static effects fall off more slowly with distance from the point of substitution, the more remote carbons will be primarily influenced by them and the net effect at the β carbons could well be larger than at the α carbons. In the deprotonation of aliphatic carboxylic acids, the charge effect at the carboxyl carbon, which is sp^2 hybridized, is probably rather high. Therefore, in this case, the upfield shift, which is expected from the change of the electronegativity of the oxygen by ionization, is not large enough to make the overall shift smaller than the downfield shift displayed by the α carbon.[20] Small downfield movements of the lysine carbonyl- and β-carbon resonances between pH 9.5 and 12.6 again suggest that at pH 9.5 the α-ammonium group is not deprotonated completely.

Conclusions

Carbon-13 nmr provides a convenient, nondestructive method for the determination of the amino acid sequence in di- and tripeptides. Essentially all of the necessary information can be obtained from the spectrum of the zwitterions, but the pH-induced shifts may also be helpful to facilitate or confirm assignments. The data here reported appear to have considerable value for interpretation of the spectra of complex polypeptides and proteins where the proton resonance lines may overlap too much to be very useful. However, it remains to be seen to what degree the secondary and tertiary structures of these compounds will influence the carbon-13 chemical shifts of the individual amino acids.

Magnetic Resonance Studies of Protein–Small Molecule Interactions. Binding of *N*-Trifluoroacetyl-D-(and L-)-*p*-fluorophenylalanine to α-Chymotrypsin

Kenneth L. Gammon, Stephen H. Smallcombe, and John H. Richards*

Contribution No. 4319 from the Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Pasadena, California 91109.
Received October 6, 1971

Abstract: Magnetic resonance studies of the interaction of *N*-trifluoroacetyl-D-(and L-)*p*-fluorophenylalanine with α-chymotrypsin have been used to investigate the molecular details of the enzyme–inhibitor interaction including the effect of pH (from 5.0 to 8.0). The principles of the technique are described. We conclude that the trifluoroacetyl group of the D isomer interacts with the catalytic locus (His-57, Ser-195) while that of the L isomer is directed toward Ser-214. The aromatic ring of both the D and L isomer resides in the hydrophobic pocket. The binding constant for the D isomer increases with neutralization of a group which has pK_a of 6.6 in the free-enzyme (presumably His-57). The dimerization of chymotrypsin strongly affects the quantitative results and has been explicitly included in the analysis.

In recent years magnetic resonance techniques have been widely and productively used to gain insight into the molecular details of interactions between proteins and small molecules (usually enzymes and their specific substrates or inhibitors).[1-4] Studies of the binding of inhibitors to α-chymotrypsin have been a particularly popular endeavor and information has been reported regarding the way in which the chemical shifts[2,5-10] or relaxation characteristics[7,11,12] of various nuclei of different inhibitors change when the inhibitor binds at the active site of α-chymotrypsin. The purpose of the present paper is to report the details of the technique and its application to a study of the binding of *N*-trifluoroacetyl-D-(and L-)-*p*-fluorophenylalanine. As a function of pH, we monitored the changes in chemical shifts of the two kinds of fluorine nuclei on interaction with the active site of α-chymotrypsin, as well as the enzyme–inhibitor dissociation constants. Dimerization of α-chymotrypsin[13-18] strongly perturbs the quantitative conclusions from these studies and was explicitly included in our analysis.

Principles of the Method. Exchange Rates. The method on which the work of this paper is based depends upon exchange between solution and the protein of the small molecule containing the nuclei whose magnetic resonance characteristics are being observed.

(1) E. W. Thomas, *Biochem. Biophys. Res. Commun.*, **24**, 611 (1966).
(2) T. M. Spotswood, J. M. Evans, and J. H. Richards, *J. Amer. Chem. Soc.*, **89**, 5052 (1967).
(3) F. W. Dahlquist and M. A. Raftery, *Biochemistry*, **7**, 3269 (1968).
(4) J. C. K. Roberts and O. Jardetsky, *Advan. Protein Chem.*, **24**, 447 (1970).
(5) E. Zeffren and R. E. Reavill, *Biochem. Biophys. Res. Commun.*, **32**, 73 (1968).
(6) B. D. Sykes, *ibid.*, **33**, 727 (1968).
(7) B. D. Sykes, *J. Amer. Chem. Soc.*, **91**, 949 (1969).
(8) J. T. Gerig and J. D. Reinheimer, *ibid.*, **92**, 3146 (1970).
(9) J. T. Gerig and R. A. Rimerman, *Biochem. Biophys. Res. Commun.*, **40**, 1149 (1970).
(10) H. Ashton and B. Capon, *J. Chem. Soc. D*, 513 (1971).
(11) J. T. Gerig, *J. Amer. Chem. Soc.*, **90**, 2681 (1968).
(12) E. W. Bittner and J. T. Gerig, *ibid.*, **92**, 5001 (1970).
(13) M. S. N. Rao and G. Kegeles, *ibid.*, **80**, 5724 (1958).
(14) F. J. Kezdy and M. L. Bender, *Biochemistry*, **4**, 104 (1965).
(15) L. D. Faller and R. E. LaFond, *ibid.*, **10**, 1033 (1971).
(16) K. C. Aune and S. N. Timasheff, *ibid.*, **10**, 1609 (1971).
(17) S. N. Timasheff, *Arch. Biochem. Biophys.*, **132**, 165 (1969).
(18) K. C. Aune, L. C. Goldsmith, and S. N. Timasheff, *Biochemistry*, **10**, 1617 (1971).

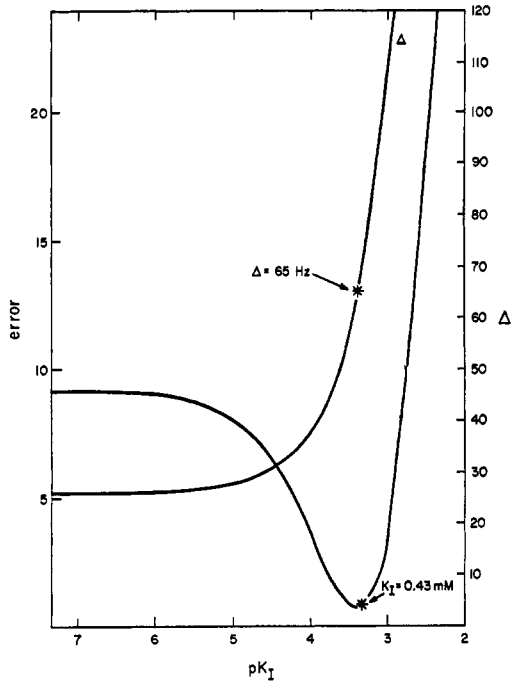

Figure 1. Dependence of "error" [see step iv in original iterative procedure] and Δ on the assumed value of pK_I. Data at pH 6.5 using $K_D = 2.0 \times 10^{-4}$ *M*.

