A STUDY OF ENZYME-INHIBITOR INTERACTIONS BY NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY.

A Thesis
submitted to the University of Glasgow
for the Degree of Doctor of Philosophy
in the Faculty of Science,

by

Howard Ashton, B.Sc., (Leicester).

Department of Chemistry,
University of Glasgow.

June, 1971.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

To my friends,

for their unfailing moral support!

Since you desire of me to know
Who's the wise man, I'll tell you who
Not he whose rich and fertile mind
Is by the culture of the arts refin'd,
Nor he that does the Science know,
Our only certainty below,
That can from Problems dark and nice
Deduce Truths worthy of a Sacrifice.

From 'The Reply' by John Norris.

(with thanks to C.McD).

ACKNOWLEDGEMENTS.

The work described in this thesis was initiated in October 1967 in the Department of Chemistry, the University of Leicester. I should like to thank Professor C.W.Rees, presently at the University of Liverpool and my supervisor, Dr.B.Capon, for their initial endeavours which allowed me the opportunity to undertake this research. To the latter I wish to express my thanks for his continued help and interest. I should also like to record my pleasure at being able to move North of the Border in the Spring of 1968 - 'a geographical shift of significant magnitude in the right direction' - for the variety of opportunities it has provided for meeting new people and widening my horizons!

My thanks are due to various members of staff and research colleagues in the department for their help and co-operation; in particular Mr.J.Gall for recording the nmr spectra and Mr.R.K.MacKenzie for his invaluable assistance in writing the Fortran V computer program.

Finally, I should like to use this opportunity to acknowledge the hard work and constant encouragement of my parents in helping me to achieve independence despite my physical disability.

Financial support for the first year of this research was kindly made available by Kingston-upon-Hull Education Committee. A Research Demonstratorship was held for the remaining period and for this I thank Professor R.A.Raphael.

made to interpret these results in terms of the nature of the binding site of this enzyme, the structure of the molecule of which is described in some detail in Part 2.

The inhibition of the α chymotrypsin catalysed hydrolysis of N-3-(2-furyl)acryloyl L-tryptophanamide by the N-trifluoroacetyl derivatives of D-tryptophan and D-phenylalanine was studied at pH 6.35, 7.67 and 8.11 by steady state kinetic methods in an attempt to verify the dissociation constants evaluated for these systems by fluorine magnetic resonance spectroscopy.

There is an upfield shift of the proton magnetic resonance signals of the N-acetyl groups of both anomers of 2-acetamido-2-deoxy-D-glucose when bound to lysozyme. The ^{19}F magnetic resonance signal of 2-trifluoroacetamido-2-deoxy-α-D-glucose, but not that of the β anomer, was also found to shift upfield in the presence of this enzyme. Analysis of the variation of the chemical shift of the α anomer with its concentration gave values of $\triangle$ 78 Hz and K_D 9.1 x 10^{-3}M. The analogous values for the α anomer of 2-acetamido-2-deoxy-D-glucose under identical conditions were 87 Hz and 2.04 x 10^{-2}M. Although both compounds bind in the active site they must be situated differently since if they were situated identically and experience identical changes in magnetic environment on binding, a much larger chemical shift would be expected for the fluoro compound.

Two new compounds, the α and β anomers of methyl-2-trifluoroacetamido-2-deoxy-D-glucoside were synthesised and characterised. Both anomers showed only a small concentration-independent downfield fluorine chemical shift in the presence of lysozyme, in contrast to the upfield shift shown by the N-acetyl resonances of the α and β anomers of methyl 2-acetamido-2-deoxy-D-glucoside. It is thought that the situation in which the N-trifluoroacetyl compound is bound leads to an unfavourable interaction with the enzyme when there is

a methoxyl group at carbon 1.

In Part 7 is described a minor ^{19}F nmr investigation of the hydrolysis of the C-F bond of α-D-glucosyl fluoride by amyloglucosidase. The disappearance of the quartet of signals in the ^{19}F spectrum of the fluoro sugar substrate is accompanied by the appearance of a signal due to the accumulation of fluoride ion in the system.

CONTENTS.

Notes.

i. Sugar nomenclature.

Throughout this thesis the following monosaccharide common names are used:

N-acetyl-D-glucosamine (abbreviated NAG) for 2-acetamido-2-deoxy-D-glucose.

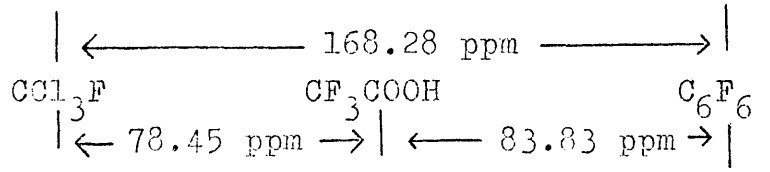
N-trifluoroacetyl-D-glucosamine (abbreviated NTG) for 2-trifluoroacetamido-2-deoxy-D-glucose.

Methyl N-trifluoroacetyl-α-D-glucosaminide for Methyl 2-trifluoroacetamido-2-deoxy-α-D-glucoside, and similarly for the β anomer.

ii. ^{19}F nmr referencing.

It has been suggested that a standard reference should be adopted for the reporting of ^{19}F chemical shifts and trichlorofluoromethane, CCl_3F, has been proposed as a suitable candidate. Shifts quoted with reference to CCl_3F are usually reported as ϕ_c values.

In this thesis ^{19}F chemical shifts have been recorded from either external CF_3COOH at 94.1 MHz or external C_6F_6 at 56.4 MHz. Conversion to ϕ_c values can be accomplished on the basis of the following relationships:



(used as external references),

and,

$$\text{Chemical shift (ppm)} = \frac{\text{Chemical shift (Hz)}}{\text{Frequency}}$$

PART 1.

THE APPLICATION OF NMR SPECTROSCOPY TO THE STUDY OF PROTEIN STRUCTURE AND PROTEIN-SMALL MOLECULE INTERACTIONS.

1.1. CHARACTERISTICS OF THE PROTON MAGNETIC RESONANCE SPECTRA OF PROTEINS.

Some eleven years elapsed after the discovery of the phenomenon of Nuclear Magnetic Resonance before the first spectrum of a protein was published (1). The spectrum obtained, at 40 MHz, was of ribonuclease in deuterium oxide and so only represented the non-exchangeable protons of the protein molecule. Although the spectrum was broad and not well resolved, resonances could be assigned to aromatic protons and to those hydrogen atoms on aliphatic carbon atoms attached only to other aliphatic carbons. The spectrum was interpreted in more detail by Jardetsky and Jardetsky (2) on the basis of their measurements of the spectra of amino acids and peptides in water and D_2O; subsequently the spectra of several other proteins were recorded in the latter solvent and trifluoro-acetic acid (5-12).

The use of NMR techniques to obtain information about the structure of proteins in solution is limited by the overlap of chemical shifts of the large number of structurally and enviromentally non-equivalent protons contained in these molecules in their native, folded conformations. Broadening of the resonances arises from magnetic dipolar interactions which, in small molecules in solution, is largely averaged out because of molecular motion. In large molecules, however, rotation is slower and so broadening becomes a dominant feature of their spectra; for native proteins of molecular weight 10^4-2×10^4 resonance half-widths can be of the order of 10 to 20 Hz. Consequently earlier studies at spectrometer resonance frequencies below 100 MHz usually resulted in only an envelope of the individual resonances being observed, although Bradbury and Sheraga (11) were able to resolve individual C2 proton resonances of the histidine residues of ribonuclease in studies at 60 MHz.

The characteristics of the proton magnetic resonance spectra of native proteins in aqueous solution are strongly

influenced by the tertiary structure of the molecule. Detailed information can be obtained about protein conformation, or interactions that alter conformation, if it is possible to resolve individual resonances, assign them to particular protons and rationalise the resonance positions in terms of protein structure.

Improving the resolution of overlapping resonances may be achieved by increasing the applied magnetic field and, concomitantly, the operating frequency of the spectrometer. This is because the separation between chemically shifted proton resonances increases linearly with the magnitude of the polarising magnetic field. Spectrometers operating at 100 MHz and 220 MHz are available and have been used to record the pmr spectra of native proteins (8,9,13-22).

However, even at these frequencies some overlap of resonances still occurs and techniques are required to dissect out the components of partially resolved groups of resonances. A somewhat novel example of one such technique has employed the fact that addition of Co^{2+} ions to a solution of hens egg white lysozyme perturbs resonances throughout the entire pmr spectrum of this enzyme recorded at 220 MHz (23). The high field spectral region has been analysed in detail and here the induced shifts show a dependence on the concentration of the metal ion that increases in a regular fashion. As a result evidence has been obtained for ten individual resonances in this region of the spectrum. Up to a concentration of 0.15 M Co^{2+} perturbations are concluded to result from the paramagnetism of this ion bound to a single site on the protein molecule. Inspection of the structure of lysozyme (as determined by X-ray spectroscopy) indicates that this binding site could be provided by the carboxyl groups of Glu-35 and Asp-52, which are known to be at the active site of this enzyme. The pH dependencies of some of the shifts indicate that the binding site for the metal ion becomes available by titration of one or more functional groups with a pK_a of about 6.0, supporting this view. Of the resonances

dissected out by this method six have been tentatively assigned to the methyl groups of Leu-8, Leu-17, Met-105 and the δ and δ methyl groups of Ileu-98. These residues are located in various regions of the lysozyme molecule and are 10Å to 14Å from Glu-35, so it appears that Co^{2+}, bound to a single site on the enzyme, can perturb resonance positions of protons throughout most of the molecule.

In the denatured state it seems that the side chains of component amino acids of proteins are largely unconstrained and exhibit chemical shifts for the various structurally non-equivalent protons that do not strongly reflect the amino acid sequence of the protein. Conformational proximity effects and nearest neighbour interactions do not appear to contribute greatly in determining the chemical shifts of side chain protons; therefore, all residues of a given amino acid produce a spectrum of resonance quite similar to that of the free amino acid so that the nmr absorption of the denatured form of the protein can be approximately represented by a superposition of the resonances of the component amino acids, weighted by the amino acid composition of the protein. McDonald and Phillips (21) have compared the pmr spectra of lysozyme, ribonuclease, pepsin, trypsin, apoferridoxin and the two small polypeptides oxytocin and ferridoxin in their extended conformations with those computed for these proteins in a random coil form, that is, when the side chains are in a solvent enviroment. The computed spectra were based on the known amino acid composition of the proteins studied. Very good correspondence between the actual and computed spectra was noted but the technique could not be used with confidence to compute the random coil spectra of proteins in enviroments that were considerably different from neutral aqueous solutions, e.g. trifluoroacetic acid; the pmr spectrum of insulin in trifluoroacetic acid has, however, been successfully analysed in detail by comparison of the spectrum with those of amino acids in this solvent (24,25).

While pmr spectra of proteins in a random coil form provide no direct information about the nature of the folded native protein molecule, they do provide a necessary reference for the measurement of spectral changes that occur when a protein reverts to a folded conformation, that is, when the protein aquires its tertiary structure upon renaturation. Proteins can be obtained in a random coil state in neutral aqueous solution by heating to above their thermal denaturation temperature. For example, lysozyme undergoes reversible thermal denaturation between 68° and 74° and at 79° it is believed to be in an extended random coil conformation, except for the four disulphide bridges in the molecule, which remain intact. Denaturation of native proteins in solution can be achieved by the addition of guanidine or urea, or by using trifluoroacetic acid as the solvent. Adequate denaturation may accompany large changes in the pH of the system.

Because proteins assume a random coil conformation upon denaturation their pmr spectra are, accordingly, often much improved. This can be of use when nmr studies are undertaken using spectrometers operating below 100 MHz when, characteristically, the spectra of native proteins are broad and not well resolved. Bradbury and King (26), in a study of non-covalent interactions in proteins, have reported the pmr spectra of eleven proteins in five denaturing solvents at 60 MHz.

1.2. SOME INVESTIGATIONS OF PROTEIN MOLECULAR STRUCTURE BY PROTON MAGNETIC RESONANCE SPECTROSCOPY.

From the foregoing it will be apparent that the technique of nmr spectroscopy has been of only limited use in the solution of problems of protein chemistry, not least of all because the spectra obtained generally consist of a number of broad resonances, making it difficult to separate the contributions from each kind of proton in the molecule. The exceptions to this pattern occur with protons attached to nitrogen atoms which do form broad, though well defined bands,

the aromatic protons of tryptophan, phenylalanine and tyrosine, the C4 protons of histidine, which can be partially separated in favourable cases, and the C2 proton of the imidazole ring of histidine. These latter resonances, when observable, are usually well defined, occuring to the low field side of the aromatic signal. This feature has made them particularly suitable as a means for study of protein behaviour; the implication of imidazole in the active site of a number of enzymes is further justification for focussing attention on these particular resonances. It has been found that, in general, the C2 protons of histidine residues can either be observed, as is the case with ribonuclease or lysozyme, or cannot be visualised at all (using a 60 MHz high resolution instrument with spectrum accumulation facilities) because of structural broadening of the resonance (27). Trypsin, chymotrypsin and their zymogens fall into the latter category and it may be that the histidine residues in the molecules of these proteins are held rigidly by an interaction of the imidazole rings with another group; the C2 histidine resonance is observed, however, in the pmr spectrum of α chymotrypsin in trifluoroacetic acid.

In 1966 Bradbury and Sheraga (11) demonstrated that the resonances of the C2 protons of the four histidine residues in ribonuclease A in D_2O could be observed at 60 MHz. When the pD of the system was varied between 5.4 and 8.0 the resonance was noted to be split into three signals. The pD dependence of the chemical shift of each of these signals was determined and from plots of shift vs pD it was possible to calculate the microscopic dissociation constant, pK, attributable to each of them. The signals having approximate pK values of 5.4 and 5.8 were tentatively assigned to the two histidine residues involved in the active site of this enzyme. The pK of the third signal was about 6.6. Subsequent experiments at 100 MHz showed four resolvable peaks (15), the chemical shift of each of them being pH dependent with pK's of 5.8, 6.2, 6.4 and 6.7 (Figures 1.1. and 1.2).

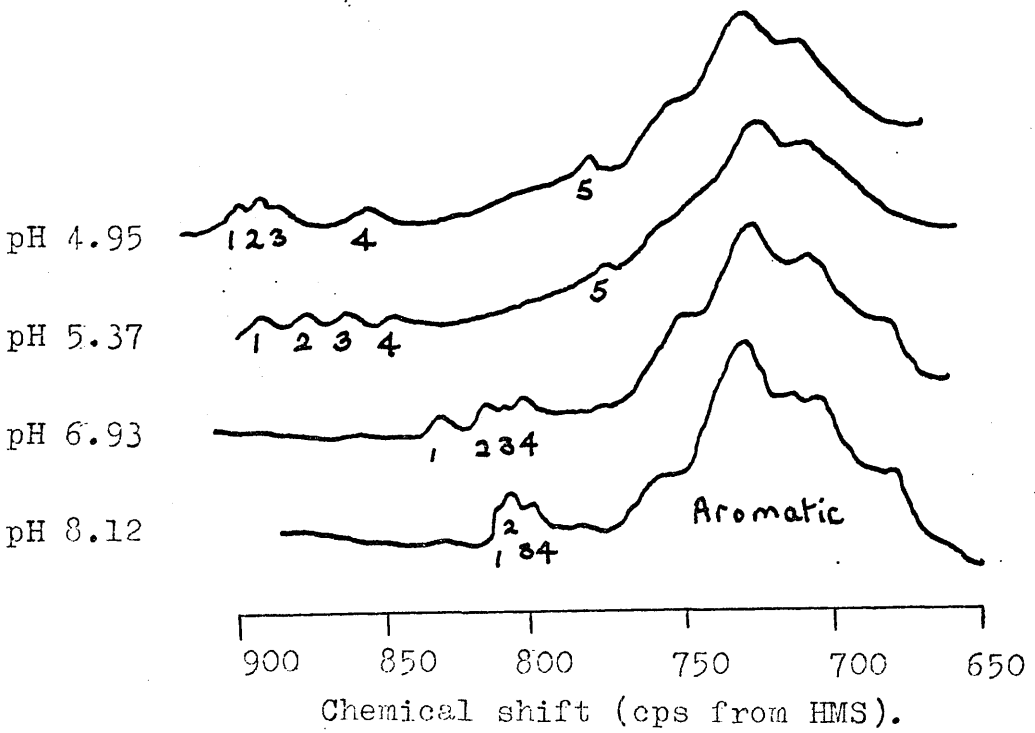


Figure 1.1 Changes in the aromatic region of the 100 MHz nmr spectrum of RNase A as a function of pH. Peaks 1-4 are the C-2 imidazole peaks of the fourhistidine residues. Peak 5 is a C-4 imidazole resonance.

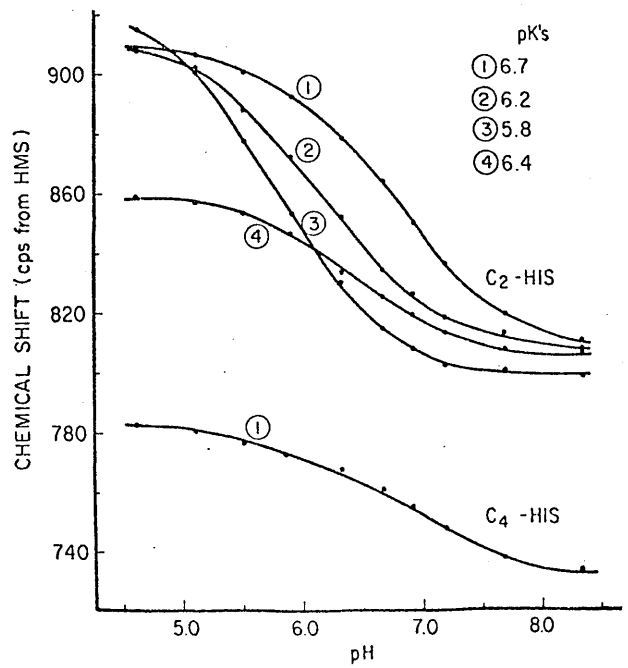


Figure 1.2 Titration curves of the C-2 peaks of the histidine residues of RNase A. Curves 1-4 correspond to peaks 1-4 of Figure 1.1.

Thus it was now possible to observe individual resonances for each of the four histidine residues in ribonuclease A, these being at positions 12,48,105 and 119 in the amino acid sequence of this enzyme (28). Meadows et al (18) assigned these four resonances in the following manner. Carboxymethylation of either His-12 or His-119 produced marked changes in the pK's of two histidine residues (those responsible for signals 2 and 3, Fig.1.1.) as determined from the pmr spectrum. Since the same two signals were affected by alkylation in both positions it was concluded that:

a. they correspond to His-12 and His-119.

b. these two residues are so close together in the three dimensional structure of the enzyme that alkylation of either one affects the pK of both.

The spatial proximity of His-12 and His-119 was originally suggested on the basis of chemical modification studies (29), and has been confirmed by X-ray diffraction methods (30,31) but the nmr evidence is independent of these findings. Consequently, the resonances having pK's of 5.8 and 6.2 (signals 2 and 3) correspond to His-12 and His-119 whereas those of pK 6.4 and 6.7 arise from the C2 protons of His-48 and His-105. In the model of ribonuclease A based on X-ray crystallography three histidine residues are on the surface of the enzyme while the fourth, His-48, is buried within the molecule. On this basis, the signal with a pK of 6.4, which is noted to be much broader than the other three signals, and the chemical shift of which, at low pH, is some 50 cycles upfield from the other resonances, was tentatively assigned to His-48. The other member of the pair, with a pK of 6.7 , could then be assigned to His-105. Exchange of the C2 proton of His-12 for deuterium eliminates signal 3 from the spectrum and so the remaining signal, pK 5.8, arises from His-119.

Nuclear Magnetic Resonance Spectroscopy has been used to determine the temperature dependence of the pK values of the histidine residues in ribonuclease, making possible calculations

of the values of free energy and of enthalpy and entropy of ionisation of these residues (32). On the basis of the results tentative proposals could be made about the extent of protein unfolding that had occurred over the temperature range studied, as well as the degree of solvation of each of the four histidine residues. Titration curves for the single histidine residue in human lysozyme have been derived at different temperatures, and a dependence of the dissociation constant on temperature shown to exist (33). Enthalpy of ionisation, ΔH, and entropy of ionisation, ΔS, could be calculated and had the values 6.4 Kcal/mole and -11.6 e.u., respectively. This larger value for ΔS here, compared with that for histidine in aqueous solution, may be taken to indicate reduced solvation of the imidazolium ion in the protein (32). On the basis of the values obtained for histidine residues in other proteins, where it is known with some certainty that the location of these residues is within the molecule, it would appear that the single histidine residue in human lysozyme is partially buried.

1.3. THE INVESTIGATION OF PROTEIN-SMALL MOLECULE INTERACTIONS BY NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY.

Probably the most successful studies on the structure of proteins have been carried out using X-ray crystallography. This technique is, however, subject to certain disadvantages and the extent to which its findings are relevant to the conformation of proteins in solution is subject to some degree of uncertainty. The first attempts to study proteins in a solvent enviroment using nmr spectroscopy produced spectra which did not lend themselves to detailed analysis and so produced little information about protein structure. The difficulties encountered in deciphering these spectra in order to obtain useful information might possibly be overcome by systematic studies of suitable model peptides (24, 34-36). Alternatively, the preparation of selectively deuterated protein and the study of denatured proteins, coupled with the use of high resolution spectrometers operating at high frequencies,

may represent other approaches to the problem.

Perhaps the most rewarding application of nmr spectroscopy to the study of proteins has resulted from its use as a means whereby conformational changes that may accompany the interaction of proteins with substrates, inhibitors, cofactors and metal ions, may be observed. Work on this topic conveniently divides itself into two categories:-

i. a study of the perturbations occurring in the nmr spectrum of the macromolecule when interactions take place between it and a small molecule or ion.

ii. the reciprocal technique, that of recording changes in the nmr spectrum of the ligand molecule produced by specific interaction with the macromolecule.

1.3.a. Method i.

This method has been extensively applied to the study of ribonuclease and lysozyme.

The observation of four low field resonances in the pmr spectrum of ribonuclease, and their assignment to the four histidine residues in the molecule of this enzyme, has already been mentioned (p.5). Use has been made of this salient feature of the spectrum to investigate the binding site of the enzyme by observing changes in the position of these resonances when the enzyme is in the presence of mononucleotide inhibitors (15). Addition of 3' cytidine monophosphate at a molar ratio as low as 1:10 produces a marked shift (3 Hz) and broadening of only one of the four histidine signals (no.2, see p.6), the other three remaining unaffected. When 5' cytidine monophosphate is added there occurs a downfield shift of signals 2 and 3. These phenomena were studied at a pH near to which binding is known to be maximal (37, 38). As the inhibitor to enzyme ratio approaches 1.0 the presence of either of these cytidine inhibitors induces all four signals to shift with varying degrees, suggesting that a conformational change in the enzyme has taken place. It is significant that signals 2 and 3, the ones most affected by the binding of the

inhibitors should later be identified as belonging to His-12 and His-119 (18),the residues known to be involved in the active site of ribonuclease (30). Further work on the binding of various inhibitors to ribonuclease A has involved recording changes in the pK values of the four histidine residues (as determined from plots of chemical shift vs pH) as well as noting shift changes in the aromatic region of the spectrum of the enzyme (39).The inhibitors studied were cytidine,5 methyl 2' cytidine,2'and 5' cytidine monophosphates and phoshate and sulphate ions. The structural information deduced from the nmr results obtained were summarised as four postulates and these,when combined with X-ray crystallographic data,made it possible to give a detailed picture of the mode of binding of these inhibitors to this particular enzyme.

A similar,though less intensive,investigation has been made of the pH dependence of the chemical shifts of the C2 proton signals of the histidine residues of ribonuclease T_1 in the presnce and absence of inhibitors (40); pK values for only two of the three histidine residues could be derived in the absence of inhibitors. The results have been interpreted in terms of interactions of the histidine residues with carboxylate anions of acidic amino acid residues. The changes which occur in the histidine C2 pmr absorption region of ribonuclease T_1 on addition of guanosine,and 2' and 3' guanosine monophosphate are indicative of the presence of these residues in the active site of this enzyme. The differences between the pmr spectra of the complexes formed by ribonuclease T_1 with 2' guanosine monophosphate and 3' guanosine monophosphate suggest that these nucleotides are bound differently to the same site,or to different sites,on the enzyme.

Staphylococal nuclease is also known to possess four histidine residues,the C2 proton signals of which can be resolved in the pH range 5.0 to 9.0.(15). This enzyme requires calcium ions for activity and addition of this cation causes a selective downfield shift of one of these signals when the pH of the system is above 6.8. This shift is attributed to an

interaction between this metal ion and the uncharged imidazole ring: under similar conditions the other three signals shift downfield only slightly. The effect of the inhibitor thymidine 5' monophosphate on the aromatic region of the pmr spectrum of this enzyme in the presence of calcium ions has been studied between pH 6.4 and 9.3. Large selective shifts of the histidine C2 proton signals of the kind noted with ribonuclease A were not observed and from this it was concluded that the two nucleases have different mechanisms of inhibitor binding.

The titration studies on the single histidine residue in human lysozyme, referred to on p.8., reveal the pK to be 7.1 at 32°, which is a higher value than that for any of the individual histidine dissociation constants for hens egg white lysozyme, ribonuclease, staphylococal nuclease and α lactalbumin. This would seem to indicate that this histidine residue must be in close proximity to a negatively charged group(s), or subject to strong hydrogen bonding with an acceptor group.

Human lysozyme is inhibited by the same oligosaccharides for which hens egg white lysozyme has an extensive binding site. Addition of di- and tri-N-acetyl glucosamine has no significant effect on the histidine titration up to pH 7.3 (33) and in the case of hens egg white lysozyme, addition of N-acetyl glucosamine results in no systematic shifting of the histidine C2 proton signals, nor is any line broadening observed (15). Both these results suggest that the single histidine residue in these enzymes is not implicated in the binding site for these inhibitors in solution.

Assignments have been made in the low field (aromatic) region of the spectrum of denatured hens egg white lysozyme by comparison with the aromatic amino acids under similar conditions. At pH 4.5 slight broadening of this part of the spectrum results on addition of high molar ratios of N-acetyl glucosamine. However, with the dimer and trimer of this compound selective effects are observed, the changes being essentially complete with a 7.5 M ratio of the disaccharide(17)

and a 2M ratio of the trisaccharide, reflecting differences in their association constants (41). The spectra obtained on saturation of the enzyme with di or tri N-acetyl glucosamine are almost superimposable, suggesting that there is an identical involvement of the tryptophan residues in the binding subsites (designated B and C) of the enzyme for the common two saccharide residues of these inhibitors (41,42). Similarly, the aromatic envelopes obtained on addition of these same two inhibitors to human lysozyme are almost indistinguishable (33) indicating an identical conformation of the enzyme with each inhibitor bound, and the presence of binding subsites for oligosaccharides similar to those in hens egg white lysozyme.

1.3b. Method ii.

The interpretation of nmr spectra is based on three parameters which characterise the absorption of radiofrequency radiation by atomic nuclei placed in a magnetic field. These are:

a. the frequency of the absorbed radiation, expressed as the chemical shifts relative to an arbitary standard absorption line.
b. the multiplicity of the lines originating from a given group of nuclei, determined by coupling between neighbouring nuclei and described by the appropriate coupling constants.
c. the decay times characterising the return of the nuclei excited by the absorption of radiation to a lower energy state, referred to as the relaxation times.

Two of the above parameters, viz. chemical shift and relaxation time, have received about equal attention in nmr studies of small molecule-macromolecule interactions. Their theories are well understood and have been developed to the point where they can be used to gain quantitative information about these phenomena.

1.3.bi. Relaxation Times.

A review of the subject of relaxation times in

this context has been published (43). Briefly, the basis of the method is that line widths in nmr spectra are sensitive to the degree and kinds of molecular motion present in the sample. Increased line widths are generally associated with diminished rates of molecular motion, particularly rotational motion. When a small molecule is bound by a larger one, increased line widths may be observed in its nmr spectrum since the motion of the bound molecule may be restricted in the complex. Zimmerman and Brittin (44) have developed a theory which deals with the case in which a small molecule exists in equilibrium between two or more phases i.e. bound and free, having different characteristic relaxation times and hence different line widths. It is predicted that the observed line widths depend on the rate of exchange of the molecules between bound and free phases, as well as on the equilibrium binding constant. For small molecules, say those of molecular weight less than 600, which have nmr spectra consisting of more than one signal, the ratio of the line widths at half height for any two signals remains unaffected by changes in the composition of system, therefore non-specific effects such as viscosity changes can be automatically corrected. Selective broadening of one or more signals can, therefore, be interpreted in terms of specific interactions involving the chemical groups to which the signals correspond. The stoichiometry of the binding process can be determined by plotting line width (or its inverse, the relaxation time) against the relative concentration of the interacting molecular species.

The application of relaxation time measurements was pioneered by Fisher and Jardetsky (45) who investigated the binding of penicillin G by serum albumin by this method. Marked broadening of the pmr signals of the penicillin was noted and a relatively large increment in the relaxation rate of the phenyl peak suggested that this portion of the molecule is in the binding site; other similar examples are to be found in the review by Jardetsky (43).

Hollis (46,47) has made a detailed study of the selective broadening effects resulting from the binding of the oxidised form of the cofactor nicotinamide adenine dinucleotide (NAD^+) by yeast alcohol dehydrogenase which were briefly reported in 1963 (48). The effects of this enzyme and equine liver alcohol dehydrogenase on the nmr spectra of both the oxidised and reduced forms of this coenzyme, as well as those of ethanol and acetaldehyde, were studied. The line broadening effects observed appear to eliminate the possibility of rigid binding of substrates with rapid turnover and suggest that the presence of the coenzyme is a requirement for binding of the substrate in the rigid mode that will account for the stereospecificity characteristics of these two enzymes.

Considerable broadening of the proton resonances of phosphonoacetamide has been noted to occur in the presence of the catalytic subunit of the enzyme aspartate transcarbamylase (49). The fact that this phenomenon cannot be observed with N-methyl phosphonoacetamide or methyl phosphonate indicates that interaction between the amide group of the methyl analogue and the enzyme limits rotation about the P-C bond in the first compound but not in the latter two. Resonance line width broadening of the methylene protons of succinate, a competitive inhibitor of the enzyme, is noted only when the system contains both the catalytic subunit and carbamyl phosphate. In this case the broadening is primarily ascribed to an increase in the lifetime of the succinate-enzyme complex, induced by the presence of carbamyl phosphate. Only analogue compounds which are not larger than carbamyl phosphate and which have a carbonyl group in addition to a phosphate moiety in the molecule are able to induce broadening of the succinate line.

Evidence for the zinc atom of carboxypeptidase A being the site of inhibitor binding has been obtained from a study of preferential broadening of certain proton resonances of competitive inhibitors such as indoleacetic acid and tertiary butyl acetic acid (50). The requirement for the substrates

of this enzyme to have a free carboxylate group in the anionic form, and the overall similarity between the specific substrate and the inhibitors, lead to the conclusion that the metal atom in carboxypeptidase A enhances the enzymic activity by binding the substrates through their free carboxylate groups, including them in its first co-ordination shell.

1.3.bii. Chemical Shifts.

In general chemical shift changes for some nuclei of a substrate or inhibitor molecule are to be expected when it is bound to an enzyme. If the small molecule exchanges rapidly between the free and complexed environments, the observed spectral positions of the affected resonances will be shifted from their corresponding positions in the spectrum of the unassociated molecule; the amount of shift is dependent upon the chemical shift of these nuclei in the complex, as well as on the fraction of the inhibitor or substrate in the system which is in the complexed state. It might be argued that chemical shift changes are more stereospecific than line width changes, in which case they provide a more sensitive probe of the active site and the nature of enzyme-substrate or enzyme-inhibitor interactions.

The technique of recording changes in the ligand spectrum produced by a specific interaction with a macromolecule has been used extensively to observe the association of N-acetylated mono and oligosaccharides with lysozyme. The methyl group protons of the acetamido side chains in these inhibitors give an intense unsplit resonance in their pmr spectra and consequently are well suited to investigation of such association phenomena based on observation of chemical shift changes in selected nuclei.

Thomas has reported chemical shifts of the order of 3 cycles to higher field for the acetamido methyl protons of α and β N-acetyl glucosamine in the presence of lysozyme (51). These shifts were attributed to the diamagnetic influence of the tryptophan residues that are known to be at the binding site of this enzyme.

The observed chemical shift of this same resonance in the presence of lysozyme is different for the α and β anomeric forms of 2-acetamido-2-deoxy-D-glucopyranose, indicating either a difference in the affinity of the enzyme for the two anomers, or different magnetic enviroments for the methyl protons in the complex. By using the methyl glycosides of these two anomeric sugars, where the conformation of each is 'frozen' (52), it has been possible to analyse the chemical shift data quantitatively and hence determine the dissociation constant, K_D, as well as the chemical shift, Δ, of the acetamido methyl protons in the enzyme-inhibitor complex for both anomers. The results, $K_{D\alpha} = 5.2 \pm 0.4 \times 10^{-2}$ M, Δ_α 0.55±0.02 ppm, and $K_{D\beta} = 3.3 \pm 0.2 \times 10^{-2}$ M, Δ_β 0.54±0.04 ppm, indicate binding to the enzyme with slightly different affinities, but with the acetamido methyl protons of both anomers experiencing identical magnetic enviroments in the complex. The effect of pH and temperature on the dissociation constant of methyl-2-deoxy-D-glucopyranoside when bound to lysozyme has been measured (53). Observation of the chemical shifts of both the acetamido and glycosidic methyl proton resonances of this inhibitor when bound to the enzyme showed that no conformational change occurs in the binding site of lysozyme up to 53° and over the pH range 2.5-10.0. In the absence of such changes it has been concluded that the ionisable groups affecting the binding constant and the chemical shift of the acetamido methyl protons are very close to the inhibitor binding site. The chemical shift of the glycosidic methyl protons is noted to be pH independent over the whole pH range studied, while that of the acetamido group shows a pH dependency and gives three pK values which may be assigned to the carboxyl residues at the binding site of this enzyme, thus: pK 6.1 to Glu-35, and pK's 4.2 and 4.7 to Asp-101 and Asp-103, respectively.

A study of the mutarotation of α and β N-acetyl-glucosamines to their equilibrium mixture by nmr spectroscopy in the presence of lysozyme has allowed the calculation of the binding constants of both anomers to a single site on the enzyme (54).

The magnitude of the dissociation constants obtained, $K_{D\alpha}$ 1.6 x 10^{-2}M and $K_{D\beta}$ 3.3 x 10^{-2}M, demonstrate that there is very little difference in the relative abilities of the anomers to be bound by the enzyme. Further, the factor of two between the relative constants $K_{D\alpha}$ and $K_{D\beta}$ represents an energy difference of only 0.4 Kcal at 31°. This, coupled with the observation that α and β N-acetyl-D-glucosamine are competitive for the same site on the enzyme, would ordinarily provide good evidence that the binding of these anomers is identical. However, the value of $\triangle$, the chemical shift of the acetamido methyl protons in the enzyme complex, is 0.51 ppm for the β anomer, but 0.68 ppm for the α form. These values clearly indicate the bound orientations of the two anomers to be different and provide a good example of the extreme sensitivity of magnetic resonance methods to environmental change when expressed as $\triangle$ values. It has been shown that the $\triangle$ value for methyl α - and methyl β -N-acetyl-D-glucosaminide bound to lysozyme are very similar (see p.16), being close to that for N-acetyl-β-D-glucosamine, suggesting that the binding orientation of these three inhibitors at the active site of this enzyme is essentially identical.

In a preliminary investigation of the binding of chitobiose and methyl β-chitobioside by lysozyme using pmr spectroscopy (55), the only effect to be observed was an enzyme induced broadening of the reducing end acetamido methyl proton resonance in chitobiose and a similar broadening of the signal of the acetamido methyl group proximal to the glycosidic end of the glycoside. This same phenomenon was again observed in a more detailed study of the association of these same two inhibitors, and of chitotriose, methyl β-chitotrioside and chitotetraose

with lysozyme (56). The broadening was ascribed to slow exchange rates since chemical shift differences were noted to exist for these nuclei between the free and enzyme bound environments. Quantitation of these shifts,under conditions of rapid exchange,made it possible to assign specific magnetic parameters to the three contiguous binding subsites on lysozyme to which acetamido pyranose residues bind. From these assignments the relative ways in which all of the inhibitors occupy these subsites have been delineated and the results found to be in general agreement with X-ray studies on crystalline lysozyme-saccharide complexes (57).

1.4. APPLICATIONS OF FLUORINE NMR SPECTROSCOPY TO THE STUDY OF BIOLOGICAL SYSTEMS.

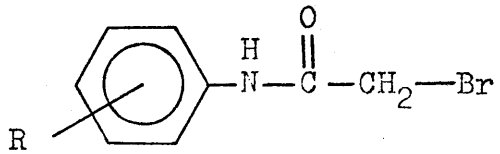
The potential of the fluorine nucleus as a probe for the study of enzyme-substrate/inhibitor interactions was first illustrated by Spotswood's observations of the chemical shift of the fluorine resonance of N-acetyl DL p-fluoro-phenylalanine in the presence of chymotrypsin (58). In citrate buffer, pD 6.0, the ^{19}F resonance of this inhibitor appears as an AA'BB'X multiplet. Resolution into two overlapping multiplets corresponding to the individual D and L isomers was achieved on addition of native chymotrypsin but not with diisopropyl phosphofluoridated chymotrypsin, lysozyme or ribonuclease. The chemical shifts, which were noted to be much larger (up to 10 Hz) than those observed for the acetyl proton resonances of N-acetyl DL phenylalanine under similar conditions, were utilised to derive a value for the enzyme-inhibitor dissociation constant according to the theory discussed in Section 3.

Study of the ^{19}F resonances of certain fluorinated surfactants dissolved in organic solvents and aqueous solution has been made as one possible approach to the problem of micelle structure (64). Zeffren and Reavill adopted the principles of this study for an investigation of the nature of the active site of chymotrypsin (65). The sharp simple one peak fluorine resonance of N-trifluoroacetyl DL phenylalanine noted in the absence of chymotrypsin was resolved into two singlets, arising from both enantiomers, on addition of the enzyme. Each of the signals had moved downfield from the position of the single resonance noted in buffer alone, while their positions when dissolved in organic solvents was well upfield of the buffer location.

Further demonstration that nmr techniques provide a direct, simple and precise means of measuring the rate constants for biological exchange reactions has been given by Sykes (62) who has studied the binding of N-trifluoroacetyl D-phenylalanine by α chymotrypsin, diisopropyl phosphofluoridated chymotrypsin and chymotrypsinogen A using relaxation time and chemical shift measurements of the equivalent fluorine nuclei of this inhibitor.

A number of attempts to gain an insight into the nature of the active site of chymotrypsin have employed the 'reporter group' concept. One of these has involved the preparation of modified α chymotrypsins which contain fluorine in the reporter group and observing the ^{19}F resonance spectrum of the resulting species (66,67). Apart from removing the nmr signal of interest from the resonances of the enzyme molecule, the method also benefits from the extreme sensitivity of the fluorine chemical shifts to the environment.

The work of Lawson and Schramm (68) has shown that organic molecules with an activated halogen atom and a larger non-polar (usually aromatic) group will inactivate α chymotrypsin irreversibly, first by forming a complex with the active site, followed by the displacement of the halogen atom by the sulphur atom of Met-192, this residue being thought to be close to a hydrophobic binding locus. On the basis of these findings the ability of a series of methyl and trifluoromethyl substituted N-phenyl α bromoacetamides to irreversibly inhibit α chymotrypsin has been studied (66).



