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

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

Contribution No. 4342 from the Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Pasadena, California 91109.
Received October 6, 1971

Abstract: A magnetic resonance technique has been developed for studying the competitive binding to proteins of two small molecules; the nmr spectrum of only one needs to be observed. This technique has been applied to study the competition between *N*-trifluoroacetyl-D-tryptophan and the L enantiomer for the active site of α-chymotrypsin from pH 5.0 to 8.0. The chemical shift for the fluorine nuclei of *N*-trifluoroacetyl-D-tryptophan bound to the enzyme is found to be the same as that for *N*-trifluoroacetyl-D-*p*-fluorophenylalanine. The binding of both D- and L-tryptophan derivatives shows a marked dependence on deprotonation of a group on the free enzyme with $pK_a = 6.6$ (presumably His-57).

The purpose of this paper is to discuss a method for studying interactions between proteins and small molecules in which the small molecule experiences no change in magnetic resonance parameters when it binds to the enzyme. This method involves competition for the active site by two species, one of which (the "sensitive" inhibitor) does show changes in chemical shift when bound This is an extension of a previously reported method[1] for studying protein–small molecule interactions. Essentially the technique involves monitoring the decrease in nmr changes observed for the "sensitive" inhibitor in the presence of enzyme as one varies the relative concentration of a second inhibitor. (This technique has been briefly and approximately described previously.[2])

We studied the binding to α-chymotrypsin of *N*-trifluoroacetyl-D-tryptophan, a "sensitive" inhibitor which does show changes in nmr parameters on binding, and *N*-trifluoroacetyl-L-tryptophan, an inhibitor which, though showing only barely perceptible nmr changes in the presence of chymotrypsin, does bind at the active site competitively with the D isomer. The binding of these two antipodes was studied over the pH range 5.0–8.0.

Principles of the Method. The main change in the earlier analytic procedures[1] involves the introduction of a new species, a complex between enzyme and competitive inhibitor, EI′. Thus

$$[E] = [E_0] - 2[EE] - [EI] - [EI'] \quad (1)$$

$$[I] = [I_0] - [EI] \quad (2)$$

$$[I'] = [I'_0] - [EI] \quad (3)$$

where

$$K_I = [E][I]/[EI] \quad (4)$$

$$K_I' = [E][I']/[EI'] \quad (5)$$

and

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

Substitution of eq 1–3 into eq 4–6 leads to

$$[EI] = \{E_0 + I_0 + K_I - EI' - 2EE - [(E_0 + I_0 + K_I - EI' - 2EE)^2 - 4I_0(E_0 - EI' - 2EE)]^{1/2}\}/2 \quad (7)$$

$$[EI'] = \{E_0 + I_0' + K_I' - EI - 2EE - [(E_0 + I_0' + K_I' - EI - 2EE)^2 - 4I_0'[E_0 - EI - 2EE)]^{1/2}\}/2 \quad (8)$$

$$[EE] = \{E_0 + K_D/4 - EI - EI' - \sqrt{K_D[(E_0 - EI - EI')/2 + K_D/16]}\}/2 \quad (9)$$

An experimentally observable chemical shift difference, δ, between the resonance for the "sensitive" inhibitor in solution in the absence and in the presence of a certain concentration of enzyme is related to the chemical shift difference, Δ, between the inhibitor free in solution and bound to the enzyme by eq 10. In principle, the

$$\delta = ([EI]/[I_0])\Delta \quad (10)$$

experimental data for the D,L racemate, consisting of observed chemical shifts, δ, as a function of appropriately varying concentrations of $[E_0]$, $[I_0]$, and $[I_0']$ could be analyzed by iterative procedures until the combination of Δ, K_I, and K_I' which gave the best linear fit of the experimental data to eq 10 was obtained. In practice, three mutually independent variables are too many to determine in such a fashion and we took the value of Δ from studies with the pure D isomer. We also used the values of the enzyme dimerization constant K_D from studies with *N*-trifluoroacetyl-*p*-fluorophenylalanine. With these two variables fixed (K_D and Δ), we determined K_I and K_I' by the following iterative procedure. (i) Assume K_I. (ii) Assume K_I'. (iii) Use the values of K_I, K_I', and K_D with eq 7, 8, and 9 to calculate values of [EI] for each experimental point. (iv) With these values of [EI] and the value of Δ, calculate δ from eq 10. (v) Calculate the root-mean-square error between these calculated values of δ and the experimentally observed chemical shifts. (vi) Repeat steps ii–v to minimize error to find the best value of K_I' for the value of K_I assumed in step i. (vii) Repeat steps i–vi to obtain the best combinations of K_I and K_I'. These iterative procedures at

(1) K. L. Gammon, S. H. Smallcombe, and J. H. Richards, *J. Amer. Chem. Soc.*, **94**, 4573 (1972).
(2) T. M. Spotswood, J. M. Evans, and J. H. Richards, *ibid.*, **89**, 5052 (1967).