If this exchange between the two environments is rapid, the observed spectrum represents the average of spectra for each of the environments weighted for the relative populations of molecules in each.[19] The condition of rapid exchange requires that $\tau_A < 1/\sqrt{2}\pi\Delta_{AB}$, where τ_A is the mean lifetime in the environment in which the molecule has the shorter lifetime (that is, $\tau_A < \tau_B$) and Δ_{AB} is the difference in chemical shifts between the two environments. Thus, if the difference in chemical shift between the two environments, Δ_{AB}, is 100 Hz, then $\tau_A < 2.3 \times 10^{-3}$ sec. As τ_A represents the lifetime of the protein–small molecule complex and the rate constant for protein–small molecule dissociation $k_{off} = 1/\tau_A$, the condition of fast exchange requires that $k_{off} > 4.4 \times 10^2$ sec,$^{-1}$ a condition which is met in the present case because the resonance of the inhibitor exchanging between solution and the enzyme appears as a single absorption. Other studies with a similar inhibitor (*N*-trifluoroacetyl-D-tryptophan) show a value for $k_{off} = 3.9 \times 10^3$ sec^{-1}.[20]

Analysis of Data. Simple Cases. Under conditions of rapid exchange, the observed chemical shift is an appropriately weighted average of that for the small molecule in solution and bound to the protein. If Δ represents the change in chemical shift between these environments, [EI] the concentration of small molecule–protein complex, [I] the concentration of small molecule free in solution, and $[I_0]$ the total small molecule concentration, then δ, the observed chemical shift, can be given by eq 1.

$$\delta = \frac{[EI]}{[I_0]}\Delta = \frac{[EI]}{[EI] + [I]}\Delta \qquad (1)$$

(19) J. A. Pople, W. J. Schneider, and K. J. Bernstein, "High Resolution Nuclear Magnetic Resonance," McGraw-Gill, New York, N. Y., 1959, p 221.

(20) S. H. Smallcombe, B. Ault, and J. H. Richards, *J. Amer. Chem. Soc.*, **94**, 4585 (1972).

In experimental determinations of δ as a function of varying EI and I concentrations one must exclude changes in chemical shift not related to specific interactions between protein and small molecule such as those arising from changes in the bulk magnetic susceptibility caused by addition of protein. For this reason one should use an internal standard which has no specific associations with the added protein (or if some association should occur, such association should cause no chemical shift change). All changes in chemical shift, δ, are then measured relative to the noninteracting, internal standard.

The association between small molecule and protein is defined by eq 2.

$$EI \rightleftharpoons E + I$$

$$K_I = [E][I]/[EI] \qquad (2)$$

The essence of the technique is to observe values of δ as a function of changing relative concentrations of E_0 and I_0 and by use of appropriate mathematical extrapolations to derive Δ and K_I from the experimental data.

Briefly, these extrapolations depend on eq 1 and 2 with, in the first approximation, the additional assumptions that protein and small molecule exist in only two forms, free in solution or complexed (eq 3 and 4).

$$[E_0] = [EI] + [E] \qquad (3)$$

$$[I_0] = [EI] + [I] \qquad (4)$$

Substitution of these stoichiometric relationships into eq 2 yields a quadratic, one of whose roots is eq 5. (The other root of the quadratic has no physical significance.)

$$[EI] = \{[E_0] + [I_0] + K_I - \sqrt{([E_0] + [I_0] + K_I)^2 - 4[E_0][I_0]}\}/2 \qquad (5)$$

Equation 5 allows computation of [EI] as a function of the experimentally determined quantities $[I_0]$ and $[E_0]$ and an assumed value of K_I. Equation 1 specifies a linear relationship between the values of $[EI]/[I_0]$ and the experimentally observed values of δ.

The mathematical analysis accordingly proceeds by the following steps: (i) assume a value of K_I; (ii) calculate the corresponding values for [EI] based on this assumed value of K_I and the experimental values of $[E_0]$ and $[I_0]$ by use of eq 5; (iii) obtain the best linear, least-squares fit, as required by eq 1, between these computed values of [EI] and the experimental values of $[I_0]$ and δ (plot δ *vs.* $[EI]/[I_0]$); (iv) sum the error between the straight line determined by the least-squares fit of step iii and the actual points; (v) repeat steps i–iv assuming different values of K_I. The final value of K_I is that which gives the smallest error in step iv and the slope of this line is taken as Δ. (Figure 1 shows the typical results of such a procedure and emphasizes that only if $K_I \sim [E_0]$ can both K_I and Δ be separately determined. If $K_I \ll [E_0]$, the value of K_I can only be set less than some maximum; Δ can, however, be accurately evaluated. If $K_I \gg [E_0]$, this experimental method essentially fails as too small a fraction of the small molecule is bound to protein to give rise usually to observable changes.)

Other parameters which vary in a linear manner with the relative amount of small molecule bound to protein can be analyzed in a similar fashion. Line

widths provide an example, though one must be careful in these cases to include exchange broadening as well as changes in the relaxation times for I in the EI complex in the analysis.[7,11,20]

The bovine α-chymotrypsin used in this work was of commercial origin and titrations[21] showed that about 80% of the protein was active; accordingly only 80% of the added protein was considered to contribute to $[E_0]$.

Complications Due to Dimerization or Oligomerization of Protein. Equation 3 imposes the condition that protein exists in only two forms, free in solution or as a protein–small molecule complex of 1:1 stoichiometry. A vexatious perturbation to the assumptions of eq 3 is the probable dimerization or oligomerization[13-18] of α-chymotrypsin in such a manner that the dimers or oligomers do not bind inhibitors either with the same affinity or overall stoichiometry as monomeric enzyme. We have incorporated this perturbation into our analysis by considering that α-chymotrypsin can dimerize with an effective dimer dissociation constant K_D (eq 6) and have further assumed that dimer

$$E \cdot E \rightleftharpoons E + E$$
$$K_D = [E]^2/[E \cdot E] \quad (6)$$

binds inhibitor with a sufficiently decreased affinity such that binding of inhibitor to dimer is negligible relative to binding of inhibitor by monomeric enzyme.