I.

Ia. R:H
Ib. R:o-CH_3
Ic. R:m-CH_3
Id. R:p-CH_3
Ie. R:o-CF_3
If. R:m-CF_3
Ig. R:p-CF_3

The reaction of the above bromoacetanilides with the enzyme involves the formation of a reversible enzyme-inhibitor complex EI, followed by an essentially irreversible displacement of the halogen atom by the nucleophilic thioether group on the enzyme to give a catalytically inactive enzyme EI;

$$E + I \underset{}{\overset{K_I}{\rightleftharpoons}} EI \longrightarrow \underline{EI}$$

The inhibition (binding) constant K_I was determined for each of the inhibitors Ia-g by conventional competitive inhibition kinetic experiments using N-glutarylphenylalanine p-nitroanilide as substrate under conditions where the formation of EI was negligible.

Table 1.1. Inhibition constants for the substituted bromoacetanilides Ia-g binding to ∝ chymotrypsin.

Substituent.	K_I mM.
a. H.	4.2
b. o-CH_3	7.3
c. m-CH_3	3.0
d. p-CH_3	0.88
e. o-CF_3	~ 30.0[1]
f. m-CF_3	7.8
g. p-CF_3	~ 10.0[1]

0.05M Tris buffer, pH 7.45, ethanol as cosolvent.

1. Estimated because of insolubility.

Table 1.2. Comparison of the inhibition constants for the methyl and trifluoromethyl substituted inhibitors.

Position of substituent.	$\frac{K_I\ CH_3}{K_I\ CF_3}$
ortho.	0.25
meta.	0.38
para.	0.09

The results given in Tables 1.1 and 1.2 show that the major effect of methyl substitution in this system is to modify the inhibition constant K_I. Substitution of an alkyl moiety in the portion of a reversible inhibitor which utilises the hydrophobic mechanism for binding usually leads to a stronger enzyme-inhibitor interaction and consequently a smaller value for K_I. With a methyl group placed in the p position binding is greatly enhanced; the binding ability of the o and m methyl compounds is not significantly different from that of the unsubstituted compound Ia (Table 1.1).

Trifluoromethylation at either the o or m ring positions would appear to decrease the binding ability of the inhibitor by a similar amount. It is the p substituted N-phenyl α bromoacetamides which are much more responsive to the exchange of a methyl for a trifluoromethyl group. The binding of the p methyl compound is about 11 times more effective than that of the trifluoromethyl analogue.

Transverse nuclear relaxation times for the enzyme active site when irreversibly inhibited with compounds Ie-g have been measured by ^{19}F nmr spectroscopy (67).

Table 1.3. Transverse relaxation times for the trifluoromethyl substituted α chymotrypsins and inhibitors.

	T_2 (sec) for CF_3 group at:		
	ortho.	meta.	para.
Enzyme-inhibitor complex (native form).	0.04	0.016	0.006
Enzyme-inhibitor complex (denatured form).	0.2	0.06	0.07
Inhibitor, Ie-g.	1.4	1.0	0.95

The decrease in the relaxation times of both the native and denatured forms of the enzyme is consistent with the molecular size of these biopolymers, with concomitant slowing of Brownian motion. It will be noted from Table 1.3 that the motion of the trifluoromethyl group is more restricted in the native enzyme derivatives than in the random denatured forms and also that this restriction becomes progressively more severe as the CF_3 group is placed o, m and then p to the acetamido linkage which holds the aromatic ring to the enzyme. The aromatic portion of inhibitors Ie-g fits into the same region of the active site of α chymotrypsin as that occupied by the tosyl group of the tosylated enzyme. Such a model implies that the CF_3 group, when in the ortho position, must be relatively exposed while when it is placed para it will be located deep in the hydrophobic pocket. The interactions between the enzyme and a p CF_3 group must be especially strong since T_2 is reduced by a factor of about 7 relative to when this same group occupies the o position.

It would appear that in the complex formed between α chymotrypsin and p-trifluoromethyl N-phenyl α bromo-acetamide the inhibitor molecule is held so tightly that it has essentially stopped rotating about its C_3 axis and that its rotational correlation time is, to a large extent, determined only by the overall Brownian motion of the protein molecule. Line broadening effects noted in a pmr study of the binding of trans cinnamate (69) and D-tryptophan (170) by α chymotrypsin also suggest that these molecules too are bound so well as to have little freedom of motion independent of the motion of the enzyme.

PART 2.

THE α CHYMOTRYPSIN MOLECULE.

2.1. THE STRUCTURE OF α CHYMOTRYPSIN.

Like trypsin, chymotrypsin is secreted from the pancreas as an inactive zymogen precursor. Unlike trypsin, however, a family of chymotrypsins results from differing degrees of proteolytic alteration of two precursors, chymotrypsinogens A and B. The former has been extensively studied and is well characterised (70,71). From a number of studies it is concluded to be a single chain polypeptide of 245 amino acid residues having a molecular weight of about 25,100. At the carboxyl terminus is an asparagine residue (72,73), while chymotrypsin has, in addition, leucine and tyrosine residues as carboxyl termini. In the amino terminal position Bettelheim found a half cystine residue (74), with additional N-terminal isoleucine and alanine residues appearing in the α chymotrypsin molecule (75,76). A detailed account of the mechanism of activation of the zymogens has been published (71) and only an outline of the steps involved will be attempted here.

The relationship between the individual members of the chymotrypsin family may be represented thus:-

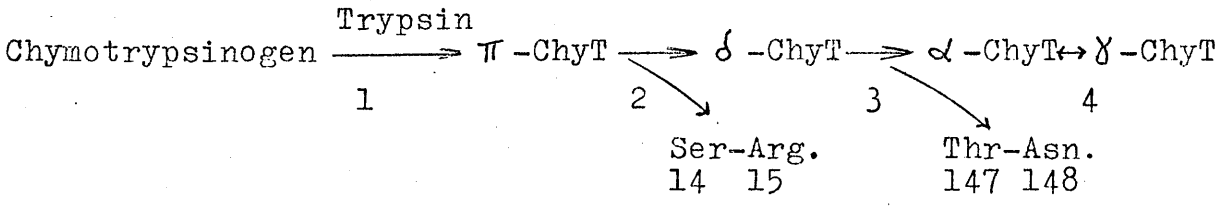


Figure 2.1.

The steps subsequent to trypsin activation do not increase chymotrypsin activity but, rather, consist of autocatalytic cleavage and conformational adjustments of the enzyme molecule. The α and γ forms of the enzyme are not in equilibrium with each other but are interconvertible under the appropriate conditions.

The processes leading to α chymotrypsin involve the sequential liberation of two dipeptides, threonyl-asparagine and seryl-arginine (Fig.2.1., steps 2 and 3). This has the effect of cutting the original single polypeptide chain of the zymogen into three shorter segments, designated the A, B and C chains, which are held together by five disulphide bridges. Figure 2.2 illustrates the relationship of the amino acid residues important in the activation process.

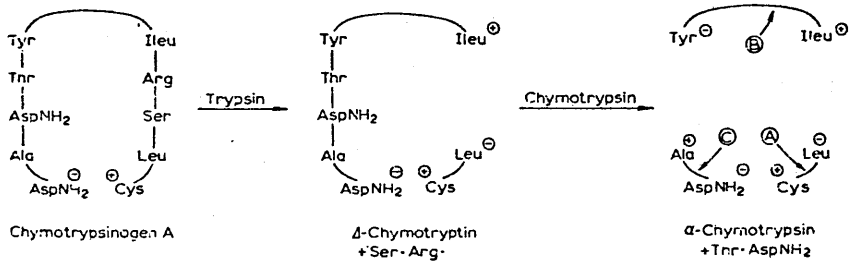


Figure 2.2

Prior to the liberation of seryl-arginine dipeptide, the activating trypsin first splits an arginyl-isoleucine bond to give π-chymotrypsin. The breaking of this bond is associated with the appearance of enzymic activity - in fact this form of the enzyme is the most active of all the known chymotrypsins. Autolysis of the second bond of one of its C-terminal sequences to give δ-chymotrypsin completes this step.

Under conditions where activation is referred to as 'slow' (77), two additional processes have time to take place. Chymotrypsinogen is slowly attacked by chymotrypsin and the δ form slowly autolyses. Two bonds, tyrosyl-threonine and asparaginyl-alanine, located in another region of the molecule, are split and a second dipeptide, threonyl-asparagine, is liberated, chymotrypsinogen A being converted into a new, activatable, form known as neo-chymotrypsin. This protein is capable of being converted to α chymotrypsin by the action of trypsin which liberates from the molecule the same dipeptide

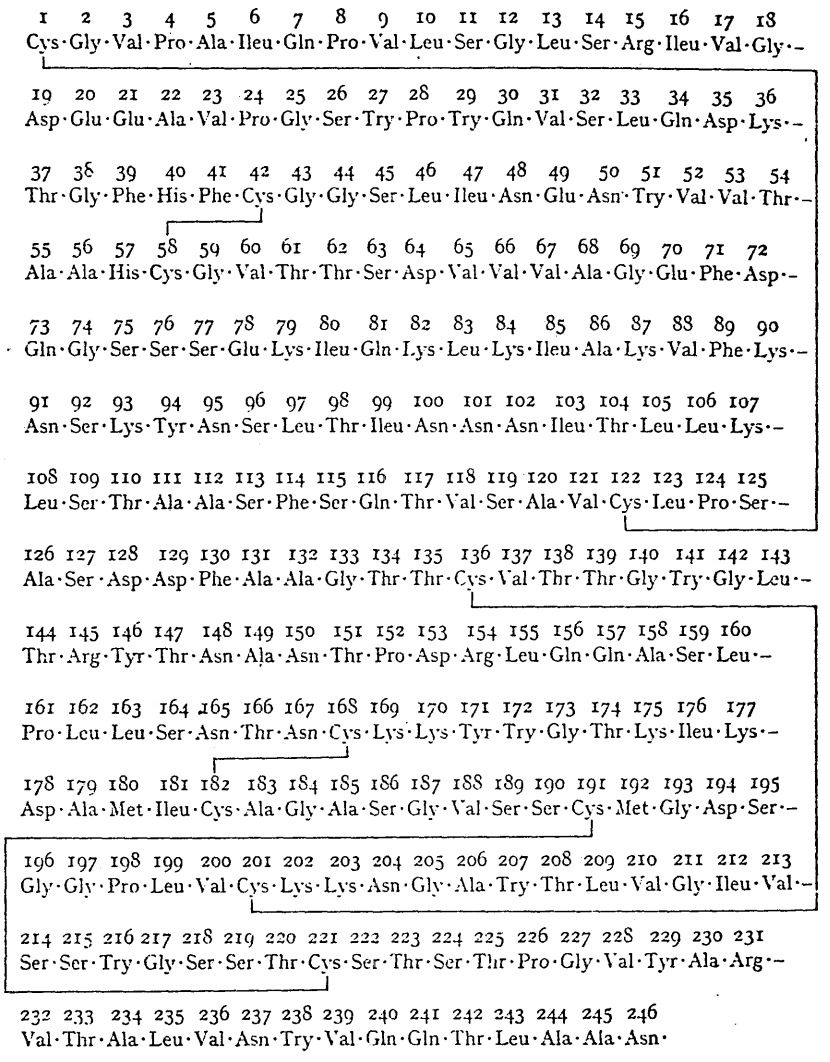


Figure 2.3. The amino acid sequence of Chymotrypsinogen A according to Hartley (81).

as that which is lost in the first step depicted in Figure 2.2.

Intensive studies have been undertaken to determine the amino acid sequence of both chymotrypsinogen A and α chymotrypsin. Meedom was able to identify the three chains constituting the molecule of the latter after oxidation with performic acid (78-80). The A chain is short, comprising 13 residues, and occupies the N-terminal position in the zymogen. Chains B and C contain, respectively, 131 and 97 residues. The C chain contains the carboxyl terminus of the precursor. The total amino acid sequence was one of the first to be determined for any enzyme (81) and has been confirmed by an independent study (82). Both the sequences published commence with the A chain; following the 'activation peptide', Ser-Arg residues 14 and 15, is the B chain which terminates with the other 'activation peptide', Thr-Asn residues 147 and 148. The two histidine residues in the α chymotrypsin molecule are located in the B chain and are closely related, structurally, through one of the five disulphide bridges.

The interpretation of the electron density map of α chymotrypsin is based on the determination of the sequence of the 245 amino acid residues of which chymotrypsinogen is composed. Figure 2.4, of the known hydrogen bonds between atoms in the main polypeptide chain, has adjacent chains drawn antiparallel from which it is apparent that the hydrogen bonding between adjacent chains is predominantly of the antiparallel pleated sheet type. However, to state that the molecule consists to any great extent of this structural feature would be an oversimplification.

Two helical regions occur in the molecule; one of them involves residues 164-171, the other being a portion of α helix at the C-terminal end. Both of these sections contain hydrogen bonds which link any particular residue to its fourth neighbour back along the chain. In the rest of the molecule hydrogen bonds to the third nearest neighbour are very common. The

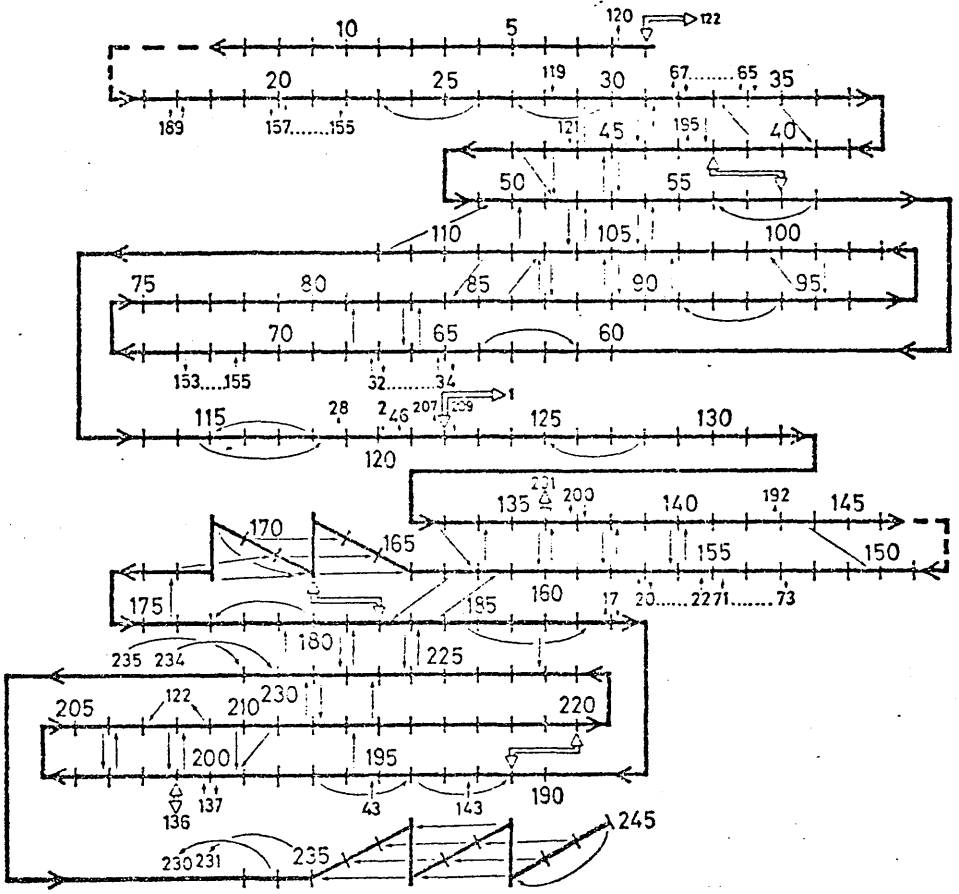


Figure 2.4. The hydrogen bonds between the main chain peptide bonds in α chymotrypsin. Disulphide bridges are indicated by the white arrows.

whole structure is made up of a number of loops in which the polypeptide chain turns back on itself, a number of antiparallel type hydrogen bonds stabilising each loop, some of which are further stabilised by the presence of disulphide bridges.

2.2. THE IMPORTANT FUNCTIONAL GROUPS OF α CHYMOTRYPSIN.

To date seven funtional groups have been implicated, in some manner, in the activity of this enzyme. They are:-

a. the hydroxymethyl group of Ser-195.
b. the imidazole group of His-57.
c. the methylmercapto group of Met-192.
d. the α amino group of the N-terminal Ile-16.
e. the indole group of a tryptophan residue.
f. the carboxyl groups of Asp-102 and Asp-194.

Evidence for the participation of a tryptophan residue(s) in the mechanism of action of α chymotrypsin is meagre and merely equates loss of one or more of these residues from the molecule with loss of enzymic activity.

X-ray crystallography and work on the pH dependent conformational changes that occur in the enzyme molecule have shed most light on the importance of the carboxyl groups of Asp-102 and Asp-194, consequently discussion of these two residues will be left until the appropriate sections.

2.2.a. Serine 195.

In common with other serine proteinases, chymotrypsin reacts with dialkyl phosphoryl compounds of structure:

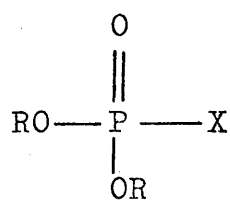


Figure 2.5

where X is a good leaving group such as phenol or a halogen atom. The reaction is accompanied by loss of activity (83,84) and the formation of one mole of phosphorylated enzyme (85-87). Hydrolysis of the phosphoryl-enzyme complex with hydrochloric acid yields O-phosphoryl serine while use of a proteolytic

enzyme to achieve hydrolysis produces O-phosphoryl serine peptides. Partial and total hydrolysis of the inactive complex formed between the enzyme and di-isopropylphosphofluoridate (DFP), in which the phosphorus atoms are radioactive, has established that it is the hydroxymethyl group of Ser-195 in the sequence:

Gly - Asp - Ser - Gly - Gly - Pro - Leu.
193 195 199

which is phosphorylated.

The phosphorylation resembles the acylation step in the reaction of chymotrypsin with an ester such as p-nitrophenyl acetate and the high stability of the phosphorylated enzyme is similar to that of trimethylacetyl chymotrypsin. Oosterbahn (88) has degraded (^{14}C)acetyl chymotrypsin with pancreatin and shown the acetyl group to be located on the hydroxyl group of the serine residue shown in the sequence above. This, together with the observation that the reaction with DFP also occurs at this same location, would appear to indicate that this single and unusually reactive serine residue is intimately involved in the enzymic function of chymotrypsin.

Other evidence to suggest that it is a hydroxyl group of a serine residue which is acylated includes the observation that the spectrum of trans cinnamoyl α chymotrypsin resembles that of an ester of serine more closely than any other possible species (89). Despite the fact that the spectrum of the cinnamoyl group of the native acyl enzyme is displaced from that of the model compound, trans cinnamoyl N-acetyl serinamide, the spectrum of the cinnamoyl group of the denatured acyl enzyme is identical with that of the model. The hydrolysis of O-cinnamoyl N-acetyl serinamide has an alkaline rate constant, in 8M urea, similar to that of the acyl enzyme, the cinnamoyl enzyme thus showing a reactivity drastically different from that of an acyl imidazole, but quite similar to that for either a seryl or tyrosyl ester.

An inactive monotosylated chymotrypsin has been converted,

via an elimination reaction, to an inactive 'anhydro' enzyme, indicating that the hydroxyl group of this serine residue plays a positive role of acyl acceptor, and not a negative one of steric hindrance (90,91).

2.2.b. Histidine 57.

The imidazole ring of a histidine residue has been considered to be a probable participant in the hydrolyses performed by chymotrypsin, initially because it is known to undergo ionisation in the general pH region where catalytic activity is also known to vary. However, although the rates of various reactions catalysed by chymotrypsin show a dependence on a group of pK_{app} about 6.5, which is close to that of free imidazole in solution, the pK_{app} of this group when bound in an enviroment such as is created within the molecule of the enzyme, might differ significantly from this value.

More direct evidence is afforded by the studies of Weil (92) who showed that photo-oxidative destruction of one of the histidine residues of chymotrypsin resulted in loss of enzymic activity and of its reactivity towards DFP.

Various investigations have been made to examine the influence on activity of modification of an imidazole group in the molecule by reaction with 2,4-dinitrofluorobenzene. Whitaker and Jandorf (93) have shown that, at pH 10.7, reaction of one of the two histidine residues with this reagent leads to inactivation of the enzyme. This result is, however, inconsistent with the findings of Massey (94) who found the reagent to have no effect on the rate of hydrolysis of N-acetyl tyrosine ethyl ester by chymotrypsin but simply increases the K_m value for the reaction. This latter study was, however, conducted at a pH lower than 10.7 and it is conceivable that the imidazole groups being acylated are not identical under differing conditions of pH.

Convincing evidence for the implication of a histidine residue at the active site of chymotrypsin comes from the work of Schoellmann and Shaw (95,96). They have demonstrated

that one mole of a bromo ketone analogue of a specific chymotrypsin substrate, α-N-p-toluenesulphonyl β-phenylalanyl bromomethane, reacts with a histidine residue to yield an inactive enzyme. Subsequently they reported the amino acid composition of a peptide from chymotrypsin containing this reactive histidine (97,98), which is consistent with the sequence of residues 55-64 of Hartley.

Incubation of chymotrypsin with methyl ^{14}C p-nitrobenzene sulphonate results in almost stoichiometric incorporation of the radioactivity in the modified enzyme, principally in the B chain (99). Diagonal paper electrophoresis indicates His-57 to be the residue which is modified and the amino acid analyses reveal that 1-amino 2-(1-methyl 4-imidazolyl) propionic acid (3 methyl histidine) is formed with loss of a corresponding amount of histidine from the enzyme derivative. Inhibition of enzymic activity by the methyl ester is prevented by addition of the competitive inhibitor β phenylpropionic acid. The interesting feature of methyl p-nitrobenzene sulphonate is that unlike phenoxymethyl chloromethyl ketone (100) and 2-phenyl 1,4-dibromoacetone (101), which have also been reported to be modifying reagents for this histidine residue, its structure is analogous to non-specific substrates for this enzyme and hence is a member of a new class of active site specific reagents.

Finally, it has been shown that the imidazole ring itself is capable of catalysing the hydrolysis of p-nitrophenylacetate, probably by way of a labile N-acetyl imidazole intermediate, in an overall reaction superficially reminiscent of the enzyme catalysed process. However, this is an inherent property of many proteins and the efficiency of catalysis is far below that of the enzymic reaction, demonstrating that the imidazole nucleus alone cannot serve as the active centre locus.

2.2.c. Methionine 192.

It has been shown (102) that the inactivation of chymotrypsin by photo-oxidation involves the destruction of a methionine, as well as a histidine residue, and the rate of inactivation is consistent with the participation of these residues in the mechanism of catalysis. Schacter and Dixon (103,a,b) demonstrated that the methionine residue which is destroyed is the one located only three residues distant from the highly reactive Ser-195. Lawson and Schramm (104) found a stoichiometric reaction to occur between chymotrypsin and p-nitrophenyl bromoacetyl α amino isobutyrate in a manner similar to that with p-nitrophenyl acetate. The derivative formed exhibited 20% of the activity of the native enzyme and would not liberate the acylating group from the protein at high pH. Analysis of the acid hydrolysates of the derivative showed that, following normal acylation of the active centre, the presence of bromine in the reagent leads to alkylation of Met-192. Use has been made of this alkylated enzyme, in conjunction with a variety of specific and non-specific substrates, to investigate the relationship between specificity changes and changes in kinetic parameters produced by alkylation of this particular methionine residue (105).

2.2.d. Isoleucine 16.

When chymotrypsinogen is acetylated it is found that all of the amino groups and the hydroxyl group of two of the four tyrosine residues in the molecule have reacted. Activation of this acetylated zymogen by trypsin produces an acetylated chymotrypsin possessing a free α amino group on the N-terminal isoleucine residue. This form of the enzyme, after reaction with di-isopropylphosphofluoridate, still retains 85% of its activity, as measured by its ability to hydrolyse N-acetyl L tyrosine ethyl ester. However, acetylation of the N-terminal isoleucine residue of this acetylated chymotrypsin is accompanied by complete loss of enzymic activity, thus establishing the necessity of this primary amino group for catalysis (106-108).

The importance of the protonated α amino group of Ile-16 is discussed in the section devoted the pH dependence of chymotrypsin catalysed reactions at alkaline pH.

2.2.e. Tryptophan residues.

As well as the loss of one histidine, approximately three tryptophan residues are destroyed during the photo-oxidation of chymotrypsin in the presence of methylene blue, with concomitant loss of esterase activity (92,109,110). There is a 50% reduction in catalytic activity and loss of one tryptophan residue when oxidation of the enzyme is conducted in the presence of partially purified horseradish peroxidase (111); a similar result is achieved using N-bromosuccinimide as the oxidising agent (112).

The indolyl group of tryptophan is chemically rather inert. The fact that destruction of one or more of these residues simply lowers the catalytic rate constant suggests that they are not directly involved in the bond breaking processes catalysed by the active site. Changes occuring in the spectral nature of the indolyl fuctions of the enzyme when the latter is acetylated are thought to arise from conformational changes in the molecule which probably result in a modification of the chemical enviroment of the tryptophan residues (113-115).

Differential UV spectra of the native enzyme and its DFP derivative appear to show that tryptophan is involved in the binding of this inhibitor. This was confirmed by isolating and analysing a tetrapeptide, Try-Leu-Val-Ser, obtained by peptic degradation, in which the inhibitor was found to be bound to the tryptophan residue (116).

2.3. HOMOLOGIES IN THE AMINO ACID SEQUENCES OF THE SERINE PROTEINASES.

Four inactive precursors (zymogens) of endopeptidases are to be found in bovine pancreatic juice. They are chymotrypsinogens A and B, trypsinogen and a component of procarboxypeptidase which resembles a chymotrypsinogen. In addition to these, porcine pancreas secretes elastase.

Comprehensive accounts of the homologies which exist in the amino acid sequences of these enzymes and the other serine proteinases have been published (117-119). An attempt will be made here to cover only the more salient features of this topic.

The serine proteinases are so called because each of these enzymes possesses a unique serine residue which will react with organophosphorus compounds with complete inhibition of enzymic activity. Despite the widely differing substrate specificities of these enzymes, it has long been suspected that they have a common catalytic mechanism. Support for this hypothesis has been provided by the recognition of homologies in their amino acid sequencies, particularly at those locations at which occur amino acid residues known to be important for catalytic activity.

Chymotrypsins A and B, which show almost identical substrate specificity, are closely related isozymes and show a very high degree (78%) of homology. Comparison of these two enzymes with trypsin and elastase shows there to be between 35% and 43% identity in their sequences.

Side chain homologies are not uniformly distributed throughout the polypeptide chain but tend to be concentrated around those residues that are important for catalysis, as well as those parts of the molecule which are internal-as judged by the apparent inaccessibilty of the amino acid side chains to water molecules in molecular models of the enzymes.

In the four enzymes, 71% of the internal position, but only 19% of those that are external, are found to be

homologous, while only 4% of the internal positions show no homology between any two enzymes of different specificities.

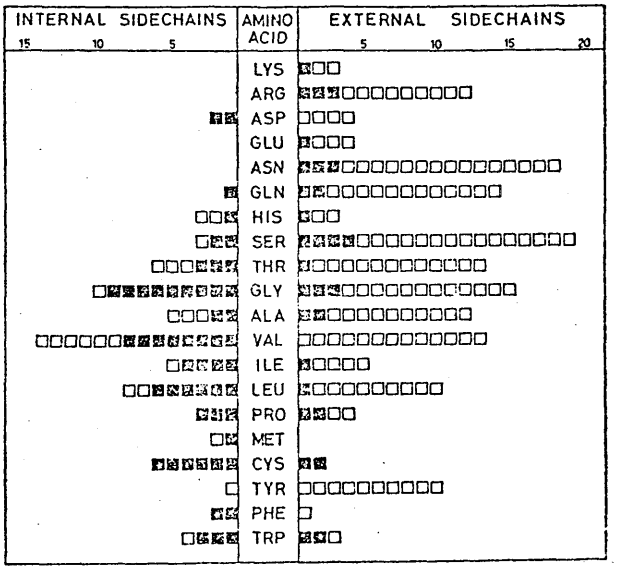


Figure 2.6. The distribution between internal and external positions of the amino acid residues in the molecule of Porcine Elastase.

Black squares: residues homologous with (identical or chemically similar to) corresponding residues in Bovine Trypsin, Chymotrypsinogen A and Chymotrypsinogen B.
White squares: residues not homologous with all three other enzymes.
'Chemically similar' residues are defined as in Figure 2.8.a.

In the amino acid sequence of each of the four enzymes, chymotrypsins A and B, trypsin and elastase, there appears a common region involving two histidine residues which are brought close together in the secondary structure by a cystine bridge. Of a sequence of 19 amino acids, including His-57, of α chymotrypsin, 16 correspond to this same sequence in trypsin, two of the three differences are merely a difference in a methylene group.

The aromatic residues at positions 39 and 41 in trypsin and chymotrypsin are replaced in elastase by alanine and threonine. It is probably significant, however, that the

positions of the two histidine residues relative to the disulphide bond are a constant feature, together with the fact that the sequence following Cys-42 and preceding Cys-58 is remarkably similar for all four enzymes (Fig.2.8).
In trypsin and α chymotrypsin this near identity extends to the complete sequence between these two half cystine residues in what is usually referred to as the 'histidine loop'. In addition, the sequence of the residues 191-201 in α chymotrypsin closely parallels that of residues 179-189 in trypsin (Fig.2.7).

Chymotrypsin:	Cys-Met-Gly-Asp-Ser-Gly-Gly-Pro-Leu-Val-Cys	
	191	201
Trypsin:	Cys-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Val-Val-Cys	
	179	189

Figure 2.7.

Striking similarity in the location of the disulphide bridges in the pancreatic proteinases has been observed (120). Four bridges common to chymotrypsinogen and trypsinogen are the same four found in porcine elastase and 52% of the residues around these bridges are identical. This leads one to expect considerable similarity in their tertiary structures and three dimensional models of trypsin and elastase have been constructed to fit the electron density map of α chymotrypsin (117). The models show that it is possible to accomodate the differences in the amino acid sequences without any serious steric problems arising. Where local differences in chain length do occur, such 'insertions' or 'deletions' are generally found at the ends of external loops which can expand or contract to accomodate such changes without disturbing the overall conformation. The models also suggest that the differences in specificity between these enzymes can be explained by relatively minor side chain substitution on the enzyme surface.

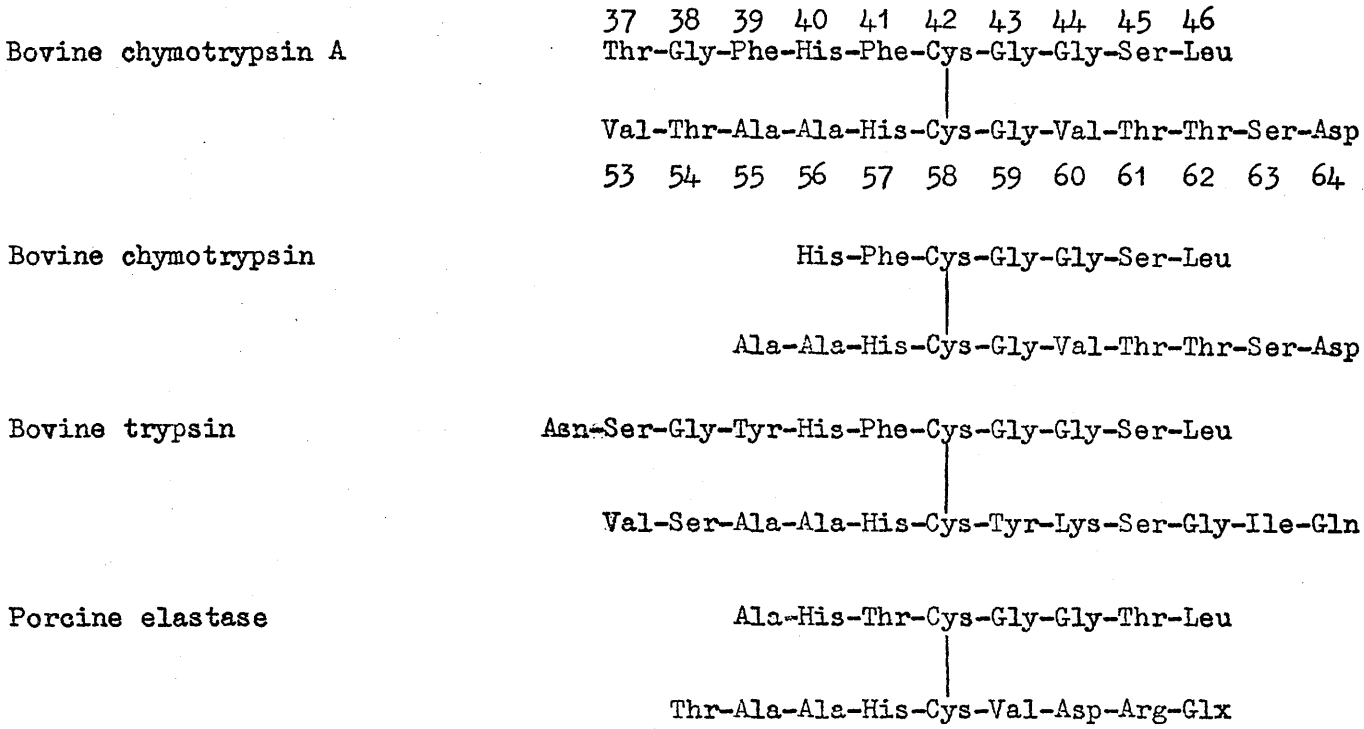


Figure 2.8. Sequences of di-histidine cystine peptides of bovine chymotrypsin A and B bovine trypsin and porcine elastase.

Figure 2.8.a. (facing page).

A comparison of the amino acid sequences of Porcine Elastase (PE),Bovine Trypsin (BT),Chymotrypsin A (BCA) and Chymotrypsin B (BCB).

The numbering sequence is that of Chymotrypsinogen A.

