntheses with varying R groups.

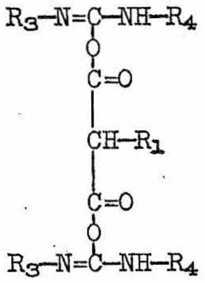
TABLE I

Reactions of Carbodiimides with Malonic Acids(1)

Malonic Acid R_1	R_2	Carbodiimide $R_3 = R_4$	Product	Yield %
H	H	C_6H_{11}	III	65
C_2H_5	H	C_6H_{11}	III	63
C_2H_5	C_2H_5	C_6H_{11}	III	50
C_6H_5	H	C_6H_{11}	III	31
H	H	i-C_3H_7†	III	60
H	H	p-tolyl	IV or V	30
C_2H_5	H	p-tolyl	IV or V	65
C_2H_5	C_2H_5	p-tolyl	III	32

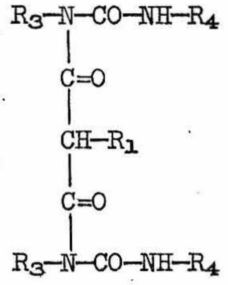
†Bose and Garratt state (2) that diisopropyl carbodiimide produces barbiturates with "equal facility from substituted and unsubstituted malonic acids."

From Table I, it is seen that whereas N,N'-di-p-tolyl carbodiimide does not form barbiturates with unsubstituted or mono-substituted malonic acids, a barbiturate is formed with a disubstituted malonic acid. The products with unsubstituted or mono-substituted malonic acids have been assigned the structures (IV) or (V).



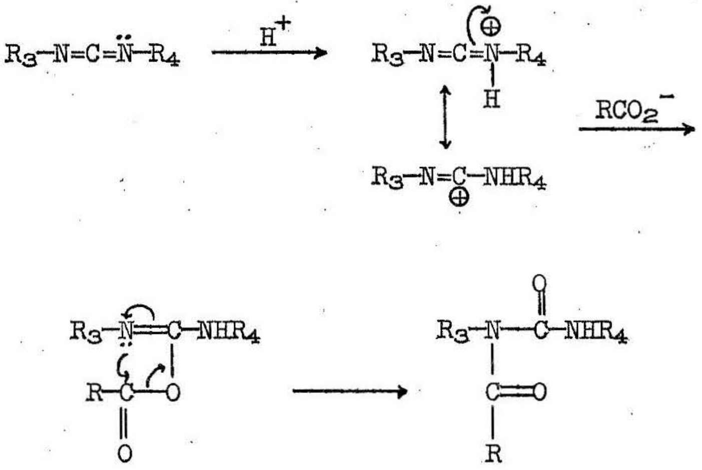
IV

(R_1 = H or C_2H_5)
(R_3 = R_4 = p-tolyl)



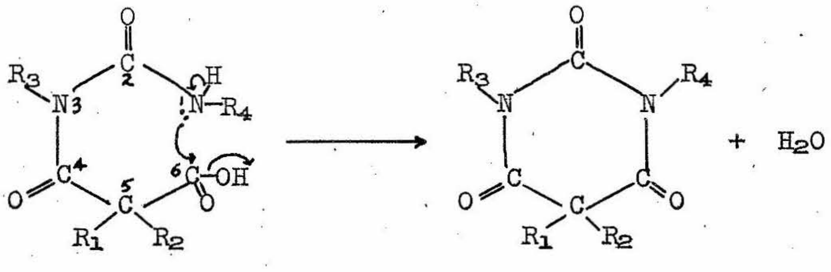
V

It is well known that N,N'-disubstituted carbodiimides react with carboxylic acids to form N-acylureas, presumably via the following mechanism (3):



VI

Bose and Garratt suggest the cyclization of the intermediate (VIa) to yield the barbiturate (III).