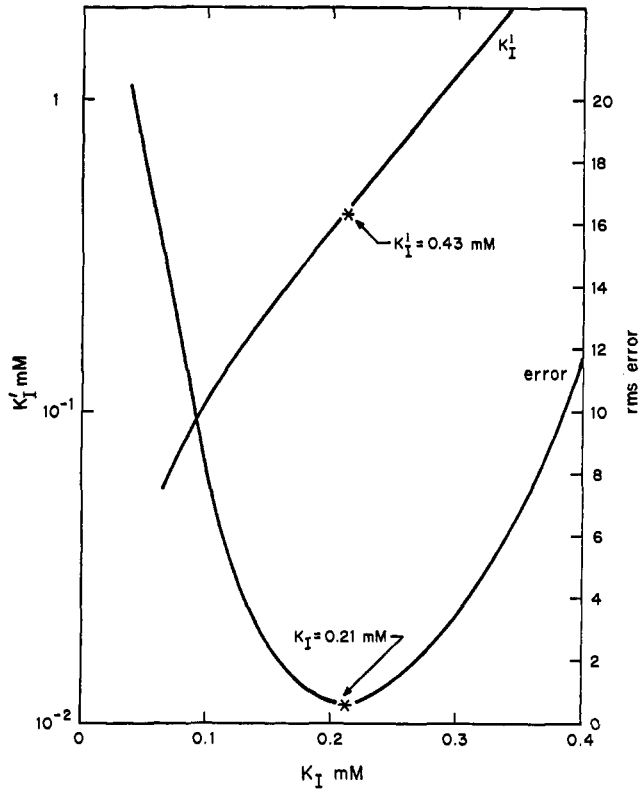

Figure 1. Relation between K_I and rms error and between K_I and K_I' at pH 5.5. The asterisks show the values picked by the computer analysis. The values of K_I = 0.21 mM and K_I' = 0.43 mM are to be compared with those of K_I = 0.18 mM and K_I' = 0.26 mM shown in Table I, the latter values being derived from analysis that assumed the same value of K_I found by independent studies with the D isomer alone.

first involve large differences in the values of K_I and K_I' assumed in steps i and ii followed by increasingly refined estimates around the region of decreasing error until values of K_I and K_I' have been chosen such that further refinement does not appreciably further minimize the error. Figure 1 shows the relationship between K_I and error and also between the best value of K_I' and K_I for one of the experiments in the present work.

An alternate approach recognizes the fact that both K_I and Δ can be evaluated from a separate set of experiments not involving the competitive inhibitor I'. Use of both K_I and Δ from such experiments (together with K_D) allows the unambiguous determination of K_I' from eq 7–10; indeed, in principle, one would need only one experimental point where both I and I' are present, though in practice more points are advisable. In spite of the greater certainty of K_I' determined in this way, the analysis of data for the case of competitive inhibition without fixing the value of K_I from an independent series of measurements serves as a useful check. For the fact that the iterative procedure applied to the data for competitive inhibition yields essentially the same value of K_I as that obtained previously from studies involving only the sensitive inhibitor (no competitive inhibition present) strengthens the correctness of the assumptions made in formulating eq 7–9 and diminishes the possibility of enzyme–inhibitor interactions other than those explicitly considered. (Such agreement between values of K_I calculated by independent means would be highly unlikely if the models on which the calculations were based were poor reflections of the physical reality.)

Results

Solutions of *N*-trifluoroacetyl-D,L-tryptophan and α-chymotrypsin exhibit two fluorine resonances, one considerably broadened and shifted downfield from its position in the absence of enzyme. The shifts and broadening are, however, considerably more pronounced if only the D isomer itself is present.