At each sequence position residues which are chemically similar,defined as;

Arg : Lys	Ser : Thr
Asp : Glu	Val : Ile
Asn : Gln	Ile : Leu
Asp : Asn	Tyr : Phe : Trp
Glu : Gln	

between any pair,are shown in capitals. Residues which are identical are underlined. Residues whose side chains are internal, i.e. judged to be inaccessible to water molecules in the molecular models of both Elastase and α Chymotrypsin, are indicated by circles round the residue numbers.

```
     (16)(17) 18  19  20  21 (22) 23  24  25  26 (27)(28)(29)(30)(31)(32)(33) 34  35  36 36A 36B 36C  37  28  39
PE : VAL-VAL-GLY-GLY-Thr-GLU-ALA-Gln-Arg-ASN-SER-TRP-PRO-Ser-GLN-ILE-SER-LEU-GLN-Tyr-ARG-Ser-Gly-Ser-SER-Trp-Ala-
BT : ILE-VAL-GLY-GLY-Tyr-Thr-Cys-Gly-Ala-ASN-THR-Val-PRO-TYR-GLN-VAL-SER-LEU-ASN ---- ---- ---- ---- ---- SER-GLY-TYR-
BCA: ILE-VAL-Asn-GLY-Glu-GLU-ALA-Val-Pro-Gly-SER-TRP-PRO-TRP-GLN-VAL-SER-LEU-GLN-Asp LYS ---- ---- ---- THR-GLY-PHE-
BCB: "  "  "  "  "  ASP "  "  "  "  "  "  "  "  "  "  "  "  "  "  "  "  Ser ---- ---- ---- "  "  "

     (40) 41 (42)(43)(44)(45)(46)(47) 48  49  50 (51)(52)(53)(54)(55) 56  57 (58) 59  60  61  62  63  64  65 65A
PE : HIS-Thr-CYS-GLY-GLY-THR-LEU-ILE-Arg-GLN-ASN-TRP-VAL-Met-THR-ALA-ALA-HIS-CYS-VAL-ASP-Arg-Glu-LEU-THR-Phe-ARG-
BT : HIS-PHE-CYS-GLY-GLY-SER-LEU-ILE-ASN-Ser-GLN-TRP-VAL-VAL-SER-ALA-ALA-HIS-CYS-Tyr-Lys-SER-GLY-ILE-Gln-VAL-ARG-
BCA: HIS-PHE-CYS-GLY-GLY-SER-LEU-ILE-ASN-GLU-ASN-TRP-VAL-VAL-THR-ALA-ALA-HIS-CYS-Gly-Val-THR-Thr-Ser-Asp-VAL ----
BCB: "  "  "  "  "  "  "  "  Ser "  ASP "  "  "  "  "  "  "  "  "  "  "  "  "  "  "  "  "  "  "  "  "  ----
PCC:                                                                                 ALA-HIS-CYS-ILE-ASN-SER-GLY-THR-SER-Arg-Thr-
     (66)(67)(68) 69  70  71  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91  92
PE : VAL-VAL-Val-GLY-GLU-His-ASN-LEU-ASN-Gln-Asn-Asn-GLY-Thr-GLU-GLN-TYR-VAL-Gly-VAL-Gln-LYS-ILE-VAL-VAL-HIS-PRO-
BT : Leu ---- ---- GLY-GLN-Asp-ASN-ILE-ASN-Val-Val-Glu-GLY-ASN-GLN-GLN-PHE-ILE-Ser-Ala-Ser-LYS-Ser-ILE-VAL-HIS-PRO-
BCA: VAL-VAL-Ala-GLY-GLU-Phe-ASP-Gln-Gly-Ser-Ser-SER-Glu-Lys-Ile-GLN-Lys-LEU-Lys-ILE-Ala-LYS-VAL-Phe-Lys-Asn-Ser-
BCB: "  "  "  "  "  "  "  "  "  Leu-Glu-THR "  ASP-Thr "  Val "  "  "  Gly "  "  "  "  "  "  PRO

     93  94  95  96  97  98  99 99A 99B 100 101 (102)(103)(104)(105)(106) 107 (108) 109 110 111 (112) 113 114 115 116 117
PE : Tyr-TRP-ASN-THR-ASP-Asp-VAL-Ala-Ala-Gly-Tyr-ASP-ILE-Ala-LEU-LEU-ARG-LEU-ALA-Gln-Ser-Val-THR-LEU-ASN-SER-Tyr-
BT : Ser-TYR-ASN-SER-ASN-THR-LEU ---- ---- ASN-ASN-ASP-ILE-Met-LEU-ILE-LYS-LEU-Lys-SER-ALA-ALA-SER-LEU-ASN-SER-Arg-
BCA: Lys-TYR-ASN-SER-Leu-THR-ILE ---- ---- ASN-ASN-ASP-ILE-Thr-LEU-LEU-LYS-LEU-Ser-THR-ALA-ALA-SER-Phe-Ser-GLN-Thr-
BCB: "  PHE-Ser-Ile "  "  VAL ---- ---- Arg "  "  "  "  "  "  "  "  ALA "  Pro "  Gln "  "  GLU "

     (118) 119 120 (121) 122 123 (124) 125 126 127 128 129 130 131 132 133 134 135 (136) 137 (138)(139)(140)(141)(142) 143 144
PE : VAL-Gln-Leu-Gly-Val-LEU-PRO-Arg-ALA-Gly-Thr-Ile-Leu-ALA-Asn-Asn-SER-Pro-CYS-Tyr-ILE-THR-GLY-TRP-GLY-LEU-THR-
BT : VAL-Ala-Ser-ILE-Ser-LEU-PRO-THR ---- SER-Cys-Ala-Ser ---- ALA-GLY-THR-Gln-CYS-Leu-ILE-SER-GLY-TRP-GLY-Asn-THR-
BCA: VAL-Ser-Ala-VAL-Cys-LEU-PRO-SER-ALA-SER-ASP-Asp-Phe-ALA-ALA-GLY-THR-Thr-CYS-Val-Thr-THR-GLY-TRP-GLY-LEU-THR-
BCB: "  "  "  "  "  "  "  "  "  Asp-GLU "  "  Pro. "  "  Met-Leu "  Ala "  "  "  "  "  Lys "

     145 146 147 148 149 150 151 152 153 154 (155) 156 157 (158) 159 (160) 161 (162) 163 164 165 166 167 (168) 169 170 170A
PE : ARG ---- THR-ASN-Gly-GLN-Leu-Ala-Gln-Thr-LEU-GLN-GLN-ALA-Tyr-LEU-PRO-Thr-VAL-Asp-Tyr-Ala-Ile-CYS-Ser-SER-Ser
BT : LYS-Ser-SER-Gly-Thr-Ser-Tyr-PRO-ASP-Val-LEU-Lys-Cys-Leu-Lys-Ala-PRO-ILE-LEU-SER-ASN-SER-Ser-CYS-LYS-SER ----
BCA: ARG-Tyr-THR-ASN-Ala-ASN-Thr-PRO-ASP-ARG-LEU-GLN-GLN-ALA-SER-LEU-PRO-LEU-LEU-SER-ASN-THR-ASN-CYS-LYS-Lys ----
BCB: LYS "  Asn-Ala-Leu-Lys "  "  "  LYS "  "  "  "  THR "  "  ILE-VAL "  "  "  ASP "  ARG "  ----

     170B 171 172 173 174 175 (176) 177 178 179 (180)(181)(182)(183)(184) 184A 185 186 187 188 188A 189 190 191 192 (193)(194)
PE : Ser-TYR-TRP-GLY-SER-Thr-VAL-LYS-ASN-Ser-MET-VAL-CYS-ALA-GLY ---- Gly-Asn-GLY-VAL-ARG-SER-Gly-CYS-GLN-GLY-ASP-
BT : ---- Ala-TYR-Pro-Gly-Gln-ILE-THR-Ser-Asn-MET-Phe-CYS-ALA-GLY-Tyr-Leu-Glu-GLY-Gly-LYS-ASP-SER-CYS-GLN-GLY-ASP-
BCA: ---- TYR-TRP-GLY-THR-LYS-ILE-LYS-ASP-Ala-MET-ILE-CYS-ALA-GLY ---- Ala-Ser-GLY-VAL ---- SER-SER-CYS-Met-GLY-ASP-
BCB: ---- "  "  "  SER-ARG-VAL-THR "  Val "  "  "  "  "  ----  "  "  "  "  ----  "  "  "  "  "  "
PCC:                                                                                                  GLU-THR-CYS-GLU-GLY-ASP-

     195 (196)(197)(198)(199)(200)(201) 202 203 204 205 206 207 208 (209)(210)(211)(212)(213)(214)(215) 216 217 217A 218 219 220
PE : SER-GLY-GLY-PRO-LEU-His-CYS-Leu-Val-ASN-GLY-Gln-TYR-Ala-Val-His-GLY-VAL-Thr-SER-PHE-Val-SER-Arg-Leu-GLY-CYS-
BT : SER-GLY-GLY-PRO-Val-VAL-CYS-Ser-Gly-Lys ---- ---- ---- ---- LEU-Gln-GLY-ILE-VAL-SER-TRP-GLY-SER ---- ---- GLY-CYS-
BCA: SER-GLY-GLY-PRO-LEU-VAL-CYS-Lys-Lys-ASN-GLY-Ala-TRP-Thr-LEU-Val-GLY-ILE-VAL-SER-TRP-GLY-SER ---- Ser-Thr-CYS-
BCB: "  "  "  "  "  "  "  Gln "  "  "  "  "  "  "  "  Ala "  "  "  "  "  "  "  "  ----  "  "  "
PCC: SER-GLY(GLY,PRO)LEU-VAL

     221 221A 222 223 224 225 226 (227)(228)(229) 230 (231) 232 233 (234) 235 236 (237)(238) 239 240 241 242 243 244 245
PE : Asn-Val-THR-Arg-LYS-PRO-Thr-VAL-PHE-THR-ARG-VAL-SER-ALA-TYR-ILE-SER-TRP-ILE-ASN-ASN-Val-ILE-ALA-SER-ASN
BT : Ala-Gln-Lys-Asn-LYS-PRO-GLY-VAL-TYR-THR-LYS-VAL-Cys-Asn-TYR-VAL-SER-TRP-ILE-Lys-GLN-THR-ILE-ALA-SER-ASN
BCA: Ser ---- THR-Ser-Thr-PRO-GLY-VAL-TYR-Ala-ARG-VAL-THR-ALA-Leu-VAL-Asn-TRP-VAL-GLN-GLN-THR-LEU-ALA-Ala-ASN
BCB: "  ----  "  "  "  "  Ala "  "  "  "  "  "  "  "  Met-Pro "  "  "  GLU "  "  "  "  "
```

2.4. CONFORMATIONAL CHANGES IN THE α CHYMOTRYPSIN MOLECULE.

Reactions catalysed by α chymotrypsin have bell-shaped pH-rate profiles (121) which are usually regarded as evidence for two ionising groups on the enzyme being important for catalysis. The mid-points of the two limbs of such a profile indicate that the reaction rate increases with the deprotonation of a group of pK_{app} about 7.0 and decreases with the deprotonation of a group on the enzyme of pK_{app} about 8.5. The former group has been identified as the imidazole nitrogen of His-57 (122-124) and the latter as the α amino group of Ile-16 (108,125).

2.4.a. The pH dependence of the catalytic reaction at alkaline pH.

The critical event in the activation of chymotrypsinogen by trypsin is the hydrolysis of the single peptide bond to produce a new N-terminal residue, that of Ile-16; the pH dependence of chymotrypsin catalysed reactions at alkaline pH is related to the events accompanying this activation process. It has also been suggested that activation is accompanied by conformational changes which establish the active site of the enzyme (126).

That there is a relationship between different conformational forms of chymotrypsin and the state of ionisation of the α amino group of Ile-16 has been demonstrated by titration experiments and circular dichroism measurements. Titration difference curves between δ chymotrypsin and chymotrypsinogen, as well as between δ chymotrypsin and this same enzyme in which the α amino group has been specifically acetylated, have served to identify this group and to demonstrate its ionisation characteristics in chymotrypsin (125,127,128).

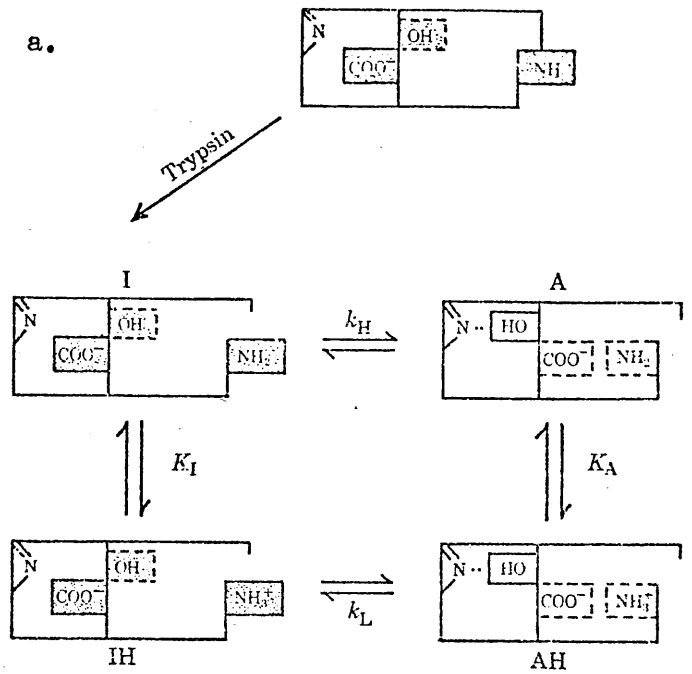
The circular dichroism characteristics of acetylated chymotrypsinogen, DFP δ chymotrypsin and of chymotrypsin in

which the α amino group of Ile-16 has been specifically acetylated show no dependence on pH in the region 7.0-10.0. However, a dependence is shown by δ chymotrypsin which indicates that a single group of pK_{app} about 8.5 controls a change in ellipticity of the molecule. It appears from these results that molecules having a positive charge on Ile-16 exist in one conformation while molecules in which this group does not, or cannot, ionise adopt a different conformation.

There is, in the alkaline pH region, a correspondence between circular dichroism of chymotrypsin and the value of K_s, the enzyme-neutral substrate dissociation constant, demonstrating that, like circular dichroism, Ks depends on the ionisation of a group of pK_{app} about 8.5. Kinetic experiments with a series of substrates have indicated that the pH dependence of α chymotrypsin catalysed reactions in the alkaline region is due to the pH dependence of K_s (129,130). All of these observations suggest that the ionisation of the α amino group of Ile-16 controls the equilibrium between two major conformations of the enzyme and it is that conformation having the α amino group of Ile-16 in a protonated state which binds substrate molecules most efficiently.

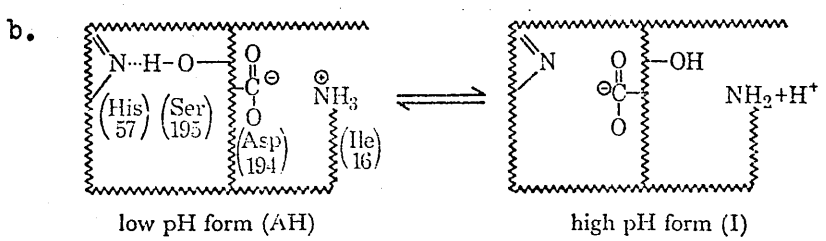
If binding of substrate molecules to the neutral conformation of the enzyme occurs more readily than to the conformation which predominates at high pH then the process of substrate binding must shift the equilibrium between the two conformations in the direction of the AH form (fig.2.9). As a result there is a conversion from a situation in which the α amino group is not protonated to one in which it is, a process that ought to be characterised by an uptake of protons from solution. For some time rapid initial proton absorption has been reported to accompany acylation reactions of α chymotrypsin (131-134) but no attempt was made to separate the effects due to binding from those due to acylation. Recent studies indicate that binding alone is sufficient to cause proton absorption by

α chymotrypsin at high pH (125,135) and that no further changes in the state of protonation of the enzyme should result from the acylation step.



Figures 2.9.a and b.

The basic pH dependent equilibria controlling enzyme conformation and the amino acid residues involved.



Proton uptake has been observed as a result of the binding of indole, benzyl alcohol, acetonitrile and N-acetyl D tryptophanamide by α chymotrypsin and of N-acetyl L tryptophanamide by δ chymotrypsin (136-138). Each of the substrates studied produced a different maximum proton uptake, indicating that the equilibrium between enzyme conformations is affected differently by different substrates. Further, the data for the dependence of hydrogen ion uptake on pH for each substrate yields two pK_{app} values, one for the free enzyme and a higher one for the enzyme-substrate complex. In order to explain the two pK_{app} values a relatively complex set of equilibria

may be formulated which include two ionisation constants for the α amino group of Ile-16 as it exists in the two conformations.

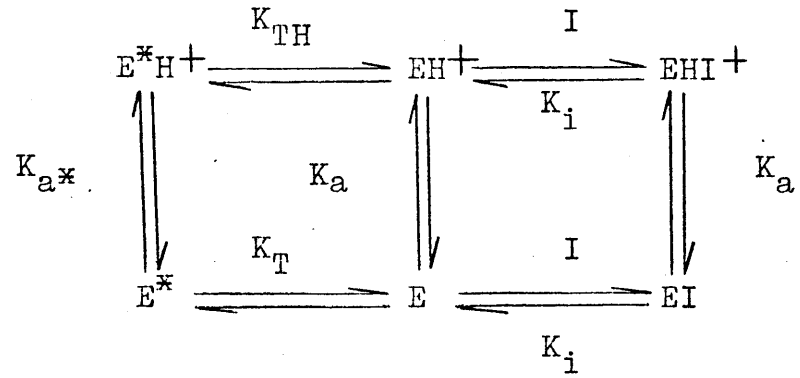


Figure 2.10.

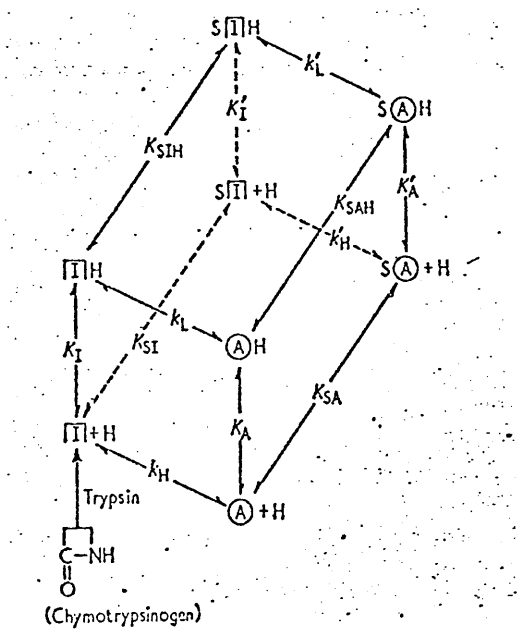


Figure 2.10.a. A proposed scheme for the formation of chymotrypsin - substrate complexes.

A and I represent active and inactive forms of the enzyme, respectively. The equilibrium constants k_L and k_L' apply to the pH 7 form of the enzyme (A) and k_H, k_H' to the pH 10.5 form (I).
The acid dissociation constants K_A and K_I refer to the same ionising group of the enzyme (that of Ile-16) in the A and I forms, respectively.

Figure 2.10 postulates two states for the enzyme. Although binding and proton absorption require that E be the major form when protonated and E^* be the major form when not protonated, the introduction of a substrate (or inhibitor) into the system transforms all of the enzyme to EHS (or EHI), explaining the pH independent catalytic steps k_2 and k_3 in the high pH region. The conformation changes are thus related to the binding of small molecules and not to the catalytic process per se.

When the conclusions about the relationship between enzyme conformation and the state of ionisation of the α amino group of Ile-16 are considered in terms of the conformation of the active site of crystalline α chymotrypsin, it becomes possible to make some important statements about the nature of the active site of chymotrypsinogen and the high pH form of α chymotrypsin.

X-ray data indicates that in α chymotrypsin at neutral pH the α amino group of Ile-16 and the carboxyl group of Asp-194 are directed away from the solvent and form an internal ion pair. In the absence of a positive charge on Ile-16 the carboxyl group will almost certainly protrude into the solvent and in doing so will interfere with substrate binding. This view is supported by the observation that the zymogen binds substrates and inhibitors much less effectively than does α chymotrypsin (139,140) and that, in its high pH form, the binding ability of the latter is much reduced compared with the neutral pH form of the enzyme (129,130,141).

$K_{m(app)}$ and k_{cat} values for the δ chymotrypsin catalysed hydrolysis of N-acetyl L tryptophan methyl ester and N-furylacroyl L tryptophanamide have been measured as a function of pH and ionic strength (142). In contrast with the α form of the enzyme, $K_{m(app)}$ showed no great increase above pH 9.0. Although the α amino group of Ile-16 is involved in the K_m - pH transitions observed with the δ form, the results indicate that its state of ionisation is not a crucial factor in determining the binding ability of this enzyme. This view

is, however, in conflict with that of Sigler, outlined in the preceding paragraph, that the active conformation of the enzyme in solution is stabilised by a salt bridge between residues 16 and 194. The observed kinetic differences between the α and δ forms of chymotrypsin at high pH suggest that the reversible inactivation of α chymotrypsin at alkaline pH may involve the participation of Tyr-146 and Ala-149. Both of these residues are present as chain termini in α chymotrypsin, but not in the δ form, as a result of the activation process.

2.4.b. Conformational changes accompanying the activation process.

Data obtained from X-ray diffraction studies make it possible to draw some conclusions about the stereochemistry of the activation process.

i. the overall tertiary structure of the zymogen is similar to that of the active enzyme, there being no gross reorganisation in the folding of the main chain, nor a significant helix-coil transition. Kraut (143) compared the low resolution electron density maps of chymotrypsinogen and δ chymotrypsin and found close correspondence between almost all of the prominent features. Correlation of these maps with that of α chymotrypsin confirmed the two locations where changes were observed as being the point of tryptic hydrolysis of peptide bond 15-16 in the zymogen and the region around Thr-147, Asn-148.

ii. the stereochemistry of the activation of the zymogen is paralleled in the active enzyme by a pH dependent structural transition. In both transitions activity depends upon the integrity of the ion pair between Ile-16 and Asp-194. The existence of this ion pair accounts for the anomalously high pK_a of the α amino group of Ile-16, compared with a value of less than

that of 8.0,usual for peptide amino groups. If the positive charge is removed from the amino group the high potential of the isolated negative charge in a medium of low dielectric constant within the molecule would require the carboxylate group of Asp-194 to seek an orientation in which it could point into the solvent. This could be accomplished by the movement of this group into the vicinity of the side chains of Ser-195 and His-57. In other words,at high pH the conformation of the active centre is similar to that in the zymogen,from which it follows that the zymogen and the high pH form of the enzyme lack activity for the same stereochemical reason.

The observed sequestering of the titratable proton of α chymotrypsin upon irreversible phosphorylation of Ser-195 probably results because the fixing of a bulky group to this residue sterically blocks the rotation that is required to bring the isolated carboxylate group to the surface upon deprotonation of the α amino group of the ion pair.

The other part of the structural transition occuring on activation is the entry of the N-terminal isoleucine residue from a surface position to the environment where the ion pair is formed. In α chymotrypsin,Ile-16 and Val-17 make several hydrophobic interactions with Thr-138,Ala-158, Leu-160 and Val-188 and presumably this energy of interaction provides the driving force for the activation transition.

2.5. X-RAY CRYSTALLOGRAPHIC STUDIES OF THE α CHYMOTRYPSIN MOLECULE.

2.5.a. The Stereochemistry of the Active Site.

A three dimensional map, at 2Å resolution, of the differences in electron density between tosyl α chymotrypsin, in which a toluene sulphonyl group is acylating the active site, and the native enzyme shows significant differences only in the vicinity of the active centre. From the data it is possible to present a picture of the juxtaposition of the important amino acids in this region of the molecule.

Sulphonation causes a movement of the γ oxygen of Ser-195, together with a displacement of Met-192 and of the side chain of His-57; the movement of the serine oxygen is consistent with a simple rotation about the α-β bond. In the tosyl derivative, the distance between the centre of the imidazole ring of His-57 and the centre of the sulphonyl group is 4.7Å, but in the native enzyme it is less than 4Å from the oxygen atom of Ser-195, making possible a direct hydrogen bond from the γ oxygen of this serine residue to the ε-nitrogen atom of His-57.

The apparent existence of a hydrogen bond from the δ nitrogen atom of His-57 to residue 102 led to a re-investigation of the nature of the latter. Although it had been identified as asparagine in chymotrypsinogen A, the amino acid sequences of several homologous enzymes, including chymotrypsinogen B, were found to have aspartic acid at this position (144). Re-examination by Hartley (145) found this to be the case with chymotrypsinogen A.

Aspartate 102 is in an internal position, being shielded from solvent molecules by the side chains of Ala-55, Ala-56, His-57, Cys-58, Tyr-94, Ile-99 and Ser-214, in such a manner that access of a hydronium ion to the carboxylate group would require some disturbance of the protein tertiary structure, One side of the imidazolium ring of His-57 is

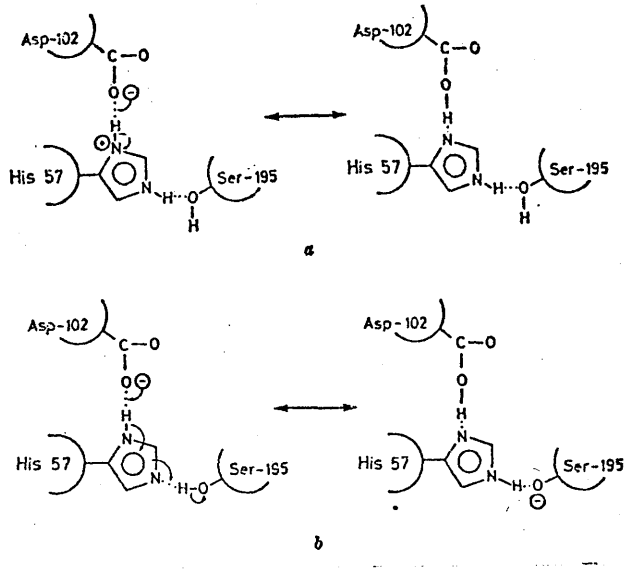


Figure 2.11. Canonical forms of the amino acids at the active centre of Chymotrypsin, representing the electronic state at a. pH 4, b. pH 8.

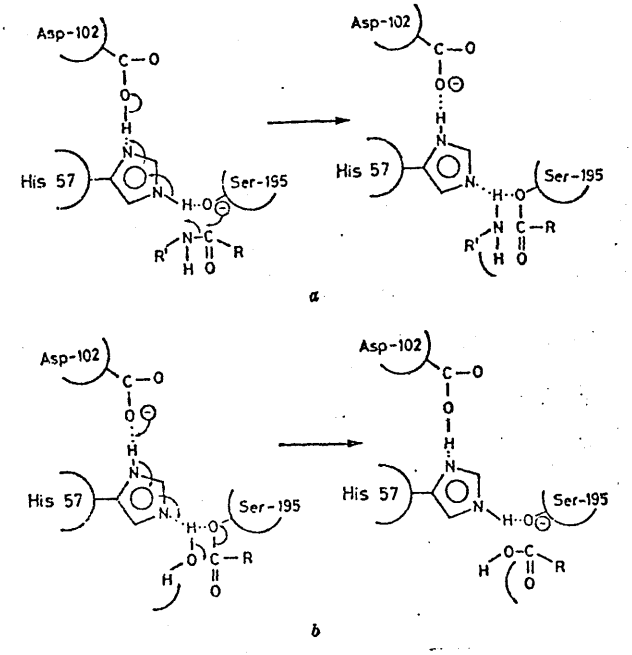


Figure 2.12. The electronic rearrangements describing possible mechanisms for the a. acylation, b. deacylation, of the active centre.

freely accessible by solvent molecules but the electrostatic interaction with Asp-102 tends to localise its excess positive charge on that side of the ring away from the solvent. The proton on the $N^{\delta 1}$ atom of His-57 is in a buried position, where it is stabilised by the hydrogen bond to Asp-102. The orientation of the C - O bond of Ser-195 is such that the serine hydroxyl group can be rotated into a position favourable for hydrogen bonding to the $N^{\epsilon 2}$ atom of His-57. The possibility of electron transfer by rearrangement of the hydrogen bonds at the active site gives rise to a 'charge-relay' system in which the negative charge originating from the buried aspartate residue may be conveyed, through a conduction channel surrounded by a non-polarisable enviroment, to the surface of the molecule. In this way the oxygen atom of Ser-195 becomes a powerful nucleophile, thus explaining its reactivity towards amides and esters (Fig.2.11 and 2.12).

The proximity of Asp-194 and Ile-16 is another feature of the α chymotrypsin molecule which is revealed by X-ray diffraction studies. The important role played by these two residues in stabilising the enzymically active conformation of the molecule has been discussed in section 2.4.

2.5.b. The Binding Site of α Chymotrypsin - the 'Tosyl Hole'.

Work directed towards determining the location and nature of the site on the α chymotrypsin molecule responsible for binding the aromatic side chain of a typical 'good' substrate for the enzyme has employed X-ray diffraction studies of complexes of inhibitors and 'virtual' substrates (products) with the enzyme at 5Å and 2.5Å resolution (146).

The X-ray structure of tosyl α chymotrypsin had shown the toluene ring of the inhibitor to be contained in a hydrophobic pocket, or slit, which has come to be known as the tosyl hole (147,148). The hole, an irregular, flattened shape with dimensions 10-12Å × 5.5-6.5Å × 3.5-4.0Å, is capable of

accomodating indole and is deep enough for tyrosine. Studies in projection suggest the slit is not deep enough to bind the aromatic groups of formyl L p-iodophenylalanine or β-(p-iodophenyl)propionate, but p-iodophenylacetate binds with the iodine atom lying deep in the pocket. The narrowness of the slit allows only one orientation of the plane of the bound aromatic side chain.

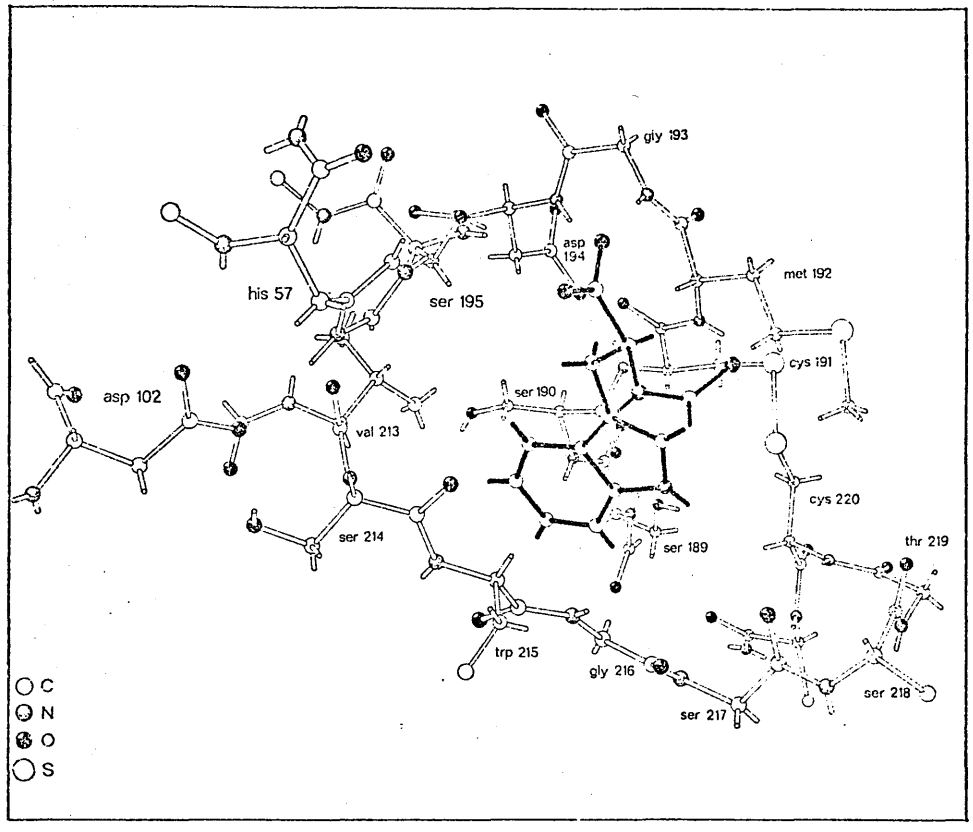


Figure 2.13. A representation of the position occupied by formyl-L-tryptophan within the active site of α chymotrypsin.

Three dimensional difference maps of formyl L tryptophan-α chymotrypsin and the native enzyme show the amido hydrogen of the virtual substrate to be directed towards the carbonyl oxygen of Ser-214 but is too far away to be forming a hydrogen bond, while the carbon atom of the carboxyl group is close to the oxygen of Ser-195. The indolyl group fits neatly into the pocket near this same serine residue. Its most prominent interactions are with peptide bonds whose planes are approximately parallel to the plane of the indolyl ring and about 4Å distant, viz: peptide bonds 190-191, 191-192 above the ring, and 215-216 below it. Formyl L phenylalanine binds in an analogous fashion, the position of the phenyl ring being between that of the five and six membered rings of tryptophan.

The substrate specificity of chymotrypsin (and trypsin) is probably best described in terms of 'lock and key' complementarity between the structure of the enzyme and that of the substrate. The binding of the side chain in the tosyl hole and the acetamido group to a 'backbone' carbonyl orients the susceptible bond in a position suitable for hydrolysis to occur. Flexible substrates, which lack either the side chain or the acylamido group, are free to take up other conformations which are unsuitable for hydrolysis. No alteration in the positions of the atoms of the backbone of the molecule has been noted upon binding of substrates or inhibitors (146). There is, however, a slight repositioning of the sulphur atom of Met-192 upon binding of dioxan, the inhibitor itself occupying a position on the enzyme nearly coincident with that of the toluene ring in tosyl α chymotrypsin. The methionine residue appears to function as a flexible, hydrophobic lid to the pocket and so might aid the binding process, a conclusion supported by the finding that chemical modification of this residue results in a reduction in the affinity of the enzyme for a particular substrate but not in the maximum rate of hydrolysis (149).

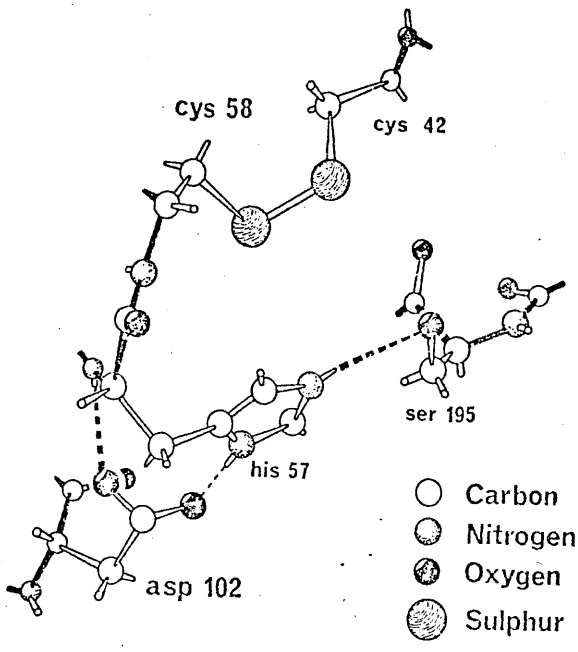


Figure 2.14. The conformation of some of the amino acid residues at the active centre of α chymotrypsin. Hydrogen bonds are indicated by dashed lines.

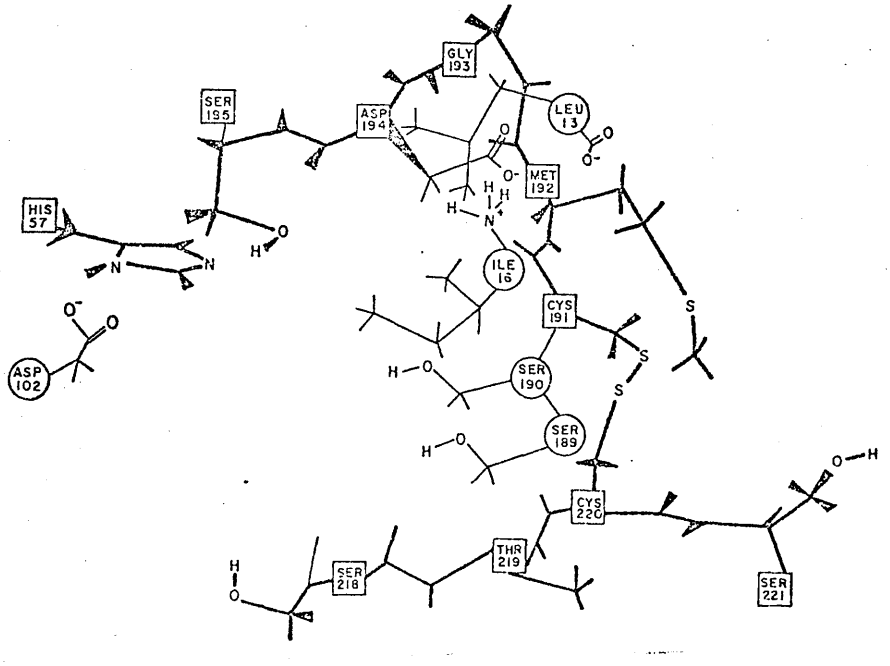


Figure 2.15. A schematic diagram of the important amino acid residues in and around the active centre of α chymotrypsin.

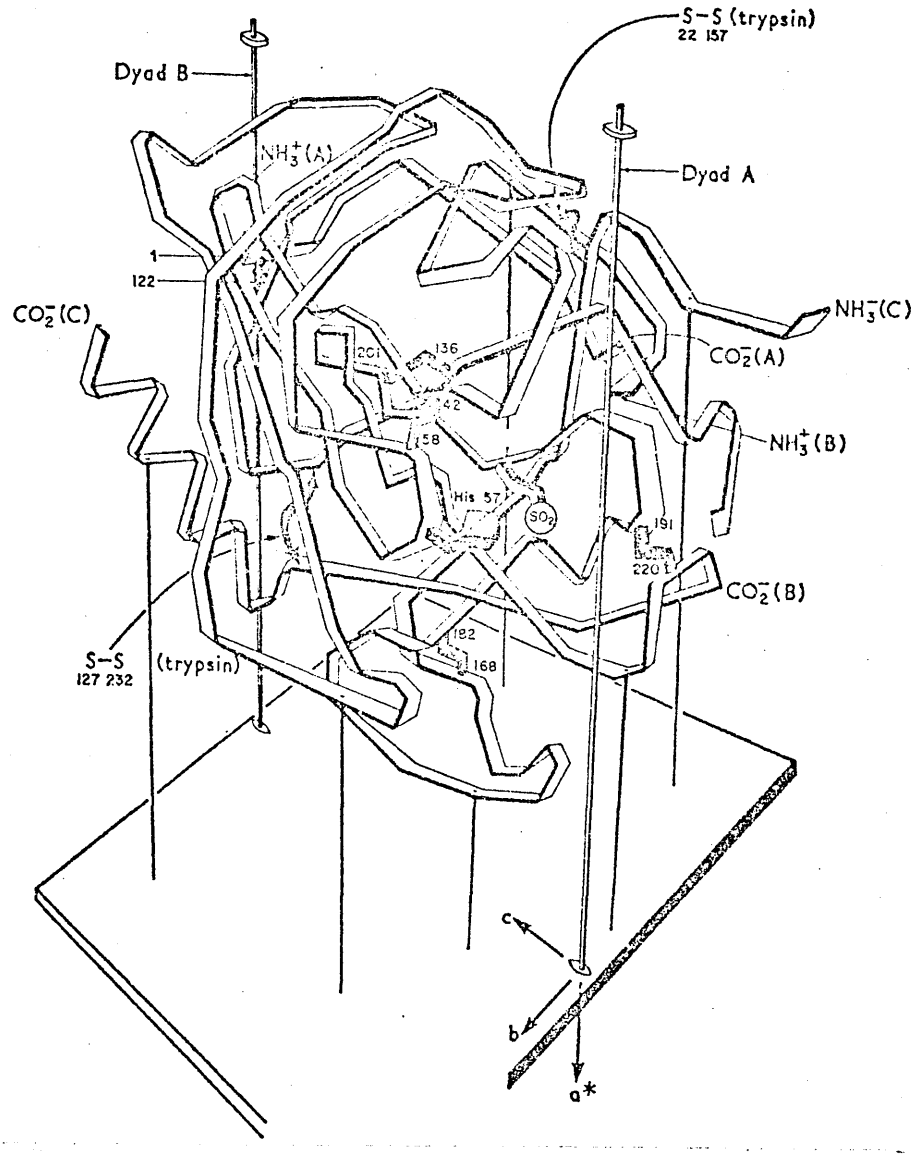


Figure 2.16. A representation of the conformation of the polypeptide chains in α chymotrypsin. The position s homologous with the locations of the two additional disulphide bridges in the trypsin molecule are indicated.

2.6. THE SUBSTRATE SPECIFICITY DIFFERENCES OF α CHYMOTRYPSIN, TRYPSIN AND ELASTASE.

Identification of the substrate specificity site of α chymotrypsin and the fact that significant amino acid homologies between this enzyme and trypsin give these two enzymes three dimensional structures that are closely analogous has made possible an explanation of their differing side chain specificities.

The trypsin molecule contains the same 'charge relay' system (145), involving Asp-102, His-57 and Ser-195, as that found in α chymotrypsin and there is also an analogous pocket near Ser-195. In trypsin this hydrophobic cavity must be designed to bind the basic side chains of lysine and arginine for which this enzyme has a particular affinity. Comparison of the amino acid sequences of trypsin (150) and chymotrypsin shows only two obvious differences in the lining of the tosyl hole of these enzymes. Met-192 of chymotrypsin becomes a glutamine residue at the pocket entrance in trypsin while Ser-189, deep in the pocket in chymotrypsin, occurs as aspartate.

When a basic amino acid side chain, such as asparagine, is located in the binding site of trypsin, a salt link can form between the guanidinium group and the carboxylate group of Asp-189. Two factors contribute to the favourable binding of such suitable side chains by trypsin. First, the strength of the salt link is enhanced by the hydrophobic character of the surrounding enviroment. Secondly, the aliphatic portions of the basic side chains contribute to the binding through hydrophobic and van der Waals interactions with the enzyme. Trypsin has three buried aspartate residues; two of these have functions identical to those of their counterparts in the chymotrypsin molecule, viz: maintaining the enzyme in an active conformation by forming an ion pair with the N-terminal isoleucine residue, the other being hydrogen bonded to the serine and histidine residues of the active site in the 'charge relay' system. The

third aspartate residue confers trypsin with a specificity for basic amino acids, as detailed above.

The three dimensional structure of elastase has been determined from an electron density map at 3.5Å resolution (151) and in many respects it resembles that of α chymotrypsin. However, two important differences occur in the tosyl hole. Residue 216, which is glycine in chymotrypsin, is replaced by valine which effectively blocks the entrance to the pocket. Gly-226 in chymotrypsin appears as a threonine residue in elastase, filling the cavity so that it can no longer be regarded as a major feature of the elastase molecule. Binding and subsequent hydrolysis of those substrates favoured by trypsin and chymotrypsin is consequently precluded, while the tosyl group of p-toluene sulphonate has been found to bind to the surface of the elastase molecule (151).

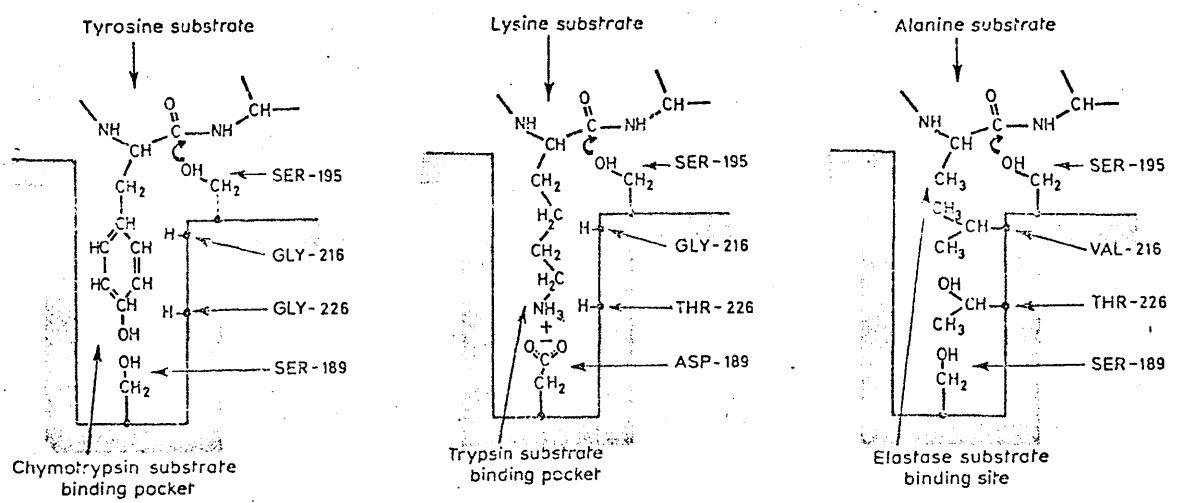


Figure 2.16.a. A diagrammatic representation of the differences between the substrate binding sites of the chymotrypsins, trypsin and elastase.

