Structure and Specific Binding of Trypsin: Comparison of Inhibited Derivatives and a Model for Substrate Binding

M. KRIEGER, L. M. KAY and R. M. STROUD

Norman W. Church Laboratory of Chemical Biology
California Institute of Technology
Pasadena, Calif. 91109, U.S.A.

(Received 21 June 1973, and in revised form 12 October 1973)

The high-resolution structure of bovine trypsin inhibited with DFP† was determined by Stroud et al. (1971 and R. M. Stroud, L. M. Kay, A. Cooper & R. E. Dickerson, Abstr. 8th Int. Congr. Biochem. 1970). The experiments reported here were designed to study the specific side-chain binding pocket of trypsin using benzamidine, which is a competitive, specific inhibitor of trypsin. High-resolution electron density synthesis and difference syntheses unambiguously identify the side-chain binding pocket, which normally recognizes and binds the side chains of arginine or lysine during proteolysis. Several important conformational differences in the protein structure are apparent between DIP- and BA-trypsins, and these are discussed with particular reference to inhibition, the binding of lysine and arginine, subsequent orientation of the target at the active site, and the enhancement of tryptic activity towards non-specific substrates seen on binding small alkyl amines or guanidines in the specific binding pocket.

The BA-trypsin structure provides a good model for the binding of real substrate side chains to trypsin during catalysis, explaining the sharp trypsin specificity for lysine or arginine side chains (Weinstein & Doolittle, 1972) and the lack of specificity for stereochemically different basic side chains. Benzamidine is shown to inhibit trypsin by steric interference with the inferred position of good substrates, even when they do not carry any side chain.

Apart from the substitution of benzamidine and DIP, the most significant differences between DIP-trypsin and BA-trypsin involve complete repositioning of the side chain of Gln192, alterations in the side chains of Asp102, His87 and Ser195 at the active site, and changes in the solvent structure around this region. The carboxyl group of Asp189, which is responsible for trypsin specificity, shows no movement on binding benzamidine. The amidinium cation of benzamidine forms a salt bridge with Asp189 in BA-trypsin; a similar salt bridge can be constructed between the side chains of model substrates with lysyl or arginyl side chains and Asp189. The γ-oxygen of Ser190 is displaced by a 120° rotation about its α-β bond on binding benzamidine and the binding pocket closes to sandwich the inhibitor ring between the peptide planes of 190–191 and 215–216. These contacts are presumably found in the enzyme-substrate complex with specific substrates.

†Abbreviations used: DFP, diisopropyl fluorophosphate; DIP, diisopropyl phosphoryl; DIP-trypsin, diisopropyl fluorophosphate-inhibited trypsin; BA-trypsin, benzamidine-inhibited trypsin. In equations the latter are further abbreviated to BAT and DIPT. The amino acid abbreviations follow standard convention, and the amino acid residue numbering is that adopted by Stroud et al. (1971).
The active site structure at pH 8.0 is discussed with particular reference to the microscopic pKₐ values of Asp102 and His57, the pKₐ of the Asp–His system, and the mechanistic consequences of these assignments.

1. Introduction

Trypsin shows its greatest activity for proteolysis between pH 7.0 and 9.0 (Northrop & Kunitz, 1932). It has been one of our goals to study the structure of trypsin within this pH range in order to minimize the possibility of examining an inactive conformation of the enzyme. Trypsin undergoes at least three pH-dependent conformational changes between pH 0.5 and 7.0, which are detectable by optical rotary dispersion (Lazdunski & Delaage, 1967). More subtle, but nonetheless mechanistically significant pH-dependent changes in the conformation, particularly around the active site, have been detected in the serine proteases using more sensitive techniques (Vandlen & Tulinsky, 1973). One of the difficulties of studying native trypsin within the most active pH range is its rapid rate of autolysis. Firmly bound inhibitors reduce the possibility of autolysis. For this reason, our first crystallographic studies were carried out on DFP-trypsin, in which the enzyme had been irreversibly inactivated by covalent modification of the reactive serine 195 using a non-specific inhibitor (DFP, Fig. 1).

Benzamidine (Fig. 1) is a competitive, specific and reversible inhibitor of trypsin, which binds with a $K_i = 1.8 \times 10^{-6}$ M (Mares-Guia & Shaw, 1965). Its amidinium

![Chemical structures](image)

**Fig. 1.** Chemical structures: (I) benzamidine; (II) diisopropyl fluorophosphate; (III) N-acetyl amino acid ethyl esters, (a) acetyl glycine ethyl ester, (b) acetyl lysine ethyl ester, (c) acetyl arginine ethyl ester; (IV) ethyl p-amidino benzoate; (V) ethyl p-amidino phenyl acetate.
cation interacts specifically with the binding pocket of trypsin in a manner that is probably similar to that of an arginine side chain on a specific substrate.