VIa

III

They attribute the difference between aromatic and aliphatic carbodiimides mainly to a steric effect (2). They state that molecular models show that when R_3 = R_4 = an aryl group in the intermediate (VIa), the molecule is much more crowded than when R_3 = R_4 = isopropyl or cyclohexyl. Furthermore, in the least hindered conformation of VIa, the least bulky group will be nearest N_1. Thus, when R_1 = R_2 = H, the carboxyl group will be in an unfavorable position for ring closure. A similar situation will prevail when R_1 = H and R_2 = C_2H_5. However, when R_1 = R_2 = C_2H_5, the substituents on C_5 will be roughly of comparable size and the carboxyl group will more readily approach N_1 within bond-forming distance.

This explanation seems to be open to some criticism:

1. The steric effects of a cyclohexyl group and a p-tolyl group with respect to hinderance of attack at the N_1 position are quite similar. In fact, the cyclohexyl group is actually somewhat more bulky.

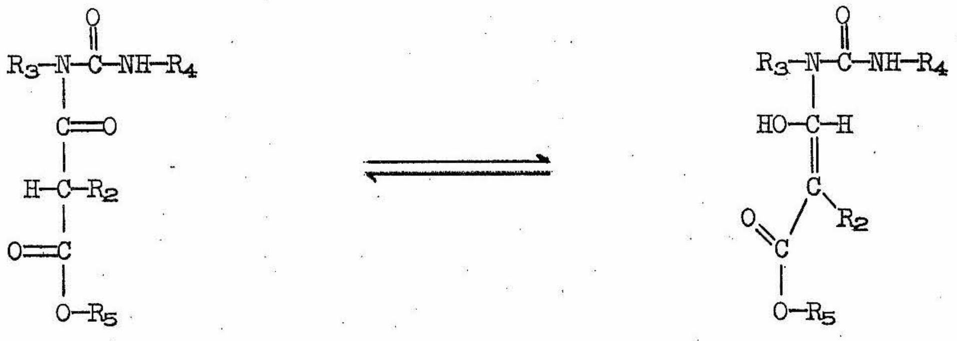
2. If the "gem-dimethyl" effect of the R_1 and R_2 substituents is important, a higher yield of barbituate from N,N'-dicyclohexyl carbodiimide with diethyl malonic acid than with the unsubstituted malonic acid should be expected. This is not the case. A possible experiment to test this effect would be to react dimethyl malonic acid with N,N'-p-tolyl carbodiimide. According to Bose and Garratt, yields should be less than with diethyl malonic acid.

An alternate explanation of the facts, based on electronic and resonance factors, is proposed:

1. When R_3 = R_4 = aryl, resonance involving N_1 (VIa) and the aryl ring will be very important in reducing the electron density at the

nitrogen, making it a poorer nucleophile and thereby hindering the cyclization reaction. Yields were significantly poorer for $R_1 = R_2 = C_2H_5$ and $R_3 = R_4 =$ p-tolyl than for $R_1 = R_2 = C_2H_5$ and $R_3 = R_4 =$ cyclohexyl.

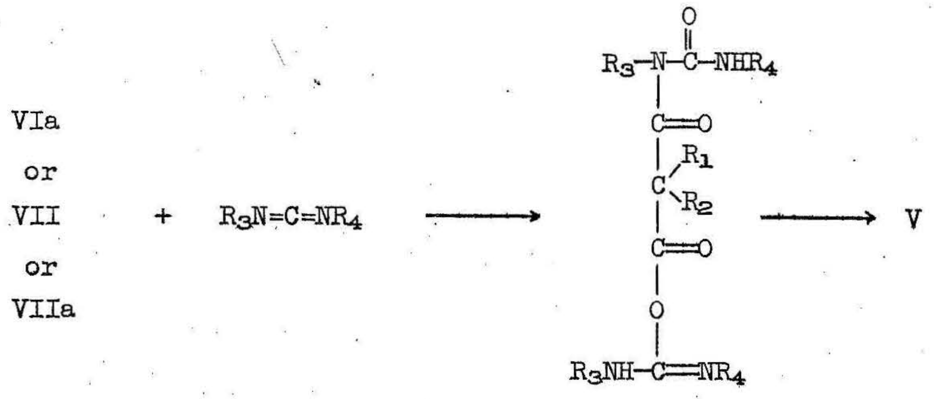
2. When $R_3 = R_4 =$ aryl, the enol form of VII (i.e., VIIa) would be stabilized by extensive electron delocalization due to an extended π system.