2.7. THE STEREOSPECIFICITY OF THE BINDING PROCESS.

There have been a great many studies designed to define the structural limits of compounds upon which chymotrypsin will act catalytically, as well as to quantify the interactions between the enzyme and substrates and inhibitors in order to gain an understanding of the criteria governing specificity. Although mechanism of hydrolysis and specificity can, in certain contexts, be divorced from one another, specificity, in the sense of the orientation of a substrate at the active site through interaction at certain loci on the enzyme molecule, must necessarily be an important factor in determining the degree of efficiency of the enzyme in hydrolysing the susceptible chemical bonds.

Study of the specific interactions of all low molecular weight compounds which may be considered structurally related to typical substrates permits an approach to 'mapping' those areas of the enzyme molecule involved in such interactions. The anticipation of the exhibition of well defined stereospecificity in interactions involving chymotrypsin derives from its experimental observation in early studies with amino acid amides (152) and from the application of Bergmann's polyaffinity theory in the analysis of such interactions (153).

Substrates of chymotrypsin can be described by the general formula $R_1CH_2R_2COR_3$ where, typically, R_1 = R'CONH, R_2 is a specific amino acid side chain and R_3 = NHR or OR. The interaction of the three groups, R_1 R_2 COR_3, distributed about an asymmetric carbon atom, with the active site of the enzyme can be considered to control the stereospecificity of the binding process, variations in the R groups being reflected in the K_m values observed. Niemann and coworkers (154) have quantitatively characterised a wide variety of compounds as being either substrates or inhibitors of chymotrypsin. They have attempted an explanation of their observations in terms of the requirements of an active site of relatively stable configuration containing three loci, designated ρ_1, ρ_2, ρ_3,

corresponding to the groups R_1, R_2, COR_3, of a typical substrate. The active site can be depicted thus:

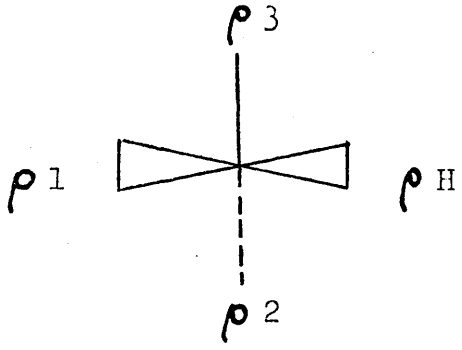


Figure 2.15

where the fourth locus, ρ_H, corresponds to the orientation of the hydrogen atom on the asymmetric carbon atom. For a given substrate there are a possible 12 different combinations at such a site for each enantiomer. The nine postulates which have been formulated for predicting the resultant of these multiple interactions have been described elsewhere (155) but some of the points are worthy of mention here:

i. only orientations which give rise to COR_3-ρ_3 interactions can lead to substrate activity, that is, to binding which can result in reaction products. Such binding modes are more effective for R_3 = OR' than for R_3 = NHR'', where R' and R'' are of comparable size; this recognises the increase in binding as well as rate of hydrolysis of esters compared with amides.

ii. all other factors being equal, a conformation which provides a larger number of interactions than another will result in greater binding energy and hence lower values for K_i or K_m. This suggests that not only more than one mode of binding of the substrate to the four loci may occur, but also that these multiple alternative binding modes may be quantitatively important in the catalytic process. If COR_3 and ρ_3 coincide then the interaction may be productive, but it will be non-productive if coincidence is with another site.

For the substrate N-acetyl L phenylalanine methyl ester, only one configuration, viz:

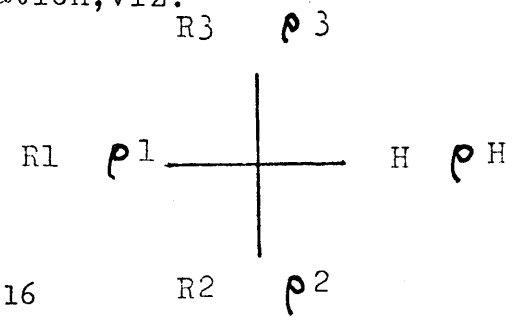


Figure 2.16

permits correct interaction at the three crucial loci of the active site. This situation cannot be achieved by the D isomer of this substrate binding to the enzyme, although several configurations with COR_3 - ρ_3 interactions are possible, as also are others having some degree of correct orientation. These are such as to suggest that hydrolysis of the D ester should be observed and this has, in fact, been reported (156), though with a k_{cat} very much lower than that for the L isomer.

Another series of studies directed at the elucidation of the requirements for stereospecificity in hydrolyses catalysed by α chymotrypsin has been carried out by Cohen and is the subject of a review (157).

Observations on the specificity and reactivity (as measured by the ratio k_{cat}/K_m) of a large number of compounds when subjected to chymotrypsin catalysed hydrolysis, have been interpreted in terms of an active site composed of four loci corresponding to various portions of the substrate molecule. Fig.2.17.a. shows the association of L methyl N-acetyl phenylalaninate, a 'natural' ester substrate, oriented in the active centre. The ar site is regarded as a cavity or fold into the surface of the enzyme in which the aromatic substituent is bound; the am site, at or near the enzyme surface, at which the α acylamido group (representing the N-terminal polypeptide of a protein substrate) associates by hydrogen bonding; the hydrolytic site, n, also near the enzyme surface, at which

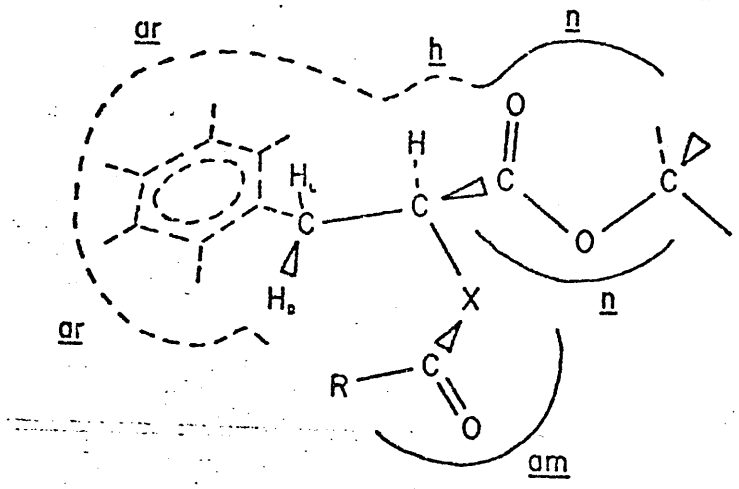


Figure 2.17.a. An illustration of the way in which the L-enantiomers of esters of α-substituted β-phenyl-propionic acids associate with the active site of α chymotrypsin.

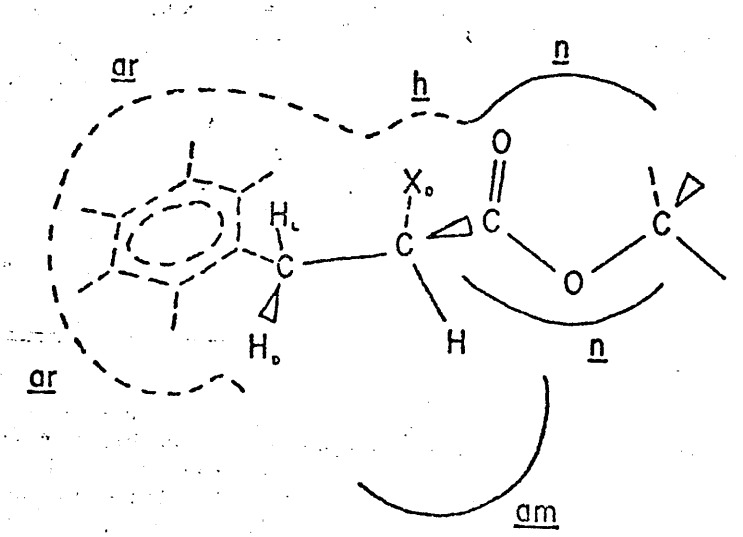


Figure 2.17.b. The association of the D enantiomers of esters of α-substituted β-phenyl propionic acids with α chymotrypsin.
When X= Cl or OH the association is effective but will fail when X= CH_3.

the group to be hydrolysed (representing the C-terminal polypeptide of a protein substrate) associates, and the h site, of restricted size, into which the α hydrogen atom may fit. The sites ρ_1, ρ_2, ρ_3, ρ_H of Hein and Niemann correspond to the am, ar, n and h sites described above.

The effective enzymic association of α substituted phenylpropionates by α chymotrypsin begins with the binding of the β phenyl group at ar, which must place the ester group at site n if hydrolysis is to occur. The L-α- acetamido and hydroxyl groups hydrogen bond at am and this results in high or moderate reactivity. The L-acetoxy carboxymethylene, methyl and chlorine groups lie at am but do not bind there and hence show low reactivity. L specificity is found in this set of compounds when the α substituents cannot fit at h. Only the small, polar groups Cl and OH can fit at h in addition to am and so allow hydrolysis of the D enantiomer (Fig.2.17.b). The methyl group, little larger than chlorine but non-polar, will not fit at h and this manifests itself in the observed L specificity of the hydrolysis of ethyl N-acetylalaninate, which proceeds with the acetamido group at am and the ester group at n; hydrolysis of the D enantiomer would require the methyl group to be at h.

Ingles and Knowles (158) sought to investigate the kinetic specificity of α chymotrypsin by preparing the p-nitrophenyl esters of the D and L enantiomers of N-acetyl tryptophan, N-acetyl phenylalanine and N-acetyl leucine as well as this same ester of N-acetyl glycine and studying the rates of deacetylation of the corresponding acyl α chymotrypsins. For those acyl enzymes of the L configuration, as the size of the amino acid side chain increases so too does the rate constant for deacetylation, a result to be expected from the increase in hydrophobic character of the substrate moiety as its bulk increases (159). The deacylation rate constants for the D series decrease in the same order. It appears, therefore, that the stereospecificity of the enzyme is closely linked to side chain specificity and these phenomena may be interpreted

in terms of a three locus model for the active site. For a specific substrate of the L configuration, the substrate moiety of the acyl enzyme is fixed at three loci:

i. by a covalent bond to the hydroxyl group of Ser-195 (A).
ii. by a hydrophobic interaction between the side chain and its corresponding locus at the active site (B).
iii. by the hydrogen bonds of the peptide link of the acyl amino group (C).

The existence of all three of these interactions will fix the susceptible carbonyl group, presumably in a position that is ideal with respect to the catalytically important side chain functions of the enzyme. The absence of one or more of these interactions will permit the substrate moiety to have a greater degree of freedom and this will be reflected in a lower deacylation rate constant.

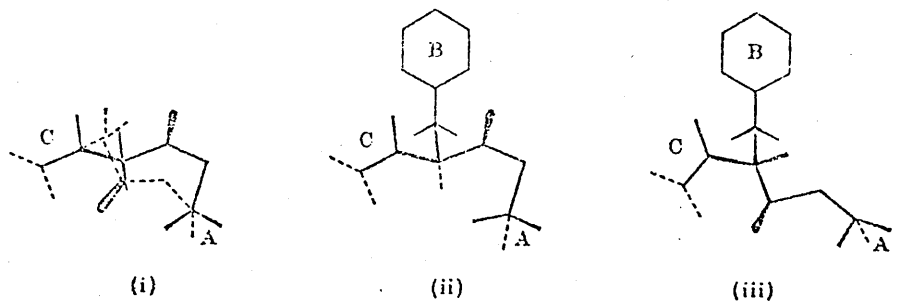


Figure 2.18. Acyl-α-chymotrypsins.
i. N-Acetylglycyl- : ii. N-acetyl-L-phenylalanyl;
iii. N-acetyl-D-phenylalanyl.

In Fig.2.18.i., where interaction of type B is not possible, rotation of the $CH_2.CO.O$ group about the fixed loci A and C can occur. Fig.2.18.ii and iii show the L and D isomers, respectively, of the N-acetyl phenylalanyl moiety of the acyl chymotrypsin in which the three interactions are present. Here the conformation of the D isomer differs from that of the L in that the susceptible carbonyl group is directed away from the centre of catalytic action.

PART 3.

THE THEORY OF NMR FAST EXCHANGE AND ASSOCIATED CHEMICAL SHIFT CHANGES.

3.1. THE QUANTITATIVE ANALYSIS OF CHEMICAL SHIFT DATA.

3.1. THE QUANTITATIVE ANALYSIS OF CHEMICAL SHIFT DATA.

To date relatively little work has been published, the object of which has been to study enzyme-inhibitor interactions by nmr spectroscopic methods utilising the chemical shifts of protons or fluorine nuclei on suitable inhibitor molecules exchanging rapidly between solution and the enzyme active site. However, that the magnitude and direction of such shifts can, potentially, give a detailed understanding of the manner in which various parts of the inhibitor molecule interact with the active site has been recognised. Equations have been developed to permit quantitative analysis of chemical shift data with a view to obtaining two useful parameters, K_D, the dissociation constant for the enzyme-inhibitor complex, and Δ the resonance-frequency shift of the complex.

Analyses performed to evaluate these parameters are, however, subject to some serious errors. In the nmr fast exchange limit, the observed chemical shift of a particular nucleus, d_{obs}, is the weighted average of the chemical shifts of the individual enviroments (bound and free) in the absence of exchange.

For the system:

$$E + I \rightleftharpoons EI$$

in rapid equilibrium,

$$K_D = \frac{[E][I]}{[EI]} \qquad (1),$$

where $[E] = E^o - [EI]$

and $[I] = I^o - [EI]$

Consequently,

$$K_D = \frac{(E^o - [EI])(I^o - [EI])}{[EI]}$$

Solving this last equation for $[EI]$ as follows:

$$E^oI^o - (EI)(I^o + E^o) + (EI)^2 = (EI)K_D$$

$$(EI)^2 - (EI)(E^o + I^o + K_D) + E^oI^o = 0$$

$$(EI) = {}^1/_2\left((E^o+I^o+K_D) \pm \sqrt{(E^o+I^o+K_D)^2 - 4E^oI^o}\right) \qquad (2).$$

gives the usual quadratic expression for a dissociation constant which also relates $[EI]$ to the known initial concentrations E^o and I^o.

For the situation in which a small molecule binds reversibly to a macromolecule, the observed chemical shift of a particular nucleus on the small molecule is given by:

$$d_{obs} = P_I d_I + P_{EI} d_{EI}$$

where P_{EI} and P_I are the fractions of small molecules bound and free in solution and d_{EI} and d_I are the chemical shifts of the bound and free species.

Since $P_{EI} = \dfrac{(EI)}{I^o}$ and $P_I = \left(\dfrac{I^o - (EI)}{I^o}\right)$

then $$d_{obs} = \left(\frac{I^o-(EI)}{I^o}\right)d_{EI} + \left(\frac{(EI)}{I^o}\right)d_I$$

When I^o is much greater than the total enzyme concentration, E^o, then:

$$\left(\frac{I^o - (EI)}{I^o}\right) \approx 1$$

so that the chemical shift of the fast exchange averaged resonance is given by:

$$d_{obs} = \left(\frac{[EI]}{I^o}\right)\triangle + d_I \qquad (3).$$

where I^o is the initial inhibitor concentration and d_I the chemical shift of the inhibitor in free solution.

Since $[EI]$ is a function of K_D and E^o it is possible to calculate K_D values from the variation of chemical shifts with changes in the concentration of inhibitor at constant initial enzyme concentration.

Typically d_{obs} is measured as a function of I^o at constant enzyme concentration. If external referencing is used d_{obs} must be corrected for any concentration dependence of the inhibitor resonance in the absence of binding (including a change in bulk diamagnetic susceptibility with changing inhibitor concentration) and d_I, measured in solution without enzyme, must be corrected for the change in bulk diamagnetic susceptibility and 'in solution' properties upon addition of the enzyme. This second factor can be particularly difficult to determine especially when remembering the sensitivity of the ^{19}F nucleus to slight changes in solvent.

Spotswood *et al* (58) derived equation (4):

$$I^o = \frac{E^o \Delta}{d} - K_D - E^o \qquad (4).$$

on the assumption that $d_{obs} < \Delta$ and where $d = d_{obs} - d_I$. The shifts are plotted according to:

$$1/d = 1/\Delta \frac{I^o}{E^o} + \frac{K_D}{\Delta E^o}$$

when a plot of I^o/E^o versus $1/d$ gives a straight line of slope Δ and intercept K_D/E^o.

Dahlquist and Raftery (51,53) have derived an equation identical with (4) in the following manner.

If the exchange lifetime is much less than $1/\Delta$ in seconds then:

$$d_{obs} = P_b \Delta$$

Therefore:

$$d_{obs} = \frac{[EI]}{I^o} \triangle$$

if d_I is taken to be zero
and:

$$[EI] = \frac{d_{obs}\, I^o}{\triangle}$$

Substitution into equation (1) leads to:

$$K_D = E^o \left[\frac{\triangle - d_{obs}}{d_{obs}}\right] - \frac{d_{obs}}{\triangle} I^o \left[\frac{\triangle - d_{obs}}{d_{obs}}\right]$$

Again assuming $d_{obs} < \triangle$ and also that K_D is of the order of I^o then:

$$I^o = \frac{E^o \triangle}{dobs} - K_D - E^o$$

when a plot of I^o versus $^1/_{d_{obs}}$ will give a straight line whose intercept is $-(K_D + E^o)$ and whose slope is $E^o \triangle$.

In general, because of viscosity and bulk susceptibility effects, d_I cannot be determined exactly from measurements on solutions containing only inhibitor, or even from solutions containing inhibitor plus some 'inert' protein. However, some approximate values of d have been calculated simply by subtracting the value without protein (d_I) from the value with protein (d_{obs}).

In order to determine K_D and $\triangle$ more accurately, Schmidt et al (49) and Sykes (62) made a 'fit' to the curve for inhibitor concentration plotted against chemical shift in a non-linear region. Analysis of the curve is based on an iterative computer procedure, using a program similar to that of Groves, Huck and Homer (63). The program works by choosing a particular value for K_D, calculating $[EI]$ according to equation (2) and then computing a least mean square fit of the nmr data to equation (3). The slope of the calculated straight line is $\triangle$ and the intercept d_I. A series of K_D values is

tested and the slope and intercept for each K_D value and the root mean square deviations of the data points from the calculated line are computed. The result for which the deviation is a minimum is taken as the best value.

Manual manipulation of the observed chemical shift data involves plotting $(d_{obs} - d_I)^{-1}$ versus I^o/E^o, which in the limit $[I] \approx I^o$ would be expected to give a straight line of slope $1/\Delta$ and intercept $K_D/E^o\Delta$. In this way an assumed value of K_D may be obtained for the computational method. The advantages of this method are that it gives a value for d_I, to which the double reciprocal plot of Spotswood is very sensitive, and also that the approximation $[I] \approx I^o$ need not be made, allowing a wider range of inhibitor concentrations to be studied.

3.2. THE QUANTITATIVE ANALYSIS OF CHEMICAL SHIFT DATA - A COMPUTER PROGRAM.

Nakano et al (59) have developed an iterative method of calculating stability constants of hydrogen bonded complexes from proton magnetic resonance data. Their approach is applicable to systems where the concentration of the two interacting species are comparable. The procedure takes into account the amount of both reactants present in the complexed form by successive approximation.

The interaction of guanosine and cytidine to form a 1:1 complex in a variety of mixed solvents has been studied by nmr (61) and the analysis of the data is based upon equations analogous to those of Nakano.

Modification of the symbolism of these equations enables a description of the interaction of enzyme and inhibitor molecules to give a 1:1 complex under conditions of rapid equilibrium. The resemblance between the equations that follow and some of those given in the section concerned with the Quantitative Analysis of Chemical Shift Data will be apparent.

For the system:

$$E + I \rightleftharpoons EI$$

in rapid equilibrium,

$$K_{assoc} = \frac{[EI]}{(E^o-[EI])(I^o-[EI])} \qquad (1).$$

The observed chemical shift, d_{obs}, of a particular nucleus on EI is the time weighted average of the chemical shift of this nucleus in the complexed and uncomplexed forms.

$$d_{obs} = \frac{[EI]}{I^o} \triangle + \frac{I-[EI]}{I^o} d_I \qquad (2).$$

Equation (2) may be rearranged so that:

$$[EI] = \frac{d_{obs} - d_I}{\triangle - d_I} \cdot I^o \qquad (3).$$

and equation (1) may be written as:

$$K_{assoc}(I^o)(E^o) - K_{assoc}[EI](I^o + E^o - [EI]) = [EI] \qquad (4).$$

Combining (2) and (3) gives:

$$E^o/d = \frac{1}{\triangle - d_I}(I^o + E^o - [EI]) + \frac{1}{K_{assoc}(\triangle - d_I)} \qquad (5).$$

Equation (5) contains two unknowns, $[EI]$ and $\triangle$, which can be calculated by means of a simple iterative procedure based on successive approximations.

A plot of E^o/d versus $(I^o + E^o)$ yields, according to equation (5), a line with a slope value approximately equal to $1/\triangle$. This can then be substituted into equation (3) to obtain the first approximation of $[EI]$. The $[EI]$ value so obtained is then used in equation (5) to calculate an improved value for the slope. These steps are repeated until two successive cycles yield essentially identical convergent values for the slope. K_{assoc} (and its reciprocal, K_{diss}) may then be calculated from the limiting slope and intercept values.

The function, F, of equation (5) may be written as:

$$F = 1/\triangle \ (I^{o}+E^{o}-[EI]) + 1/K_{assoc} \triangle - E^{o}/d \qquad (6).$$

where $1/\triangle$, labelled BARDEL in the program, is equal to $\frac{1}{\triangle - d_I}$.

and $\frac{1}{K_{assoc} \triangle}$ i.e. KCEBAR, is equivalent to $\frac{1}{K_{assoc}(\triangle - d_I)}$.

Also, $d = d_{obs} - d_I$.

Substituting equation (3) for $[EI]$ into equation (6):

$$F = 1/\triangle \ (I^{o}+E^{o}(1-d/\triangle)) + 1/Kassoc \triangle - E^{o}/d \qquad (7).$$

the gradient of this function being approximately equal to BARDEL and the intercept to KCEBAR.

A least squares adjustment of the function, which has two variables, I^{o} and d, and two adjustable parameters, K_{assoc} and $\triangle$, is carried out according to the method of Wentworth (60).

K_{dissoc} is computed according to the relationship

$$K_{diss} = \frac{1}{K_{assoc}}$$

and $\triangle$ (BIGDEL) according to

$$BIGDEL = \frac{1}{BARDEL}$$

A test for satisfactory convergence of the values of K_{diss} and $\triangle$ is written into the program-lines 137-139.

The enzyme-inhibitor complex concentration, COMCON, is calculated for each inhibitor concentration according to equation (3).

The program also calculates $[EI]$ according to the usual quadratic expression for a dissociation constant, utilising the convergent value for K_{diss}. Taking these values for $[EI]$ at each I^o, 'calc dobs' is then computed according to :

$$d_{obs} = \frac{[EI]}{I^o}\Delta + d_I$$

The program is written in FORTRAN V and the computations were executed on a Univac 1108 machine at the National Engineering Laboratory, East Kilbride, Glasgow.

The INPUT Sequence.

Estimated BARDEL - the gradient of the function
Estimated KCEBAR - the intercept of the function
E^o - the initial enzyme concentration

are read in first followed by the experimental values of

observed shift of inhibitor in the presence of enzyme - DELOBS (cps).
observed shift of the inhibitor in the buffer solution alone - DELREF (cps).
the initial concentration of the inhibitor - INHCON,

in sets of three for each experimental point.

After each set of data points a negative DELOBS value with dummy DELREF and INHCON (free format) should be punched. Data for a series of runs should be terminated by 999.

A 'Fortran V' Computer Program to calculate K_D, $\triangle$,and the enzyme-inhibitor concentration,EI.

```
C   THIS PROGRAM CALCULATES THE ASSOCIATION CONSTANTS FOR ENZYME
C   INHIBITOR COMPLEXES FROM CHEMICAL SHIFTS AT DIFFERENT INHIBITOR
C   CONCENTRATIONS, AND CONSTANT ENZYME CONCENTRATION.
C   THE EQUATIONS OF NAKANO (J.PHYS.CHEM. 71,3954,/1967/) ARE FITTED
C
C   BY THE LEAST SQUARES METHOD OF WENTWORTH.
C
C
      INTEGER RUNNO,P
      DIMENSION DOBS(10),EI(10)
      DIMENSION B(2,2),D(2,2),C(2)
      DIMENSION BIGDEL(10),DELOBS(10),DELREF(10),INHCON(10),BARDEL(10)
      DIMENSION KASSOC(10),DELDIF(10)
      DIMENSION KDISSC(10)
      DOUBLE PRECISION INHCON,L,KASSOC,KDISSC
      DOUBLE PRECISION KCEBAR
      IMPLICIT DOUBLE PRECISION (B-F,W)
      DIMENSION COMCON(10)
C
C
C
C   THIS IS THE INPUT SECTION.
C
C
C   THE ESTIMATED VALUES OF BARDEL - 1.0/(DEL(COMPLEX) - DEL(INHIBITOR))
C   KCEBAR - 1.0/(KC*(DEL(COMPLEX) - DEL(INHIBITOR)) APPROPRIATE TO THE
C   EQUILIBRIUM
C
C
C                INHIBITOR + ENZYME = COMPLEX
C
C
```

continued...

```
C   AND ENZYME CONCENTRATION - ENZCON - ARE READ IN.
C
C   BARDEL IS GIVEN APPROXIMATELY BY THE GRADIENT OF THE FUNCTION,
C   AND KCEBAR BY THE INTERCEPT.
C
C   THEN THE EXPERIMENTAL VALUES OF OBSERVED SHIFT OF INHIBITOR IN THE
C   PRESENCE OF ENZYME - DELOBS, INHIBITOR ALONE - DELREF, AND THE
C   CONCENTRATION OF THE INHIBITOR - INHCON, IN SETS OF THREE FOR EACH
C   EXPERIMENTAL POINT. AFTER EACH SET OF DATA POINTS, A NEGATIVE DELOBS
C   WITH DUMMY DELREF AND INHCON - (FREE FORMAT) - SHOULD BE PUNCHED.
C   AFTER ALL RUNS ARE COMPLETE, THE NUMBER 999 SHOULD BE PUNCHED.
C
C
C
C   THE FORMAT ON INPUT IS FREE.
C
C
C
      1 READ(5,1CCC) RUNNO
1CCC FORMAT(I3)
1CC1 FORMAT(  )
     IF(RUNNO.EQ.999) GO TO 9CCC
     READ(5,1CC1) BARDEL(1),KCEBAR,ENZCON
     I=1
   2 READ(5,1CC1) (DELOBS(I),DELREF(I),INHCON(I))
     IF(DELOBS(I).LT.C.C) GO TO 3
     I=I+1
     GO TO 2
   3 NUMOB=I-1
     KDISSC(1)=1.C/KCEBAR*BARDEL(1)
     KASSOC(1)=1.C/KDISSC(1)
     ERSHIF=C.1
     WDLDIF=1.C/ERSHIF
```


continued...

```
66*             ERRCON=C.CC5
67*             WINCON=1.C/ERRCON
68*             DO 33 I=1,NUMOB
69*            33 DELDIF(I)=DELOBS(I)-DELREF(I)
70*             P=1
71*            5 CONTINUE
72*             DO 7 I=1,2
73*             C(I)=C.C
74*             DO 7 J=1,2
75*            7 B(I,J)=C.C
76*             DO 4 I=1,NUMOB
77*             FUNCTO=ENZCON*BARDEL(P)+INHCON(I)*BARDEL(P)-
78*            1INHCON(I)*DELDIF(I)*BARDEL(P)*BARDEL(P)+
79*            2KASSOC(P)*BARDEL(P)-ENZCON/DELDIF(I)
80*    C
81*    C      B IS BARDEL.
82*    C
83*             DFBYDB=ENZCON+INHCON(I)-2.C*INHCON(I)*DELDIF(I)
84*            1*BARDEL(P)+KASSOC(P)
85*    C
86*    C      K IS KASSOC.
87*    C
88*             DFBYDK=BARDEL(P)
89*    C
90*    C      I IS INHCON.
91*    C
92*             DFBYDI=BARDEL(P)-DELDIF(I)*BARDEL(P)*BARDEL(P)
93*    C
94*    C      D IS DELDIF.
95*    C
96*             DFBYDD=-INHCON(I)*BARDEL(P)*BARDEL(P)+ENZCON/(DELDIF(I)
97*            1*DELDIF(I))
98*    C
```

continued...

```
 99*    C   THIS SECTION CARRIES OUT A GENERALISED LEAST SQUARES ADJUSTMENT
100*    C   OF A FUNCTION WITH TWO VARIABLES AND TWO ADJUSTABLE PARAMETERS.
101*    C   AS FAR AS POSSIBLE,THE NOMENCLATURE IS THAT OF WENTWORTH,
102*    C   J.CHEM.ED.42,96,/1965/).
103*    C
104*    C
105*           L=DFBYDI*DFBYDI/WINCON+DFBYDD*DFBYDD/WDLDIF
106*    C
107*    C
108*           B(1,1)=B(1,1)+DFBYDB*DFBYDB/L
109*           B(1,2)=B(1,2)+DFBYDB*DFBYDK/L
110*           B(2,2)=B(2,2)+DFBYDK*DFBYDK/L
111*           C(1)=C(1)+DFBYDB*FUNCTO/L
112*           C(2)=C(2)+DFBYDK*FUNCTO/L
113*         4 CONTINUE
114*    C
115*    C
116*           B(2,1)=B(1,2)
117*    C
118*    C
119*           DETERM=B(1,1)*B(2,2)-B(1,2)*B(2,1)
120*    C
121*    C
122*           D(1,1)=B(2,2)/DETERM
123*           D(1,2)=-B(2,1)/DETERM
124*           D(2,1)=D(1,2)
125*           D(2,2)=B(1,1)/DETERM
126*    C
127*    C
128*           DELBAR=D(1,1)*C(1)+D(1,2)*C(2)
129*           DELKAS=D(2,1)*C(1)+D(2,2)*C(2)
130*    C
131*    C
```

continued...

```
132*          BARDEL(P+1)=BARDEL(P)-DELBAR
133*          KASSOC(P+1)=KASSOC(P)-DELKAS
134*          KDISSC(P+1)=1.C/KASSOC(P+1)
135*    C
136*    C
137*          IF(ABS(KDISSC(P+1)-KDISSC(P)).LE.ABS(KDISSC(P)*C.CCC1).
138*         1AND
139*         2.ABS(BARDEL(P+1)-BARDEL(P)).LE.ABS(BARDEL(P+1)*C.CCC1))
140*         3 GO TO 9191
141*          P=P+1
142*          IF(P.GT.9) GO TO 9999
143*          GO TO 5
144*    C
145*    C
146*    C
147*       9191 WRITE(6,4CCC) P
148*       4CCC FORMAT(' CONVERGENCE WAS SATISFACTORY AFTER ` I2, ' ITERATIONS.`)
149*            P=P+1
150*       9999 CONTINUE
151*            WRITE(6,2CC1) RUNNO
152*       2CC1 FORMAT(//' HISTORY OF RUN NUMBERED ` I3,' IS AS FOLLOWS - '//////)
153*            WRITE(6,2165)
154*       2165 FORMAT(/////////)
155*            DO 9 I=1,NUMOB
156*          9 COMCON(I)=DELDIF(I)*BARDEL(P)*INHCON(I)
157*            DO 99 I=1,P
158*         99 BIGDEL(I)=1.C/BARDEL(I)
159*            DO 1C I=1,NUMOB
160*            EG=ENZCON+KASSOC(P)+INHCON(I)
161*            ROOT2=(EG*EG-4.C*ENZCON*INHCON(I))
162*            IF (ROOT2.LT.C.C) GO TO 1C
163*            ROOT=DSQRT(ROOT2)
164*            EI(I)=C.5*(EG-ROOT)
```

continued...

```
165*            DOBS(I)=DELREF(I)+EI(I)/INHCON(I)*BIGDEL(P)
166*          1C CONTINUE
167*            WRITE(6,2CC3)
168*       2CC3 FORMAT(T1C, 'IO (MOLE/LITRE)'T35,'DELTA (HZ.)'T6C,'CALC. DOBS.')
169*            WRITE(6,2CC4)  (INHCON(N),DELOBS(N),DOBS(N),N=1,NUMOB)
17C*       2CC4 FORMAT(T12,D12.6,T37,D12.6,T62,D12.6)
171*            WRITE(6,2CC5)
172*       2CC5 FORMAT(///)
173*            WRITE(6,2CC6)
174*       2CC6 FORMAT(T1C,'CALC. COMPLEX CONC.'T35,'(MOLE/LITRE)'T6C,'CALCUL. EI')
175*            WRITE(6,2CC7)  (COMCON(N),EI(N),N=1,NUMOB)
176*       2CC7 FORMAT(T12,D12.6,T62,D12.6)
177*            WRITE(6,2CC5)
178*            WRITE(6,2CC8)
179*       2CC8 FORMAT(T1C,'BIGDEL (HZ.)'T35,'KASSOC.'T6C,'KDISSC.')
18C*            WRITE(6,2CC9)  (BIGDEL(N),KDISSC(N),KASSOC(N),N=1,P)
181*       2CC9 FORMAT(T12,D12.6,T37,D12.6,T62,D12.6)
182*            WRITE(6,2121)
183*       2121 FORMAT(1H1)
184*            GO TO 1
185*       9CCC STOP
186*            END
```

PART 4.

A FLUORINE NMR INVESTIGATION OF THE BINDING OF THE N-TRIFLUOROACETYL DERIVATIVES OF D-TRYPTOPHAN AND D-PHENYLALANINE BY α CHYMOTRYPSIN.

4.I.i. Experimental.

Melting points were determined using a Kofler-Reichert hot stage apparatus and are uncorrected.

Optical rotations were measured using a Perkin-Elmer 141 Polarimeter and 10 cm. cell at the sodium D line and operating temperature of the instrument.

Yields are quoted as percentages of the theoretical maximum.

Buffers were constructed using 'AnalaR' grade reagents and glass distilled water. The pH of each was measured at the temperature at which it was to be used with a Radiometer Type 26 apparatus equipped with scale expander. A Radiometer G2026 glass electrode and type K401 calomel electrode were employed and the pH meter was standardised using solutions of pH 4,7 and 9 constructed from commercial buffer powders.

α chymotrypsin was a product of Seravac Laboratories Ltd.,grade IIA,lot 385 A and was used without further purification. Titration of the enzyme with N-trans cinnamoyl imidazole according to the method of Bender (89) indicated it to be 86% active,based on a molecular weight of 25,000.

4.1.ii. Syntheses.

a. N-trifluoroacetyl D-tryptophan.

D-tryptophan (2.0g,0.01 mole) was dissolved in trifluoroacetic acid (11.5 ml) at room temperature to give an olive green solution. After cooling to 0^{o},anhydrous ether (13 ml) was added,followed by trifluoroacetic anhydride (2 ml) after further cooling to -10^{o} in an acetone/Drikold bath. The reaction mixture was stirred at about -10^{o} for approximately 10 minutes during which time a cream coloured precipitate formed. After removing the cold bath stirring was continued for a further ½ hour prior to distillation of the solvent under reduced pressure. The resultant red/yellow solid was dissolved in the minimum of cold ethanol and twice this volume of water was then added to give a clear amber solution. Reducing the volume of this solution in vacuo gave an off-white precipitate which was filtered off at the pump;more material was obtained by further reducing the volume of the aqueous ethanol filtrate. The precipitate was pressed free of solvent,dried under vacuum and recrystallised twice from benzene.

Yield:1.5 g (50%).
m.p. 162^{o}-163^{o} (lit m.p. for the L isomer 162-163^{o} (188)).

b. N-trifluoroacetyl D-phenylalanine.

D-phenylalanine (3.3g,0.013 mole) was dissolved, with stirring at room temperature,in trifluoroacetic acid (18 ml) to give a colourless solution. After cooling to -10^{o} and adding trifluoroacetic anhydride (6.4 ml) the reaction mixture was stirred at this temperature for ½ hour when a white precipitate formed. After removing the cold bath stirring was continued for a further ½ hour whereupon the precipitate gradually redissolved. Removal of the solvent

in vacuo left an off-white residue. This was dissolved in the minimum of hot benzene and the solution decolourised with animal charcoal. The material obtained upon cooling was filtered off and recrystallised twice from benzene. After drying in vacuo, 1.3g (28%) of the trifluoroacetylated amino acid were recovered.

m.p. 119-120°.

$[\alpha]_D^{25}$ -17.5° (c 3.3 in EtOH).

(lit m.p. 120-122°, $[\alpha]_D^{25}$ -17.2° (c 2 in EtOH) (189)).

Conversion of these triflouroacetylated amino acid derivatives to their respective sodium salts was achieved by treating an aqueous suspension with an exact equivalent of sodium bicarbonate. After evolution of carbon dioxide had ceased the solution was quickly frozen and lyophilised to give white, slightly hygroscopic, low density solids.

4.1.iii, Nuclear Magnetic Resonance.

Spectral measurements were performed on a Varian Associates HA-100 Spectrometer operating in the frequency sweep mode at 94.1. MHz and 33°-34°. The ^{19}F resonance spectrum of the three equivalent fluorine nuclei in the compounds studied consists of a sharp, simple singlet and was recorded on a 50 c.p.s. sweep width. The spectrometer was 'locked' on to an external capillary* of trifluoroacetic acid held concentric with the axis of the nmr tube. The same capillary was used for all samples of any particular 'run' in order to minimise susceptibility errors.

Each sample was placed in a constant temperature water bath at 36° prior to being transfered to the probe of the spectrometer. At least three spectra of each sample were recorded after allowing to equilibrate at the ambient temperature of the probe.

Chemical shifts were determined by electronic counting of the difference between the sweep frequency and the manual oscillator frequency using a Hewlett-Packard counter, type 5212A. By moving the recorder pen downfield to the visually determined centre of a signal it was possible to obtain concordant shift values with an error of ± 0.1 Hz.

* 3.5 mm 'Dual Purpose' type, supplied by NMR Specialities, New Kensington, Pennsylvania, USA.

4.2. RESULTS.

Data relating to the α chymotrypsin-inhibitor systems investigated by ^{19}F nmr spectroscopic methods are presented in the following Tables, 4.1-4.9, in computer output format.

Notes.

DELTA (HZ)- is the observed ^{19}F chemical shift of the inhibitor in the presence of $2 \times 10^{-3}M$ α chymotrypsin.

CALC.DOBS- is the ^{19}F chemical shift for each inhibitor concentration calculated according to equation 3, section 3.1, using the convergent value of BIGDEL and the relevant value of [EI]. The latter is calculated according to the usual quadratic expression (equation 2, section 3.1) using the convergent value for KDISSC and is tabulated as CALCUL.EL.