In studying BA-trypsin, we hoped to learn more about the highly specific binding pocket of trypsin, how it functions in co-operation with the active site, and how certain small molecules shut down tryptic catalysis, while others serve to enhance the catalytic rate for substrates with small, non-polar side chains.

2. Experimental

(a) Preparation of crystals of inhibited trypsin

(i) Benzamidine-inhibited trypsin

Bovine trypsin (150 mg) (Mann Fine Chemical Company) and benzamidine hydrochloride (5-7 mg) (Aldrich Chemical Company) were dissolved in 3-0 ml cold 0-05 M-Tris buffer (pH 8-17) and stored at 5°C for 30 min. 500 mg of anhydrous MgSO₄ were added slowly with constant stirring at 5°C. The solution was clarified by centrifugation and sterilized by passing it through a Millipore filter into 3 sterile vials. The vials were stoppered and stored at 5°C.

Tiny needles were observed in the solution after about 1 month; these crystals grew to small, but usable size (about 0-4 mm × 0-2 mm × 0-2 mm) in about 2½ months.

(ii) Diisopropyl phosphoryl trypsin

Bovine trypsin was inhibited with DFP according to the procedure of Cunningham (1954).

150 mg of DIP-trypsin was dissolved in 3-75 ml distilled water at room temperature and the pH was adjusted to 7-5 with 0-05 N-NaOH. 1-25 ml of a solution of MgSO₄ (15 g anhydrous salt dissolved in 50 ml water) was added slowly with stirring. The solution was clarified by centrifugation and sterilized by passing it through a Millipore filter into 2 or 3 sterile vials. The vials were stoppered and stored at room temperature. Rod-shaped crystals of usable size usually took at least 2 months, and often much longer, to grow.

(iii) State of the trypsin in the crystals

Preparations of DIP-trypsin and BA-trypsin crystals are quite similar. In order to retard autolysis of the reversibly inhibited trypsin, BA-trypsin solutions were refrigerated. It was then necessary to raise the precipitant (MgSO₄) concentration by about 33% above that used in the DIP-trypsin crystallizations. In both cases, the pH was between 7-5 and 8-2, in the region where the uninhibited enzyme would be most active. Furthermore, evidence summarized by Stroud et al. (1971) leads us to believe that the structure of DIP-trypsin at pH 7-5 is very similar to native trypsin at pH 7-5 in gross conformation and state of protonation.

Chromatography of solutions prepared from crystals of DIP-trypsin and BA-trypsin indicates that both contain mixtures of inhibited α and β-trypsins and small amounts of other partially cleaved molecules. Although the ratios vary, they tend toward a 1:1 ratio of α to β-trypsin, a mixture richer in α-trypsin than the commercial trypsin or the DIP-trypsin from which the crystals were prepared. β-Trypsin contains a single polypeptide chain of 223 amino acids, while α-trypsin has a single tryptic scission between Lys145 and Ser146, and shows almost identical activity to α-trypsin (Schroeder & Shaw, 1968).

(b) X-ray data collection

Complete 3-dimensional X-ray diffraction data from BA-trypsin and DIP-trypsin were collected to a resolution of 2-7 Å. Data for DIP-trypsin were collected on an extensively modified Supper diffractometer using Ni-filtered CuKα X-radiation. A graphite monochromator designed by Dr Sten Samson was built and incorporated into the system for collection of BA-trypsin data (Stroud et al., 1974). In all, 32 crystals were used for DIP-trypsin and 6 for BA-trypsin. The difference in number is due to the reduced rate of crystal decay in the monochromatized beam.
Fig. 2. The $(2 \sin \theta)^2$ dependence of the average values of $(\bigcirc) |F_{\text{DBT}}|$, $(\bigcirc) |F_{\text{DBT-TE}} - F_{\text{DPT}}|$, $(\square) |F_{\text{D BT}} - F_{\text{DPT}}|$, and the estimated errors $(\delta F)$ for DIP-trypsin data collected from 2 different crystals, one with a monochromator and one without $(\bigtriangleup)$ ($\pm 2.5\%$ from the mean value).

It was surprising to find that the mean difference in diffracted amplitudes $\langle \delta F \rangle$ between BA-trypsin and DIP-trypsin ($\delta F = F_{\text{D AT}} - F_{\text{DPT}}$) was greater than the mean difference between DIP-trypsin and a thallium derivative previously used in phase analysis. The value of $\langle \delta F \rangle/\langle |F_p| \rangle$ was 21.6%, at least 6 to 7 times the expected errors from all sources, and was roughly constant out to 2.7 Å, indicating that quite specific changes were occurring in the enzyme between highly isomorphous crystals (Fig. 2).