Recent evidence, in fact, suggests that the active site is involved in dimerization[14] which would lead to reduced binding of inhibitor to dimer. (Indeed, proflavin apparently binds only to the monomeric form of α-chymotrypsin at pH 6.2.[15]) Moreover, self-association of the enzyme increases markedly at lower pH with maximum association near pH 4.[16]

Dimerization can be included in the mathematical treatment by modification of eq 3 to account for the decrease in the free enzyme concentration caused by dimerization[22] (eq 7). Substitution of this expression

$$[E_0] = [E] + [EI] + 2[E \cdot E] \quad (7)$$

(with use of eq 6) into eq 2 allows [EI] to be calculated as a function of $[E_0]$, $[I_0]$, K_I, and K_D. This substitution produces eq 8 which can be solved for EI in terms of

$$EI^3 - EI^2[K_I + 2I_0 + E_0 - 2(K_I^2/K_D)] + EI(2E_0I_0^2 + K_II_0) - E_0I_0^2 = 0 \quad (8)$$

an assumed value of K_D by Newton's method, though this method often leads to a root which has no physical significance. A more efficient approach is to use eq 9

$$EI = [(E_0 + I_0 + K_I - 2EE) - \sqrt{(E_0 + I_0 + K_I - 2EE)^2 - 4I_0(E_0 - 2EE)}]/2 \quad (9)$$

(21) J. R. Knowles and J. M. Preston, *Biochim. Biophys. Acta*, **151**, 290 (1968).

(22) Though oligomers may be formed which contain more than two units of monomer, this analysis does not explicitly include such possibilities, but lumps together all self-association into this single, effective "dimerization" constant. Also, this treatment will eliminate as a source of possible error perturbations caused by the 20% enzymatically inactive protein present in the commercial chymotrypsin used in this work, for the concentration of this inactive protein will vary linearly with that of $[E_0]$. Many other models are possible (such as dimer binding 2 mol of inhibitor but each having a different Δ). Both for simplicity and because of our reliance on the most recent data of Faller,[15] for example, that proflavin probably binds only to monomer, we have selected a model in which oligomeric forms of chymotrypsin do not bind inhibitor competitively with monomer. We recognize the ambiguity of this assumption which essentially reflects the present imperfect understanding of perturbations caused by chymotrypsin oligomerization.

and 10 and solve them iteratively until, for a given K_I

$$EE = \frac{K_D/4 + E_0 - EI - \sqrt{K_D[(E_0 - EI)/2 + K_D/16]}}{2} \quad (10)$$

and K_D, self-consistent values of EI are obtained.

However, direct analysis of experimental data by the steps previously described with variations in both K_I and K_D does not produce a uniquely satisfactory solution because there are too many independent variables. A modified analytical procedure was therefore adopted based on two sets of experimental data. In one, values of δ are observed while $[I_0]$ is varied with $[E_0]$ being held constant; in the other, values of δ are determined while $[E_0]$ is varied and $[I_0]$ is held constant and of such magnitude that the enzyme is not in all cases saturated with inhibitor.

Analysis of the resulting data is facilitated by definition of K_{Iapp} and consideration of the following relationships.

$$K_{Iapp} = \frac{([E] + 2[EE])[I]}{[EI]} = \frac{([E_0] - [EI])([I_0] - [EI])}{[EI]} \quad (11)$$
$$= \frac{[E][I]}{[EI]} + \frac{2[EE][I]}{[EI]} = K_I + \frac{2[EE][I]}{[EI]}$$

Substitution for [EE] from eq 6 leads to

$$K_{Iapp} = K_I + 2\frac{K_I}{K_D}[E] = K_I\left(1 + 2\frac{[E]}{K_D}\right) \quad (12)$$
$$= K_I + 2\frac{K_I^2}{K_D} \cdot \frac{EI]}{[I]} \quad (13)$$

Recall, from eq 1, that

$$[EI] = \delta[I_0]/\Delta \quad (14)$$

Thus, if one knows Δ and has a set of experimental points (which most conveniently consist of values of δ as a function of changing $[E_0]$ at constant $[I_0]$), one can then calculate a value of K_{Iapp} for each experimental point. From eq 14 one can also calculate [EI]/[I] and then plot K_{Iapp} *vs.* [EI]/[I]. This plot will have K_I as the intercept on the ordinate and $2K_I^2/K_D$ as the slope, from which K_D can then be determined.

How can one extract values of I_I, K_D, and Δ from the experimental data when determination of K_I and K_D by the procedure of the previous paragraph depends on knowing Δ and determination of Δ by the procedure of varying K_I and getting the best agreement to eq 1 depends on knowing K_D? Briefly it turns out that in experiments where $[E_0]$ is varied at constant $[I_0]$, the value of K_D is relatively insensitive to the value of Δ assumed. Conversely, in experiments where $[I_0]$ is varied while $[E_0]$ is held constant, the value of Δ chosen by the computer analysis is relatively insensitive to the value of K_D.

Thus, in practice the two sets of experimental data are analyzed by the following iterative procedures. First, a reasonable range of values of K_D is obtained from a set of data at constant $[I_0]$ and varying $[E_0]$ by an analysis which determines K_D while the assumed value of Δ is varied over a wide range (for example, 50–250 Hz at pH 5). The resulting values of K_D cover a

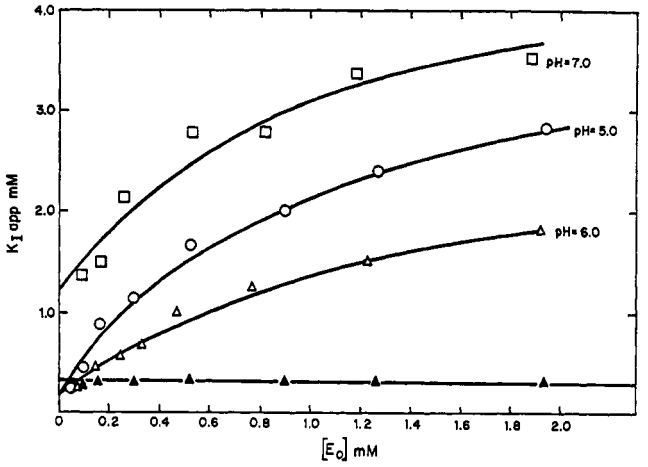

Figure 2. Dependence of K_{Iapp} on $[E_0]$ when self-association is neglected at pH 7.0 (□), 5.0 (○), and 6.0 (△). When dimerization is included, K_I does not depend on $[E_0]$ (▲ are for pH 5.0 and $K_D = 5 \times 10^{-5}$ *M*).

relatively narrow range (for example, 3×10^{-5}–2×10^{-4} *M* at pH 5).