CALC.COMPLEX CONC.- is the value for [EI] as calculated according to equation 3, section 3.2, utilising DELTA, d_I and the convergent value of BIGDEL.

The initial values of BIGDEL and KDISSC are estimates of these parameters obtained from the slope and intercept of a plot of E/d vs $(I^o + E^o)$ - compare the data accompanying Figure 4.1 with Table 4.9.

Table 4.1. Data for the binding of the sodium salt of N-trifluoroacetyl D-tryptophan by α chymotrypsin, pH 6.12.

IO (MOLE/LITRE)	DELTA (HZ.)	CALC. DOBS.
.500000-001	.308600+003	.308326+003
.400000-001	.309600+003	.309594+003
.350000-001	.310200+003	.310291+003
.300000-001	.310500+003	.311045+003
.200000-001	.312600+003	.313589+003
.150000-001	.315100+003	.315418+003
.100000-001	.321400+003	.317954+003

CALC. COMPLEX CONC. (MOLE/LITRE)	CALCUL. EI
.166095-002	.158500-002
.150593-002	.150456-002
.143395-002	.145153-002
.129554-002	.138599-002
.108515-002	.119468-002
.102148-002	.104791-002
.102979-002	.839018-003

BIODEL (HZ.)	KASSOC.	KDISSC.
.206333+003	.480000+002	.208333-001
.180928+003	.869203+002	.115048-001
.180623+003	.788692+002	.126792-001
.180620+003	.788867+002	.126764-001
.180620+003	.788867+002	.126764-001

DELREF (d_I) = 302.7 ± 0.1 Hz.

Table 4.2. Data for the binding of the sodium salt of N-trifluoroacetyl D-tryptophan by α chymotrypsin, pH 6.34.

IO (MOLE/LITRE)	DELTA (HZ.)	CALC. DOBS.
.400000-001	.312200+003	.312097+003
.350000-001	.312600+003	.312567+003
.300000-001	.313000+003	.313288+003
.200000-001	.315600+003	.315511+003
.150000-001	.317200+003	.317343+003
.100000-001	.320600+003	.320155+003

CALC. COMPLEX CONC. (MOLE/LITRE)	CALCUL. EI
.174761-002	.171090-002
.168519-002	.167484-002
.155145-002	.162860-002
.149795-002	.148215-002
.133746-002	.135665-002
.119479-002	.115507-002

BIODEL (HZ.)	KASSOC.	KDISSC.
.128205+003	.169565+002	.589744-001
.112215+003	-.982730+003	-.101757-002
.112153+003	.154460+003	.647416-002
.112153+003	.154560+003	.646999-002
.112153+003	.154560+003	.646999-002

DELREF (d_I) = 307.2 Hz.

Table 4.3. Data for the binding of the sodium salt of N-trifluoroacetyl D-tryptophan by α chymotrypsin, pH 6.89.

10 (MOLE/LITRE)	DELTA (HZ.)	CALC. DOBS.
.500000-001	.310200+003	.310089+003
.400000-001	.311100+003	.310952+003
.350000-001	.311400+003	.311570+003
.300000-001	.311900+003	.312318+003
.200000-001	.314300+003	.314397+003
.150000-001	.315800+003	.315899+003
.100000-001	.318900+003	.317911+003

CALC. COMPLEX CONC. (MOLE/LITRE)	CALCUL. EI
.153604-002	.150567-002
.144827-002	.141579-002
.132483-002	.135755-002
.121786-002	.128660-002
.107523-002	.108592-002
.929852-003	.937992-003
.789963-003	.735686-003

BIGDEL (HZ.)	KASSOC.	KDISSC.
.189753+003	.642683+001	.155598+000
.182310+003	.978173+002	.102231-001
.182267+003	.628064+002	.159219-001
.182267+003	.628093+002	.159212-001

DELREF (d_I) = 304.5 Hz.

Table 4.4. Data for the binding of the sodium salt of N-trifluoroacetyl D-tryptophan by α chymotrypsin, pH 7.43.

IO (MOLE/LITRE)	DELTA (HZ.)	CALC. DOBS.
.500000-001	.309600+003	.309409+003
.400000-001	.310400+003	.310425+003
.350000-001	.310900+003	.311005+003
.300000-001	.311300+003	.311700+003
.200000-001	.313400+003	.313590+003
.150000-001	.314800+003	.314917+003
.100000-001	.317700+003	.316552+003

CALC. COMPLEX CONC. (MOLE/LITRE)	CALCUL. EI
.146952-002	.143952-002
.133796-002	.134309-002
.126217-002	.128144-002
.114458-002	.120724-002
.962562-003	.100238-002
.846676-003	.855885-003
.721242-003	.661243-003

BIGDEL (HZ.)	KASSOC.	KDISSC.
.206186+003	.457547+002	.218557-001
.191394+003	.535046+002	.186900-001
.191337+003	.528892+002	.189074-001
.191337+003	.528895+002	.189073-001

DELREF (d_I) = 304.0 Hz.

Table 4.5. Data for the binding of the sodium salt of N-trifluoroacetyl D-tryptophan by α chymotrypsin, pH 7.72.

IO (MOLE/LITRE)	DELTA (HZ.)	CALC. DOBS.
•500000-001	•315500+003	•315246+003
•400000-001	•315600+003	•315780+003
•350000-001	•316100+003	•316207+003
•300000-001	•316300+003	•316488+003
•200000-001	•317400+003	•317469+003
•150000-001	•318000+003	•318116+003
•100000-001	•319000+003	•318913+003
•500000-002	•320600+003	•319918+003

CALC. COMPLEX CONC. (MOLE/LITRE)	CALCUL. EI
•134625-002	•126077-002
•110393-002	•115232-002
•106018-002	•108541-002
•969304-003	•100721-002
•794290-003	•803636-003
•656299-003	•668011-003
•504846-003	•499022-003
•306273-003	•283334-003

BIODEL (HZ.)	KASSOC.	KDISSC.
•132450+003	•433908+002	•230464-001
•148579+003	•357523+002	•279702-001
•148560+003	•349925+002	•285775-001
•148560+003	•349926+002	•285774-001

DELREF (d_I) = 315.5 Hz.

Table 4.6. Data for the binding of the sodium salt of N-trifluoroacetyl D-tryptophan by α chymotrypsin, pH 8.02.

IO (MOLE/LITRE)	DELTA (HZ.)	CALC. DOBS.
.400000-001	.321400+003	.321393+003
.350000-001	.321500+003	.321548+003
.300000-001	.321800+003	.321814+003
.200000-001	.322210+003	.322084+003
.150000-001	.322300+003	.322292+003
.100000-001	.322600+003	.322618+003
.500000-002	.322800+003	.322863+003

CALC. COMPLEX CONC. (MOLE/LITRE)	CALCUL. EI
.551529-003	.550606-003
.493556-003	.498766-003
.441850-003	.443126-003
.326530-003	.318658-003
.249128-003	.248764-003
.172353-003	.172903-003
.893101-004	.902898-004

BIGDEL (HZ.)	KASSOC.	KDISSC.
.322581+003	.945122+001	.105806+000
.319115+003	.963162+001	.103825+000
.319113+003	.962972+001	.103845+000
.319113+003	.962972+001	.103845+000

DELREF (d_I) = 317.0 ± 0.1 Hz.

Table 4.7. Data for the binding of the sodium salt of N-trifluoroacetyl D-tryptophan by α chymotrypsin, pH 8.12.

IO (MOLE/LITRE)	DELTA (HZ.)	CALC. DOBS.
.400000-001	.320000+003	.319991+003
.350000-001	.320100+003	.320095+003
.300000-001	.320300+003	.320218+003
.200000-001	.320600+003	.320733+003
.150000-001	.321100+003	.321232+003
.100000-001	.321400+003	.321366+003
.500000-002	.321900+003	.321740+003

CALC. COMPLEX CONC. (MOLE/LITRE)	CALCUL. EI
.704859-003	.703405-003
.644144-003	.643427-003
.587366-003	.577711-003
.415072-003	.425485-003
.328925-003	.336703-003
.238862-003	.237535-003
.129220-003	.126088-003

BIGDEL (HZ.)	KASSOC.	KDISSC.
.256410+003	.139286+002	.717949-001
.255378+003	.138048+002	.724385-001
.255378+003	.138053+002	.724359-001

DELREF (d_I) = 315.4 ± 0.1 Hz.

Table 4.8. Data for the binding of the sodium salt of N-trifluoroacetyl L-tryptophan by α chymotrypsin, pH 6.34.

IO (MOLE/LITRE)	DELTA (HZ.)	CALC. DOBS.
.400000-001	.311300+003	.311277+003
.350000-001	.311400+003	.311374+003
.300000-001	.311600+003	.311577+003
.200000-001	.311600+003	.311700+003
.150000-001	.311700+003	.311822+003
.100000-001	.312000+003	.311950+003
.500000-002	.312200+003	.312087+003

CALC. COMPLEX CONC. (MOLE/LITRE)	CALCUL. EI
.419447-003	.416832-003
.376936-003	.374398-003
.331590-003	.329649-003
.226728-003	.232412-003
.174297-003	.179467-003
.124701-003	.123291-003
.651844-004	.635829-004

BIGDEL (HZ.)	KASSOC.	KDISSC.
.344828+003	.693780+001	.144138+000
.352871+003	.665724+001	.150212+000
.352845+003	.665153+001	.150341+000
.352845+003	.665153+001	.150341+000

DELREF (d_I) = 307.6 Hz.

Table 4.9. Data for the binding of the sodium salt of N-trifluoroacetyl D-phenylalanine by α chymotrypsin, pH 6.34.

IO (MOLE/LITRE)	DELTA (HZ.)	CALC. DOBS.
.400000-001	.321000+004	.320946+004
.350000-001	.321200+004	.321218+004
.300000-001	.321300+004	.321326+004
.200000-001	.321900+004	.322080+004
.150000-001	.322700+004	.322549+004
.100000-001	.323100+004	.323099+004
.500000-002	.323800+004	.323753+004

CALC. COMPLEX CONC. (MOLE/LITRE)	CALCUL. EI
.967354-003	.955433-003
.884909-003	.886340-003
.807961-003	.812184-003
.604597-003	.624389-003
.519403-003	.506927-003
.368254-003	.368196-003
.203364-003	.202082-003

BIGDEL (HZ.)	KASSOC.	KDISSC.
.185185+003	.230769+002	.433333-001
-.997848+003	.184414+002	.542258-001
.215816+004	.177524+002	.563303-001
.182046+004	.248855+002	.401840-001
.181940+004	.234255+002	.426886-001
.181940+004	.234263+002	.426871-001

DELREF (d_I) = 316.5 ± 0.1 Hz.

Table 4.10. A Summary of the Parameters describing the binding of some N-trifluoroacetyl amino acids by ∝ chymotrypsin.

pH.	$K_D \times 10^2 M$.[a]	Δ Hz.[b]	K_I (M).[c]
6.12	1.27	180.6	-
6.34	0.65	112.2	ca. 10^{-3}
	15.0[d]	352.8[d]	-
	4.27[e]	181.9[e]	ca. 3.10^{-2}
6.89	1.59	182.3	-
7.43	1.89	191.3	-
7.72	2.86	148.6	ca. 10^{-2}
8.02	10.4	319.1	-
8.12	7.24	255.4	-

a. Dissociation constant of the ∝ chymotrypsin-inhibitor complex determined from ^{19}F resonance results at 94.1 Hz.
b. Difference in chemical shift between the fully bound and unbound inhibitor.
c. Kinetically determined inhibition constant.
d. Values for N-trifluoroacetyl L-tryptophan.
e. Values for N-trifluoroacetyl D-phenylalanine. All other values relate to N-trifluoroacetyl D-tryptophan.

Figure 4.1. Plot of the ^{19}F Chemical Shift data (derived from Table 4.9) for the sodium salt of N-trifluoroacetyl D-phenylalanine in the presence of α chymotrypsin (2×10^{-3}M), pH 6.34, according to the method of Nakano, Nakano and Higuchi (59).

Slope $= 1/\triangle$.

Intercept $= K_D \times$ slope.

slope $= 0.0054$,

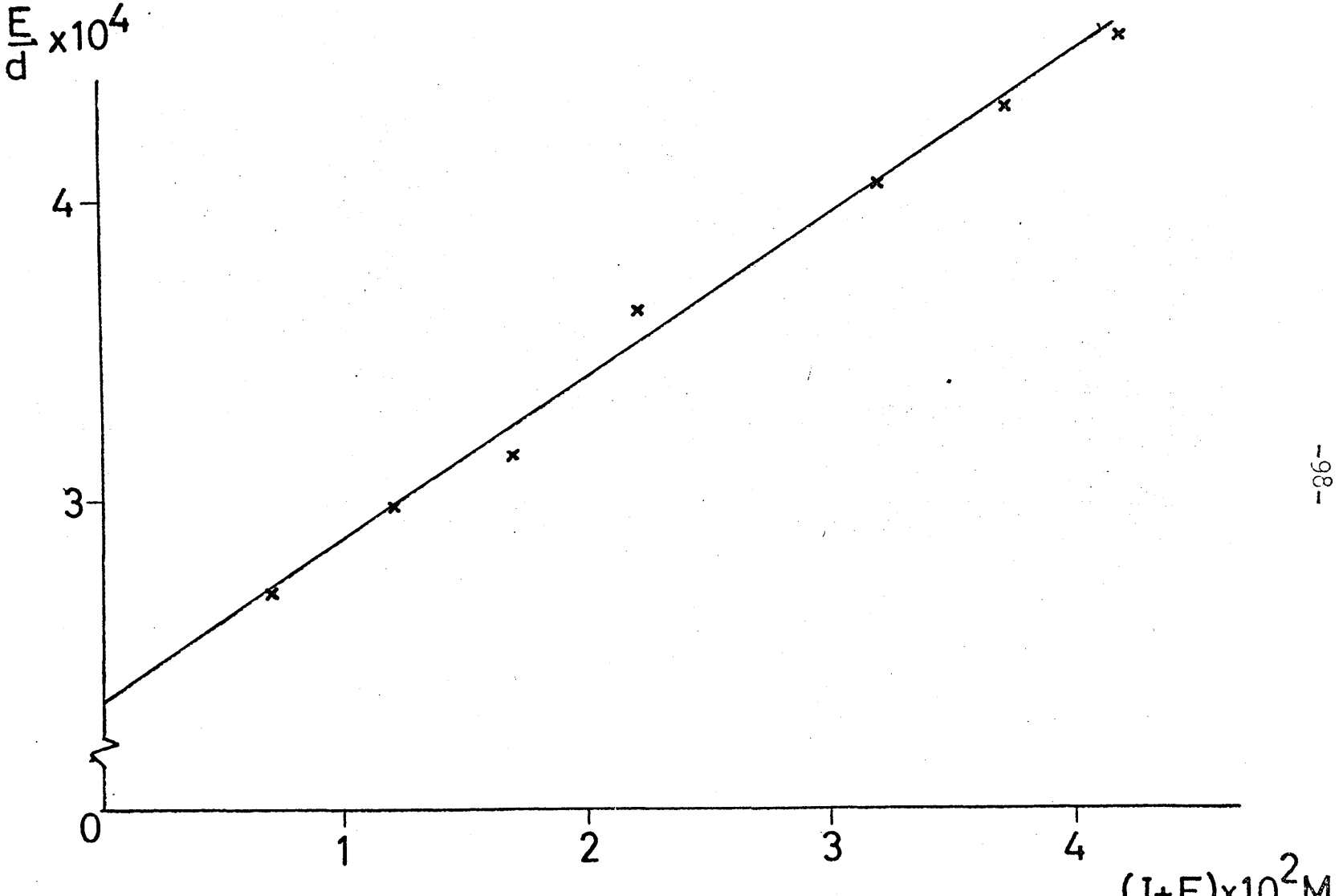
$\triangle = 185.2$ Hz.

intercept $= 0.000234$

$K_D = 4.33 \times 10^{-2}$M.

The above slope and intercept values were used to compute K_D and $\triangle$ values for this system using the computer program described in Part 3. The results are presented in computer output form in Table 4.9.

Figure 4.1.



4.3. DISCUSSION.

4.3.i. Control of pH.

In a preliminary investigation of the binding of N-trifluoroacetyl D-tryptophan by α chymotrypsin at pH 6.56 certain anomalies were noted in the measured chemical shift values. These were particularly apparent in those systems to which enzyme had not been added; little or no variation of the ^{19}F chemical shift with inhibitor concentration should have been apparent in these systems. However, after thermal equilibration the downfield shifts of these samples, in which the concentration of the fluoro compound varied from 0.05 M to 0.01 M, were noted to have increased 0.5 Hz to 1.5 H_Z over their initial values and which in themselves were not consistent.

Experiments were then performed to ascertain whether the buffers to be used in the binding study were capable of maintaining their pH in the presence of this inhibitor. For each buffer, ionic strength 0.8, two 0.5 ml samples were prepared, 0.05 M and 0.01 M in inhibitor. After thermal equilibration at 36° the pH's of the solutions were measured using a Radiometer Type TTTlc pH meter equipped with scale expander. A Radiometer glass electrode, type G2221C, and dropping calomel electrode type K1301 were employed as this combination of electrodes made it possible to measure the pH of 0.1 ml of solution. A decrease of between 0.15 and 0.78 pH units was noted for the 0.05 M solutions and of 0.05 to 0.32 pH units for those of 0.01 M.

The results suggest that the anomalous chemical shifts referred to above arise for two reasons. Firstly because of pH instability and secondly because the samples were not thermally equilibrated prior to recording their spectra. The effects of pH instability will be made immediately apparent in the anomalous chemical shifts noted for samples which have not been thermally

equilibrated prior to recording their spectra. It was found to be possible to overcome these problems by using the sodium salt of the amino acid inhibitor and thermally equilibrating the samples at 36°. In this way the chemical shifts of a set of samples were concordant and their pH's remained constant to within ± 0.05 units.

The two previous ^{19}F nmr investigations of the binding of N-trifluoroacetyl amino acid derivatives by α chymotrypsin have been concerned with phenylalanine. Zeffren and Reavill (65) reported that in 0.1 M citrate buffer, pH 6.0, the fluorine resonance of both enantiomers of N-trifluoroacetyl DL-phenylalanine shifted downfield in the presence of the enzyme. However, in the absence of α chymotrypsin the fluorine chemical shifts recorded were as follows:

N-trifluoroacetyl DL-phenylalanine ($\times 10^2$ M).	^{19}F Chemical Shift, ppm (Hz)[a]
8.00	3.302 (310.7)
4.00	3.292 (309.8)
2.00	3.286 (309.2)
1.00	3.290 (309.6)
0.30	3.285 (309.1)

a. from an external reference of trifluoroacetic acid.

The shifts will be seen to vary randomly, the highest and lowest differing by as much as 1.6 Hz and it is very probable that the lack of consistency in the values has resulted because of pH instability. Attempts to record accurate chemical shifts in the presence of α chymotrypsin will also be subject to serious error for the same reason (and this is suggested by the lack of consistent incremental change in the published chemical shift data) and consequently will invalidate the associated chemical shift differences upon which the evaluation of K_D and Δ

depend.

Nevertheless, analysis of the chemical shift data was performed, according to the method of Spotswood (58) (see part 3), and values of K_D, 3.4×10^{-2}M, Δ 1.75 ppm (164.7 Hz) were calculated for the more tightly bound enantiomer (later identified as the D form) and K_D, 5.1×10^{-2}M Δ 1.18 ppm (111.0 Hz) for the L isomer. However, the dissociation constants can only be regarded as accurate to within a factor of two since the presence of both isomers decreases the effective enzyme concentration experienced by either one.

These results provide an interesting contrast to those derived by Sykes for the D isomer of N-trifluoroacetyl phenylalanine binding to α chymotrypsin in 0.1 M tris-HCl buffer, pH 7.8 (62).

Although employing a different method of data analysis, under these conditions the ^{19}F signal was observed to shift upfield, Δ being 0.43 ppm (40.5 Hz) and K_D 4.9×10^{-2}M. However, inspection of the published plot of $(d_{obs}-d_I)^{-1}$ vs I^o/E^o for this system reveals that the observed chemical shift differences are very small (62). Thus, the variation in shift for a variation in inhibitor concentration of 4×10^{-3}M to 3.6×10^{-2}M was less than 1.0 Hz, with an experimental error of $\pm$ 0.1 Hz, making the calculation of meaningful parameters based on such small shift differences open to question.

In a similar experiment, performed with solutions of N-trifluoroacetyl D-phenylalanine of comparable concentration to those employed by Sykes, the difference in pH between the two extremes of inhibitor concentration in the presence of α chymotrypsin (2×10^{-3}M) in 0.1 M Tris-HCl buffer of nominal pH 7.88 was over one unit. Furthermore, the difference in chemical shift in the presence and absence of the enzyme was 4.5 $\pm$ 0.2 Hz for all

concentrations of inhibitor and varied randomly, the signal moving to lower field, relative to its position in buffer alone, upon addition of the enzyme.

The ^{19}F signal of the sodium salt of N-trifluoroacetyl D-phenylalanine at pH 7.96 ± 0.01 moved to lower field in the presence of α chymotrypsin (2×10^{-3}M) but the shift was constant at 3.2 ± 0.2 Hz over the concentration range 5×10^{-3}M to 4×10^{-2}M. At pH 6.34 ± 0.05 there was a downfield shift of the ^{19}F signal in the presence of the enzyme which varied with inhibitor concentration (Table 4.9). The computed K_D for this system, 4.27×10^{-2}M, is in quite good agreement with the inhibition constant of 3×10^{-2}M determined kinetically for the hydrolysis of N-furylacroyl L-tryptophanamide.

4.3.ii. Some derivatives of Tryptophan and Phenylalanine as inhibitors of α chymotrypsin.

The ^{19}F nmr data pertaining to the binding of the sodium salt of the N-trifluoroacetyl derivative of D-tryptophan to α chymotrypsin in the pH range 6.12 to 8.12 is given in Tables 4.1-4.9. Table 4.10 summarises the dissociation constant, K_D, and Δ, the total chemical shift difference between the bound and free forms of the inhibitor, for these systems, calculated using the computer program described in Part 3.

The large increase in K_D at ca. pH 8 is similar to that reported by Johnson and Knowles for the binding of N-acetyl D-tryptophan binding to α chymotrypsin, as measured by equilibrium dialysis (171). From Dixon plots (pK_I vs pH) they found K_I to be dependent upon the ionisation of a group on the enzyme of $pK_a \sim 7.3$.

As the pH rises above neutrality the anions of the N-acetyl amino acid inhibitor (which have a pK_a of 3.63) are repelled by the binding site because there is developing here an actual or net negative charge of $pK_a \sim 7.3$, this repulsion overcoming other attractive binding forces. Calorimetric studies of the hydrocinnamate- α chymotrypsin system (172) have indicated that a group in the active site is 'largely neutralised at pH 7.1 but largely negative at pH 7.9'. α chymotrypsin, however, has no amino acid residues which could conceivably have a pK_a of 7.3. Consequently it is thought that the ionising group on the enzyme having this pK_a is, below neutrality, a positively charged species which, on loosing a proton, leaves the active site with a net negative charge. Translating this hypothesis into fact involves, as one possibilty, the close juxtaposition of two ionising groups at the active site of the enzyme, one of which carries a negative charge above pH 6, while the other, of $pK_a \sim 7.3$, has a negative charge in its acid form. The two most likely candidates are the carboxyl group of Asp-194 and the imidazole group of His-57.

At pH 6.34 both the D and L isomers of the sodium salt of N-trifluoroacetyl tryptophan show a concentration dependent downfield shift of their ^{19}F signal in the presence of α chymotrypsin. In the case of the L isomer, however, the shift differences are rather too small (less than 5 Hz) to allow the determination of a reliable dissociation constant; computer analysis of the shift data gives K_D as 1.5×10^{-1}M, which is about 23 fold greater than the corresponding value of 6.5×10^{-3}M for the more tightly bound D isomer.

The K_I values determined by Foster and Niemann (173) for the acetyl derivatives of D- and L-tryptophanate binding to α chymotrypsin at pH 6.9 and 7.9 are as follows:

	pH 6.9	pH 7.9
Acetyl L-tryptophanate	2.0±0.3 mM	10±2 mM
Acetyl D-tryptophanate	1.3±0.3 mM	7.5±1.5 mM

from which it will be noted that the K_I values at pH 7.9 are significantly greater than those at pH 6.9, a tenfold increase in acidity of the reaction medium causing a five fold increase in the affinity of the catalytic site for these anionic inhibitors. The inhibitors are present as negatively charged species at these two pH's and it follows that the increase in affinity may be ascribed to substantial protonation at pH 6.9 of a negatively charged group at or near the active site which, at pH 7.9, is protonated to a much lesser extent.

It should be noted that the K_D values obtained by the direct nmr method for the sodium salts of the N-trifluoroacetyl derivatives of D-tryptophan and D-phenylalanine are of the order of 10^{-2}M (Table 4.10) and might be compared with those obtained for similar inhibitors of α chymotrypsin by steady state kinetic techniques. Thus, at pH 7.9 K_I for D-tryptophanamide binding to α chymotrypsin is 4.0±1.0 mM, identical with that reported for the N-trifluoroacetyl derivative of this inhibitor (173) and similar to that of 2.7×10^{-3}M for N-acetyl D-tryptophanamide at this same pH (174). In contrast, K_I for the benzoyl and p-methoxybenzoyl derivatives of D-tryptophanamide at pH 7.9 are much smaller

being, respectively, 0.7 ± 0.2 mM and 0.6 ± 0.2 mM (173), the observed order of affinity, benzoyl > trifluoroacetyl > acetyl, suggesting a van der Waals interaction between the enzyme and inhibitor.

One of the characteristics of α chymotrypsin is its marked affinity for aromatic amino acids. This affinity is, however, much greater for the indole side chain (cf. tryptophan) than for those of either benzyl (cf. phenylalanine) or p-hydroxybenzyl (cf. tyrosine) and this is illustrated by the following K_I values (at pH 7.9) for the same derivative of the three different amino acids;

Acetyl D-tryptophanamide	2.7×10^{-3}M	(174)
Acetyl D-phenylalaninamide	14×10^{-3}M	(175)
Acetyl D-tyrosinamide	12×10^{-3}M	(176)

4.3.iii. Attempts to elucidate the nature of the binding site of α chymotrypsin from nmr chemical shift data.

One of the unique features of nmr spectroscopy is that it allows observation of proteins in solution and consequently is a potentially powerful tool for investigations directed towards gaining a better understanding of the 'in vivo' characteristics of these particular biopolymers.

For a small molecule exchanging rapidly between solution and the binding site of an enzyme, it is the magnitude and direction of the chemical shift of a nucleus, or group of equivalent nuclei, on that molecule when bound to the enzyme which are likely to provide the most useful starting point for an understanding of the physico-chemical

nature of this site. Changes occuring in these two parameters upon altering the conditions of the system under investigation (temperature, ionic strength and pH of the medium) may then allow identification of any resultant conformational transitions occuring in the enzyme molecule.

It has been concluded from the observation of a downfield shift of the ^{19}F signal of N-trifluoroacetyl D-phenylalanine in the presence of α chymotrypsin at pH 6.0, and an upfield shift of this same signal at pH 7.8, that there is a 'major change in the magnetic environment of the fluorine nuclei of the bound inhibitor with pH' (177). However, the results obtained with this inhibitor in this investigation do not reveal any change in the direction of the fluorine chemical shift, either at pH 7.88 or pH 7.96, and although the shifts are small they were always to lower field than their corresponding position in buffer alone. If any alteration in the conformation of the chymotrypsin molecule does occur on going to more alkaline pH it is almost certainly not as significant as previously supposed and is reflected more in a marked reduction in the ability of the enzyme to bind this inhibitor at ca. pH 7.8 compared with ca. pH 6.0, than in a major change in the magnetic environment experienced by the fluorine nuclei.

An interesting approach to answering the question of what the nature of the binding site environment of α chymotrypsin might be (at least, at pH 6.0) was made by Zeffren and Reavill (65). They attempted to correlate the magnitude and direction of the ^{19}F chemical shifts of N-trifluoroacetyl DL-phenylalanine in citrate buffer, pH 6.0, in the presence and absence of the enzyme with those noted in a variety of organic solvents, the idea arising out of

similar applications of ^{19}F nmr spectroscopy to a study of micelle formation.

Evidence of micelle formation in soap solutions has been obtained from the concentration dependence of proton chemical shift (178) and spin relaxation time of water and solute proton resonance (179,180).

The shielding of the fluorine nucleus, being more susceptible to the environment than the shielding of the proton, favours the use of salts containing fluorine in studies such as these because they show a stronger concentration dependence of their ^{19}F chemical shift. Consequently, fluorinated colloidal electrolytes exhibit greater changes in chemical shift during micelle formation. Such systems are of interest because they have potential as models for investigations directed towards gaining an understanding of the nature of the environment experienced by small molecules binding to protein macromolecules, specifically, substrate and inhibitor molecules binding to the active site of enzymes.

Muller and Birkhahn (64) investigated micelle formation by soaps of the type $CF_3(CH_2)nCOONa$, where n is 8, 10 and 11, in organic solvents and in aqueous solution above and below the critical micelle concentrations of the various surfactants. By comparing the ^{19}F chemical shifts of the micelles with those obtained for the soaps it was concluded that the medium surrounding the CF_3 group when in the micelles has characteristics about midway between those of water and hydrocarbon.

An important aspect of such an investigation is the ability to use the chemical shift data of the CF_3 group to define a parameter Z, which approximately represents the

degree of non-aqueous character of the medium surrounding this group within the micelle.

The process of micelle formation may be represented by the somewhat oversimplified relationship;

$$mS \rightleftharpoons Sm$$

where S represents the detergent ion and m the aggregation number, which is assumed to be large compared with unity and to have the same value for all micelles in a given solution. Then;

$$Z = \frac{\left[d(Sm)-d(S)_{aq}\right]}{\left[d(S)_{org.}-d(S)_{aq}\right]}$$

where d(S) and d(Sm) are, respectively, the chemical shift of the monomeric soap ions and the shift for the ions in the micelle.

If the nature of the environment experienced by the ligand molecule is to be described semi-quantitatively in terms of the hydrophobic characteristics of non-polar solvents then the applicability of such an analysis to systems like those with which we are concerned here is dependent upon the magnitude and direction of the chemical shift of the bound inhibitor being similar to those observed when it is in the presence of these solvents.

The attempt of Zeffren and Reavill to apply this method, however, illustrates one of its limitations as a suitable model. Somewhat surprisingly they found the position of the inhibitor fluorine signal in ethanol, dioxan, benzene and benzene-dioxan mixtures, to be always upfield from its buffer location, whereas the effect of the

enzyme was to move the signal to lower field. Furthermore, the magnitude of the downfield shift calculated for 100% complexing of the inhibitor by the enzyme is greater than the upfield shifts observed in the presence of dioxan and benzene.

Two possible explanations for the enzyme induced downfield shifts referred to above are;

i. the fluorine nuclei are situated in an environment where they are adjacent to anisotropic groups (such as aromatic rings).

ii. on the basis that any one of the organic solvents employed approaches a hydrophobic environment, then the environment experienced by the trifluoroacetyl group when associated with the enzyme is more polar than when it is in citrate buffer alone.

This second hypothesis is given support by the observation that the ^{19}F chemical shift of N-trifluoroacetyl DL-phenylalanine in sodium chloride solution, pH 6.0, is in the downfield direction. Extrapolation of the linear variation of shift with NaCl concentration indicates that the D isomer, if dissolved in a theoretical 9.75 M NaCl solution, would yield a downfield shift of the same magnitude as would be observed if 100% of the inhibitor were complexed by α chymotrypsin. An apparently highly polar environment of acyl chymotrypsin has been indicated by fluorescence and ultra-violet spectral studies.

Of the two explanations given by Sykes for the upfield shift of the fluorine signal of N-trifluoroacetyl D-phenylalanine binding to α chymotrypsin at pH 7.8, one is the same as i. above. The other is given, somewhat ambiguously, as 'the relative polar or hydrophobic environment

experienced within the active site'.

It is anticipated that the polar environment might be created by an electric field gradient of an ionisable residue in the active site, possibly resulting from the proposed protonation of His-57 below ca. pH 7.3 (171), which would be responsible for effecting a change in the relative environment of this site. Assuming that the very small chemical shift differences recorded for this system are significant, and this is doubtful, then the change in the sign of Δ with change in pH from 6.0 to 7.8 could possibly be reconciled in terms of the ionisation of a group on the enzyme of $pK_a \sim 7.3$. If, as seems likely, there is no change in the direction of the chemical shift on going to more alkaline pH then the small shifts are more consistent with changes in the active site which preclude the binding of the inhibitor molecule, rather than of it binding in a different environment.

Mention is made in section 1.4 of Spotswoods investigation of the binding of N-acetyl DL-p-fluorophenylalanine by α chymotrypsin. Upon addition of the enzyme the ^{19}F resonance multiplet was resolved into two overlapping multiplets corresponding to the individual D and L isomers. The observed separation is due primarily to a downfield shift of the ^{19}F resonances of the D isomer, those of the L isomer showing no significant change in position, implying that the latter is much less strongly bound. Similar results were obtained for the corresponding m compound. It will be seen from the following data that there is a close correspondence in the Δ values for these inhibitors, the magnetic environment of a p fluorine atom apparently being similar to that experienced by one placed m on the aromatic ring;

	Δ (Hz).	K_D.
N-acetyl D p-fluorophenylalanine.	-83 ± 5	$(6 \pm 2) \times 10^{-3}$M.
N-acetyl D m-fluorophenylalanine.	-85 ± 7	$(11 \pm 5) \times 10^{-3}$M.

Interpretation of the shifts to lower fields as evidence for the binding of the aromatic ring in a hydrophobic pocket at or near the active site is in accordance with the suggested models for the binding site (156,181), and is supported by the electron density map of tosyl α chymotrypsin (148).

In contrast, the ^{1}H signal of the acetyl protons of N-acetyl D-phenylalanine, under similar conditions, was found to shift upfield ca. 10 Hz in the presence of the enzyme yet Spotswood regards this as being 'a similar result' to those obtained for the analogous fluoro compounds and in doing so concludes that both the aromatic ring and the acetyl group bind in a hydrophobic pocket at the active site!

Trans-cinnamic acid is a good competitive inhibitor of α chymotrypsin and the stability of the Michaelis complex formed between the cinnamate ion and the enzyme under a wide range of conditions has made it a suitable system for investigation by magnetic resonance techniques (69). The pmr spectrum of the ligand molecule undergoes line broadening and differential shifting of the signal positions in the presence of α chymotrypsin. As both the vinyl and aromatic portions of this acid molecule are stereochemically rigid, spin-spin coupling constants are thought not to change when it is bound to the enzyme. Consequently changes in the positions of the major peaks of the spectrum reflect variations in the chemical shifts induced by the enzyme.

At pH 7.4 the aromatic proton signals of the anion show an upfield shift of 0.6-0.8 ppm while the vinyl protons are only slightly affected. No spectral changes (apart from some residual line broadening) occur when α chymotrypsin in which the active site has been chemically modified, e.g. by phosphorylation or reaction with α bromoacetanilides, is used in place of the native enzyme so that the active site must be implicated in the cause of these spectral effects. The upfield shifts experienced by the aromatic ring protons in the complex are of too large a magnitude to ascribe to any influence other than a magnetic field associated with the induced 'ring current' of an aromatic ring on the enzyme (182). Accepting that local interaction between the cinnamic acid molecule and such an aromatic ring is responsible for the major portion of the observed chemical shifts then a number of arrangements of two aromatic rings can be visualised which are, semi-quantitatively, consistent with the chemical shift data. One of these places the rings about 4Å apart, which must involve near contact as space filling models indicate a benzene ring to be approximately 3.1Å thick. Such a situation could exist if the two hydrocarbon components interacted by a water exclusion mechanism.

Interaction between D and L tryptophan and α chymotrypsin at pH 6.6-6.9 is indicated by appreciable line broadening of the aromatic and alkyl portions of the pmr spectra of these amino acids (170). As with the trans-cinnamate ion, the aromatic protons show an apparent upfield shift relative to the vinyl singlet but, in contrast, the effect is small and not easily quantified. These nmr observations have been interpreted in terms of the ρ area theory of substrate binding (see section 2.7) in which the ρ_2 locus is capable of strong hydrophobic interactions

with amino acid side chains of the aromatic type. Alkylation of Met-192 with N-bromoacetyl 3-trifluoromethyl aniline, for example, is thought to place the aromatic ring of this alkylating agent in the ρ_2 site (105) and in fact when the enzyme is so modified the pmr spectrum of D-tryptophan is unperturbed. This suggests that the only point of interaction between α chymotrypsin and this amino acid is specifically the ρ_2 binding site. However, it should be borne in mind that such a modified enzyme may be so conformationally different from its native counterpart that the normal hydrophobic binding site may be no longer accessible to a suitable ligand molecule.

4.3.iv. The nature of the α chymotrypsin binding site — some thermodynamic approaches.

An early thermodynamic study of α chymotrypsin-inhibitor complex formation by Doherty and Vaslow (183) showed the enthalpy and entropy of binding to vary over a wider pH range than could be expected simply on the basis of the ionisation of one or two functional groups. These results were among the first to indicate the possibility of a conformational change taking place in the enzyme molecule upon binding an inhibitor. Conformational changes have subsequently been shown to take place when α chymotrypsin complexes with methyl hippurate and naphthalene on the basis of the very large decrease in heat capacity accompanying these highly exothermic processes (184). Similar effects have been observed at pH 6.9 with a series of inhibitors based upon the benzene ring bearing different functional groups viz: benzene, toluene, benzyl alcohol, benzaldehyde and the benzoate ion, where $\triangle C_p$ had a large

negative value, in contrast with being zero when the inhibitors were tertiary amyl alcohol and 3-pentanone (185). This implies that the proposed conformational change brought about by the aromatic inhibitors does not take place in the absence of an aromatic ring and suggests that a charge transfer complex may be involved in triggering this change. The aliphatic alcohol and ketone quite possibly bind at a location which is somewhat different from that with which the aromatic inhibitors combine. The conformational change refered to in this context is not the same as that which occurs at alkaline pH and which is associated with an absorption of a proton by α chymotrypsin during complex formation (see section 2.4). Shiao has investigated the thermodynamic changes associated with the binding of benzoate, hydrocinnamate, β-naphthoate and N-acetyl L-tryptophan by α chymotrypsin at pH 7.8 and of hydrocinnamate, phenol, indole and N-acetyl D-tryptophan at pH 5.6.(186). Here also, because the enthalpy and entropy changes accompanying the binding of the inhibitors to the enzyme at pH 7.8 are, in general, large and negative, a significant part of these changes are thought due to ligand induced conformational changes in the enzyme molecule. Similar, though much smaller, changes in enthalpy and entropy occuring at pH 5.6 would indicate that the conformational change at this pH is not so pronounced, though not neccessarily absent. These changes would appear to parallel the changes in the catalytic activity of α chymotrypsin, which has been shown to be maximal at pH 7.8 yet very small at pH values below 5. However, as the rate of the proposed conformational changes are probably slow they most likely do not participate in the normal catalytic processes of the enzyme.

For simple hydrocarbons of low polarity a simple extraction mechanism is liable to play a prominent part in the formation of α chymotrypsin-inhibitor complexes. The active site of the enzyme is known to contain an area of considerable hydrocarbon character. Thus hydrophobic binding could serve as a potential energy reservoir for the lowering of the high potential energy barrier associated with the removal of the hydrocarbon portion of the inhibitor molecule from an aqueous medium to the enzyme.