An estimate of the expected error in a single measurement, arising from all sources, was made by comparing data sets from 2 different crystals of DIP-trypsin, one set taken with Ni-filtered CuKα radiation ($F_{\text{PM}}$) and the other with monochromatized CuKα X-rays ($F_{\text{PM}}$). The mean value of these differences, shown in Fig. 2, gives a mean residual,

$$R = \frac{\sum |F_{\text{PM}} - F_{\text{PS}}|}{\sum |F_{\text{PM}}|},$$

of 5.1% after corrections and scaling (including all zero intensities) or $\pm 2.5\%$ from the mean value of $F_p$.

Data were corrected for absorption, time-dependent decay, geometry, monochromator polarization, etc., and scaled together to obtain $F_{\text{D AT}}$ and $F_{\text{DPT}}$ according to the procedure of Stroud et al. (1974). Final scaling of the two data sets was accomplished with a 2-parameter ($K'$, $B'$) exponential function derived by least-squares minimization of $E'(s) = [K' \exp (B' s^2) \langle |F_{\text{D AT}}| \rangle_s - \langle |F_{\text{DPT}}| \rangle_s]^2$, computed from 20 zones of $s = 2 \sin \theta$. The cell dimensions of BA-trypsin and DIP-trypsin crystals were identical within experimental error. The space group and cell dimensions are:

- space group: $P2_12_12_1$
  - $a = 54.8 \pm 0.05$ Å
  - $b = 58.7 \pm 0.05$ Å
  - $c = 67.6 \pm 0.05$ Å.

(c) Electron density calculations

A difference electron density synthesis was calculated using phases $\phi_p$ and figures of merit, $m$, previously determined for DIP-trypsin (Stroud et al., 1974) using the expression

$$\Delta \rho(r) = \frac{1}{V} \sum_{r} m \Delta F_r \exp \{2\pi i r \cdot \mathbf{s} + \phi_p\},$$

where $\Delta F_r = \langle |F_{\text{D AT}}| - |F_{\text{DPT}}| \rangle_r$. Reflections where $F_{\text{D AT}}$ or $F_{\text{DPT}}$ were less than 2σ were left out of the synthesis, and the map was calculated at intervals of 0.01 in $y$ and $z$, and 0.02 in $x$, making the finest interval 0.6 Å.

The difference map was contoured at intervals of 0.05 electrons/Å³, omitting the zero
| Map               | $\langle |F_{\text{calc}}| \rangle$ (e) | $\sigma$ (e) | Calculated† $\langle \Delta \rho \rangle$ (e Å$^{-2}$) | Observed‡ r.m.s. error (e Å$^{-2}$) | Observed highest noise (e Å$^{-2}$) | Observed highest peak (e Å$^{-2}$) |
|------------------|---------------------------------|--------------|-----------------------------------------------------|-----------------------------------|----------------------------------|---------------------------------|
| BA-trypsin—DIP-trypsin | 84.7                           | 2.3          | 0.069                                               | 0.059                             | 0.17                             | 2.5                             | 0.75                            | 11                             |
| DIP-trypsin      | 573.0                           | 21.0         | 0.38                                                | —                                 | —                                | —                               | —                               | —                             |

† $\Delta F$: $\langle \Delta \rho \rangle = \frac{1}{2V^2} \sum_{hkl} \Delta F^2 (2 - m^2)$,

$F_{\text{diff}}$: $\langle \Delta \rho \rangle = \frac{1}{V^2} \sum_{hkl} F^2_{\text{diff}} (1 - m^2)$,

(after Henderson & Moffat, 1971).

‡ The observed root mean-square density error is based on a relatively featureless region of the map.

§ s.d., the electron density given as a multiple of the calculated r.m.s. error.
and first contour in both the negative and positive density, using an approximate (statistically determined) absolute scale factor (Stroud et al., 1974). Five sections of the map in the region of benzamidine binding are shown in Plate I. The root mean-squared error \( \langle d\rho^2 \rangle \) in this map is 0.009 electrons/Å³, and a comparison of the difference and DIP-trypsin electron density maps may be found in Table I. The highest uninterpreted “noise” density peak corresponds to 2.6 times the calculated value for the standard deviation of difference density in the map, indicating that all the observed electron density changes above the noise have been accounted for.