VII
(R_5 = H)

VIIa
(R_5 = H)

This effect should be much greater for $R_3 = R_4 =$ aryl than when $R_3 = R_4 =$ alkyl. In addition, this enolization is possible only if R_1 and/or R_2 = H. Models show that the enol form is now quite hindered from cyclization.

The reaction in competition with the cyclization reaction is probably the following:



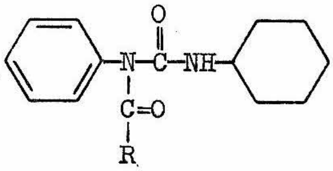
VIII

Thus, the enol form (VIIa) could proceed via this pathway to V, one of the proposed structures for the product from N,N'-di-p-tolyl carbodiimide and mono- or unsubstituted malonic acids.

It is interesting to note that the $R_3{-}NH{-}\underset{|}{C}{=}N{-}R_4$ moiety is used to activate the carboxyl group in peptide synthesis. It is an excellent leaving group and thus might be the leaving group (rather than -OH) in the cyclization reaction.

One piece of evidence supporting the argument of enol stabilization by an extended π system is that reactions of carbodiimides (II) (where $R_3 = R_4 = (CH_3)_2{-}N{-}C_6H_4{-}$) with α,β-unsaturated acids give highly colored N-acylurea derivatives, whereas the products with saturated acids are colorless (4).

An interesting experiment would utilize an unsymmetrical carbodiimide. Khorana (5) found that the reaction of N-cyclohexyl,N'-phenyl carbodiimide (II) ($R_3 = C_6H_5$, $R_4 = C_6H_{11}$) with a carboxylic acid gave only product IX.



X

If Bose and Garratt are correct in their interpretation, reaction of such a carbodiimide (e.g., R_3 = p-nitrophenyl, $R_4 = C_6H_{11}$ with malonic acid should give a yield of the barbiturate comparable to that of the dicyclohexyl case, whereas this proposition would predict a lower yield of the barbiturate and a significant yield of the alternate product (V).

REFERENCES

1. A. K. Bose and S. Garratt, Tetrahedron, 19, 85 (1963).
2. A. K. Bose, S. Garratt and J. J. Pelosi, J. Org. Chem., 28, 730 (1963).
3. H. G. Khorana, Chem. Rev., 53, 145 (1953).
4. F. Zetzsche and G. Roettger, Ber., 72b, 1599 (1939).
5. H. G. Khorana, J. Chem. Soc., 2081 (1952).

PROPOSITION III

A novel method of modifying polypeptide structure is proposed. This technique, involving the use of a solid-phase support, would permit the "synthesis" of polypeptides which are analogous to biologically-active compounds (hormones, etc.) but contain a partially different sequence. This method is discussed in terms of a proposed modification of the structure of porcine β-MSH.

* * * * * * * *

Modification of the amino acids existing in the primary sequence of biologically-active oligo- and polypeptides is often employed in the investigation of the relationship between the structure of these compounds and their activity. A great deal has been accomplished with regard to the chemical modification of amino acid side chains in such peptides. This method, however, is essentially limited to amino acids which possess side chains having functional groups ($-NH_2$, $-CO_2H$, $-OH$, etc.). Smaller peptides containing a partial sequence of larger biologically-active polypeptides have been synthesized. In some cases, these peptides have partial biological activity. By modification of these peptides, much useful information can be gained with respect to the nature of "essential" amino acids in the sequence of the polypeptide from which the peptides are derived.