Evidence for a two point hydrophobic attachment of water analogue substrates in the enzyme complex arises out of the observation of a dependence of K_I upon the size of the inhibiting molecule (187) i.e. when complex formation with hippurate esters was compared with hydrocarbon inhibitors and derivatives of hydrocarbons. Compounds like methyl hippurate may be represented by R-P-R', where R' is an aromatic residue, P a polar central portion and R an aliphatic hydrocarbon residue. R' will bind at that site usually occupied by aromatic type inhibitors. Phenones, which have a similar R-P-R' structure and hence capable of a hydrophobic two point attachment, do in fact show the same size dependence as true substrates for the enzyme.

PART 5.

AN ATTEMPT TO DETERMINE INHIBITION CONSTANTS USING STEADY STATE KINETIC METHODS.

5. AN ATTEMPT TO DETERMINE INHIBITION CONSTANTS USING STEADY STATE KINETIC METHODS.

5.1. Introduction.

Verification of the K_{diss} values obtained by the direct ^{19}F nmr spectroscopic method for the binding of the sodium salts of N-trifluoroacetyl D tryptophan and D phenylalanine by α chymotrypsin was sought by employing steady state kinetic techniques to evaluate $K_{m(app)}$, and hence K_I for these systems.

Among the list of specific substrates for which the kinetic pathway for this particular enzyme has been elucidated is N-3-(2-furyl)acryloyl-L-tryptophan methyl ester. The α chymotrypsin catalysed hydrolysis of an ester (or other substrate) is a three step process:

$$E + S \underset{}{\overset{K_s}{\rightleftharpoons}} ES \xrightarrow{k_2} \underset{+\ P_1}{ES} \xrightarrow{k_3} E + P_2$$

The steps which follow the reversible complex formation have been characterised and most of the data pertaining to these steps has been analysed in terms of the turnover constant k_{cat}, $(k_2k_3/(k_2+k_3))$ and the apparent binding constant K_m, $(K_sk_3/(k_2+k_3))$; neither of these parameters is necessarily related to the individual rate constants in a simple manner. However, for amide substrates it is well known that K_m and k_{cat} measure the simple constants K_s and k_2.(160).
N-3-(2-furyl)acryloyl L tryptophan amide was selected as the substrate for this investigation as it has been shown that this particular amide possesses a chromophore which is ideal for monitoring the hydrolytic reaction of the primary amide bond (161).

5.2.i. Experimental.

N-3-(2-furyl)acryloyl-L-tryptophanamide (Cyclo Chemical Company, California, Lot No.R-6215) was recrystallised from ethyl acetate-n hexane and had m.p. 174-176.

∝ chymotrypsin was from the same lot as that which was used in the nmr spectroscopic studies i.e. Seravac IIA, lot 385A, and was used without further purification.

Stock substrate solutions were prepared using BDH acetonitrile. Buffer solutions were prepared from AnalaR grade reagents and glass distilled water; all were of ionic strength 0.8 and pH's were measured at the temperature at which the hydrolyses were performed (approximately 33°) being the operating temperature of the 100 MHz nmr spectrometer.

5.2.ii. Kinetic Procedure.

Kinetics of hydrolysis were determined using a Cary 14 recording spectrophotometer (0-0.1, 1-2 absorbance slide wires) equipped with thermostated cell compartments. Hydrolyses were performed at about 33° and followed at 340 mμ against a buffer blank. Cells of 10 mm path length were used and the instrument slit width set at 0.9 mm.

The method employed in conducting a typical hydrolysis experiment is described:

2.5 ml of the buffer solution was equilibrated in the sample compartment of the spectrophotometer. 5-50 μl of the acetonitrile stock solution of the amide (5×10^{-2}M) and 10-45 μl of acetonitrile were added and recording commenced. Within one minute 50 μl of stock enzyme solution (1.8×10^{-5}M) were added and recording re-commenced within 30 seconds. The final substrate concentration in a total reaction mixture volume of 2.6 ml was between 9.6×10^{-4}M and 0.96×10^{-4}M, the

percentage of acetonitrile being maintained at 1.92%.

The experiments performed to investigate the ability of the sodium salts of either N-trifluoroacetyl D-tryptophan or N-trifluoroacetyl D-phenylalanine to inhibit the α chymotrypsin catalysed hydrolysis of the amide substrate were conducted in a similar manner. To these systems were added 50 µl or 100 µl of a stock inhibitor solution (0.13M) in the appropriate buffer, keeping the final volume at 2.6 ml. The rates of hydrolysis were then recorded as before, employing the same operating conditions.

The change in molar absorptivity, $\Delta\epsilon$, was determined for every 'run' at two different substrate concentrations by allowing these reactions to proceed to completion.

5.3. The Analysis of Kinetic Data.

The general equation for the velocity, v, of an enzyme catalysed reaction is:

$$v = \frac{k_{cat}\,[E]\,[S]}{K_m + S} \qquad (1).$$

where $$k_{cat} = \frac{k_2 k_3}{k_2 + k_3} \qquad (2).$$

and $$K_m = K_s \cdot \frac{k_3}{k_2 + k_3} \qquad (3).$$

For amides, acylation is the rate limiting step, i.e. $k_3 > k_2$

so that equation (2) reduces to $k_{cat} = k_2$ and equation (3) to $K_m = K_s$. The experiments described above were conducted under conditions where $[S] > [E]$. The steady state rates of the reactions studied were measured for a fraction of the overall time period of the reaction, hence:

$$v = \frac{\Delta \text{abs}}{(\Delta\epsilon \times \Delta t)} \qquad (4).$$

where Δ abs is the change in optical density observed during time interval Δt seconds and $\Delta\epsilon$ is the total change in molar absorptivity between the amide and the acid and is calculated according to:

$$\Delta\epsilon = \frac{\Delta \text{abs}}{c \times l}$$

With [S] being of the order of 10^{-4}M and the rate, v, having values of about 10^{-7} seconds, the reciprocals 1/S and 1/v required to construct Lineweaver-Burk plots will be seen to be rather unwieldy numbers. Although reasonably good straight lines were obtained in such plots (see figures 5.1 and 5.2) they would have benefitted from additional data points corresponding to higher substrate concentrations; the long extrapolations of these lines to intercept the 1/v and 1/S axes must neccessarily introduce considerable errors into the corresponding parameters V_{max} and $K_{m(app)}$. Graphical representation of the experimental reaction rates in the form of Michaelis-Menten profiles demonstrated that the rate data utilised for the Lineweaver-Burk plots fell on the linear portion of the expected hyperbolic curves. In other words, the kinetic experiments had been performed under conditions where $K_m \gg [S]$.

5.4. The Evaluation of K_I from the reaction velocity, v.

From the Michaelis-Menten equation $v = \frac{V_{max}}{1 + K_m/S}$ (4),

$$V_{max} = v + v.K_m/S$$

and hence;

$$V_{max}/v = 1 + K_m/S$$

and;

$$1/v = K_m/V_{max}.1/S + 1/V_{max}$$

Similarly, the rate of an enzyme catalysed reaction in the presence of a competitive inhibitor, v_{inhib}, may be written;

$$v_{inhib} = \frac{V_{max}}{1 + K_m/S\ (1+I/K_I)} \qquad (5).$$

where I is the concentration of inhibitor and K_I the inhibition constant. Thus, the effect of a competitive inhibitor is to produce an apparent increase in K_m by the factor $(1 + I/K_I)$ so that K_m increases without limit as I is increased.

Combining equations (4) and (5);

$$v/v_{inhib} = 1 + I/K_I \qquad (6)$$

Difficulties encountered in attempts to evaluate K_m almost certainly stem from the need to make long extrapolations of the Lineweaver-Burk plots, neccessary to obtain approximate values for this parameter for computation. In some cases the intercept of the extrapolated line with the 1/S axis was very close to the origin and the estimated value was also subject to an error of $\pm 10^{-3}$. In view of this it was decided to calculate the dissociation constant K_I according to equation (6) using the experimentally measured rates for the inhibited and uninhibited hydrolyses at the same substrate concentration.

The α chymotrypsin catalysed hydrolysis of the N-3-(2-furyl)acryloyl derivatives of L tryptophan and L phenylalanine has been investigated by Brot and Bender at pH 8.05 in 0.1 M Tris buffer at 25° (162). Plotting the approximate values for v and S, extracted from the published Lineweaver-Burk plots for the hydrolysis of these two compounds gave straight line Michaelis-Menten profiles indicating that they had followed the kinetics of hydrolysis only over the linear (second order) portion. This neccessitated long extrapolation of the Lineweaver-Burk plots (which were defined by four data points for the phenylalanine compound and only three for the tryptophan derivative) and therefore invoking large errors in the corresponding K_m values. That this is so is indicated by the published values for this kinetic parameter, being $1.4 \pm 0.4 \times 10^{-3}$M and $4.5 \pm 1.5 \times 10^{-3}$M for the tryptophan and phenylalanine derivatives, respectively.

To perform hydrolysis experiments such as these to evaluate K_m and V_{max}, in which K_m turns out to be greater than that of the substrate concentrations employed, is not a particularly valid exercise. Reliable values for these parameters can only be obtained if the reaction rates at high substrate concentration are approaching V_{max} when the data points in this part of the Michaelis-Menten profile will help define the corresponding straight line Lineweaver-Burk plot close to where it intercepts the 1/v axis and, by further extrapolation, the 1/S axis.

Accepting that the large errors in the two K_m values given above arise out of attempts to manipulate inadequate data, K_m for the phenylalanine derivative is seen to be about three times greater than that for the tryptophan compound and not, as Brot and Bender maintain, ten times as large.

A similar trend is found for four derivatives of N-acetyl L phenylalanine compared with the same derivatives of N-acetyl L tryptophan (163):

Derivative.	N-acetyl L tryptophan. $K_m \times 10^5$M.	N-acetyl-L phenylalanine $K_m \times 10^5$M.
Amide.	730.	3700.
Ethyl ester.	9.7.	88.
Methyl ester.	9.5.	150.
p-nitrophenyl ester.	0.2.	2.4.

However, although the k_{cat} values calculated by Brot and Bender for the α chymotrypsin catalysed hydrolysis of the N-furylacroyl derivatives of tryptophan and phenylalanine ($6.9 \pm 0.4 \times 10^{-2}$, $4.22 \pm 1.22 \times 10^{-2}$ sec^{-1}) are quite close to

those of the corresponding N-acetyl amino acid amides ($4.37 \pm 0.9 \times 10^{-2}$, 4.6×10^{-2} sec^{-1}, respectively), demonstrating that the affinity of α chymotrypsin for these compounds is governed only by differences in the amino acid side chain, their interpretation of the larger K_m value for the N-furylacryloyl derivative of phenylalanine being indicative of better binding of derivatives of this amino acid by α chymotrypsin over those of tryptophan is a wrong one, for the following reason.

The Michaelis constant, K_m, is equivalent to the substrate concentration at which the reaction velocity is half maximal. If the affinity of one substrate for the active site of an enzyme is greater than that of another then the smaller is the amount of this substrate that is required to attain a given reaction velocity. Hence, as a measure of binding ability, the greater the value of K_m the less the affinity of this substrate for the enzyme, and vice-versa. Brot and Bender's observation of a higher K_m for the phenylalanine compound binding to α chymotrypsin, compared with that of a similar derivative of tryptophan, means that the active site has a lesser affinity for phenylalanine side chains, which is the converse of the conclusion drawn by these workers.

5.5. Results.

Kinetic Data for the α chymotrypsin catalysed hydrolysis of N-3-(2-furyl)acryloyl L-tryptophanamide (FATA) inhibited by the sodium salts of the N-trifluoroacetyl derivatives of D-tryptophan and D-phenylalanine.

Table 5.1. Data for the hydrolysis of FATA inhibited by the sodium salt of N-trifluoroacetyl D-tryptophan, pH 6.35.

$[S] \times 10^4$M.	$v \times 10^7$sec.[a]	$v_{inhib} \times 10^7$sec.[a]	$K_I \times 10^3$M.[b]
9.62	5.6	1.8	1.21
7.69	4.6	1.4	1.13
5.77	4.0	1.0	0.8
3.85	2.8	0.7	0.8

[I] 2.5×10^{-3}M.
Temperature 33.4°.
$\Delta\epsilon$ 541.1

Table 5.2. Data for the hydrolysis of FATA inhibited by the sodium salt of N-trifluoroacetyl D-phenylalanine, pH 6.35.

$[S] \times 10^4$M.	$v \times 10^7$sec.[a]	$v_{inhib} \times 10^7$sec.[a]	$K_I \times 10^2$M.[b]
9.62	4.8	4.2	3.6
7.69	4.4	3.9	3.8
5.77	3.6	3.1	3.1
3.85	2.7	2.4	3.1

[I] 5.0×10^{-3}M.
Temperature 33.3°.
$\Delta\epsilon$ 549.8

Table 5.3. Data for the hydrolysis of FATA inhibited by the sodium salt of N-trifluoroacetyl D-tryptophan, pH 7.67.

$[S] \times 10^4 M$.	$v \times 10^7$ sec.[a]	$v_{inhib} \times 10^7$ sec.[a]	$K_I \times 10^2 M$.[b]
9.62	8.2	5.6	1.09
7.69	6.8	4.5	0.96
5.77	5.7	3.7	0.90
3.85	3.8	2.7	1.25

$[I]$ $5.0 \times 10^{-3} M$.
Temperature 34.0°.
$\triangle\epsilon$ 562.4

Table 5.4. Data for the hydrolysis of FATA inhibited by the sodium salt of N-trifluoroacetyl D-tryptophan, pH 8.11.

$[S] \times 10^4 M$.	$v \times 10^7$ sec.[a]	$v_{inhib} \times 10^7$ sec.[a]
9.62	6.5	6.4
7.69	5.4	5.4
5.77	4.3	4.3
3.85	2.7	2.7

$[I]$ $2.5 \times 10^{-3} M$.
Temperature 33.8°.
$\triangle\epsilon$ 611.1

In this particular system the presence of the inhibitor had no effect on the rate of hydrolysis, in keeping with the ^{19}F nmr demonstration of marked reduction in the ability of the enzyme to bind this inhibitor at ca. pH 8.0. (see table 4.10).

For footnotes to Tables 5.1-5.4 see overleaf.

Footnotes to Tables 5.1 - 5.4.

a. v and v_{inhib} calculated according to equation (4), section 5.3, using $\triangle\epsilon$ values given at the foot of each table.

b. Inhibition constant K_I calculated using equation (6), section 5.4.

Figure 5.1. The Lineweaver-Burk Plot for the α chymotrypsin catalysed hydrolysis of N-3-(2-furyl) acryloyl L-tryptophanamide inhibited by the sodium salt of N-trifluoroacetyl D-tryptophan.

pH 6.35, temperature 33.4°.
[S] 3.85-9.61 x 10^{-4}M.
[E] 1.8 x 10^{-5}M.
[I] 2.5 x 10^{-3}M.

For the uninhibited reaction $-1/K_m$ is approximately -1.6×10^{-3}M giving a K_m value of 0.625×10^{-3}M.

For the inhibited reaction $-1/K_m$ is approximately -0.3×10^{-3}M giving a K_m value of 3.3×10^{-3}M.

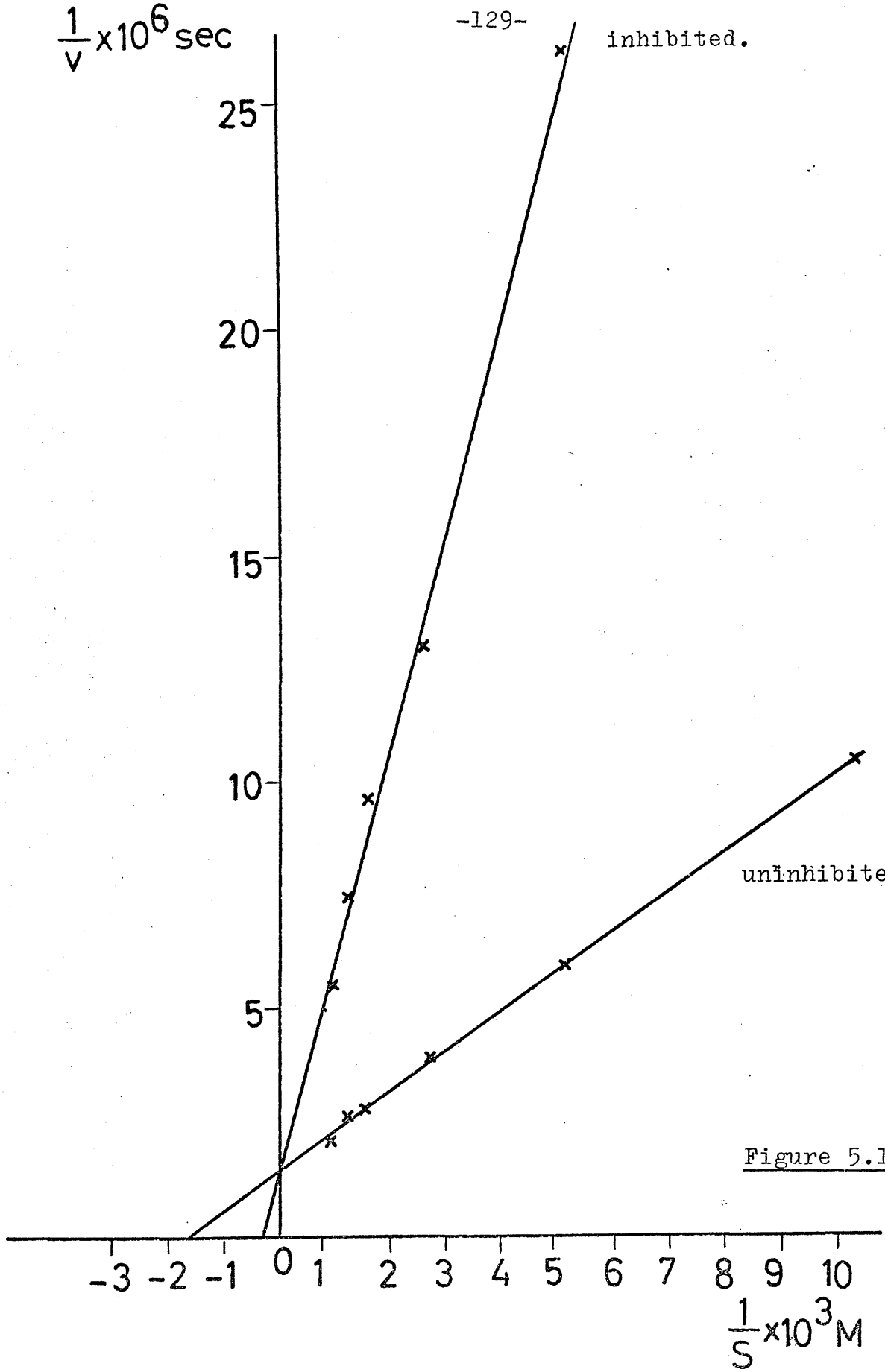

$\frac{1}{v} \times 10^6$ sec
25
20
15
10
5
-3 -2 -1 0 1 2 3 4 5 6 7 8 9 10
$\frac{1}{S} \times 10^3$ M
inhibited.
uninhibited.

Figure 5.1.

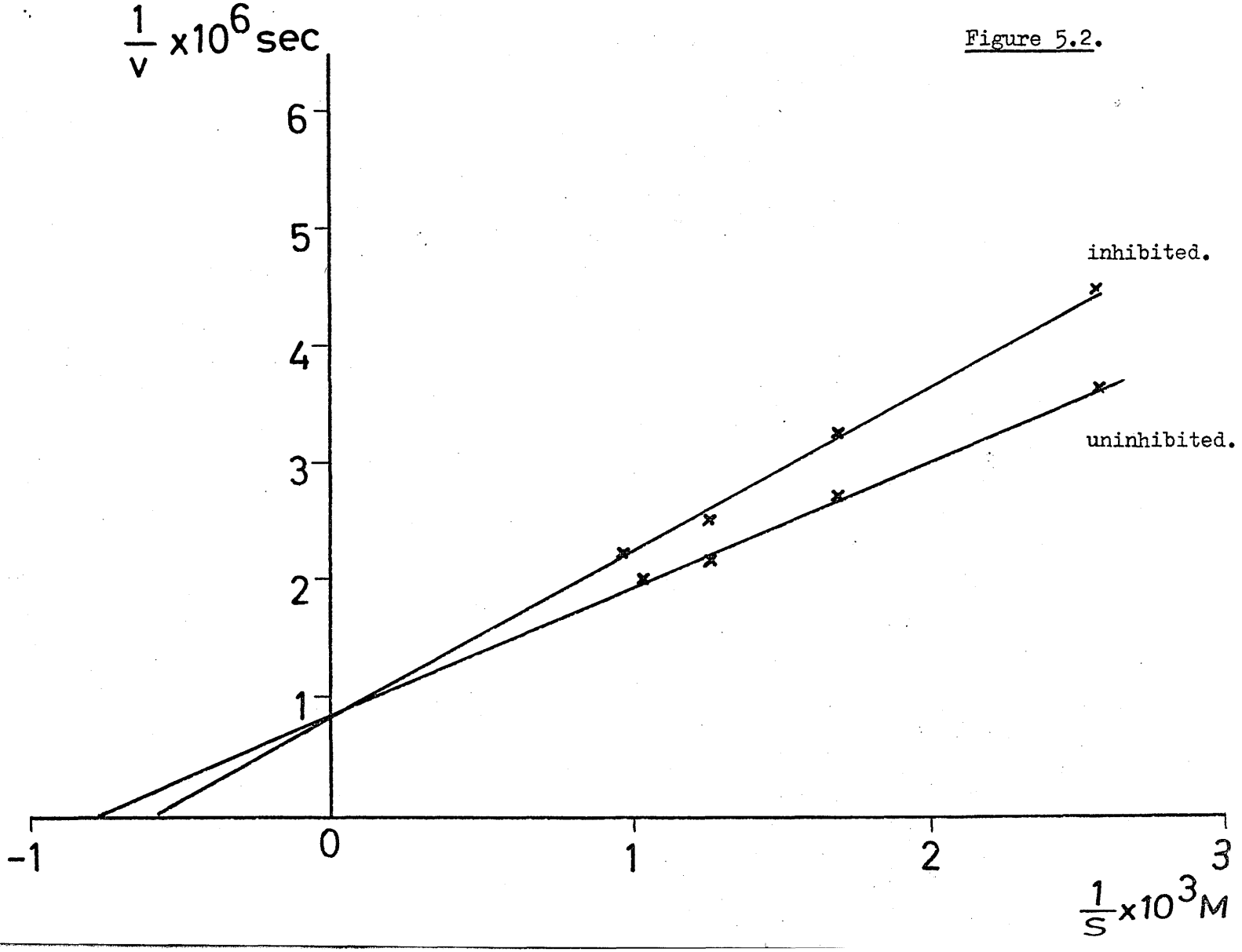
Figure 5.2. The Lineweaver-Burk Plot for the α chymotrypsin catalysed hydrolysis of N-3-(2-furyl) acryloyl L-tryptophanamide inhibited by the sodium salt of N-trifluoroacetyl D-phenylalanine.

pH 6.35, temperature 33.3^{o}.
[S] $3.85\text{-}9.62 \times 10^{-4}$M.
[E] 1.8×10^{-5}M.
[I] 5.0×10^{-3}M.

For the uninhibited reaction $-1/K_m$ is approximately -0.8×10^{-3}M giving a K_m value of 0.125×10^{-3}M.

For the inhibited reaction $-1/K_m$ is approximately -0.6×10^{-3}M giving a K_m value of 0.166×10^{-3}M.

Figure 5.2.



PART 6.

A PROTON AND FLUORINE NMR SPECTROSCOPIC STUDY OF THE BINDING OF SOME MONOSACCHARIDE INHIBITORS BY LYSOZYME.

6.1. EXPERIMENTAL.

6.1.i. Syntheses.

a. A Route to the synthesis of Methyl N-trifluoroacetyl α -D-glucosaminide.

i. The Preparation of Methyl α -D-glucosaminide (II).

Methyl N-acetyl- α -D-glucosaminide (2.5 g) were dissolved in 60 ml of water and barium hydroxide octahydrate (6.5g) added and the whole refluxed for 24 hours. TLC analysis of the reaction mixture after this time showed no starting material to be present ($nPrOH:H_2O$, 2:1, silica stationary phase). The suspension was filtered and the filtrate saturated with carbon dioxide. The barium carbonate precipitate was filtered off through glass fibre paper and the procedure repeated until no more $BaCO_3$ was precipitated. The pH of the filtrate was adjusted to about 4.0 by the careful addition of $\underline{N}$ H_2SO_4 and the resultant precipitate of $BaSO_4$ filtered off, again using glass fibre paper. The clear colourless filtrate was lyophilised to give 1.5g (73%) of a clear, pale brown syrup the I.R. spectrum of which did not show the amide bands characteristic of the starting material.

ii. The Preparation of Methyl N-trifluoroacetamido-2-deoxy-1,3,4,6-tetra-O-trifluoroacetyl- α -D-glucosaminide (III).

To the 1.5g of the amino sugar (II) were added sodium trifluoroacetate (0.8g) and trifluoroacetic anhydride (12.5 ml). A vigorous reaction ensued and the two solids quickly dissolved to give a straw coloured solution. A double surface condenser fitted with silica gel drying tube was attached to the reaction flask and after the initial reaction had subsided the reaction

was refluxed gently overnight. The resultant clear yellow solution was allowed to cool and the excess trifluoroacetic anhydride removed in vacuo leaving a slightly turbid orange syrup. This syrup was repeatedly extracted with hot carbon tetrachloride, the extracts being quickly filtered and pooled. Removal of the organic solvent in vacuo gave 1.54g (34%) of a pale green syrup.

I.R. (neat syrup); 1720 cm^{-1} (amide I), 1550 cm^{-1} (amide II), 1780 cm^{-1} (ester carbonyl), approx.1200 cm^{-1} (broad band, C-F stretch).

iii. The Preparation of Methyl N-trifluoroacetyl-α-D-glucosaminide (IV).

O-detrifluoroacetylation of III was achieved by treating the syrup with an excess of anhydrous methanol for approximately 48 hours. Removal of the methanol in vacuo gave a colourless, semi-crystalline solid which was recrystallised from n-propanol to give a white crystalline compound. This was filtered off at the pump and dried in vacuo to give 350mg (45%) of white needles.

m.p. 192-193°.

$[\alpha]_D^{25}$ + 84.6° (c 0.8 in water)

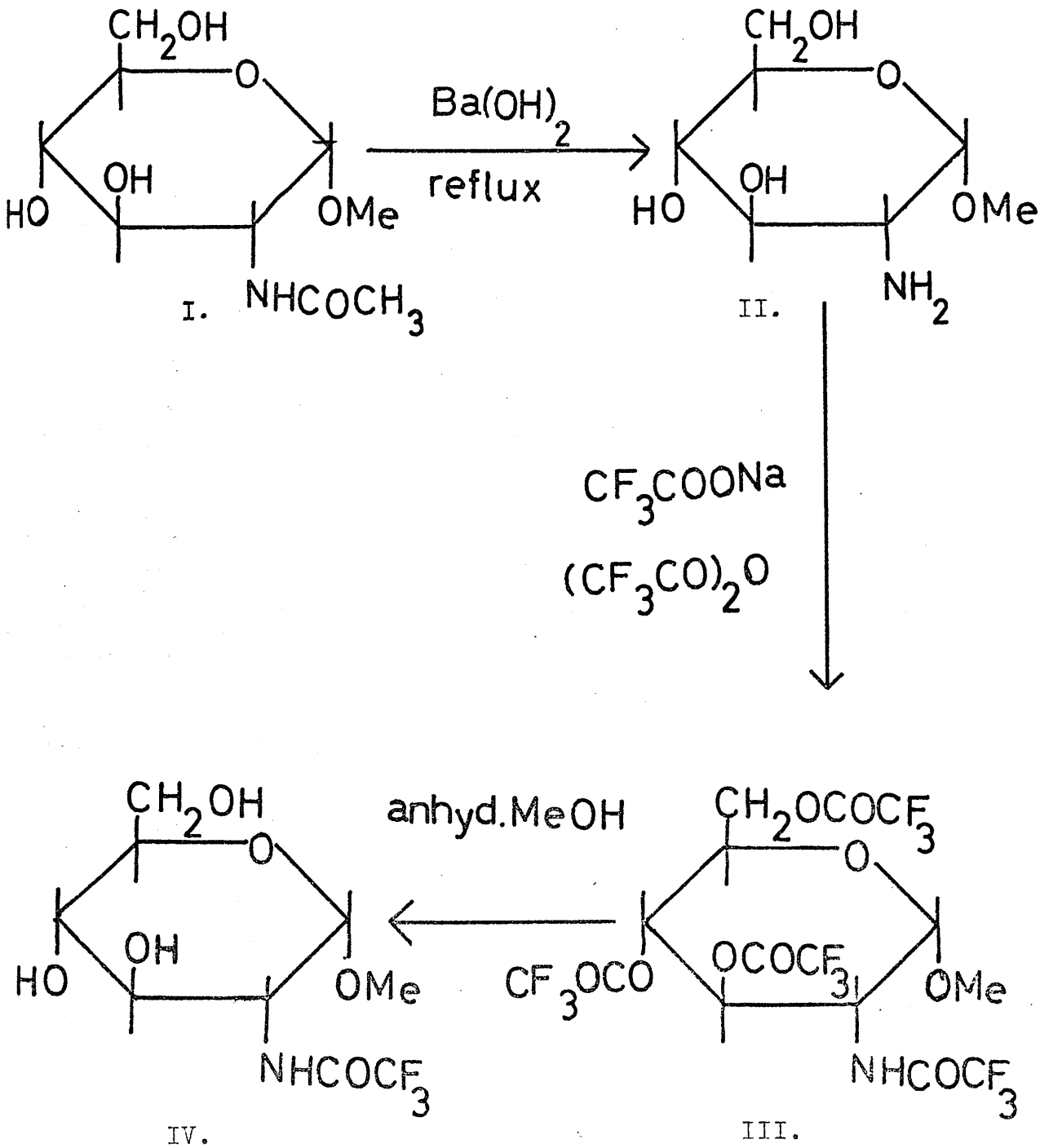
I.R. (KBr disc): 1710 cm^{-1} (amide I), 1555 cm^{-1} (amide II) 1190 cm^{-1} (C-F stretch).

N.M.R. (d_5-pyridine): singlet 3.4 ppm (3H, OCH_3), doublet centred on 5.2 ppm, coupling constant 4 Hz (1H, anomeric proton).

Found: C 37.16, H 5.06, N 4.87, F 19.8.

$C_9H_{14}O_6NF_3$ requires: C 37.37, H 4.84, N 4.84, F 19.72.

A Route for the Synthesis of Methyl N-trifluoroacetyl α D-glucosaminide.



b. A Route to the Synthesis of Methyl N-trifluoroacetyl-β-D-glucosaminide.

i. The Preparation of 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-α-D-glucopyranosyl chloride (VI).

This was prepared from N-acetyl-D-glucosamine and acetyl chloride according to the literature method (190).
Yield, 17.4g (72%) from 15g NAG.
m.p. 126°, lit. m.p. 125-127° (190).

ii. The Preparation of Methyl-2-acetamido-tri-O-acetyl-2-deoxy-β-D-glucopyranoside (VII).

The glucopyranosyl chloride (VI) was converted to (VII) by treating it with anhydrous methanol and freshly prepared silver carbonate according to the published method (191). Silver carbonate, as well as silver oxide, when used in a Koenigs-Knorr reaction, lead to a Walden inversion giving the β-D anomer from a stable α-D-Oacyl glycosyl halide.
A total of 9.8g (26%) of (VII) was obtained from 17.4g of the chloro sugar (VI) as prepared in i. above and 20g of commercial glucopyranosyl chloride (Koch Light Laboratories Ltd.).

m.p. 158°, lit. m.p. 163° (191).
TLC: nPrOH:H_2O, 2:1, silica stationary phase. Single spot, R_f same as that of an authentic sample.
I.R. (Nujol): 1640 cm^{-1} (amide I), 1550 cm^{-1} (amide II), 1735 cm^{-1} (ester carbonyl), 2900 cm^{-1} ($-OCH_3$).

N- and O- deacetylation of (VII), followed by total trifluoroacetylation of the resultant amino sugar was achieved by employing methods similar to those used for the α anomer. O-detrifluoroacetylation was performed as

before, using anhydrous methanol. After allowing to stand for about 48 hours the solvent was removed in vacuo and the off-white residue dissolved in the minimum of hot n-propanol. On standing this solution yielded a somewhat glutinous white material which proved difficult to free of solvent. The jelly-like residue was first dried in warm air to give a brittle white compound which was ground to a coarse powder and dried in vacuo.

TLC: n-PrOH:H_2O, 2:1, silica stationary phase. Only a single spot revealed after spraying with concentrated sulphuric acid and heating.

m.p. 223-224°.

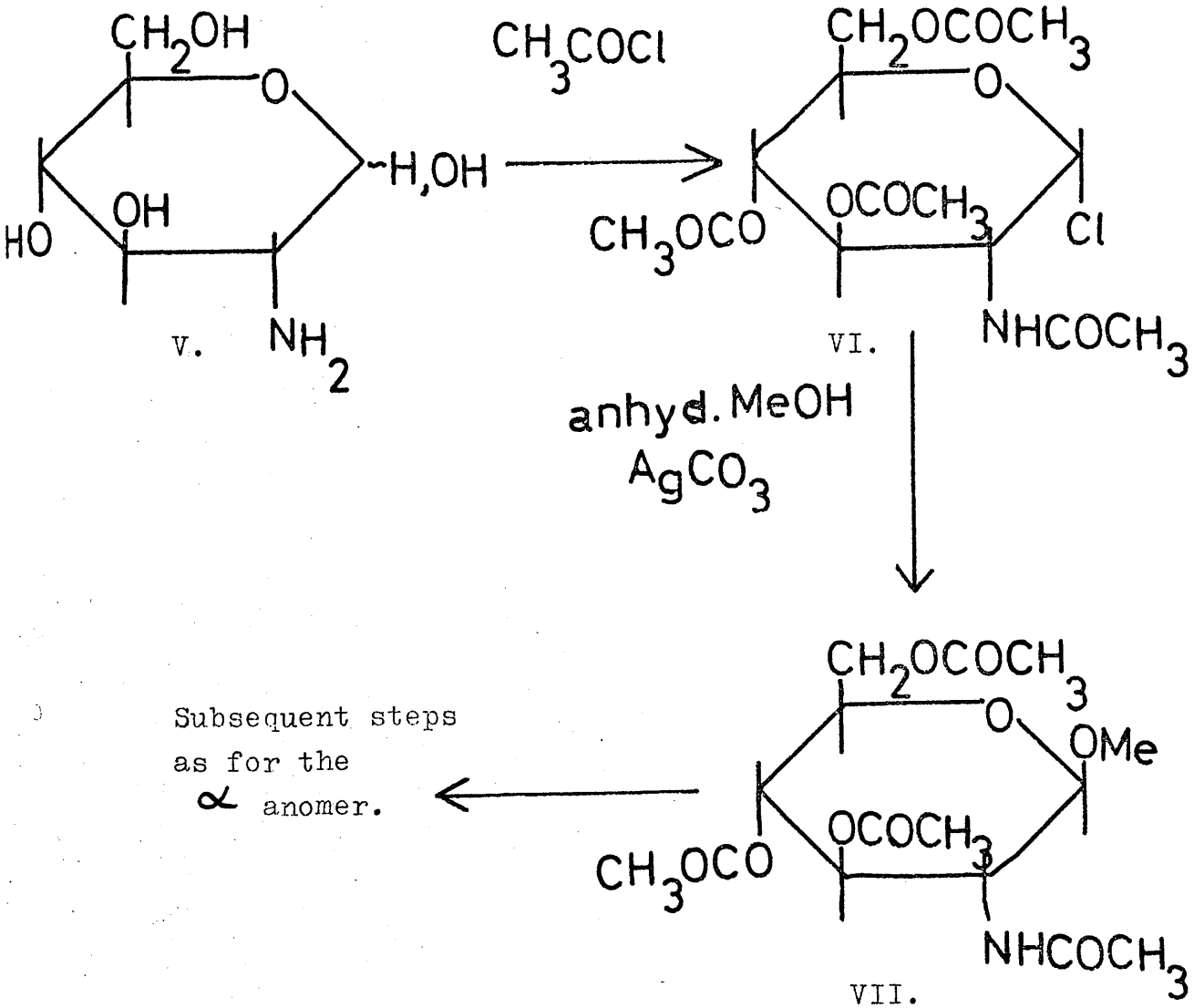
$[\alpha]_D^{25}$ -28° (c 0.5 in water).

N.M.R. (d_5-pyridine): singlet 3.5 ppm (3H, OCH_3), doublet centred on 5.0 ppm, coupling constant 8 Hz (1H, anomeric proton).

Found: C 37.01, H 4.96, N 4.80, F 19.5.

$C_9H_{14}O_6NF_3$ requires: C 37.37, H 4.84, N 4.84, F 19.72.

A Route for the Synthesis of Methyl N-trifluoroacetyl β D-glucosaminide.



6.1.ii. Experimental Method.

N-trifluoroacetyl-α-D-glucosamine (NTG), m.p. 199-202, and N-acetyl D-glucosamine (NAG), m.p. 214-215, were kindly supplied by Dr.B.Capon.

The α and β anomers of the methyl glycosides of NTG were synthesised according to the methods described in section 6.1.i.

Tetra-N-acetyl chitotetraose (NAG-4), mol.wt. 830, was isolated from a mixture of chitin oligosaccharides by R.L.Foster (165).

α chymotrypsin was obtained from Seravac Laboratories Ltd., grade IIA, lot 385A.

A stock solution of hens egg white lysozyme (Boeringer Corporation, 3.9×10^{-3}M, based on a mol.wt. of 14388) was prepared by dissolving the enzyme in glass distilled water and had a pH of 4.42.

The samples for NMR spectroscopy were prepared as follows:

Two quantities of the particular inhibitor sugar were accurately weighed out and each dissolved in 0.5 ml. of the stock enzyme solution to give solutions of approximately 8×10^{-2}M and 6×10^{-2}M. These samples were then placed in a constant temperature water bath at 36° for about $1\frac{1}{2}$ hours. This allowed those samples containing either NTG or NAG to reach mutarotational equilibrium at a temperature approximating the ambient probe temperature of the spectrometer (33°-34°). The spectra were recorded following thermal equilibration in the probe.

Dilution of 250 µl of these initial samples with a similar volume of stock enzyme solution gave solutions containing half the original concentration of inhibitor.

Repetition of this procedure made it possible to study systems over the concentration range $\sim 8 \times 10^{-2}$M to $\sim 7.5 \times 10^{-3}$M using the minimum of inhibibitor. After dilution each sample was thermally equilibrated prior to recording the spectrum.

A corresponding series of samples containing inhibitor alone at the same concentrations were constructed in a similar fashion.

6.1.iii. Nuclear Magnetic Resonance.

^{19}F nmr spectral measurements of NTG and the α and β anomers of the methyl glycosides of NTG were performed on a Varian Associates HA-100 Spectrometer operating at 94.1 MHz and 33°-34°. Chemical shifts were recorded from an external capillary of trifluoroacetic acid to which the spectrometer was 'locked', the same capillary being used for all samples.

The proton magnetic resonance spectrum of the acetamido methyl protons of NAG was recorded on the same instrument operating at 100 MHz. The water resonance, from which the chemical shifts were measured, was used as the 'lock' signal.