Accordingly a series of experiments at each half-unit from pH 5.0 to 8.0 was done with *N*-trifluoroacetyl-D-tryptophan (concentration varied from 0.8 to 22 mM) and α-chymotrypsin (concentration, corrected for inactive protein, held constant at 1.92 mM) using 2 mM *N*-trifluoroacetylglycine as the internal standard.[3]

The resulting data were analyzed as described previously also using values for K_D obtained in the earlier work,[1] and lead to the results collected in Table I.

Table I. Enzyme–Inhibitor Dissociation Constants, K_I and K_I', for α-Chymotrypsin and *N*-Trifluoroacetyl-D-(and L)-tryptophan[a]

pH	Buffer (concn, 0.1 M)	K_I, mM	K_I', mM
5.0	Citrate	0.26	0.46
5.5	Citrate	0.18	0.26
6.0	Citrate	0.16	0.26
6.5	Citrate	0.30	0.61
7.0	Tris	0.59	2.12
7.5	Tris	1.60	*b*
8.0	Tris	4.24	*b*

[a] The initial enzyme concentration was 1.92 mM and the inhibitor concentration varied between 0.7 and 22 mM. [b] Too large for measurement by our methods.

The fluorine nuclei of the *N*-TFA group of *N*-trifluoroacetyl-D-tryptophan experience a downfield shift of about 70 Hz (−0.74 ppm) which is essentially independent of pH, while K_I increases markedly above pH 6.

Using *N*-trifluoroacetyl-L-tryptophan with *N*-trifluoroacetylglycine as internal standard, we found that the L isomer does not shift significantly in the presence of enzyme ($\Delta \leq -5$ Hz). However, studies with the D,L racemate indicate that the L isomer does compete with the D isomer for the active site.

Accordingly, the L isomer can be considered as a competitive inhibitor of the D isomer and the chemical shifts observed when racemate interacts can be analyzed as described previously in terms of $K_{I(D)}$ and $K_{I(L)}$ (in this case I' is the L isomer) and K_D. (In these experiments, δ is actually taken as the splitting between the resonances of the D and L isomers.)

Table I shows the values of $K_{I(L)}$ derived by analytical procedures using K_D obtained previously,[1] and

(3) The possibility that *N*-trifluoroacetylglycine could bind competitively with *N*-trifluoroacetyl-D-tryptophan was explicitly studied at pH 6.0, where the concentrations of *N*-trifluoroacetylglycine were varied from 1–80 mM with an α-chymotrypsin concentration of 1.9 mM and *N*-trifluoroacetyl-D-tryptophan concentration of 2.3 mM. Little change in the relative positions of the *N*-trifluoroacetylglycine and *N*-trifluoroacetyl-D-tryptophan resonances were noted suggesting that K_I for *N*-trifluoroacetylglycine was about 6 mM or more than 40 times that for *N*-trifluoroacetyl-D-tryptophan. (Only when the concentrations of *N*-trifluoroacetylglycine approach 80 mM are clearly discernible changes in the chemical shift of 2.3 mM *N*-trifluoroacetyl-D-tryptophan observed. Therefore, we considered that *N*-trifluoroacetylglycine caused only a negligible perturbation when used as an internal standard at a concentration of 2 mM.)

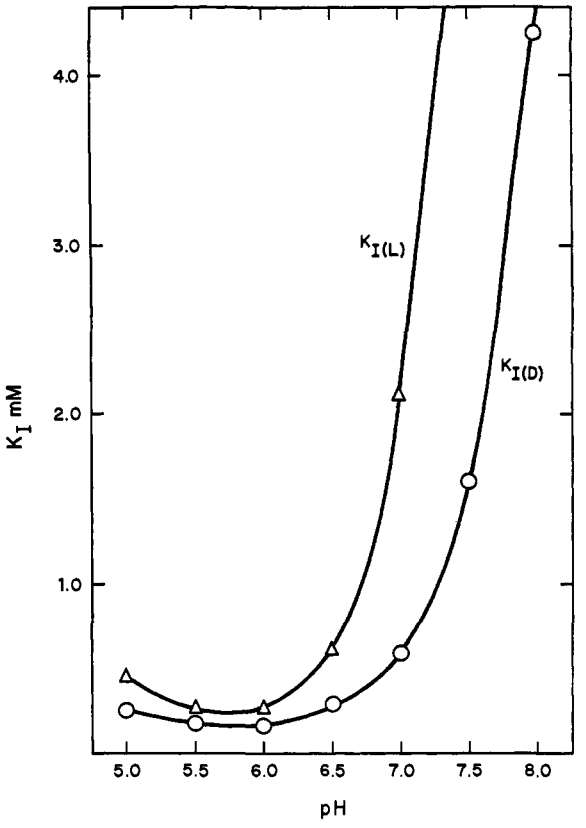

Figure 2. Dependence of binding constants $K_{I(L)}$ and $K_{I(D)}$ on pH.