A method is proposed in which some amino acids in the biologically-active polypeptide itself may be replaced by other amino acid residues. In effect, this would result in the "synthesis" of polypeptides. This method will be discussed in terms of a specific example.

β-Melantocyte-stimulating hormone (β-MSH) is isolated from posterior pituitary glands. The sequence of this octadecapeptide isolated from hogs is given below (1,2):

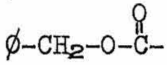
asp^{1}-glu^{2}-gly^{3}-pro^{4}-tyr^{5}-lys^{6}-met^{7}-glu^{8}-his^{9}-

-phe^{10}-arg^{11}-try^{12}-gly^{13}-ser^{14}-pro^{15}-pro^{16}-lys^{17}-asp^{18}

It has been shown (3) that at least some of the hormonal activity is located in the 8-13 hexapeptide sequence. This peptide was synthesized and had some melantocyte-stimulating activity (about 1/10,000 that of the hormone itself). It has also been shown with peptides of this type that the arginine-11 residue is extremely important, if not essential, for activity (4). The synthesis of bovine β-MSH (containing a seryl residue instead of glutamic acid-2) has been achieved (5).

Figure 1 contains a scheme proposed for the "synthesis" of a polypeptide which is identical in structure to β-MSH except that the 11th and 12th residues (arg-try) are replaced by two other residues. Variation of these residues should provide more information regarding the importance of the arginine-11 residue. Most of the reaction steps are analogous to those used so successfully by Merrifield (6) in his step-by-step syntheses of peptides. The sequence of reaction is discussed below.

1. Amino Group Protection: The free amino groups on the polypeptide must first be protected by a residue which is stable under the conditions required for the removal of the t-butyloxycarbonyl- group

(t-BOC) (Step 8). Because these conditions are relatively mild (HCl in acetic acid) the carbobenzoxy- group may suffice here (i.e., X =  $\phi\text{-}CH_2\text{-}O\text{-}\overset{\overset{O}{\|}}{C}\text{-}$ ≡ CBZ).

2. Carboxyl Group Protection: The novelty of the method proposed is that the carboxyl groups of the polypeptide are protected by esterification with a highly chloromethylated polymeric solid-support. This step is analogous to the first step in Merrifield's synthesis of peptides (6) in which the desired carboxy-terminal residue is bound by means of a "benzyl ester" linkage to the polymer (a chloromethylated copolymer of styrene and divinyl benzene) during all of the subsequent synthetic steps. The added advantage of this type of protection is that it provides a "handle" by which the peptide fragments (in later steps) are retained on the solid-support.

3. Enzymatic Cleavage: It is proposed (for this example) that α-chymotrypsin be employed to cleave the polypeptide (bound to the solid-support) and thus liberate the dipeptide, arginyl-tryptophan. Although there is no precedent for the action of a proteolytic enzyme on a polypeptide bound in this way, it is quite possible that conditions could be found for which this cleavage would be successful. If this proteolytic cleavage does succeed, the three bonds which are indicated by arrows should be split. This is the cleavage-pattern observed during sequence studies on this hormone. The dipeptide, arginyl-tryptophan, should be split out, while the other peptides will remain bound to the solid-support.

4. Formation of the tyr^5-lys^6 Bond: This bond formation could be effected with the use of dicyclohexylcarbodiimide (DCD) (see ref. 6).

Formation of a phenylalanyl^{10}-glycine^{13} bond should be prevented by steric restrictions imposed by the polymer-support.

5, 6 and 7. Insertion of a Dipeptide: These three steps involve the addition of an N-protected dipeptide (t-BOC-A-B) to the amino group of glycine-13 (step 5), removal of the t-BOC group (step 6) and formation of the phe^{11}-A^{12} bond with DCD (step 7). These steps are strictly analogous to those employed by Merrifield (6).