Molecular movements that were less than the resolution (2.7 Å) were calculated from the difference map using vector moments evaluated for pairs of equal positive and negative features, and by analysis of the difference between the density map, computed using terms

\[
(2|F_{\text{Bat}}| - |F_{\text{DIP}}|) m \exp (i\phi_{\text{DIP}}),
\]

and the density map for DIP-trypsin obtained previously (Stroud et al., 1974). Peaks in the difference map were integrated and their electron equivalents evaluated assuming that the benzamidine peak in the difference map contained its full complement of 64 electrons (Table 2). Actually, this peak represents (benzamidine—solvent) density; however, an approximation of this kind is probably more accurate than any estimate of the absolute scale of the protein. It also accounts for the factor \( \langle m^2 \rangle /2 \), which reduces apparent electron densities in difference maps (Henderson & Moffat, 1971).

**Table 2**

Intensities of some difference map peaks scaled to the integrated benzamidine density with its full complement of 64 electrons

<table>
<thead>
<tr>
<th>Peak region</th>
<th>Positive peak electrons †</th>
<th>Negative peak electrons</th>
</tr>
</thead>
<tbody>
<tr>
<td>His57</td>
<td>6.9</td>
<td>8.9</td>
</tr>
<tr>
<td>Cystine bridge (191–220)</td>
<td>8.0</td>
<td>9.6</td>
</tr>
<tr>
<td>Gln192</td>
<td>22.8</td>
<td>18.0</td>
</tr>
<tr>
<td>Benzamidine</td>
<td>64.0</td>
<td>—</td>
</tr>
<tr>
<td>DIP</td>
<td>—</td>
<td>90.8</td>
</tr>
</tbody>
</table>

† Relative peak contents were evaluated by integration and scaled so that the benzamidine peak corresponds to 64 electrons.

An estimate of the error in a shift calculated from vector moments was made by computing an “error vector moment.” The error vector moment for a pair of positive and negative difference peaks is equivalent to the vector moment for the pair, except that the root mean-squared electron density error rather than the observed electron density is integrated over the peak volumes. The expected error in a shift was evaluated by dividing the error vector moment by the number of electrons, \( N \), in the group which moves:

\[
\text{shift error} = \frac{\text{“error vector moment”}}{N} = \frac{\int_{r_1} [\rho(r)]^2 r \, dr + \int_{r_1} -[\rho(r)]^2 r \, dr}{N}. 
\]

Plate I shows an analysis of part of the difference density in a view showing the relevant region of the molecular model of DIP-trypsin to which a skeletal model of benzamidine has been added simultaneously. Superposition was accomplished with a half-silvered mirror (Richards, 1968). Newly ordered solvent molecules whose densities lie only partially in these sections are indicated by \( W_1 \) and \( W_2 \).
3. Results

(a) Interpretation of the map

The whole difference map, which is remarkably clean, allows a detailed interpretation of peaks with an integrated electron content down to about three electrons (corresponding to a spherical peak of diameter 2.7 Å, with density of $3 \times \langle d^2 \rangle^{1/3}$; Table 1).

The large positive peak (BA) shows how benzamidine binds to trypsin in the specific binding pocket. Evidently, the uniquely high specificity of trypsin and the binding energy of benzamidine, about 6.2 kcal/mol (Mares-Guia & Shaw, 1967), lead to one-site binding in contrast to other pancreatic serine proteases. α-Chymotrypsin, for example, often binds small substrate analogues at multiple sites (A. Tulinsky, personal communication). The benzamidine density reaches a maximum of about +0.45 e/Å$^3$ in the centre of the phenyl ring, and shows quite distinct necking in the region of the C$_{11}$–C$_{17}$ bond. Solvent molecules are displaced from the binding pocket when benzamidine is bound, so that this density represents benzamidine minus solvent. The broad negative peak beneath benzamidine is due to a displacement of partially ordered solvent from the binding pocket. The amidinium cation

![Diagram of the DIP-trypsin electron density map showing the position of the DIP inhibitor bound to the γ-oxygen of Ser195. The first contour represents 0.25 e/Å$^3$. The calculated noise is between 1 and 2 contour levels, and the mean figure of merit was 0.71 out to 2.7 Å resolution. Atomic positions are labeled. The lower isopropyl group enters the binding pocket, while the upper one is not visible, presumably because of thermal motion.](image-url)
of benzamidine interacts with the negatively charged carboxyl group of Asp189 at the back of the binding pocket. N(2) of benzamidine lies 2.9 Å from one of the carboxyl oxygens. The large negative region in the map is due to the subtraction of DIP density. The ring atom C(4) in benzamidine lies on a region of zero electron density, and there is a scarp at the top of the ring around C(15)-C(4). One of the isopropyl groups on DIP lies very close to the region of C(15)-C(4)-C(2) (Fig. 3), so