Chemical shifts were determined by electronic counting of the difference between the sweep frequency and the manual oscillator frequency using a Hewlett-Packard counter.

The ^{19}F nmr spectrum of mutarotated NTG in water consists of two singlets of equal intensity, 312.4 Hz and 339.8 Hz downfield from an external reference of trifluoroacetic acid. The assignment of each of these signals to the appropriate anomer of NTG has been achieved on the basis of evidence which supports the fact that this compound crystallises from n-propanol in the α form (165). A 60 MHz spectrum of freshly crystallised NTG in d_6-DMSO (plus approximately 5% D_2O) shows a doublet with a coupling of 3 Hz centred on δ 5.10 ppm. This integrates for one proton and by analogy with the splitting of the C-1 proton of methyl α-D-glucosaminide is assigned to the C-1 equatorial proton of α NTG. Mutarotation of the sample provides the pmr spectrum with a second doublet, with a coupling of 7 Hz

centred on δ 4.60 ppm and by similar comparison this can be assigned to the C-1 axial proton.

A freshly dissolved sample of NTG has a ^{19}F nmr spectrum in which the intensity of the signal to lowest field (339.8 Hz), after about 5 minutes, is seen to be greater than that of the second signal. As mutarotational equilibrium is attained their intensities equalise. Hence, the resonance at 339.8 Hz (A) may be ascribed to the trifluoroacetyl group of the α anomer while that at 312.4 Hz (B) belongs to this same group in β NTG.

In contrast, the intense unsplit resonance in the ^{1}H nmr spectrum of mutarotated NAG has been assigned to the methyl group protons of the acetamido group side chain. Upon addition of lysozyme this singlet is resolved into two signals, both of which are seen to move upfield from the original position in the absence of the enzyme.

The proton chemical shifts for the α and β anomers of NAG and the ^{19}F chemical shifts for the anomers of NTG and its methyl glycosides with varying inhibitor concentration in the presence and absence of lysozyme are set down in Tables 6.1 to 6.3.

6.2. RESULTS.

In the following Tables, 6.1-6.3 give the ^{1}H nmr chemical shift data for the binding of the α and β anomers of N-acetyl-D-glucosamine, and the ^{19}F nmr chemical shift data for the binding of the α and β anomers of N-trifluoroacetyl-D-glucosamine and its methyl glycosides to the enzyme, lysozyme. Table 6.4 summarises the K_D and Δ values for the complexes formed between these and some related monosaccharides and the enzyme.

Table 6.1. Proton Chemical Shift Data for the binding of the α and β anomers of N-acetyl-D-glucosamine by lysozyme, pH 4.42.

Mutarotated NAG x 10^2M.	^{1}H Chemical Shift (Hz)[a]				
		α anomer.		β anomer.	
8.5	268.9[b]	272.1[c]	3.2[d]	270.3[c]	1.4[d]
6.22	268.8	272.8	4.0	270.4	1.6
4.25	268.5	274.3	5.8	271.2	2.7
3.11	268.5	275.1	6.6	271.5	3.0
2.13	268.5	276.7	8.2	272.5	4.0
1.56	268.5	277.8	9.3	273.0	4.5
1.07	268.5	278.7	10.2	273.4	4.9
0.78	268.8	279.3	10.5	273.9	5.1

a. Experimental error ± 0.1 Hz.

b. Signal position, upfield from the water resonance, in the absence of lysozyme.

c. Signal position in the presence of 3.9×10^{-3}M lysozyme.

d. Upfield chemical shift difference.

Table 6.2. ^{19}F Chemical Shift Data for the binding of the α and β anomers of N-trifluoroacetyl-D-glucosamine by lysozyme, pH 4.42.

Mutarotated NTG x 10^2M.	^{19}F Chemical Shift (Hz)[a]. α anomer (d_A).			β anomer (d_B).		
9.44	339.8[b]	337.7[c]	-2.1[d]	312.4[b]	316.0[c]	3.6[e]
7.26	339.8	336.3	-3.5	312.3	316.0	3.7
4.72	339.8	334.0	-5.8	312.2	316.1	3.9
3.63	339.7	332.0	-7.7	312.1	316.0	3.9
2.36	339.6	329.9	-9.7	312.0	316.3	4.3
1.82	339.5	328.0	-11.5	312.2	316.4	4.2
1.18	339.5	326.1	-13.4	312.1	316.2	4.1
0.91	339.5	325.2	-14.3	312.2	316.1	3.9
1.45 + 3.76 x 10^{-3}M chymotrypsin.	340.3					
1.18 + 4.88 x 10^{-3}M NAG-4.	340.7			315.5		

a. Experimental error $\pm$ 0.1 Hz.
b. Signal positions, downfield from external trifluoroacetic acid, in the absence of lysozyme.
c. Signal positions in the presence of 3.9 x 10^{-3}M lysozyme.
d. Upfield chemical shift difference.
e. Downfield chemical shift difference.

Table 6.3. ^{19}F Chemical Shift Data for the binding of the α and β anomers of the methyl glycosides of N-trifluoroacetyl-D-glucosamine by lysozyme, pH 4.42.

i. Methyl N-trifluoroacetyl α-D-glucosaminide.

Methyl α NTG x 10^2M.	^{19}F Chemical Shift (Hz)[a].		
8.2	345.8[b]	341.7[c]	4.1[d]
6.6	345.8	341.6	4.2
4.1	346.0	341.4	4.6
3.3	345.9	341.5	4.4
2.05	345.9	341.4	4.5
1.65	345.9	341.4	4.5
1.03	346.0	341.3	4.7
1.21+3.76 x 10^{-3}M chymotrypsin.		343.8	

(continued overleaf)...

Table 6.3. (continued):-

ii. Methyl N-trifluoroacetyl β-D-glucosaminide.

Methyl β NTG x 10^2M.	^{19}F Chemical Shift (Hz)[a]		
7.55	314.4[b]	310.2[c]	4.2[d]
6.14	314.6	310.2	4.4
3.78	314.9	310.3	4.6
3.07	314.8	310.4	4.4
1.89	315.0	309.8	5.2
1.54	315.1	309.7	5.4
0.95	315.4	309.8	5.6
0.77	315.4	309.6	5.8
1.25+4.0 x 10^{-3}M chymotrypsin.		313.1	

a. Experimental error ± 0.1 Hz.

b. In the presence of 3.9 x 10^{-3}M lysozyme.

c. Signal position in the absence of lysozyme.

d. Chemical shift difference, downfield from external trifluoroacetic acid.

Table 6.4. Collected K_D and $\triangle$ values for some monosaccharides binding to lysozyme.

	Monosaccharide.	$K_D \times 10^2$M.	$\triangle$ ppm (Hz).	pH.	solvent.
1.	Methyl α NAG [a]	3.9 ± 0.5	0.65 ± 0.08 (61.2)	5.3	D_2O
	Methyl β NAG [a]	2.9 ± 0.4	0.68 ± 0.1 (64.0)	5.3	D_2O
2.	Methyl α NAG [b]	5.2 ± 0.4	0.55 ± 0.02 (51.8)	5.5	0.1 M citrate
	Methyl β NAG [b]	3.3 ± 0.5	0.54 ± 0.04 (50.8)	5.5	0.1 M citrate
3.	α NAG [c]	2.4 ± 0.4	0.78 ± 0.08 (73.4)	5.2	H_2O
	β NAG [c]	2.3 ± 0.5	0.30 ± 0.04 (28.2)	5.2	H_2O
4.	α NAG [d]	1.6 ± 0.1	0.68 ± 0.03 (64.0)	5.5	0.1 M citrate
	β NAG [d]	3.3 ± 0.2	0.51 ± 0.03 (48.0)	5.5	0.1 M citrate
5.	α NAG [e]	2.04	0.99 (87.0)	4.42	H_2O
	β NAG [e]	1.1	0.35 (33.3)	4.42	H_2O
6.	α NAG [f]	Complex K_D	0.97 (91.6)	4.42	H_2O
	β NAG [f]	2.81	0.48 (45.5)	4.42	H_2O
7.	α NAG [g]	1.87	0.92 (86.3)	4.42	H_2O
	β NAG [g]	1.27	0.37 (34.2)	4.42	H_2O
8.	α NTG [e]	0.914	0.82 (78.0)	4.42	H_2O
	α NTG [g]	0.66	0.80 (75.5)	4.42	H_2O

See overleaf for footnotes.

Footnotes to Table 6.4.

a. Reference 166.

b. Reference 52. Data plotted according to I° vs 1/d.

c. Reference 164. Data plotted according to I°/E° vs 1/d.

d. Reference 54.

e. This work. Data plotted according to I°/E° vs 1/d.

f. This work. Data plotted according to I° vs 1/d.

g. This work. Parameters computed using the program described in section 3.2, based on estimates obtained from a plot of E/d vs $(I^{\circ} + E^{\circ})$.

Figure 6.1. Plot of the Chemical Shift data (from Table 6.1) for the acetamido methyl resonances of α and β NAG at mutarotational equilibrium in the presence of 3.9×10^{-3}M lysozyme, pH 4.42, according to the method of Dahlquist and Raftery (54).

Intercept $= -0.022 = -(K_D + E^o)$

so complex $K_D = 0.022 - 0.0039$,

$= 2.81 \times 10^{-2}$M.

Slope $= E^o \Delta$,

so $\Delta_\beta = 45.5$ Hz,

and $\Delta_\alpha = 91.6$ Hz.

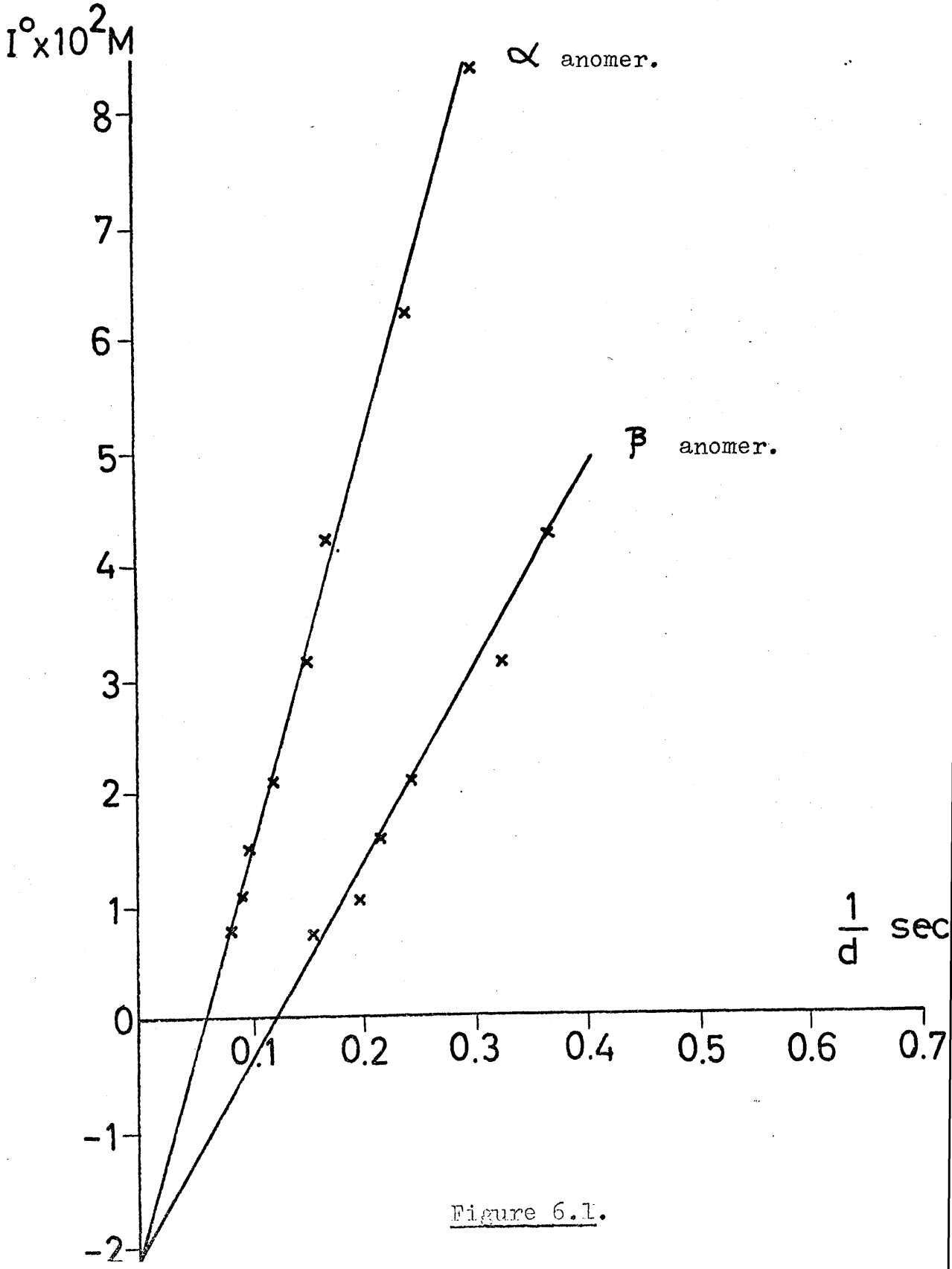

I°x10²M
8
7
6
5
4
3
2
1
0
-1
-2
α anomer.
β anomer.
1/d sec
0.1
0.2
0.3
0.4
0.5
0.6
0.7

Figure 6.1.

Figure 6.2. Plot of the Chemical Shift data for the acetamido methyl resonances of α and β NAG (from Table 6.1) and the trifluoromethyl fluorine resonance of α NTG (from Table 6.2) at mutarotational equilibrium in the presence of 3.9×10^{-3}M lysozyme, pH 4.42, according to the method of Sykes and Parravano (164).

Slope $= 1/\triangle$.
Intercept $= K_D/E^o\triangle$.

For α NAG:
slope $= 0.0115$,
$\triangle = 87.0$ Hz.
intercept $= 0.06$
$K_D = 2.04 \times 10^{-2}$M.

For β NAG:
slope $= 0.03$
$\triangle = 33.3$ Hz.
intercept $= 0.085$
$K_D = 1.1 \times 10^{-2}$M.

For α NTG:
slope $= 0.0128$
$\triangle = 78.0$ Hz.
intercept $= 0.03$
$K_D = 9.14 \times 10^{-3}$M.

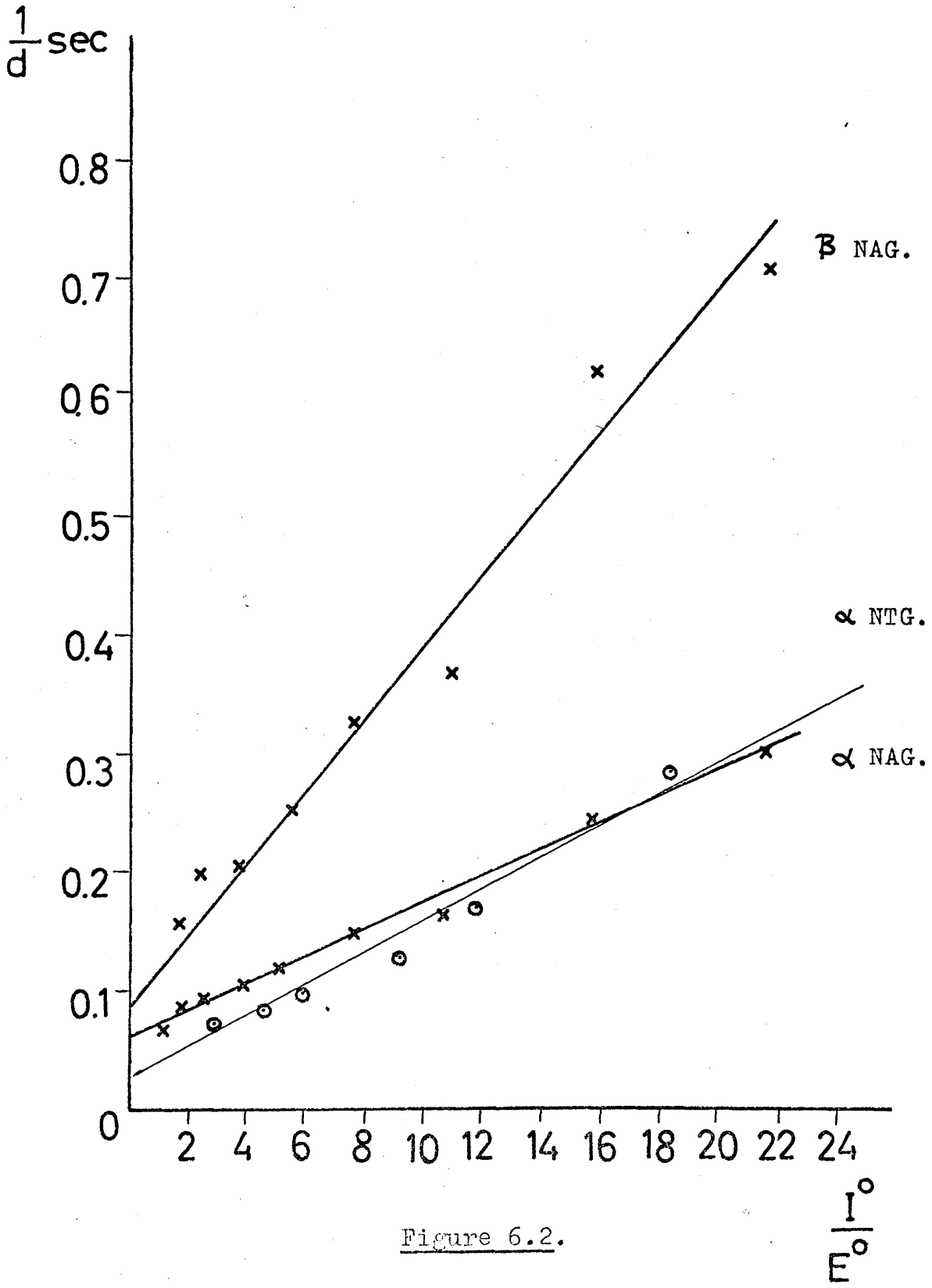

1/d sec
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
4
6
8
10
12
14
16
18
20
22
24
I°/E°
β NAG.
α NTG.
α NAG.

Figure 6.2.

Figure 6.3. Plot of the Chemical Shift data for the acetamido methyl resonances of α and β NAG (from Table 6.1) and the trifluoromethyl fluorine resonance of α NTG (from Table 6.2) at mutarotational equilibrium in the presence of 3.9×10^{-3}M lysozyme, pH 4.42, according to the method of Nakano, Nakano and Higuchi (59).

Slope $= 1/\triangle$.
Intercept $= K_D \times$ slope.

For α NAG:
slope $= 0.0112$,
$\triangle = 90.0$ Hz.
intercept $= 0.0002$
$K_D = 1.79 \times 10^{-2}$M.

For β NAG:
slope $= 0.028$,
$\triangle = 35.7$ Hz.
intercept $= 0.0003$
$K_D = 1.07 \times 10^{-2}$M.

For α NTG:
slope $= 0.0126$,
$\triangle = 79.4$ Hz.
intercept $= 0.00008$
$K_D = 6.35 \times 10^{-3}$M.

The above slope and intercept values were used to compute K_D and $\triangle$ values using the program described in Part 3. The results are given in Table 6.4.

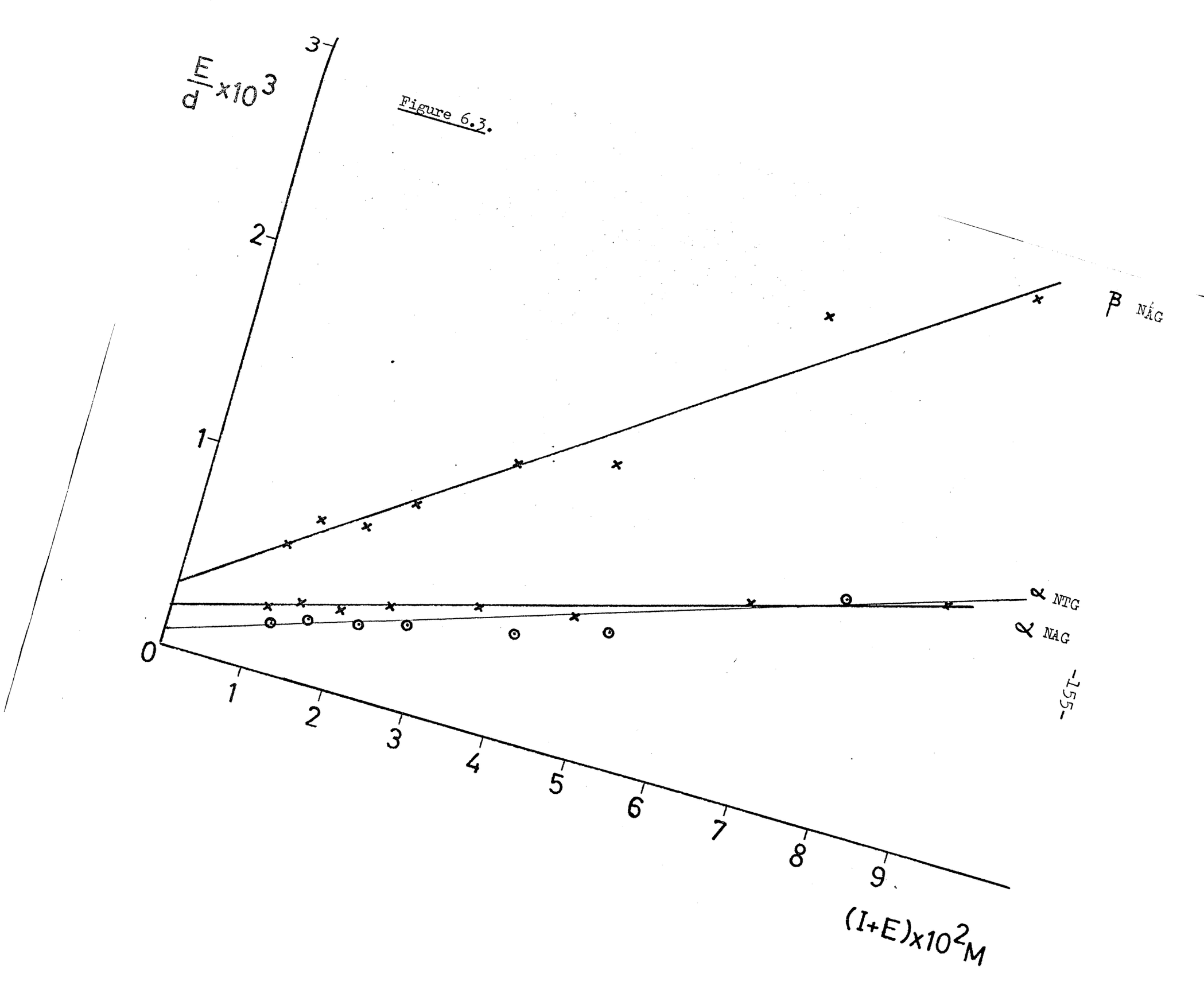

$\frac{E}{d} \times 10^3$
3
2
1
0
1
2
3
4
5
6
7
8
9
$(I+E) \times 10^2 M$
β NÁG
α NTG
α NAG

Figure 6.3.

6.3. DISCUSSION.

The binding of the α and β anomers of N-acetyl-D-glucosamine (NAG) by lysozyme has been investigated by several workers using proton magnetic resonance techniques. The observation of separate resonances for the α and β anomers in the presence of lysozyme is indicative of binding of the two anomeric forms to the enzyme with different affinities, or that they bind in a manner such that their acetamido side chains do not occupy the same position on the enzyme, or both. These conclusions are possible whether or not the two anomers bind to different sites or compete for the same site on the enzyme surface. That the latter is true has been demonstrated by the addition of N-acetyl-(d_3)-α-D-glucopyranose to an equilibrated mixture of NAG and lysozyme (52). The observed chemical shifts of both anomers were decreased (and also, therefore, the percentage of each bound) by approximately equivalent amounts, indicating that the two anomers do compete for the same site on the enzyme.

Table 6.1 shows the incremental changes in the chemical shifts of the acetamido methyl group protons for the anomers of NAG with variation in inhibitor concentration at pH 4.42, the shifts for the α anomer being larger than those for the β. The chemical shifts are all upfield relative to the water resonance. Evaluation of K_D and Δ were made according to three different methods and the values obtained for these parameters are compared in Table 6.4 with those already published. Figure 6.1 shows the results plotted according to I^o vs $1/d$, where I^o represents the total amount of NAG present i.e. the sum of the concentrations

of the two anomers present at equilibrium. As expected, the two lines are seen to intersect on the I^{o} axis. This intercept is related to the dissociation constant for the enzyme-inhibitor complex but because it is the same for both anomers the value of 2.81×10^{-2}M represents a complex K_D which cannot be separated into the individual contributions made by each anomer. Values for Δ, the chemical shift of the lysozyme bound form of each inhibitor are 91.6 Hz (0.97 ppm) and 45.5 Hz (0.48 ppm) for the α and β anomers respectively.

The similarity of the relative computed dissociation constants for the α and β anomers of NAG, $K_{D\alpha}$ 1.87×10^{-2}M and $K_{D\beta}$ 1.27×10^{-2}M (Table 6.4, data set 7), combined with the observation that the factor of two between $K_{D\alpha}$ and $K_{D\beta}$ obtained by Dahlquist and Raftery at pH 5.5 (Table 6.4, data set 4) represents a difference of only 0.4 kcal at 31^{o}, might ordinarily be regarded as good evidence that the binding of the two anomers is identical. However, the extreme sensitivity of the magnetic resonance method to environmental changes (expressed as Δ values, the chemical shifts of selected nuclei in the enzyme complex) reveals that the bound orientations of α NAG and β NAG are different. At pH 4.42 Δ_α 86.3 Hz (0.92 ppm) and Δ_β 34.3 Hz (0.37 ppm) differ by a similar factor to those reported by Sykes at pH 5.2 (164) viz: Δ_α 73.4 Hz (0.78 ppm), Δ_β 28.2 Hz (0.30 ppm). Dahlquist and Raftery's values of 64 Hz (0.68 ppm) and 48 Hz (0.51 ppm) for Δ_α and Δ_β respectively at pH 5.5 do not differ as much but are still indicative of a different mode of binding for each of the anomers (54).

The ^{19}F chemical shifts observed for α and β N-trifluoroacetyl-D-glucosamine binding to lysozyme at pH 4.42 are given in Table 6.2. In the case of the α anomer

these shifts are upfield from an external reference capillary of trifluoroacetic acid and show a dependence on inhibitor concentration.

When tetra-N-acetyl chitotetraose (NAG-4) was added at a concentration 1.2 times that of the lysozyme the ^{19}F signals of the α and β anomers appeared 0.7 Hz and 3.9 Hz downfield from their original position in the absence of the enzyme. These downfield shifts presumably arise from a medium or bulk susceptibility effect since downfield shifts were also recorded in the presence of α chymotrypsin. The upfield shift of the α anomer which is suppressed by NAG-4 must therefore arise from binding in the active site. Further indication that the site competitively occupied by α and β NAG is the same one to which α NTG binds is supplied by the observation that tri-N-acetyl chitotriose (NAG-3), which is known to bind to lysozyme about 10^4 fold more strongly than NAG, also completely eliminates the binding of α NTG, whereas mutarotated NAG only slightly reduces the recorded shift differences (165).

The variation in the chemical shift of the α NTG fluorine signal with inhibitor concentration was analysed by the method of Sykes (166) and values of 78 Hz for the difference in chemical shift between the bound and unbound forms of the inhibitor and 9.1×10^{-3}M for the dissociation constant were calculated. The analogous values for α NAG, determined under indentical conditions, were 87 Hz and 2.04×10^{-2}M. The similarity of the $\triangle$ values suggests that although α NTG is bound in the active site it is situated differently from α NAG since if they were situated identically and experiencing identical magnetic environments in the binding site a much larger shift should be found with the fluoro compound.

The upfield shift of the methyl acetamido resonances of α and β NAG, and of the corresponding trifluoromethyl signal of α NTG when bound to lysozyme, is due to the increased magnetic shielding of these protons (or fluorine atoms) in the complex. The most likely causes of these shifts are the proximity of the CH_3 (or CF_3) group to;

i. an aromatic side chain of the enzyme (167)

or ii. an electric field of an ionisable group nearby (168).

The only difference between α and β NAG lies in the configuration of the hydrogen atom and hydroxyl group at C-1 and it is certainly this feature which results in the differing dissociation constants and bound orientations for the two anomers. The binding energy of α NAG has been accounted for in terms of the interactions of the acetamido side chain and the C-3 hydroxyl group with tryptophan residues 62 and 63. Since α NAG is also capable of utilising these same interactions and since it binds better than β NAG it is probably able to do so because it can form an additional bond to the enzyme via its C-1 hydroxyl group. A hydrogen bond between this group and the main chain -NH- of Val-109 would also allow the trifluoroacetyl group of α NTG to occupy a position on the enzyme similar to that occupied by the acetamido methyl group protons of α NAG, adjacent to Try-108 in sub-site C.

The crystallographic studies of Blake (57) on the association of NAG with lysozyme have shown that in the crystalline complexes the orientations of the two anomers differ with respect to the position of their pyranose rings, the acetamido side chains making the same contacts with the enzyme in both cases. This does, however, result in the acetamido methyl group of β NAG occupying a position in which it is less magnetically shielded by Try-108.

The ^{19}F signal of the β anomer of NTG moves downfield ca. 4 Hz in the presence of lysozyme and the shifts are independent of inhibitor concentration indicating that this anomer has little affinity for the enzyme. The inability of β NTG to bind to lysozyme must be due, in part at least, to the fact that it cannot form a hydrogen bond to the enzyme via its C-1 hydroxyl group. The formation of a stable complex may also be sterically prevented by the diequatorial arrangement of the trifluoroacetamido group.

The methyl glycosides of NAG, in which the ~~conformation~~ configuration of each anomer is 'frozen', have been shown to compete for the same site on the enzyme surface and that this site is identical with that occupied by α NAG (52). From the data in Table 6.4 (sets 1 and 2) it will be seen that these two glycosides bind to lysozyme with slightly different affinities but with the acetamido methyl groups of both occupying identical magnetic environments, as indicated by the close correspondence between Δ_α and Δ_β. On the other hand the glycosidic methyl groups of these compounds have been noted to experience different magnetic environments when bound to the enzyme. Thus, the methyl signal of the β anomer was shifted to lower fields but apparently not affected in the α glycoside.

The value Δ, 0.51 ppm, for β NAG, being almost the same as that for the α and β anomers of methyl NAG binding to lysozyme under identical conditions (0.1 M citrate buffer, pH5.5), suggests a similar mode of binding for these three sugars to the same site on the enzyme, whereas Δ of 0.68 ppm for α NAG indicates that, although this inhibitor is competitive for this same site, its acetamido methyl group protons experience a different magnetic environment.

Methyl α NTG shows only a concentration independent downfield shift of ca. 4.4 Hz in the presence of lysozyme, in marked contrast to methyl α NAG which shows an upfield shift similar to that shown by the free sugar (51,54,55). The larger dissociation constants for the methyl glycosides of NAG compared with α and β NAG are thought to reflect a steric interaction between the additional methyl group and Asp-52 and Val-109. Similar unfavourable interactions must also occur between these residues and the methoxyl group at C-1 of both anomers of methyl NTG and these probably require the trifluoroacetamido group to adopt an orientation which effectively precludes the binding of these compounds to the enzyme.

In conclusion, Kent and Dwek have recently reported an investigation of the binding of N-fluoroacetyl-α-D-glucosamine to hens egg white lysozyme employing fluorine magnetic resonance methods (169). At pH 7 in the presence of the enzyme this compound shows a downfield shift of 13 Hz, in contrast to the upfield shifts observed here with N-trifluoroacetyl-α-D-glucosamine. The ABX type quartet of the fluoroacetyl compound, arising from the non-equivalence of the $-CH_2F-$ protons, is still observable but is accompanied by line broadening. The change in chemical shift upon addition of the enzyme is thought to be attributable to a fast exchange between free and bound forms of the inhibitor, but a dependence of shift on inhibitor concentration was not reported.

PART 7.

A MINOR FLUORINE NMR STUDY OF THE HYDROLYSIS OF α-D-GLUCOSYL FLUORIDE BY AMYLOGLUCOSIDASE.

7.1. Introduction.

The existence of starch degrading enzymes in various species of Aspergillus and other moulds has long been known. Studies on the amylolytic enzymes from these sources showed that while some preparations liquefied starch with the production of maltose and dextrins (α -amylase type action), in others there was a rapid production of glucose. Barker and Carrington reported that a cell-free extract of Aspergillus niger gave glucose as the major product of its action upon starch (192) and the enzyme responsible has subsequently been purified (193,194); it belongs to a general group of enzymes known as the glucamylases or amyloglucosidases.

The hydrolysis of a variety of amylosaccharides by glucamylases is multichain and consists of the step-wise removal of single glucose units from the non-reducing ends of the chains (195). The point of cleavage of the α 1-4 glucosidic link when glucamylase acts on maltose has been investigated by conducting hydrolyses in $H_2{}^{18}O$. Isolation of the glucose as its phenylosazone, followed by pyrolysis of the latter, gave carbon dioxide which mass-spectrometry showed to be not enriched in oxygen-18 (196). This result demonstrates that the hydrolysis proceeds by the fission of the C1-O bond, that between the anomeric carbon atom and the glycosidic oxygen, a route general for hydrolases and transglycolases.

The Glucamylase catalysed hydrolysis of fluoro-monosaccharides.

Fluorine analogues of metabolites are usually observed to either inhibit a particular biochemical pathway or to be utilised by enzyme systems with the fluorine atom intact. Only rarely have there been reports of C-F bond cleavage (197). Barnett et al., however, have observed the hydrolysis of α-D-glucosyl fluoride by an extract of rat intestinal mucosa (198). With a pH optimum of 6.6, hydolysis of this substrate was complete in 30 minutes, the glucose rapidly entering the cell to be metabolised while fluoride was found to accumulate in the incubation medium. This α-D-glucosyl fluoride hydrolase activity was confirmed to be due to α-D-glucosidase, rather than the associated mutarotase, as it could not use β-L-arabinosyl fluoride as substrate. Furthermore, tris inhibited the α-glucosidase and the hydolase activities to the same extent but had no inhibiting effect on mutarotase. The glucosyl fluoride was also found to be hydrolysed by Agidex, a fungal amylogluco-sidase.

α-D and β-D-glucopyranosyl-, α-D and β-D-galacto-pyranosyl, α-D-mannopyranosyl and α-D-xylopyranosyl fluorides have been hydrolysed specifically by the respective glycosidases from several sources (199). These glycosyl fluorides obey the normal specificity rules for the glycosidases in that the enzyme is specific for both the glycosyl moiety and the configuration at C1.

It has been proposed that the initial step in the glucosidase catalysed hydrolysis of α-D-glucosyl fluoride involves the protonation of the fluorine atom and the elimination of hydrogen fluoride to leave either a carbonium ion or an enzyme-glycoside covalent complex (198). This mechanism, in which the fluorine atom takes the place of the oxygen in the normal glycoside substrate, requires that the specificity requirements of the glycosidases with

glycosyl fluorides should be the same as with glycosides. Hence, all glycosidases using protonation as the initial step should hydrolyse the respective glycosyl fluoride and this has been shown to be the case (199).

A complete understanding of the mechanism of action of a glycosidase requires elucidation of the initial product of its action. One method used to infer the nature of the initial hydrolysis product involves the determination of the rotation of the product of glycosidase action under conditions in which mutarotation is kept to a minimum. Substrates suitable for this purpose must be rapidly hydrolysed by the enzyme and have an optically inactive aglycone and some independent method of assaying the extent of hydrolysis. In this way the initial product of hydrolysis of α-D-glucosyl fluoride by Agidex, a fungal (Aspergillus) glucamylase, was shown to be β-D-glucose (200).

7.2. Experimental.

7.2.i. Syntheses.

a. The Preparation of 2,3,4,6-tetra-O-acetyl α-D-glucosyl fluoride (I).

D-glucose pentaacetate (30g) was dissolved, in portions with stirring, in anhydrous liquid HF (about 75 ml), which had been cooled in polythene vessel to -30° (solid CO_2/acetone). The vessel was sealed with a tight fitting rubber bung through which passed a short length of polythene tubing carrying a silica gel drying tube, and kept at this temperature for about half an hour. The resulting solution was then left to warm to room temperature before being poured directly into a mixture of chloroform, ice and water. The organic layer was washed several times with water and dried over $MgSO_4$. Evaporation of this solution produced a clear syrup which crystallised on standing. The product was recrystallised from ethanol and dried in vacuo. Yield 9.1g (30%).

m.p. 107-108°, $[\alpha]_D^{30}$ + 89° (c 1, $CHCl_3$).

(lit.m.p. 108°, $[\alpha]_D^{20}$ + 90.1°. (201)).

^{19}F nmr: $CHCl_3$ solution, 56.4 MHz, C_6F_6 internal reference.

Chemical shift 971.0 Hz (17.2 ppm) upfield.

J_{FH_1} 52.95 Hz.

J_{FH_2} 22.80 Hz.

Lit. ^{19}F nmr data (202): $CHCl_3$ solution, 56.4 MHz, CF_3COOH external reference.

Chemical shift 71.0 ppm

J_{FH_1} 52.8 Hz.

J_{FH_2} 23.8 Hz.

b. The Preparation of α-D-glucosyl fluoride.

Deacetylation of (I), prepared as above, was achieved using the published method (203).

Tetra-O-acetyl-α-D-glucosyl fluoride (4.5g) was suspended in absolute methanol (4.5 ml) and 0.5 ml of a freshly prepared 1% sodium methoxide solution added. The acetylated fluoro sugar dissolved within a few minutes to give a light green solution having a noticeable odour of methyl acetate. After standing at room temperature for ½ hour crystallisation of the glucosyl fluoride began and which was enhanced by rubbing with a glass rod. The crystals were filtered off, dissolved in the minimum of hot absolute methanol and anhydrous ether added until turbid. Upon cooling in ice the required compound was precipitated. Yield 1.8g, (70%).

m.p. Yield 1.8g, (70%).

(lit. m.p. 125-128° (decomp.), $[\alpha]_D^{30}$ + 95.2° (c 1, water).

I. (lit.m.p. 118-125° (decomp.), $[\alpha]_D^{18}$ + 96.7° (203)). and the complete absence of the carbonyl stretch at 1740 cm^{-1} in the I.R spectrum of (I).

^{19}F nmr: aqueous solution, 56.4 MHz, C_6F_6 external reference.

Chemical shift 963.0 Hz (17.1 ppm) upfield.

J_{FH_1} 54.35 Hz.

J_{FH_2} 26.75 Hz.

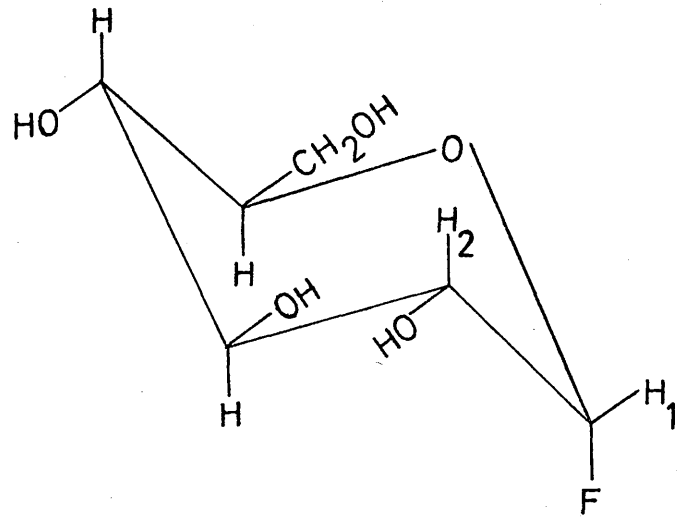
Lit. ^{19}F nmr data (204): aqueous solution, 56.4 MHz, CCl_3F external reference.