8 and 9. Removal and Deprotection: The removal of the N-protected polypeptide from the polymer may be effected with anhydrous HBr in acetic acid (6). Although these conditions might suffice also for removal of the X = CBZ groups, a subsequent deprotection step could be employed.

The product of this synthesis would be an analogue of β-MSH containing two new residues in the 11th and 12th positions.

Although there would undoubtedly be several practical difficulties involved in the techniques described above, the novelty of the method, its advantages over more conventional synthetic techniques (e.g., few isolations of intermediates) and the wide scope possible for this technique, make it an intriguing proposal.

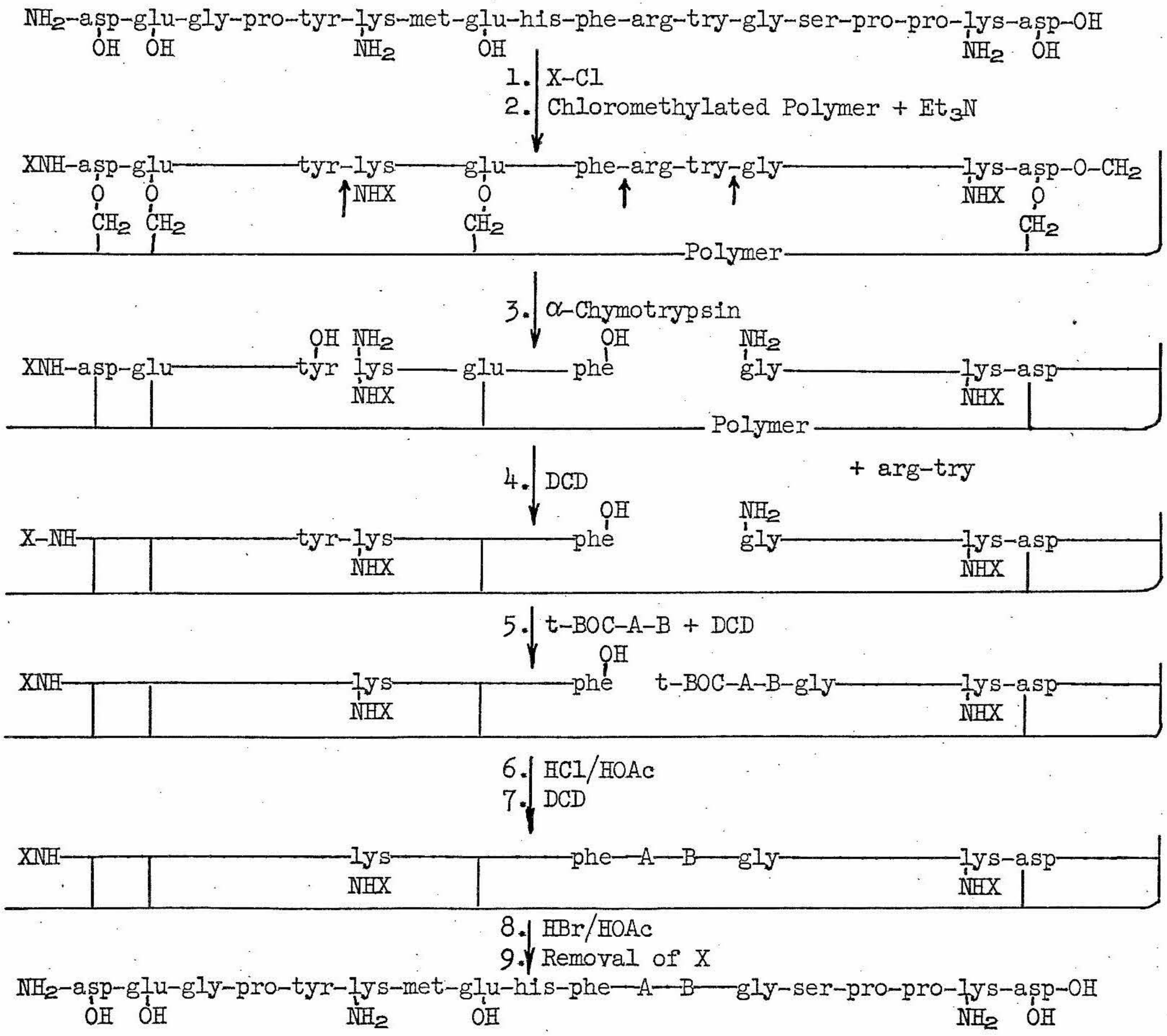

NH2-asp-glu-gly-pro-tyr-lys-met-glu-his-phe-arg-try-gly-ser-pro-pro-lys-asp-OH
OH
NH2
1. X-Cl
2. Chloromethylated Polymer + Et3N
XNH-asp-glu
tyr-lys
glu
phe-arg-try-gly
lys-asp-O-CH2
NHX
O
CH2
Polymer
3. α-Chymotrypsin
tyr
phe
gly
lys-asp
+ arg-try
4. DCD
X-NH
5. t-BOC-A-B + DCD
t-BOC-A-B-gly
6. HCl/HOAc
7. DCD
XNH
phe—A—B—gly
8. HBr/HOAc
9. Removal of X
NH2-asp-glu-gly-pro-tyr-lys-met-glu-his-phe—A—B—gly-ser-pro-pro-lys-asp-OH

Figure 1

Synthesis of a Modified Porcine β-MSH

REFERENCES

1. J. I. Harris and P. Roos, *Nature*, **178**, 90 (1956).
2. I. I. Geschwind, C. H. Li and L. Barnati, *J. Am. Chem. Soc.*, **79**, 620 (1957).
3. H. Kappeler, *Helv. Chim. Acta*, **44**, 476 (1961).