A set of data of observed values of δ at constant $[E_0]$ and varying $[I_0]$ are now analyzed by the first procedure (eq 1, 9, and 10) by including values of K_D covering the previously determined range (3×10^{-5}–2×10^{-4} *M* at pH 5). This analysis yields values of Δ which are relatively insensitive to the value of K_D assumed (for example 64–68 Hz at pH 5). The value of K_I derived by this analysis does, however, vary widely, depending on the assumed value of K_D.

The important point is that this second analysis yields a narrow range of Δ values which can now be used to reevaluate K_{Iapp} and K_D from the data for constant $[I_0]$ and varying $[E_0]$. The values of K_D so obtained (4.8×10^{-5}–5.2×10^{-5} *M* at pH 5) are essentially constant over the allowed range of Δ (64–68 Hz at pH 5). This value of K_D can now be used in reanalysis of the data at constant $[E_0]$ and varying $[I_0]$ to determine K_I.

In practice this iterative analysis of the two sets of experimental data converged rapidly because one set of data (those at constant $[I_0]$ and varying $[E_0]$) allows a reasonably accurate determination of K_D but is insensitive to the relationship between K_I and Δ, whereas the other set of data (those at constant $[E_0]$ and varying $[I_0]$) allows a reasonably accurate determination of Δ, but is insensitive to the relationship between K_D and K_I. Thus, the two sets of data effectively complement each other and, together with the analytic procedure outlined above, quickly lead to a unique determination of Δ, K_I, and K_D which are well within the limits of accuracy of the experimental data.

Results

A solution of *N*-trifluoroacetyl-D,L-*p*-fluorophenylalanine and α-chymotrypsin exhibits two peaks for both the acetyl and aromatic fluorine nuclei. The aromatic fluorine absorption was collapsed from the multiplet characteristic of the X part of an AA'BB'X spectrum to a singlet by heteronuclear noise decoupling of the aromatic protons. For both the acetyl and aromatic resonances, one of each pair is shifted downfield and broadened relative to the other whose position and line width do not change appreciably from that of the absorption of the racemate in solution in the absence of enzyme. The amount of splitting between the resonances of the bound and unbound absorptions and broadening of the bound signal both increase as the concentration of enzyme relative to inhibitor increases.

Experiments using *N*-trifluoroacetyl-L-*p*-fluorophenylalanine confirm, as in previous studies with derivatives of *p*-fluorophenylalanine[2] and tryptophan,[2,3] that the D isomer exhibits the larger downfield shift. Addition of pure L enantiomer to the racemic mixture in the presence of enzyme decreases the amount of splitting observed and allows determination[2,23] of K_I for the L isomer to be about 3×10^{-3} *M* at pH 6.0 which exceeds K_I for the D isomer (about 3×10^{-4} at pH 6.0) by about 10. At other pH values $K_{I(L)}$ is also about ten times $K_{I(D)}$. Thus, the L isomer binds sufficiently less strongly than the D isomer that the L isomer serves as an effective internal standard and values of δ used in the analytic procedures described previously can, in the present case, be taken as the splittings between the two absorptions of the two enantiomers.

Data were collected under essentially two sets of experimental conditions: (i) constant $[E_0]$ and varying $[I_0]$ and (ii) constant $[I_0]$ and varying $[E_0]$. In case i the enzyme concentration $[E_0]$ was 1.9 m*M* (corrected for inactive enzyme) and $[I_0]_D$ (the concentration of the D isomer) was varied from 0.5 to 20 mM. A total of eight points were determined at each half pH unit between pH 5 and 8. From pH 5 to 7, 0.1 M citrate buffer was used; from pH 7 to 8, 0.1 M Tris buffer was used. Experiments with both buffers at pH 7 showed no changes dependent on the nature of the buffer.

In case ii binding curves were determined at pH 5.0, 6.0, and 7.0 in which $[I_0]_D$ was held constant ~1.5 m*M* and $[E_0]$ was varied from 1.9×10^{-3} to 5×10^{-5} *M*. Eight points were again determined at each pH.

In case i splittings of both acetyl and aromatic fluorines were recorded, though for the aromatic fluorines only about four (those at higher $[I_0]_D$) were useful because of the lower intensity of the aromatic fluorine signal relative to that of the acetyl signal and because the aromatic fluorine broadens considerably more than the acetyl fluorines. In case ii only the splittings of the acetyl fluorines were recorded.

The splittings of the acetyl fluorines as a function of $[E_0]$ and $[I_0]_D$ from case i and ii were analyzed as described previously.

Figure 2 demonstrates the importance of dimerization at pH 5.0, 6.0, and 7.0. At all these pH's, $\Delta \sim -65$ Hz. At pH 5.0 for example, if one uses this value of Δ, K_{Iapp} is found to vary by an order of magnitude (from about 0.3 to 3 m*M*) as $[E_0]$ varies from 0.05 to 1.9 m*M* when the possibility of enzyme dimerization is neglected (open circles of Figure 2). When the possibility of dimerization is included in the analysis, K_{Iapp} does not depend on $[E_0]$ and has a constant value of 0.35 m*M* (open triangles of Figure 2). The enzyme dimer dissociation constant (K_D) from this analysis is found to be 5×10^{-5} *M* at pH 5.0. Incidentally, the amount of monomeric protein thus calculated to be present agrees with that which would be present if one used

(23) S. Smallcombe, K. L. Gammon, and J. H. Richards, *J. Amer. Chem. Soc.*, **94**, 4581 (1972).

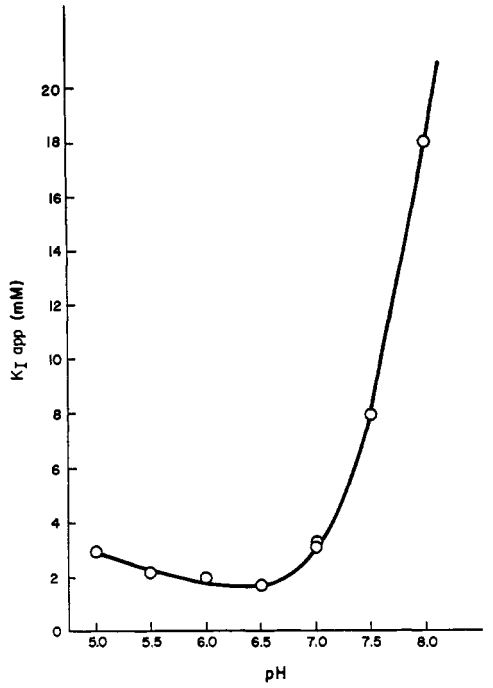

Figure 3. Variation in $K_{I_{app}}$ (dimerization neglected) as a function of pH.