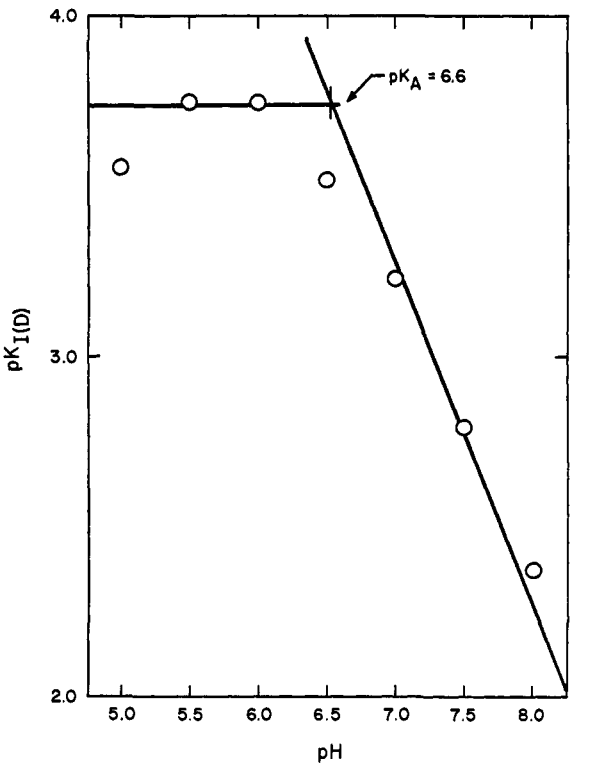

Figure 3. Plot of $pK_{I(D)}$ *vs.* pH showing dependence of $pK_{I(D)}$ on a group of the free enzyme with $pK_a = 6.6$.

$K_{I(D)}$ and Δ obtained earlier from studies with pure D isomer. Analysis of these experimental data by the iterative procedure which lets $K_{I(D)}$ vary while taking fixed values of K_D and Δ led to values of $K_{I(D)}$ and $K_{I(L)}$ which differed only slightly from those shown in Table I.

In Figure 2, the binding constants $K_{I(D)}$ and $K_{I(L)}$ are plotted as a function of pH. In Figure 3, $pK_{I(D)}$ is plotted against pH and shows a sharp decrease [that is, $K_{I(D)}$ itself increases) which is caused by neutralization of a group in the free enzyme with $pK_a = 6.6$.

Recent, independent work[4] on the binding of *N*-trifluoroacetyl-D-tryptophan to chymotrypsin has been reported with values of K_I varying from 6.5 m*M* at pH 6.33 to 72 m*M* at pH 8.12 and of Δ varying from 112 Hz at pH 6.33 to 255 Hz at pH 8.12. In this work no account was taken of enzyme oligomerization and this, and other unjustified approximations in collecting and analyzing the data, account for the large discrepancies between the magnitudes of the values obtained in the two cases and the apparent variation in Δ with pH which is, in fact, not a real characteristic of the *N*-trifluoroacetyl-D-tryptophan system.

Discussion

The foregoing results can be summarized. (i) When bound to the active site of chymotrypsin, the fluorine nuclei of *N*-trifluoroacetyl-D-tryptophan are shifted downfield 70 Hz (−0.74 ppm) relative to their position in solution. (ii) In contrast, the fluorine nuclei of *N*-trifluoroacetyl-L-tryptophan experience only a negligible change in chemical shift ($\leqslant -5$ Hz) on being bound to the enzyme. (iii) The chemical shifts for the bound inhibitors (both D and L isomers) are insensitive to changing pH (over the range 5–8). (iv) The enzyme–inhibitor dissociation constant for the L isomer (K_I') is larger than that for the D isomer (K_I) by a factor which increases from 2 at pH 5–6 to 4 at pH 7. Both dissociation constants show sharp increases dependent upon neutralization of a group in the free enzyme with $pK = 6.6$; the dissociation constant for the L isomer rises more sharply than that for the D isomer.