4. M. Bodansky, M. A. Ondetti, B. Rubin, J. J. Piala, J. J. Fried, J. T. Sheehan and C. A. Birkheimer, *Nature*, **194**, 485 (1962).
5. R. Schwyzer, B. Iselin, H. Kappeler, B. Riniker, W. Rittel and H. Zuber, *Helv. Chim. Acta*, **46**, 1975 (1963).
6. R. B. Merrifield, *Science*, **150** N° 3693, 178 (1965) and references therein.

PROPOSITION IV

Experiments are proposed which would further elucidate the mechanism of the reaction of lead tetra-acetate with primary amines.

* * * * * * * *

Acott and Beckwith (1) have recently reported that the reaction of primary amides (I) in benzene or benzene-acetic acid with lead tetra-acetate yields the appropriate N-acetylamine (II) and dialkyl urea (III).

$$\underset{\text{I}}{\text{RCONH}_2} \xrightarrow[\phi\text{H} \text{ or } \phi\text{H-HOAc}]{\text{Pb(OAc)}_4} \underset{\text{II}}{\text{RNHAc}} + \underset{\text{III}}{\text{RNHCONHR}}$$

The authors propose that the reaction proceeds via formation and rearrangement of an acyl nitrene:

$$\text{I} + \text{Pb(OAc)}_4 \longrightarrow \text{RCO}\ddot{\text{N}}\text{:} + \text{Pb(OAc)}_2 + 2\ \text{HOAc}$$

$$\begin{tikzcd}[column sep=small] & & & \text{II} \\ \text{RCO}\ddot{\text{N}}\text{:} \arrow[r] & \text{R-N=C=O} \arrow[r, "\text{HOAc}"] & \text{RNHCO}_2\text{Ac} \arrow[ur] \arrow[dr] & \\ & & & \text{III} \end{tikzcd}$$

Their evidence for this path (similar to that proposed for the Curtius reaction (2,3)) is as follows:

1. The products obtained in this reaction are similar to those yielded in a Curtius reaction carried out in acetic acid (4,5).