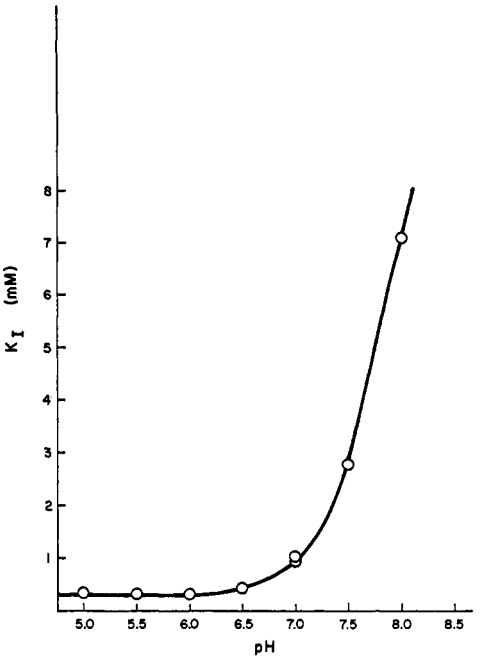

Figure 4. Variation in K_I (dimerization included) as a function of pH.

the values for the dimerization constant, $K_D = 9.6 \times 10^{-4}$ M, and for the trimerization constant (for $E_3 \rightleftharpoons E \cdot E + E$), $K_T = 2.9 \times 10^{-4}$ M, previously reported at pH 6.2.[13,15]

The manner in which the original iterative procedure (eq 1, 9, and 10) yields best values for K_I and the dependence of Δ on K_I are shown in Figure 1 for data for the shifts for the acetyl fluorines of *N*-trifluoroacetyl-D-*p*-fluorophenylalanine at pH 6.5 using $K_D = 2.0 \times 10^{-4}$ M.

Table I collects case ii results at three pH values (5.0, 6.0, and 7.0) and shows how the values of K_I and K_D vary with pH. These data are based on shifts of the acetyl fluorines.

Table I. Enzyme–Inhibitor and Enzyme–Dimer Dissociation Constants from Analysis of Type ii Experiments Using α-Chymotrypsin and *N*-Trifluoroacetyl-D-*p*-fluorophenylalanine[a]

pH	K_I	K_D
5.0	3.6×10^{-4}	5×10^{-5}
6.0	3.2×10^{-4}	9×10^{-5}
7.0	1.2×10^{-3}	4×10^{-4}

[a] The inhibitor concentration was ~1.5 mM and the enzyme concentration was varied from 0.05 to 1.94 mM.

Table II collects case i results between pH 5.0 and 8.0 and shows how $K_{I_{app}}$ (neglecting enzyme dimerization) and K_I (including enzyme dimerization) vary with pH. Figures 3 and 4 display these results graphically. The slight discrepancy between the values of K_I at a given pH from Table I and Table II arises from the two experimental approaches and their related analytical procedures. Table I comes from case ii data (constant $[I_0]_D$, varying $[E_0]$) and analysis by eq 12, 13, and 14. Table II comes from case i data (constant $[E_0]$, varying $[I_0]$) and analysis by eq 1, 9, and 10.

Table II. Enzyme–Inhibitor Dissociation Constants for the Chymotrypsin and *N*-Trifluoroacetyl-D-*p*-fluorophenylalanine[a]

pH	Buffer (concn, 0.1 M)	$K_{I_{app}}$,[b] mM	K_I,[c] mM	pK_I
5.0	Citrate	2.94	0.35	3.46
5.5	Citrate	2.19	0.31	3.50
6.0	Citrate	0.93	0.33	3.49
6.5	Citrate	1.64	0.43	3.37
7.0	Citrate	3.03	0.91	3.04
7.0	Tris	3.21	0.96	3.02
7.5	Tris	7.69	2.79	2.55
8.0	Tris	18.0	7.08	2.15

[a] The enzyme concentration was 1.94 mM and the inhibitor concentration varied between 0.5 and 20 mM. [b] Calculated assuming all active enzyme is available for binding. [c] Calculated assuming only enzyme monomer binds inhibitors.

In all these cases and at all pH's studied, the acetyl fluorines of the D isomer experience a downfield shift of 65 Hz (0.69 ppm), when bound to the enzyme. The shifts of the aromatic fluorine for the D isomer were evaluated using those values of K_I and K_D previously determined by analysis of the data for the acetyl fluorines of the D isomer and found to be -115 ± 15 Hz (1.22 ppm) at all pH's. [Analysis of case i data for the aromatic fluorines, without using the K_I value from the data for the acetyl fluorines, gave essentially the same results, though the errors are greatly increased because only some of the experimental points could be used.]

The shifts (Δ) experienced by the aromatic and acetyl fluorines on binding of *N*-trifluoroacetyl-L-*p*-fluorophenylalanine were determined at pH 6.0 using pure L isomer and measuring shifts from the deuterium in D_2O which was used as lock. The data were analyzed using a K_I of 3×10^{-3} M and lead to the following results for the L isomer: for the acetyl fluorines Δ is less than -5 Hz, for the aromatic fluorine $\Delta \sim -120$ Hz, essentially the same as the value for the D isomer.

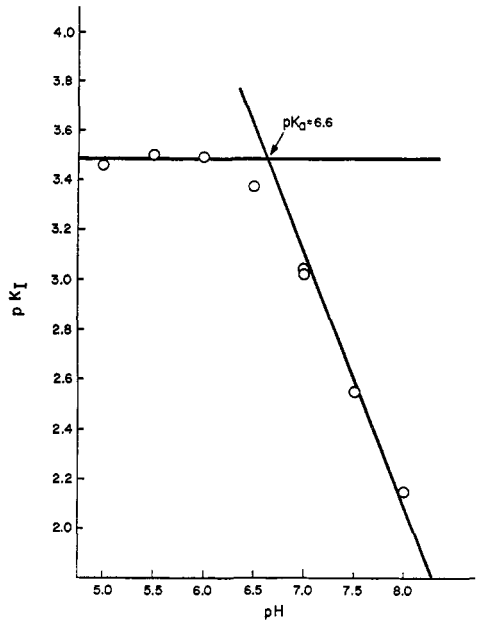

Figure 5. pK_I *vs.* pH. The intersection of the two linear segments indicates a pK_a for the free enzyme of 6.6.

Discussion

The foregoing results can be summarized. (i) The chemical shift of the acetyl fluorines of *N*-trifluoroacetyl-D-*p*-fluorophenylalanine on binding to α-chymotrypsin moves downfield relative to its position in solution by 65 Hz ($\Delta = -65$ Hz). (ii) In contrast, the acetyl fluorines of the L isomer are shifted on binding to enzyme by less than -5 Hz. (iii) The aromatic fluorine of *both* D and L isomers experiences a downfield shift of about 115 Hz on binding. (iv) None of the shifts in i–iii show any observable change with pH. (v) Self-association of protein plays a strongly perturbing role in this system and increases at more acidic pH. (vi) After correction for dimerization, K_I is independent of pH in the acidic range but as the pH is raised, K_I increases in a manner consistent with ionization of a group on the free enzyme with $pK_a = 6.6$ (Figure 5).