Stereochemistry of Binding. As discussed previously for the binding of *N*-trifluoroacetyl-D-(and L)-*p*-fluorophenylalanine,[1] the experimental facts summarized in i-iii suggest that the *N*-TFA group of the D and L isomers occupy different loci on the enzyme; for both isomers the tryptophan ring occupies the hydrophobic pocket and the α hydrogen points toward the γ-methylene group of methionine-192. For the D isomer the *N*-TFA group is situated in the catalytic locus near His-57 and Ser-195 and the carboxylate carbon is directed toward Ser-214.[5] These two groups interchange loci when the L isomer binds.

The similarity of binding between the *N*-TFA derivatives of tryptophan and *p*-fluorophenylalanine is supported experimentally by the very similar shifts experienced by the fluorine nuclei of the *N*-TFA groups when these inhibitors bind. Both D isomers experience a downfield shift of about 70 Hz (*N*-trifluoroacetyl-D-*p*-fluorophenylalanine, -65 ± 5 Hz,[1] and *N*-trifluoroacetyl-D-tryptophan, -70 ± 5 Hz), whereas neither L isomer shows a significant shift ($\Delta \leqslant -5$ Hz). The previous discussion[1] of the origin of the observed downfield shift of *N*-trifluoroacetyl-D-(and L)-*p*-fluorophenyl-

(4) H. Ashton and B. Capon, *J. Chem. Soc. D*, 513 (1971).

(5) T. A. Seitz, R. Henderson, and D. M. Blow, *J. Mol. Biol.*, 46, 337 (1969).

alanine derivatives accordingly applies also to the shifts observed in this work on *N*-trifluoroacetyl-D-(and L)-tryptophan.

The values of Δ of both the D and L isomer show no marked pH dependence (in contrast to the inhibition constants). This behavior results most likely from the absence of significant binding of these inhibitors to deprotonated chymotrypsin. Thus, even at pH 8 inhibitor binds significantly only to protonated enzyme so that the *N*-TFA groups show the same chemical shift changes on binding (Δ) as they do at more acidic pH.

pH Dependence of $K_{I(D)}$ and $K_{I(L)}$. The binding of both D and L isomers shows a strong dependence on deprotonation of a group on the enzyme with pK_a = 6.6. This is analogous to the behavior of *N*-trifluoroacetyl-D-*p*-fluorophenylalanine and for reasons given earlier, neutralization of His-57 is the most plausible origin of the effect.

A second point of interest is the stronger dependence shown by the L isomer (relative to the D isomer) on the state of protonation of the enzyme; $K_{I(L)}$ rises more rapidly to higher values than does $K_{I(D)}$ when the group of pK_a = 6.6 is deprotonated. Similar differences between $K_{I(D)}$ and $K_{I(L)}$ have been observed for *N*-acetyl-D-(and L)-tryptophan.[6] If this group being deprotonated is, in fact, His-57, the greater dependence of $K_I[L]$ on a group with pK_a = 6.6 can be rationalized. We have previously discussed the stereochemistry of binding as involving insertion of the carboxylate carbon of the L isomer into the catalytic locus which includes His-57 and Asp-102. (In the case of the D isomer, the *N*-TFA group occupies this region of the active site.) Deprotonation of His-57 will give the catalytic locus an overall negative charge (from Asp-102) which will more strongly repel the carboxylate carbon of the L isomer than it will the *N*-TFA group of the D isomer. As a result, binding of the L isomer might be expected to reflect more strongly than binding of the D isomer the state of ionization of His-57.

In contrast to the behavior of anionic inhibitors [such as *N*-trifluoroacetyl-*p*-fluorophenylalanine,[1] *N*-trifluoroacetyl-D-(and L)-tryptophan, *N*-acetyl-D-(and L)-tryptophan[6]], neutral inhibitors (such as *N*-acetyl-D-tryptophanamide[7] or *N*-trifluoroacetyl-D-tryptophanamide[8]) show no significant dependence of K_I on a group in the free enzyme with pK_a = 6.6.