2. Formation of III appears to be suppressed when neat acetic acid is used as the solvent.

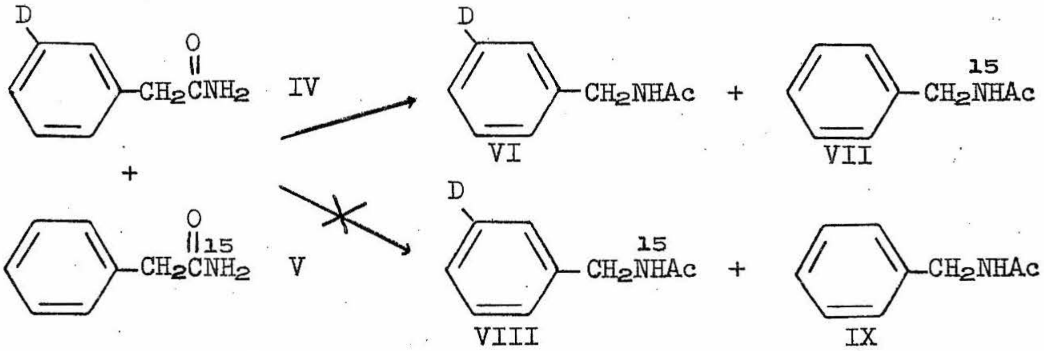
3. The authors claim to have spectrophotometrically detected the formation of alkyl isocyanate. No experimental data are given, however.

The mechanism described above appears to be a reasonable one. Three experiments, which have helped to elucidate the mechanisms of the Curtius and Hofmann reactions are proposed. The results of these experiments should provide further evidence in support of the similarity between the reaction in question and the Curtius reaction.

Intramolecularity of the Rearrangement Step

1) The lead tetra-acetate reaction should be carried out with an optically active amide (e.g., one enantiomer of α-phenylpropionamide). If the rearrangement is intramolecular, retention of configuration should be observed. Retention was observed in a Curtius reaction with a similar substrate (6).

2) Reaction of a mixture of IV and V with lead tetra-acetate should give VI and VII but no VIII or IX if the rearrangement is intramolecular. A similar experiment was performed with the Hofmann reaction (7).



In the two experiments above, retention of configuration (expt. 1) and lack of exchange (expt. 2), although they would not prove intramolecularity, would at least indicate that the migrating group is not separated from the rest of the molecule long enough for an exchange reaction to occur.

Reaction of the Isocyanate

3) In a study of the reaction of isocyanates with carboxylic acids, Fry (8) determined, by means of a labelling experiment, that the CO_2 produced during the production of acylamines and dialkylureas came from the isocyanate. A similar experiment, involving the reaction of $R-C^{14}ONH_2$ and lead tetra-acetate is proposed. The CO_2 evolved should contain all the C^{14} if this reaction proceeds in the same manner.

These three proposed experiments should help to elucidate the mechanism of the reaction of lead tetra-acetate with primary amines and provide support for the similarity of this reaction to the Curtius reaction.

REFERENCES

1. B. Acott and A. L. J. Beckwith, Chem. Comm., N° 8, 161 (1965).
2. J. D. Roberts and M. C. Caserio, Basic Principles of Organic Chemistry, W. A. Benjamin, Inc., New York (1964), pp. 655-656.
3. P. A. S. Smith, in Molecular Rearrangements, P. de Mayo, Ed., Interscience, New York, N. Y. (1963), vol. 1, pp. 457-593.
4. R. G. Arnold, J. A. Nelson and J. J. Verbanc, Chem. Rev., **57**, 47 (1957).
5. P. A. S. Smith, Org. Reactions, **3**, 337 (1946).
6. L. W. Jones and E. S. Wallis, J. Am. Chem. Soc., **48**, 169 (1926).
7. T. J. Prosser and E. L. Eliel, ibid., **79**, 2544 (1957).
8. A. Fry, ibid., **75**, 2686 (1953).