Stereochemistry of Binding. The experimental facts summarized in i–iii suggest that binding of these inhibitors occurs in such a way that the trifluoroacetyl groups of the two antipodes occupy different loci on the enzyme, whereas the *p*-fluorophenyl rings of both antipodes occupy the same locus, presumably the hydrophobic pocket.[24]

Thus, the *N*-TFA groups of the D and L isomers experience quite different environments when these two antipodes bind to chymotrypsin. The most likely explanation is that the carboxylate and *N*-TFA groups exchange positions. In the case of the L isomer, inhibitor is bound to the enzyme in the normal way, with C-1 (the carboxylate carbon) in the catalytic locus[25] near Ser-195 and His-57 and the *N*-TFA group in the direction of Ser-214.[24] (Our evidence does not, however, indicate significant interaction between the protein and this *N*-TFA group). When the D isomer binds, these two groups interchange binding loci; now the *N*-TFA group occupies the catalytic locus and the carboxylate carbon is directed toward Ser-214. These interactions are outlined schematically in Figure 6.

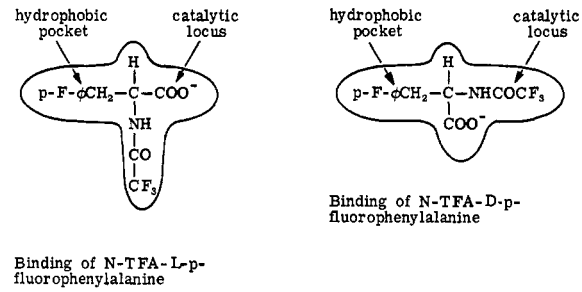

Figure 6.

That the difference in binding loci for the substituents attached to C-2 should involve interchange of C-1 and the *N*-TFA group seems reasonable on the following two grounds. The strongest binding arises from interaction of an aromatic (or, somewhat less favorably, a large hydrophobic) side chain with the hydrophobic pocket; indeed, this interaction serves as the basis for the specificity of chymotrypsin.[26] Steric constraints of the active site, especially interference between the α hydrogen of the inhibitor and γ-methylene group of methionine-192,[24] prevent acceptance of groups significantly larger than hydrogen in the locus normally occupied by the C-1 hydrogen. Thus the aromatic ring and hydrogen substituents of the α carbon have essentially fixed binding loci. The only way therefore to accommodate D and L antipodes is by interchanging the binding loci of the other two substituents of the α carbon, the *N*-TFA and the carboxylate groups.

Origin of Chemical Shift Changes in Binding. The major effect which influences chemical shift parameters for ^{19}F nuclei in different environments is the paramagnetic term in Ramsey's equation.[27] Saika and Slichter[28] have developed a quantitative theory for ^{19}F chemical shifts and these shifts have been related[29] to changes in ionic character, bond order, and degree of hybridization of the bond involving the fluorine atom in question. In general, an increase in electronegativity of the atom attached to fluorine decreases the ionic character of the fluorine, which thereby increases the paramagnetic contribution to shielding, *i.e.*, causes a downfield shift.

However, in our work the observed changes in chemical shift more likely result from changes in external environment (solution or active site of the enzyme) than they do from changes in the hybridization of the carbon–fluorine bonds. In this regard Evans[30] has studied the effects of solvent on the shielding of ^{19}F nuclei and suggested that an increase in the paramagnetic shielding term caused by the presence of polarized solvent molecules can, in some cases, contribute importantly to the observed chemical shift. More recently, Emsley and Phillips[31] have quantitatively studied the effect of the external environment on the chemical shift of fluorine. They have related these changes to external factors such as diamagnetic susceptibility, van der Waals interactions, and electric field effects. At-

(24) T. A. Seitz, R. Henderson, and D. M. Blow, *J. Mol. Biol.*, **46**, 337 (1969).
(25) J. Hein and C. Niemann, *J. Amer. Chem. Soc.*, **84**, 4495 (1962).
(26) J. R. Knowles, *J. Theor. Biol.*, **9**, 213 (1965).
(27) Reference 19, p 168.
(28) A. Saika and C. P. Slichter, *J. Chem. Phys.*, **22**, 26 (1954).
(29) M. Karplus and T. P. Das, *ibid.*, **34**, 1683 (1962).
(30) D. F. Evans, *J. Chem. Soc.*, 877 (1960).
(31) J. W. Emsley and L. Phillips, *Mol. Phys.*, **11**, 437 (1966).

tribution of an observed change in chemical shift in a quantitative manner to these various possible perturbations is not generally possible. Zeffren[32] studied the chemical shift of the ^{19}F nuclei of *N*-trifluoroacetyl-D-phenylalanine and found that increasing salt concentration causes downfield shifts; moreover the efficacy of a given salt concentration in causing downfield shifts is markedly higher in solvents of low dielectric constant than in water. On the other hand, changing to a more hydrophobic solvent (one of lower dielectric constant) than water (without added salt) resulted in upfield shifts. Simple acceptance of these results as having a direct bearing on environmental effects on ^{19}F chemical shifts is complicated by the failure to separate from the observed shifts those that arise just from the changes in bulk diamagnetic susceptibility; such changes have, of course, been eliminated in the present work by use of suitable internal standards.

At the catalytic locus of the active site of chymotrypsin the *N*-TFA group of *N*-trifluoroacetyl-D-*p*-fluorophenylalanine should experience some of the positive charge of the protonated His-57 residue which could give rise to the observed downfield shift on binding. This charge can be transmitted to the fluorines either electrostatically or by hydrogen bonding between the imidazole of His-57 and the carbonyl oxygen or, less likely, the amide nitrogen of the *N*-TFA group. Another possible origin of the downfield shift is the electric field created between the positive imidazolium ring of His-57 and the carboxylate anion of the inhibitor molecule. When the L isomer binds, its *N*-TFA group, not being in the catalytic locus, will not experience the positive charge of His-57 and its attendant effects; accordingly, change in the chemical shift is observed.

The chemical shift of the *N*-TFA group of the D isomer does not change even when the pH of the solution has been raised to 8, as the inhibitor binds essentially only to the protonated form of the enzyme.