As seen in Figure 2, the values for both $K_{I(D)}$ and $K_{I(L)}$ seem to show a slight increase at increasingly acidic pH. [Incidentally, the value for $K_{I(L)}$ is very sensitive to the value used for $K_{I(D)}$ so that the apparently greater increase in $K_{I(L)}$ relative to $K_{I(D)}$ at pH 5 may be an artifact of the choice of $K_{I(D)}$.] When the formation of chymotrypsin dimers is not included in the analysis, similar behavior is apparent also in $K_{I(D)}$ *vs.* pH plots for *N*-trifluoroacetyl-D-*p*-fluorophenylalanine[1] and *N*-acetyl-D-(and L)-tryptophan.[6] In the present analysis, however, enzyme dimerization was included. Therefore, though the increase in K_I at acidic pH could result from our analysis having allowed for insufficient dimerization, the observed increase could be real and reflect possibly a conformational change in the enzyme[9] which manifests itself more clearly with tryptophan derivatives than it did with derivatives of *p*-fluorophenylalanine.

The relative values of the inhibition constants for the two antipodes [$K_{I(D)} \sim 2 \cdot K_{I(L)}$, at acidic pH] agree well with the general observation that D isomers bind more strongly than do L isomers by about 400 cal mole^{-1}.[10]

Conclusion

Binding on *N*-trifluoroacetyl-D-(and L)-tryptophan to α-chymotrypsin parallels the same magnetic resonance characteristics previously encountered for *N*-trifluoroacetyl-D-*p*-fluorophenylalanine except that $K_{I(D)}$ and $K_{I(L)}$ are sufficiently close that one antipode binds competitively with the other, a situation for which analytical techniques have been developed and discussed. The binding constants are both sharply increased with deprotonation of a group on the enzyme with pK_a = 6.6 (presumably His-57); $K_{I(L)}$ is more strongly affected than $K_{I(D)}$.

Experimental Section

***N*-Trifluoroacetyl-D-tryptophan.**[11] Dry D-tryptophan (2.04 g, 0.01 mol) was dissolved in anhydrous trifluoroacetic acid (11.2 ml) with stirring. Anhydrous ethyl ether (13 ml) was added to the solution. The reaction mixture was cooled to 0° in an ice bath. Trifluoroacetic anhydride (2 ml) was added dropwise over a period of 10 min. After about 5 min a precipitate formed. The mixture was allowed to stand another 15 min and then filtered. The filtrate was washed with water and dried thoroughly. The crude product was recrystallized twice from toluene to yield 19 g (63%) of white needles; mp 162–163° (lit.[12] mp 163–164°).

N-Trifluoroacetyl-L- and- D,L-tryptophan were prepared in the above manner: mp L 162–163° (lit.[12] mp 163–164°); D,L 154–155°. *Anal.* Calcd for $C_{13}H_{11}O_3NF_3$: C, 51.99; H, 3.49; N, 9.39. Found: C, 52.12; H, 3.70; N, 9.33.

***N*-Trifluoroacetylglycine.**[13] Dry glycine (0.75 g, 0.10 mol) was dissolved in anhydrous trifluoroacetic acid (60 ml) and the solution was cooled to −10° in an ice–salt bath. Trifluoroacetic anhydride (17.6 ml) was then added dropwise with stirring. After all the anhydride had been added, the bath was removed and the mixture was stirred for an additional 30 min at room temperature. The mixture was then stripped to dryness on a rotary evporator. The residue was dissolved in ethyl ether (200 ml), filtered, and concentrated to half volume on a rotary evaporator. Hot toluene (150 ml) was added to the solution and the remaining ether was removed at room temperature. The crystalline product was collected by suction filtration and recrystallized once from toluene; mp 116–116.5° (lit. mp 116°).

Nmr samples were prepared and run as previously described.[1]

Acknowledgment. This work was supported by a grant from the U. S. Public Health Service (GM16424).

(6) C. H. Johnson and J. R. Knowles, *Biochem. J.*, **101**, 56 (1966).
(7) C. H. Johnson and J. R. Knowles, *ibid.*, **103**, 428 (1967).
(8) K. L. Gammon, S. H. Smallcombe, and J. H. Richards, unpublished results.
(9) A. R. Fersht, presented at Cold Spring Harbor Symposium on Protein Structure and Function, June, 1971.
(10) J. R. Knowles, *J. Theor. Biol.*, **9**, 213 (1965).
(11) F. Weygand and R. Geiger, *Chem. Ber.*, **89**, 647 (1956).
(12) E. E. Schallenberg and M. Calvin, *J. Amer. Chem. Soc.*, 77, 2779 (1955).
(13) F. Weygard and E. Csender, *Angew. Chem.*, **64**, 136 (1952).