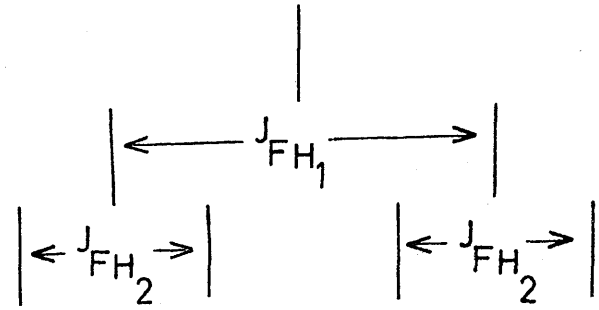
Chemical shift 149.7 ppm.

J_{FH_1} 53.0 Hz.

J_{FH_2} 26.8 Hz.

The Evaluation of F-H Coupling Constants for α-D-glucosyl fluoride.





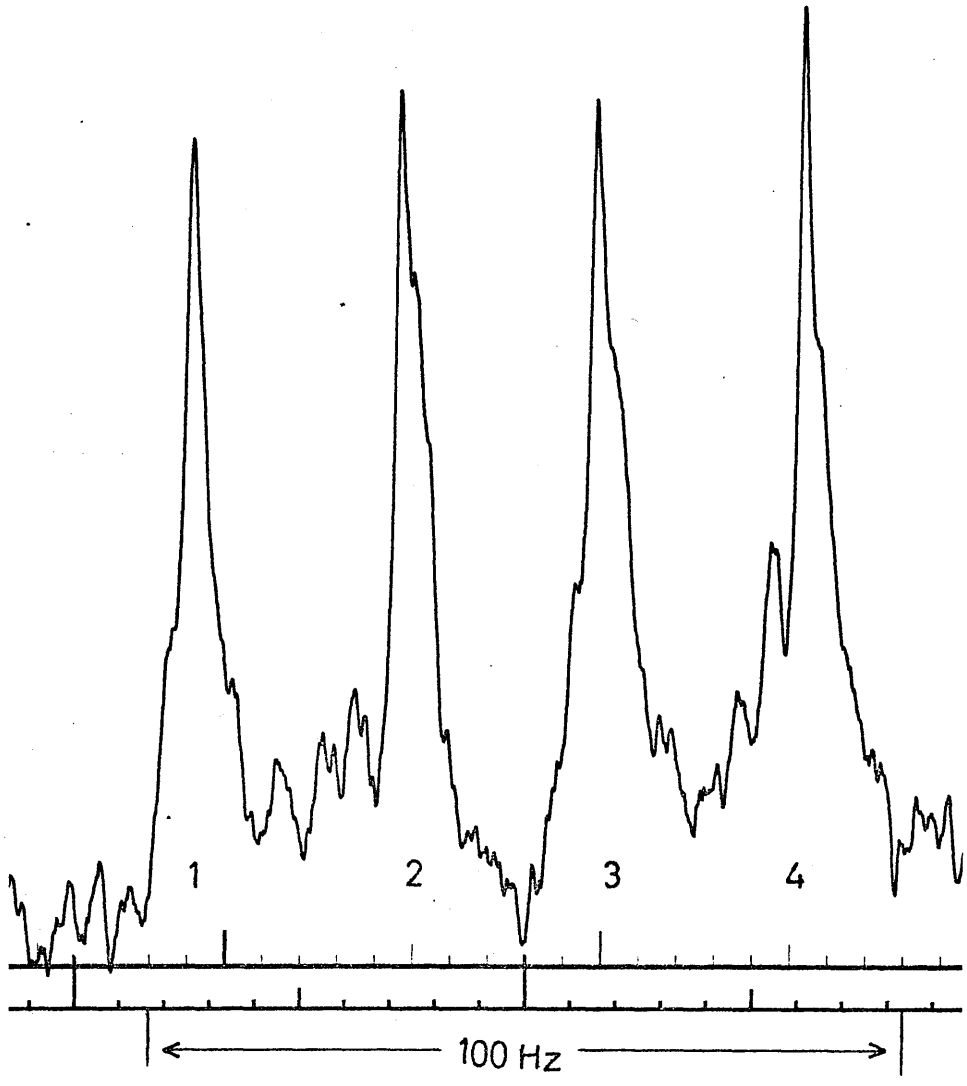
Chemical shift of outer signals, D_o: $J_{FH_1} + J_{FH_2}$

Chemical shift of inner signals, D_i: $J_{FH_1} - J_{FH_2}$

$$J_{FH_1} = \frac{D_o + D_i}{2}$$

$$J_{FH_2} = \frac{D_o - D_i}{2}$$

The ^{19}F NMR Spectrum of an aqueous solution of α-D-glucosyl fluoride, recorded at 56.4 MHz on a 500 Hz sweep width.



Chemical shifts of signals 1-4, upfield from an external reference of C_6F_6, are as follows:

1. 1003.5 Hz
2. 976.0 Hz
3. 950.0 Hz
4. 922.6 Hz

J_{FH_1} = 54.35 Hz

J_{FH_2} = 26.75 Hz.

b. The experiments described here constituted a first attempt to use ^{19}F nmr spectroscopy as a means of investigating enzyme-substrate/inhibitor interactions in solution. The enzyme in question, amyloglucosidase from Aspergillus oryzae (Koch Light Laboratories Ltd.), was used as a crude preparation, no attempt being made to purify it for the nmr studies.

Enzyme solutions were prepared by dissolving approximately 100mg of the crude material in 10 ml of the appropriate buffer and left to stand overnight at 4°. The clear supernatant was decanted from any insoluble material and made up to 25 ml with buffer solution. 0.5 ml of this enzyme solution was then added to 0.5 ml of a solution of α-D-glucosyl fluoride in the same buffer and the ^{19}F spectrum recorded at intervals.

The ^{19}F nmr spectra were recorded on a 500 Hz sweep width using a Varian Associates HA-60 spectrometer operating at 56.4 MHz and 30°. Chemical shifts were measured upfield from an external capillary of C_6F_6 by electronic counting of the difference between the sweep frequency and the manual oscillator frequency.

7.3. Results.

^{19}F nmr Data.

i. 0.1 M sodium acetate buffer, pH 4.36.

A 0.54 M solution of α-D-glucosyl fluoride had a ^{19}F nmr spectrum characterised by the following parameters;

Initial signal positions:		Positions after 60 mins:	
	998.7 Hz.		999.5 Hz.
	970.5 Hz.		970.2 Hz.
	944.3 Hz.		945.1 Hz.
	918.3 Hz.		918.2 Hz.
Chemical shift	958.5 Hz.	Chemical shift	959.5 Hz.
J_{FH_1}	53.3 Hz.	J_{FH_1}	53.3 Hz.
J_{FH_2}	27.1 Hz.	J_{FH_2}	27.1 Hz.

In the presence of the enzyme preparation the ^{19}F spectrum of a 0.8 M solution of α-D-glucosyl fluoride had the following initial parameters;

	996.6 Hz.
	968.5 Hz.
	943.1 Hz.
	915.2 Hz.
Chemical shift	955.9 Hz.
J_{FH_1}	53.5 Hz.
J_{FH_2}	27.9 Hz.

The intensity of the quartet of signals had decreased significantly 75 minutes after mixing the enzyme and substrate solutions and had disappeared altogether after leaving overnight. It was then possible to locate a ^{19}F signal 1621.2 Hz upfield from the reference corresponding to fluoride ion.

ii. 0.1 M sodium phosphate buffer, pH 7.2.

The parameters of the ^{19}F nmr spectrum of a 0.8 M solution of α-D-glucosyl fluoride were;

Initial signal positions:		Positions after 160 mins:	
	1004.3 Hz.		1004.3 Hz.
	976.8 Hz.		976.9 Hz.
	950.5 Hz.		950.9 Hz.
	923.3 Hz.		923.2 Hz.
Chemical shift	963.8 Hz.	Chemical shift	963.6 Hz.
J_{FH_1}	53.65 Hz.	J_{FH_1}	53.4 Hz.
J_{FH_2}	27.35 Hz.	J_{FH_2}	27.4 Hz.

In the presence of the enzyme preparation the ^{19}F nmr spectrum of a 0.3 M solution of α-D-glucosyl fluoride had the following initial parameters;

	994.1 Hz.
	965.0 Hz
	940.8 Hz.
	913.1 Hz.
Chemical shift	953.6 Hz.
J_{FH_1}	52.6 Hz.
J_{FH_2}	28.4 Hz.

After 80 minutes the quartet of signals had completely disappeared, a singlet being observed 1578.0 Hz from the reference, arising from the fluoride ion.

7.4. Discussion.

The disappearance of the ^{19}F resonances of α-D-glucosyl fluoride in the presence of the amyloglucosidase preparation must occur as a result of the enzyme catalysed hydrolysis of the C-F bond of the substrate since the signals do not decrease in intensity when the substrate is dissolved in buffer alone, either at pH 4.36 or pH 7.2. These results independently confirm Barnetts observation of the hydrolysis of α-D-glucosyl fluoride by an extract of rat intestinal mucosa and by the fungal amyloglucosidase, Agidex (198). Furthermore, the accumulation of fluoride ion in the system could be detected by the ^{19}F nmr method.

Use of an impure amyloglucosidase in these nmr studies makes it impossible to give anything more than a qualitative description and explanation of the results because it is not known what percentage of the total protein in the enzyme preparation was specifically active amyloglucosidase. Consequently, although chemical shift differences are noted when the substrate is in the presence of the 'enzyme' (about 3.6 Hz at pH 4.36, 10 Hz at pH 7.2), this might arise from specific binding of the α-D-glucosyl fluoride in the active site or to indiscriminate binding to inactive protein, or a combination of both. In addition to the chemical shift changes, there are also changes in both coupling constants on addition of the 'enzyme'. These are more pronounced for J_{FH_2} which, at pH 4.36, increases by 0.8 Hz and at pH 7.2 by 1.0 Hz. The change in J_{FH_1} at pH 4.36 is also 0.8 Hz but this is a decrease in the value in buffer alone while at pH 7.2 this same coupling

constant shows a slight increase of 0.2 Hz. These changes might ordinarily be indicative of an alteration in the conformation of the fluoro sugar upon binding to the enzyme. In view of the impurity of the enzyme preparation the cause of these coupling constant changes, like those for the chemical shift changes, must remain in some doubt.

It was hoped to persue this investigation further and to this end attempts were made to prepare β-D-glucosyl fluoride and 1-fluoro-maltose but attempts to obtain the free sugars from their acetylated derivatives met with no success. Furthermore, pure amyloglucosidase could not be obtained commercially and much time and effort was consumed in trying to purify the enzyme from an impure source known as Diazyme (supplied by Kingsley Chemicals Ltd., 73-76 Jermyn St., London, S.W.1.). The method employed was based on that of Pazur and Ando (194). The 100mg of material obtained after column chromatography on DEAE-cellulose hydrolysed starch with glucose as the sole product, determined by running a paper chromatogram of the hydrolysate (n-butanol:pyridine:water, 6:4:3 and spraying with ammoniacal silver nitrate solution). It had an elemental composition C 45.65, H 6.95 and N 8.99 and left no residue after ignition. Comparison of the relative abilities of Diazyme and the 'purified enzyme' to hydrolyse p-nitrophenyl-α-D-glucoside at pH 4.7 and 35° indicated that the purification procedure had yielded an enzyme material whose purity on this basis was only 7.5 fold better than that of the starting material.

The amyloglucosidase catalysed hydrolysis of α-D-glucosyl fluoride has been shown to proceed with inversion of configuration at C1.(200). The acid catalysed methanolysis of this compound was followed by observing the change in optical rotation of a known weight of the fluoro sugar in 0.01 M methane sulphonic acid at 30°. Assuming total conversion of the

α-D-glucosyl fluoride to a mixture of α and β methyl glycosides, from the final rotation value it was calculated that the anomeric mixture consisted of 87.2% β anomer and 12.8% α anomer indicating that here too inversion of the configuration at C1 was the major feature of the hydrolysis of the C-F bond.

PART 8.

POSTSCRIPT.

8.1. Nmr spectroscopy-a solution to a problem?

Whilst nmr spectroscopy is now established as a powerful physical method of structure analysis in organic chemistry, its application to structural problems in biochemistry is only just beginning to gain momentum, principally because the molecular structures involved are much more complex and therefore accompanied by greater difficulties of data acquisition and interpretation. However, with the advent of spectrometers operating at 100 MHz and 220 MHz, the last five years have seen significant advances in this field, though even with resolutions of this order the pmr spectrum of a protein is still a complex entity.

There are two general types of problem in protein chemistry that can, in principle, be solved by high resolution nmr spectroscopy. The first embraces the problems of the three dimensional structure of proteins in solution; the second is concerned with the structure of the binding sites and the mode of bonding of small molecules to these sites. The more successful investigations of proteins by nmr spectroscopy have been directed towards the solution of problems of the second type. This may be achieved by focussing attention on either the spectrum of the enzyme-ligand complex or on that of the ligand molecule only. In view of the complexity of protein spectra it is usually more profitable to adopt the latter approach and to record and interpret changes occuring in the nmr spectrum of the ligand upon binding to the protein. These can manifest themselves either as a change in the chemical shift of a particular resonance, reflecting an alteration in the magnetic environment experienced by the nucleus giving rise to this resonance, and/or a selective broadening (increase

in relaxation rate) of the resonance, attributable to decreased motional freedom of part of the ligand molecule. If chemical shift changes are observed to occur which can be shown to be due specifically to the binding of the ligand molecule in the binding site of the enzyme then it is a relatively simple task to quantify the binding phenomenon in terms of two parameters, K_D and Δ, respectively, the dissociation constant for the complex and the chemical shift of the fully bound ligand.

Thus far the technique has helped us to observe and quantify a phenomenon that might well have been known to occur on the basis of other investigations; it can also indicate that one ligand is bound more or less strongly than another, perhaps similar molecule, under identical or differing conditions. An important aspect of the technique lies in its potential to provide an independent means of verifying whether the Michaelis-Menten parameter for a substrate binding to an enzyme is, in fact, a true binding constant. For α chymotrypsin, however, this is not particularly feasible, at least for the trifluoroacetyl derivatives of the L-amino acid substrate, since the observed ^{19}F chemical shifts for these compounds are too small to allow reliable data analysis to be performed.

The problem arises when an attempt is made to describe the nature of the environment of the binding site on the basis of the magnitude and direction of the observed chemical shifts. The most pronounced spectral changes, whether they affect chemical shift or line width, would be expected to occur in lines originating from groups that are interacting directly

and so it should be possible, therefore, to infer the structure of the complex from the differential changes in the nmr spectrum. Unfortunately the situation is not. this simple for proximity is not the sole factor in determining the magnitude of the spectral changes and the magnitude and distribution of local charges, the relative orientations of the different groups and the group polarisabilities must also be taken into account. Where the method is of use is in those situations where changes can be seen to occur either in the direction or magnitude of the recorded shifts upon altering the conditions of the system, e.g pH. It is then possible to make some comment about the relative magnetic environments being experienced by the bound molecule resulting from, say, a change in the conformation of the protein molecule.

It is at this point that high resolution nmr spectroscopy and X-ray diffraction must join forces in any attempts directed towards gaining a better understanding of the physico-chemical nature of the binding site. Between them they are the only two methods that have sufficient information content to define the structure of a protein molecule and both are non-destructive. However, whereas nmr spectroscopy allows direct detection of an interaction between neighbouring groups, X-ray diffraction allows only an inference based on proximity. The latter technique can only be used for looking at a 'static' situation-of the geometry of a crystalline complex grown under one set of conditions. It must be borne in mind that the conformation of the protein molecule in such a complex may be significantly different from that of the molecule in a similar complex in solution.

Information relating to the binding of NAG and related compounds by lysozyme in solution, obtained by nmr methods, is in fair agreement with that from X-ray diffraction studies on similar, crystalline, complexes. These had, however, revealed some conformational change in the protein during binding, with the small molecules being bound in some unexpected ways. Consequently, interpretations made only on the basis of nmr binding data would not always be correct, making it essential to have some crystallographic data as well. In other words, the two techniques must not be regarded as being mutually exclusive in solving these structure problems but as complementary to one another.

8.2. Some comments on the applicability of fluorine nmr spectroscopy to a study of enzyme-inhibitor interactions.

The recognition of the potential of the fluorine nucleus as a probe for the investigation of protein structure in solution and the study of the binding of small molecules by these biopolymers will very likely result in a gradual increase in the volume of published material in this particular sector of an expanding interest in enzymes and their catalytic properties. To date, however, only a handful of papers have appeared which have been concerned with the application of ^{19}F nmr spectroscopy in this context. The two methods of approach are;

i. either incorporate fluorine into the protein molecule, that is, modification by covalent attachment of a small fluorinated probe moiety,

or ii. observe changes in the ^{19}F nmr spectrum of a specifically fluorinated ligand molecule upon binding to the protein.

Valuable though the technique might be, providing, for example, a means of observing protein conformational changes in solution, there are some practical aspects which must not be overlooked when considering systems potentially suitable for investigation by ^{19}F nmr spectroscopy.

First, it is obviously desirable that the fluorine resonance should show a change in chemical shift, though in a few instances it may be possible to make some comment about the absence of such a change on introducing the ligand or altering the conditions of the system. An important characteristic of ^{19}F chemical shifts is their greater magnitude compared with those of the proton

so that should shift changes occur they ought to be significant enough to allow easy measurement. To offset this the response for a given number of fluorine nuclei, is about 20% weaker than for the same number of hydrogen nuclei, which means that for the signal intensity to be sufficiently strong the fluoro compound must be present at higher concentrations. Theory states that for a given concentration of protein the smaller the amount of ligand present the greater will be the chemical shift of the fluorine nuclei in the complex as this shift is the weighted average of the shift of the signal in the bound and free forms, the percentage of the former being greater when the ligand:biopolymer ratio is reduced.

A balance of sorts must be found between the concentration of the fluoro compound which will give signals of suitable intensity and that for which the resultant complex concentration yields significant shift changes. In achieving this, problems of solubility are likely to arise. A suitable compromise involves using a fluorine probe possessing three equivalent fluorine atoms, e.g. the trifluoroacetyl group, when the signal intensity can be enhanced without concomittant increase in concentration.

It is preferable to avoid the use of compounds in which the fluorine signal appears as a multiplet due to coupling with adjacent protons, when accurate chemical shift measurements of one or more component resonances could present difficulties.

In view of the well documented examples of the affinity of α chymotrypsin for a variety of aromatic compounds it was proposed to extend the ^{19}F nmr investigation of the binding of the N-trifluoroacetyl derivatives of D-tryptophan and D-phenylalanine by

α chymotrypsin to include the following compounds; trifluoroacetamide, p-fluorobenzamide, p-fluorobenzyl alcohol, 5-fluoro-indole 3-acetic acid and the o, m and p isomers of fluorophenylacetic and fluorobenzoic acids. Lack of time and instrumentation difficulties precluded any serious work being accomplished with these compounds but it was apparent that, interesting though they may be as potential fluorine probes, their use in this context would be fraught with most of the difficulties outlined above, particularly line multiplicity and signal intensity.

BIBLIOGRAPHY.

1. Saunders, Wishnia and Kirkwood. J.Amer.Chem.Soc., 79, 3289 (1957).
2. Jardetsky and Jardetsky. J.Amer.Chem.Soc., 79, 5322 (1957).
3. Jardetsky and Jardetsky. J.Biol.Chem., 233, 383 (1958).
4. Takeda and Jardetsky. J.Chem.Phys., 26, 1348 (1957).
5. Bovey, Tiers and Filipovich. J.Polymer Sci., 38, 73 (1959).
6. Wishnia and Saunders. J.Amer.Chem.Soc., 84, 4235 (1962).
7. Kowalsky. J.Biol.Chem., 237, 1807 (1962).
8. Mandel. J.Biol.Chem., 240, 1586 (1965).
9. Mandel. Proc.Nat.Acad.Sci.(US)., 52, 736 (1964).
10. Kowalsky. Biochemistry, 4, 2382 (1965).
11. Bradbury and Sheraga. J.Amer.Chem.Soc., 88, 4240 (1966).
12. Kowalsky and Boyer. Fed.Proc., 19, 337 (1960).
13. Saunders and Wishnia. Ann.N.Y.Acad.Sci., 70, 870 (1958).
14. Sternlicht and Wilson. Biochemistry, 6, 2881 (1967).
15. Meadows, Markley, Cohen and Jardetsky. Proc.Nat.Acad.Sci.(US)., 58, 1307 (1967).
16. McDonald and Phillips. J.Amer.Chem.Soc., 89, 6332 (1967).
17. Cohen and Jardetsky. Proc.Nat.Acad.Sci.(US)., 60, 92 (1968).
18. Meadows, Jardetsky, Epand, Ruterjans and Sheraga. Proc.Nat.Acad.Sci.(US)., 60, 766 (1968).
19. Kurland and Davis. J.Amer.Chem.Soc., 90, 2700 (1968).
20. Wuthrich, Schulman and Peisach. Proc.Nat.Acad.Sci.(US)., 60, 373 (1968).
21. McDonald and Phillips. J.Amer.Chem.Soc., 91, 1513 (1969).
22. Ferguson and Phillips. Science, 157, 257 (1967).
23. McDonald and Phillips. Biochem.Biophys.Res.Comm., 35, 43 (1969).
24. Bak, Pedersen and Sundby. J.Biol.Chem., 242, 2637 (1967).
25. Bak, Dambmann, Nicolaison, Pedersen and Bhacca. J.Mol.Spect., 26, 78 (1968).

26. Bradbury and King. Aust.J.Chem.,22,1083 (1969).
27. Bradbury and Wilairat.
Biochem.Biophys.Res.Comm.,29,84 (1967).
28. Smith,Stein and Moore. J.Biol.Chem.,238,227 (1963).
29. Crestfield,Stein and Moore. J.Biol.Chem.,238,2413 (1963).
30. Kartha,Bello and Harker. Nature (London),213,862 (1967).
31. Wyckoff,Hardman,Allewell,Inagami,Johnson and Richards.
J.Biol.Chem.,242,3984 (1967).
32. Roberts,Meadows and Jardetsky. Biochemistry,8,2053 (1969).
33. Cohen. Nature (London),223,43 (1969).
34. Nakamura and Jardetsky.
Proc.Nat.Acad.Sci.(US).,58,2212 (1967).
35. Morlino and Martin. J.Amer.Chem.Soc.,89,3107 (1967).
36. Markley,Meadows and Jardetsky. J.Mol.Biol.,27,25 (1967).
37. Hammes and Schimmel. J.Amer.Chem.Soc.,87,4665 (1965).
38. Ross,Mathias and Rabin. Biochem.J.,85,145 (1962).
39. Meadows,Roberts and Jardetsky. J.Mol.Biol.,45,491 (1969).
40. Ruterjans,Witzel and Pongs.
Biochem.Biophys.Res.Comm.,37,247 (1969).
41. Chipman,Grisaro and Sharon. J.Biol.Chem.,242,4388 (1967).
42. Johnson and Phillips. Nature (London),206,761 (1965).
43. Jardetsky. Adv.Chem.Phys.,7,499 (1964).
44. Zimmerman and Brittin. J.Phys.Chem.,61,1328 (1957).
45. Fischer and Jardetsky. J.Amer.Chem.Soc.,87,3237 (1965).
46. Hollis,Bolen and Kellum.
Biochem.Biophys.Res.Comm.,22,135 (1966).
47. Hollis. Biochemistry,6,2080 (1967).
48. Jardetsky,Wade and Fischer.
Nature (London),197,183 (1963).
49. Schmidt,Stark and Baldeschweiler.
J.Biol.Chem.,244,1860 (1969).
50. Navon,Schulman,Wyluda and Yamane.
Proc.Nat.Acad.Sci.(US).,60,86 (1968).

51. Thomas. Biochem.Biophys.Res.Comm., 24,611 (1966).
52. Raftery,Dahlquist,Chan and Parsons. J.Biol.Chem.,243,4175 (1968).
53. Dahlquist and Raftery. Biochemistry,7,3277 (1968).
54. Dahlquist and Raftery. Biochemistry,7,3269 (1968).
55. Thomas. Biochem.Biophys.Res.Comm., 29,628 (1967).
56. Dahlquist and Raftery. Biochemistry,8,713 (1969).
57. Blake,Mair,North,Phillips and Sarma. Proc.Royal Soc.(London),B167,365.
58. Spotswood,Evans and Richards. J.Amer.Chem.Soc.,89,5052 (1967).
59. Nakano,Nakano and Higuchi. J.Phys.Chem.,71,3954 (1967).
60. Wentworth. J.Chem.Educ.,42,96 (1965).
61. Schweighardt,Mol and Li. J.Mag.Res.,2,35 (1970).
62. Sykes. J.Amer.Chem.Soc.,91,949 (1969).
63. Groves,Huck and Homer. Chem.Ind.(London),915 (1967).
64. Muller and Birkhahn. J.Phys.Chem.,71,957 (1967).
see also:
Muller and Birkhahn. J.Phys.Chem.,72,583 (1968).
Haque. J.Phys.Chem.,72,3056 (1968).
Arrington,Clouse,Doddrell,Dunlap and Cordes. J.Phys.Chem.,74,665 (1970).
65. Zeffren and Reavill. Biochem.Biophys.Res.Comm., 32,73 (1968).
66. Bittner and Gerig. J.Amer.Chem.Soc.,92,2114 (1970).
67. Bittner and Gerig. J.Amer.Chem.Soc.,92,5001 (1970).
68. Lawson and Schramm. Z.Physiol.Chem.,97,332 (1963).
69. Gerig and Reinheimer. J.Amer.Chem.Soc.,92,3146 (1970).
70. Green and Neurath in Neurath and Bailey (Eds), 'The Proteins' Vol.2,part B,p.1057. Acad.Press,N.Y. 1954.
71. Desneuelle in 'The Enzymes',Boyer,Lardy and Myrbäck (Eds) Vol.4,partA,p.93. Acad.Press,N.Y. 1960.

72. Pechere, Dixon, Maybury and Neurath. Biochem. Biopys. Acta, 31, 259 (1959).
73. Rovery, Gabeloteau, Vernejoul, Guldoni and Desneuelle. Biochem. Biophys. Acta, 32, 256 (1959).
74. Bettelheim. J. Biol. Chem., 212, 235 (1955).
75. Rovery, Fabre, Desneuelle. Biochem. Biophys. Acta, 10, 481 (1953)
76. Rovery, Fabre, Desneuelle. Biochem. Biophys. Acta, 12, 547 (1953)
77. Rovery, Poilroux, Yoshida and Desneuelle. Biochem. Biophys. Acta, 23, 608 (1957).
78. Meedom. Acta. Chem. Scand., 10, 881 (1956).
79. Meedom. Biochem. Biophys. Acta, 30, 429 (1958).
80. Meedom. Biochem. Biopys. Acta, 31, 260 (1959).
81. Hartley. Nature (London), 201, 1284 (1964).
82. Meloun et. al. Biochem. Biophys. Acta, 130, 543 (1966).
83. Hartley and Kilby. Nature (London), 166, 784 (1950).
84. Balls and Jansen. Adv. Enzymol., 13, 321 (1952).
85. Fleisher, Jandorf, Summerson and Norton. Fed. Proc., 9, 171 (1950).
86. Jansen and Balls. J. Biol. Chem., 194, 721 (1952).
87. Jansen, Curl and Balls. J. Biol. Chem., 190, 557 (1951).
88. Oosterbahn, Adrichem and Cohen. Biochem. Biophys. Acta, 63, 204 (1962).
89. Bender, Schonbaum and Zerner. J. Amer. Chem. Soc., 84, 2540 (1962).
90. Strumeyer, White and Koshland. Proc. Nat. Acad. Sci. (US)., 50, 931 (1963).
91. Weiner, White, Hoare and Koshland. J. Amer. Chem. Soc., 88, 3851 (1966).
92. Weil, James and Buchert. Arch. Biochem. Biophys., 46, 266 (1953).
93. Whitaker and Jandorf. J. Biol. Chem., 223, 751 (1956).
94. Massey and Hartley. Biochem. Biophys. Acta, 21, 361 (1956).

95. Schoellmann and Shaw. Biochem.Biophys.Res.Comm., 7,36 (1962).
96. Schoellmann and Shaw. Biochemistry,2,252 (1963).
97. Ong,Shaw and Schoellmann. J.Amer.Chem.Soc.,86,1271 (1964).
98. Ong,Shaw and Schoellmann. J.Biol.Chem.,240,694 (1965).
99. Nakagawa and Bender. J.Amer.Chem.Soc.,91,1566 (1969).
100. Stevenson and Smillie. J.Mol.Biol.,12,937 (1965).
101. Schramm. Biochem.Z.,342,139 (1965).
102. Ray,Lathon,Katsoulis and Koshland. J.Amer.Chem.Soc., 82,4743 (1960).
103. Schacter and Dixon. Biochem.Biophys.Res.Comm.,9,132(1962).
103a. Schacter,Halliday and Dixon. J.Biol.Chem.,238,3134 (1963).
103b. Schacter and Dixon. J.Biol.Chem.,239,813 (1964).
104. Lawson and Schramm. J.Amer.Chem.Soc.,84,2017 (1962).
105. Kezdy,Feder and Bender. J.Amer.Chem.Soc.,89,1009,(1967).
106. Havsteen and Hess. J.Amer.Chem.Soc.,85,791 (1963).
107. Labouesse,Oppenheimer and Hess. Biochem.Biophys.Res.Comm., 14,318 (1964).
108. Ghelis,Labouesse and Labouesse. Biochem.Biophys.Res.Comm., 29,101 (1967).
109. Weil. J.Cellular Comp.Physiol.,47,Supp.1,144 (1956).
110. Weil and Seibles. Arch.Biochem.Biophys.,54,368 (1955).
111. Wood and Balls. J.Biol.Chem.,213,297 (1955).
112. Viswanatha and Lawson. Arch.Biochem.Biophys.,93,128 (1961)
113. Labouesse,Havsteen and Hess. Proc.Nat.Acad.Sci.(US), 48,2137 (1962).
114. Havsteen,Labouesse and Hess. J.Amer.Chem.Soc.,85,796 (1963).
115. Wootton and Hess. J.Amer.Chem.Soc.,84,440 (1962).
116. Simon. Experimentia,18,150 (1962).
117. Hartley. Phil.Trans.Royal Soc.(London),257,part B,p.77, (1970) and references therein.

118. Shotton and Hartley. Nature (London), 225, 802 (1970).
119. Sorm, Holeysovsky, Mikes and Tomasek. Coll. Czech. Chem. Comm., 30, 2103 (1965).
120. Hartley, Brown, Kauffman and Smillie. Nature (London) 207, 1157 (1965).
121. Bender and Kezdy. Ann. Rev. Biochem., 34, 49 (1965).
122. Weil, James and Buchert. Arch. Biochem. Biophys., 46, 266 (1953).
123. Ray and Koshland. Brookhaven Symp. Biol., 13, 135 (1960).
124. see reference 96.
125. Oppenheimer, Labouesse and Hess. J. Biol. Chem., 241, 2720 (1966).
126. Neurath, Rupley and Dreyer. Arch. Biochem. Biophys., 65, 243 (1956).
127. Moon, Sturtevant and Hess. J. Biol. Chem., 240, 4204 (1965).
128. Havsteen and Hess. Biochem. Biophys. Res. Comm., 14, 313 (1964).
129. Himoe and Hess. Biochem. Biophys. Res. Comm., 23, 234 (1966).
130. Himoe, Parks and Hess. J. Biol. Chem., 242, 919 (1967).
131. Gutfreund and Sturtevant. Proc. Nat. Acad. Sci. (US)., 42, 719 (1956).
132. Erlanger, Castleman and Cooper. J. Amer. Chem. Soc., 85, 1872 (1963).
133. Fahrney and Gold. J. Amer. Chem. Soc., 85, 349 (1963).
134. Keizer and Bernhard. Biochemistry, 5, 4127 (1966).
135. Bender, Gibian and Whelan. Proc. Nat. Acad. Sci. (US)., 56, 833 (1966).
136. McConn, Ku, Odell, Czerlinski and Hess. Science, N.Y, 161, 274 (1968).
137. Wedler and Bender. J. Amer. Chem. Soc., 91, 3894 (1969).

138. Bender and Wedler. J.Amer.Chem.Soc.,89,3052 (1967).
139. Vaslow and Doherty. J.Amer.Chem.Soc.,75,928 (1953).
140. Weiner and Koshland. J.Biol.Chem.,240,2764 (1965).
141. Himoe,Brandt and Hess. J.Biol.Chem.,242,3963 (1967).
142. Valenzuela and Bender. Proc.Nat.Acad.Sci.(US).; 63,1214 (1969).
143. Kraut,Wright,Kellerman and Freer. Proc.Nat.Acad.Sci.(US).,58,304 (1967).
144. Smillie,Furka,Nagabhushan,Stevenson and Parkes. Nature (London),218,343 (1968).
145. Blow,Birktoft and Hartley. Nature (London),221,337 (1969).
146. Steitz,Henderson and Blow. J.Mol.Biol.,46,337 (1969).
147. Sigler,Blow,Matthews and Henderson. J.Mol.Biol.,35,143 (1968).
148. Matthews,Sigler,Henderson and Blow. Nature (London),214,652 (1967).
149. Koshland,Strumeyer and Ray. Brookhaven Symp.Biol., 15,101 (1962).
150. Walsh,Kauffman,Kumar and Neurath. Proc.Nat.Acad.Sci.(US),51,301 (1964).
151. Watson and Shotton. Phil.Trans.Royal Soc.(London)., 257,part B,p.111 (1970).
152. Neurath and Schwert. Chem.Rev.,46,69 (1950).
153. Bergmann. Harvey Lectures Ser.,31,37 (1936).
154. see Cunningham in 'Comprehensive Biochemistry',Vol.16, Elsevier Publishing Company 1965,p.128-129,references 64-67,71-74,77-80,164-166,168-173 and 175.
155. see reference 154,pp.129-134.
156. Hein and Niemann. J.Amer.Chem.Soc.,84,4495 (1962).
157. Cohen. Trans.New York Acad.Sci.,31,705 (1969), and references therein.

158. Ingles and Knowles. Biochem.J.,104,369 (1967).
159. Knowles. J.Theoret.Biol.,9,213 (1965).
160. Bender,Clement,Gunter and Kezdy.
J.Amer.Chem.Soc.,86,3697 (1964).
161. Bernhard,Lau,Noller. Biochemistry,4,1108 (1965).
162. Brot and Bender. J.Amer.Chem.Soc.,91,7187 (1969).
163. Zerner,Bond and Bender. J.Amer.Chem.Soc.,86,3674 (1964).
164. Sykes and Paravano. J.Biol.Chem.,244,3900 (1969).
165. Foster. Ph.D. Thesis,University of Leicester 1969.
166. Sykes. Biochemistry,8,1110 (1969).
167. Johnson and Bovey. J.Chem.Phys.,29,1012 (1958).
168. Buckingham,Shaefer and Schneider.
J.Chem.Phys.,32,1227 (1960).
169. Kent and Dwek. Biochem.J.,121,11P (1971).
170. Gerig. J.Amer.Chem.Soc.,84,2681 (1968).
171. Johnson and Knowles. Biochem.J.,101,56 (1966).
172. Canady and Laidler. Canad.J.Chem.,36,1289 (1958).
173. Foster and Niemann. J.Amer.Chem.Soc.,77,3365 (1955).
174. Huang and Niemann. J.Amer.Chem.Soc.,73,1541 (1951).
175. Foster,Huang and Niemann.
J.Amer.Chem.Soc.,74,105 (1952).
176. Thomas,MacAllister and Niemann.
J.Amer.Chem.Soc.,73,1548 (1951).
177. Roberts and Jardetsky. Adv.Protein Chem.,24,512 (1970).
178. Clifford and Pethica.
Trans.Faraday Soc.,60,1483 (1964).
179. Clifford. Trans.Faraday Soc.,61,1276 (1965).
180. Lawson and Flaut. J.Phys.Chem.,69,3204 (1965).
181. Hein and Niemann. J.Amer.Chem.Soc.,84,4495 (1962).
182. Emsley,Feeney and Sutcliffe. 'High Resolution Magnetic Resonance Spectroscopy',Pergamon Press Inc., New York 1965,p.22.

183. Doherty and Vaslow. J.Amer.Chem.Soc.,74,931 (1952).
184. Hymes,Robinson and Canady. J.Biol.Chem.,240,134 (1965).
185. Hymes,Cuppett and Canady. J.Biol.Chem.,244,637 (1969).
186. Shiao. Biochemistry,9,1083 (1970).
187. Wildnauer and Canady. Biochemistry,5,2885 (1966).
188. Weygand and Geiger. Chem.Ber.,89,647 (1956).
189. Fones and Lee. J.Biol.Chem.,210,227 (1954).
190. Organic Syntheses,46,1. Ed.E.J.Corey.
191. Methods in Carbohydrate Chemistry,Vol.II p.334.
Ed. Wolfrom and Whistler,Academic Press,1963.
192. Barker and Carrington. J.ChemSoc.,3588 (1953).
193. Barker and Fleetwood. J.Chem.Soc.,4857 (1957).
194. Pazur and Ando. J.Biol.Chem.,234,1966 (1959).
195. Barker,Bourne and Fleetwood. J.Chem.Soc.,4865 (1957).
196. Fleetwood and Weigel. Nature (London),196,984 (1962).
197. Goldman. J.Biol.Chem.,240,3434 (1965).
198. Barnett,Jarvis and Munday. Biochem.J.,103,699 (1967).
199. Barnett,Jarvis and Munday. Biochem.J.,105,669 (1967).
200. Barnett and Jarvis. Biochem.J.,105,9P (1967).
201. Brauns. J.Amer.Chem.Soc.,45,833 (1923).
202. Hall and Manville. Chem.Ind.(London),468 (1967).
203. Helferich,Bauerlein and Wiegand. Annalen,447,27 (1926).
204. Hall and Manville. Canad.J.Chem.,47,1 (1969).

The Sources from which some of the Figures are taken.

Figure 1.1. Reference 18.
Figure 1.2. Reference 18.
Figure 2.2. Reference 154.
Figure 2.3. Reference 81.
Figure 2.4. Blow, Biochem.J.,112 261 (1969).
Figure 2.6. Reference 118.
Figure 2.8.a. Reference 118.
Figure 2.9.a. Reference 117.
Figure 2.9.b. Reference 117.
Figure 2.10.a. McConn, Fasman and Hess, J.Biol.Chem.,39 551 (1969).
Figure 2.11. Reference 145.
Figure 2.12. Reference 145.
Figure 2.13. Reference 146.
Figure 2.14. Reference 145.
Figure 2.15. Kasserra and Laidler, Canad.J.Chem.,47 4031 (1969).
Figure 2.16. Reference 147.
Figure 2.16.a. Shotton, New Scientist, March 19,1970,p.546.
Figure 2.17.a. Reference 157.
Figure 2.17.b. Reference 157.
Figure 2.18. Reference 158.