The downfield shift seen on binding for the aromatic fluorine of both D and L inhibitors cannot be simply the result of their entering an environment which is more hydrophobic than solution, for this causes an upfield shift. There are at least three possible origins. The shift could arise from formation of a hydrogen bond between the aromatic fluorine and Ser-189 which lies deep in the hydrophobic pocket.[24] Or, the downfield shift could be a response to the salt bridge between Asp-194 and Ile-16 which lies on the other side (toward the interior of the protein) of the polypeptide chain that forms the hydrophobic pocket. Or, the inhibitor experiences increasingly constrictive van der Waals interactions in the hydrophobic pocket (enzymatic bear hug).

Dependence of K_I and K_D on pH. The value of K_I for the D isomer increases markedly with titration of a group which, in the free enzyme, has a pK_a of 6.6[33] (see Figure 5 where pK_I is plotted against pH). If the possibility of self-association is not allowed, the apparent pK_a is raised to about 6.9.[34] As we have been unable to observe experimentally any binding to the deprotonated form of the enzyme, we cannot obtain K_I for this species or pK_a for the EI complex.

A considerable body of evidence suggests that the group with a pK_a between 6.5 and 7.0 on the free enzyme is His-57.[35,36] There is only one other histidine in chymotrypsin (His-40) and though its pK_a is not known with certainty, it is generally assumed to be lower than that of His-57; moreover no clear chemical evidence for major perturbations, either chemical or physical, caused by neutralization of His-40 have been reported.

The evidence of this paper does not require that it be His-57 which is titrated with $pK_a = 6.6$, but this possibility provides a ready rationalization for the results, for, when His-57 is neutralized, the catalytic locus would take on a negative character[34] created by the carboxylate anion of Asp-102 which lies adjacent to His-57.[24] This negative character of the active site would give rise to electrostatic repulsion of the anionic inhibitor. In contrast, neutral inhibitors, such as *N*-acetyl-D-tryptophanamide[37] and *N*-trifluoroacetyl-D-tryptophanamide,[23] do not show significant dependence of K_I on a group in the free enzyme with $pK_a = 6.5$–7.0. Although neutralization of His-40 could, by changing the overall charge of the active site, also lead to increased values of K_I for carboxylate anion inhibitors, the remoteness of His-40 relative to the proximity of His-57 should sharply weaken this effect.

The magnitude of K_D shows a dependence on pH in these studies analogous to that observed by other methods,[16,17] becoming more significant at more acidic pH (for example, K_D decreases from 4×10^{-4} *M* at pH 7.0 to 5×10^{-5} *M* at pH 5.0; see Table I).

A somewhat unusual aspect of the values of K_I found in this work is the difference between that for the D and L isomers; *N*-trifluoroacetyl-D-*p*-fluorophenylalanine is bound about ten times more strongly than the L isomer, while, in general, D and L isomers bind with roughly similar affinities.[26] At this time we have no completely convincing molecular explanation for this observation.

Conclusion

Binding of *N*-trifluoroacetyl-D-(and L)-*p*-fluorophenylalanine to the active site of chymotrypsin involves insertion of the aromatic ring of both the D and L inhibitor into the hydrophobic pocket where the *p*-fluorine substituent experiences a downfield shift caused by hydrogen bond formation to Ser-184 or interaction with the Asp-194–Ile-16 salt bridge. The *N*-TFA group of the D isomer resides near the catalytic locus; its fluorines experience a downfield shift caused probably by interaction, either electrostatic or hydrogen bonding, with the protonated His-57–Ser-195 system. The TFA group of the L isomer is located away from the catalytic locus, probably being directed toward Ser-214; its fluorines experience no significant change of chemical shift on binding.

The variation in K_I from these experiments can most conveniently be ascribed to neutralization of His-57 which we determine to have $pK = 6.6$ on the free enzyme.

Dimerization (or more extensive self-association) of α-chymotrypsin is an important perturbation which has

(32) E. Zeffren, *Arch. Biochem. Biophys.*, **137**, 291 (1970).
(33) M. Dixon and E. C. Webb, "Enzymes," Academic Press, New York, N. Y., 1964, p 116, *et seq.*
(34) C. H. Johnson and J. R. Knowles, *Biochem. J.*, **101**, 56 (1966).
(35) G. P. Hess, J. McCoon, E. Ku, and G. McConkey, *Phil. Trans. Roy. Soc. London, Ser. B*, **257**, 80 (1970).
(36) M. L. Bender, G. E. Clement, F. J. Kezdy, and H. d'A Meck, *J. Amer. Chem. Soc.*, **86**, 3680 (1964).
(37) C. J. Johnson and J. R. Knowles, *Biochem. J.*, **103**, 428 (1967).

been included in the analysis and becomes increasingly important at acidic pH.

Experimental Section

N-Trifluoroacetyl-D,L-p-fluorophenylalanine. D,L-p-Fluorophenylalanine (10 g, 0.036 mol) was dissolved in anhydrous trifluoroacetic acid (60 ml). The solution was cooled to 0° in an ice–salt bath and trifluoroacetic anhydride (18 ml) was added dropwise with stirring over a 10 min period. After 2 hr at 0°, trifluoroacetic acid and anhydride were removed by distillation under vacuum (~1 mm) at 25°. The solid residue was dissolved in ether and the solution filtered and dried over anhydrous sodium sulfate. After reduction of the solution to one-third the original volume, the product crystallized. Recrystallization from *n*-hexane–ether yielded 5.2 g (50%) of white crystalline solid, mp 143–144°. *Anal.* Calcd for $C_{11}H_9O_3NF_4$: C, 47.32; H, 3.25; N, 5.02. Found: C, 47.17; H, 3.02; N, 4.63.

D,L-p-Fluorophenylalanine Methyl Ester Hydrochloride. To a solution of thionyl chloride (10 ml) in absolute methanol (45 ml) in a Dry Ice–acetone bath was added D,L-p-fluorophenylalanine (10.0 g). The mixture was refluxed for 2 hr and solvent and excess reagent were removed on a rotary evaporator (20–30° (20 mm)). The residue was dissolved in a minimum amount of absolute methanol and the solution filtered. Addition of anhydrous ether caused precipitation of white crystals which were collected by filtration and dried over sodium hydroxide, yield 13.1 g (96%).