PROPOSITION V

Diazoacetyl-L-phenylalanine chloromethyl ketone (DAPCK) and its peptide analogues are proposed as trifunctional irreversible inhibitors for α-chymotrypsin. These reagents could potentially form stable covalent bonds at two places in the active site.

* * * * * * *

Reaction of selective reagents at the active site of α-chymotrypsin has lead to the identification of a serine residue (1) and one histidine residue (2) as probable catalytic groups (i.e., groups located at the active center in the ρ_3 locus of the active site)[†] involved in the action of the enzyme. In addition, it has been shown (3,4) that one methionine residue is probably important in the binding of substrates to the active site.

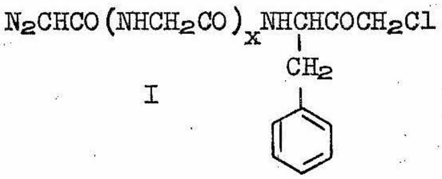
Although studies with selective reagents for methionine were performed with bifunctional[‡] inhibitors which, at least in one case (4), were designed such that the reactive portion should have reacted with an amino acid residue in the ρ_1 locus,[†] the actual locus position of this residue is in doubt. Knowles (5) found that kinetic results on a periodate-oxidized chymotrypsin (which had one methionine residue converted to a sulfoxide) were most consistent with the location of this residue in the ρ_2 locus.

[†] See the General Introduction of this thesis for definitions of active site loci.

[‡] One binding function and one reactive function.

A trifunctional irreversible inhibitor (i.e., an inhibitor with one function capable of binding non-covalently at ρ_2 and two functions capable of forming covalent bonds in the ρ_1 and ρ_3 loci, respectively, of the active site) should provide an excellent tool for investigating the nature of the amino acid residues in the ρ_1 locus.

The proposed reagents are of the form:



$N_2CHCO(NHCH_2CO)_xNHCH(CH_2C_6H_5)COCH_2Cl$

I

x = 0 = DAPCK

The simplest reagent of this kind (I, x = 0) would be diazoacetyl-L-phenylalanine chloromethyl ketone. This compound, because of its similarity to N-tosyl-L-phenylalanine chloromethyl ketone (TPCK), the selective histidine reagent of Schoelmann and Shaw (2), should in a "dark" reaction alkylate the histidine-57 residue located in the ρ_3 locus. Photolysis of the now irreversibly inhibited enzyme should generate a carbene, a highly reactive species which has the potential of reacting with amino acid residues in the ρ_1 locus. Westheimer (6) has found that acid hydrolysis of the modified protein obtained by photolysis of diazoacetyl chmotrypsin yields O-carboxymethylserine, and 1-carboxymethylhistidine, O-carboxymethyltyrosine and at least three other unidentified minor products.

The ρ_1 locus has been shown (part II--this thesis) to be an extensive area containing many amino acid residues. Thus the peptide analogues of I ($x > 0$) might profitably be employed to further investigate the nature of the groups in this locus.

Because of the trifunctionality of these reagents, if attack at both ρ_3 and ρ_1 occurs, much useful information would be gained concerning the spatial arrangement (at least in one dimension) of groups present in the active site of the enzyme.

REFERENCES

1. N. K. Schaffer, S. C. May and W. H. Summerson, J. Biol. Chem., 206, 201 (1954).
2. G. Schoellman and E. Shaw, Biochemistry, 2, 252 (1963).
3. H. J. Schramm and E. Lawson, J. Am. Chem. Soc., 84, 2017 (1962); Z. Physiol. Chem., 332, 97 (1963).
4. G. Gundlach and F. Turba, Biochem. Z., 335, 573 (1962).
5. J. R. Knowles, Biochem. J., 95, 180 (1965).
6. J. Shafer, P. Baronowsky, R. Laursen, F. Fin and F. H. Westheimer, J. Biol. Chem., 241, 421 (1966).