N-Trifluoroacetyl-D,L-p-fluorophenylalanine Methyl Ester. D,L-p-Fluorophenylalanine methyl ester hydrochloride (13.1 g) was placed in a separatory funnel with saturated sodium bicarbonate (400 ml) and cooled with ice, and ethyl acetate (400 ml) was added. The mixture was shaken vigorously and the phases were separated. The aqueous phase was extracted a second time with ethyl acetate (200 ml). The two ethyl acetate fractions were combined, dried over anhydrous sodium sulfate, filtered, and evaporated to dryness on a rotary evaporator leaving a gummy residue; crystallization was not attempted. The gummy residue was dissolved in anhydrous ethyl acetate (300 ml) and fresh trifluoroacetic anhydride (10 ml) was added dropwise with stirring. The solution was allowed to stand at room temperature for 30 min with occasional stirring. Additional trifluoroacetic anhydride (8 ml) was then added with stirring and the solution allowed to stand for another 30 min. The ethyl acetate solution was extracted with 10% aqueous sodium bicarbonate (300 ml), dried over anhydrous sodium sulfate, and filtered, and the ethyl acetate evaporated. The resulting gummy residue was dissolved in a minimum amount of ethyl ether and this solution added to five volumes of *n*-hexane. The solution was evaporated on a rotary evaporator to give a white powder (12.2 g, 76%, mp 59–61°).

N-Trifluoroacetyl-L-p-fluorophenylalanine. N-Trifluoroacetyl-D,L-p-fluorophenylalanine methyl ester (12.2 g) was dissolved in methanol (80 ml). Aliquots (10 ml) of this solution were added to a solution of α-chymotrypsin (250 mg) in water (750 ml). The pH was maintained at 7.5 by addition of 0.203 *N* NaOH and all the ester was added over a period of 1 hr. The pH remained essentially constant after 98.5 ml of base had been added (corresponding to hydrolysis of 96% of the original L isomer). The aqueous solution was extracted twice with ethyl acetate (400 ml) and lyophilized to dryness. The residue was dissolved in a minimum amount of water and precipitated by addition of concentrated hydrochloric acid to pH 2. The white solid was filtered with suction and washed on the filter with cold 0.01 *N* HCl. Then it was dissolved in a minimum amount of methanol, filtered, and precipitated by slow addition of 0.01 *N* HCl. The fine needle-shaped crystals were filtered, washed with cold 0.01 *N* HCl, and dried in a vacuum desiccator with NaOH pellets overnight; yield 2.08 g, mp 133–134°. *Anal.* Calcd for $C_{11}H_9O_3NF_4$: C, 47.32; H, 3.25; N, 5.02. Found: C, 47.14; H, 3.05; N, 4.74.

p-(N,N,N-Trimethylammonium)cinnamic Acid Iodide p-Nitrophenyl Ester. The titrant was prepared essentially by the method of Knowles and Preston[21] with several modifications suggested by Parker.[38] p-(N,N,N-Trimethylammonium)cinnamic acid iodide (3.3 g) was mixed with chloroform (40 ml) and triethylamine (1.0 g, 1.38 ml). Ethyl chloroformate (1.1 g, 0.97 ml) was added and the mixture was stirred for 15 min. A second portion of chloroform (15 ml) was added to the stirred suspension. Some solid material still remained. p-Nitrophenol (1.4 g) was then added and the mixture refluxed for 2 min. The resulting yellow solution was cooled and allowed to stand at room temperature overnight. A precipitate formed which was removed by filtration. The filtrate was evaporated to dryness, the residue dissolved in hot dimethylformamide (30 ml), and the solution allowed to cool. Ether (6 ml) was added and, after standing 1 hr, the solution was filtered. Additional ether (30 ml) was added and the solution allowed to stand at 0° overnight yielding an orange crystalline material. This was collected by filtration and then redissolved in warm dimethylformamide (60 ml) and ether (30 ml) was added. After several hours, no solid having separated, additional ether (150 ml) was added. Some solid separated slowly and the mixture was allowed to stand overnight during which time more solid crystallized. The resulting yellow solid (about 0.7 g) was collected by filtration, dried, and redissolved in dimethylformamide (50 ml), and ether (30 ml) was added. A white precipitate formed which was removed by filtration. The resulting filtrate was allowed to stand overnight at 0° during which time product in the form of fine canary yellow needles separated. This was collected by filtration, dried in a stream of dry nitrogen, and stored in the freezer. Titrant prepared in this manner is of high purity (~97%) and may be stored for many months without detectable decomposition. A 5 m*M* solution of titrant in acetone was used for assays of enzymatic activity.

α-Chymotrypsin. Three times recrystalilzed, salt-free α-chymotrypsin was obtained from Sigma Chemical Co. and stored in a freezer when not in use. Its activity, as determined below, was ~80%.

Enzymatic Activity. The activity of the α-chymotrypsin was determined by the method of Knowles and Preston[21] using, as titrant, p-(N,N,N-trimethylammonium)cinnamic acid iodide p-nitrophenyl ester. Absorbances were measured on a Cary 14 spectrometer at 400 nm, and the final pH measured on a Radiometer pH meter. Calculations assumed $pK_a = 7.04$ for p-nitrophenol and ϵ 18,320 for p-nitrophenolate anion. The enzymatic activity was determined before samples were placed in the probe and after observation. The activities agreed within experimental error and were consistently 80 ± 2%.

Nmr Samples. Samples for nmr measurements were prepared in 0.1 *M* citrate buffer for pH 5.0–7.0 and 0.1 *M* Tris buffer for pH 7.0–8.0. Appropriate amounts of α-chymotrypsin and inhibitor were accurately weighed into a 5-ml graduated centrifuge tube. Buffer (2 ml) was added and the mixture was stirred and centrifuged to dissolve solid. Appropriate volumes of buffer at various pH's were then added to give a final volume of 4 ml and a pH within 0.05 of the reported values. After a final centrifugation to remove any undissolved protein (only minimal amounts of material do not dissolve), the solutions were filtered into 12-mm nmr tubes. The tubes were equilibrated in a constant temperature bath at 32° (the probe temperature) for about 8 min before being transferred to the nmr probe. The pH of representative samples was checked after observation of the nmr spectrum; it had never changed significantly.

Determination of Spectra. Spectra were recorded on a Varian HA-100-15 (or, in a few cases of control experiments, XL-100-15) spectrometer. Frequency sweep was provided by a Fabri-tek 1062 time averaging computer driving a very stable voltage-controlled oscillator. The number of scans accumulated varied between 2 and 32, depending on the concentration of inhibitor. The spectrometer (HA-100-15) was field-frequency locked on a small capillary of hexafluoroacetone held concentrically by a Teflon vortex plug inside the 12-mm tubes.

Acknowledgment. This work was supported by a grant from the U. S. Public Health Service (GM-16424). One of us (K. L. G.) is also grateful for the support of an NSF Fellowship. The titrant was synthesized by R. B. Moon, who also carried out some early magnetic resonance studies on analogous systems. M. Hunkapiller and M. McMillan also contributed valuable synthetic assistance.

(38) L. Parker, personal communication.