**Fig. 4.** The structure around the active site and binding pocket in BA-trypsin. Benzamidine and solvent molecules are indicated by right-handed shading. Parts of the trypsin molecule in this region which adopt different conformations in BA-trypsin and DIP-trypsin are shown in left-handed shading. The general "sandwiching" effect of pocket closure is not shown. In the crystal structure this movement is due to repositioning of the chain between residues 190 and 191 and the cystine bridge 191-220. Other changes occur around the left-hand side of Asp102. This view is approximately 90° from Plate I and Fig. 3, about an almost vertical axis. This view has been adopted as "standard" and rotated to match the structure of α-chymotrypsin (Birktoft & Blow, 1972) for ease of comparison. Asp189 is dotted and forms a salt bridge with benzamidine. This and the other computer-plotted figures (5 and 6) were drawn by a modified version of ORTEP (Johnson, 1965).

**Fig. 5.** The structure around the active site in DIP-trypsin. Right-handed shading picks out the DIP group attached to Ser195. One isopropyl group partly fills the neck of the binding pocket approximately in the same position as C(15)-C(4)-C(2) in BA-trypsin (Fig. 4).
that this region in the difference map (Plate I) corresponds to the difference between part of the firmly bound benzamidine and an isopropyl group vibrating about the same general location. The sharp concave scarp between benzamidine and DIP corresponds to the region of maximal difference density gradients between them.

The positive region associated with Gln192 is matched by a negative region behind the sections shown in Plate I. These peaks define the complete reorientation of

![Chemical structures](image)

**Fig. 8.** A view of benzamidine in the specific binding pocket oriented 90° about a vertical axis from Fig. 4.

<table>
<thead>
<tr>
<th></th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Atom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzamidine</td>
<td>0.537</td>
<td>0.260</td>
<td>0.764</td>
<td>BC1</td>
</tr>
<tr>
<td></td>
<td>0.538</td>
<td>0.233</td>
<td>0.756</td>
<td>BC2</td>
</tr>
<tr>
<td></td>
<td>0.523</td>
<td>0.224</td>
<td>0.740</td>
<td>BC3</td>
</tr>
<tr>
<td></td>
<td>0.511</td>
<td>0.246</td>
<td>0.729</td>
<td>BC4</td>
</tr>
<tr>
<td></td>
<td>0.510</td>
<td>0.288</td>
<td>0.736</td>
<td>BC5</td>
</tr>
<tr>
<td></td>
<td>0.523</td>
<td>0.275</td>
<td>0.753</td>
<td>BC6</td>
</tr>
<tr>
<td></td>
<td>0.551</td>
<td>0.288</td>
<td>0.781</td>
<td>BC7</td>
</tr>
<tr>
<td></td>
<td>0.563</td>
<td>0.253</td>
<td>0.792</td>
<td>BN1</td>
</tr>
<tr>
<td></td>
<td>0.550</td>
<td>0.290</td>
<td>0.786</td>
<td>BN2</td>
</tr>
<tr>
<td>Active-site solvent</td>
<td>0.498</td>
<td>0.188</td>
<td>0.667</td>
<td>BW1</td>
</tr>
<tr>
<td></td>
<td>0.510</td>
<td>0.174</td>
<td>0.605</td>
<td>BW2</td>
</tr>
<tr>
<td></td>
<td>0.476</td>
<td>0.154</td>
<td>0.643</td>
<td>BW3</td>
</tr>
<tr>
<td>Ser190</td>
<td>0.550</td>
<td>0.327</td>
<td>0.765</td>
<td>190CA</td>
</tr>
<tr>
<td></td>
<td>0.554</td>
<td>0.336</td>
<td>0.760</td>
<td>190CB</td>
</tr>
<tr>
<td></td>
<td>0.551</td>
<td>0.349</td>
<td>0.743</td>
<td>190OG</td>
</tr>
<tr>
<td>Gln192</td>
<td>0.604</td>
<td>0.208</td>
<td>0.699</td>
<td>192CB</td>
</tr>
<tr>
<td></td>
<td>0.592</td>
<td>0.184</td>
<td>0.696</td>
<td>192CG</td>
</tr>
<tr>
<td></td>
<td>0.578</td>
<td>0.190</td>
<td>0.674</td>
<td>192CD</td>
</tr>
<tr>
<td></td>
<td>0.554</td>
<td>0.184</td>
<td>0.673</td>
<td>192NE2</td>
</tr>
<tr>
<td></td>
<td>0.590</td>
<td>0.197</td>
<td>0.660</td>
<td>192OE1</td>
</tr>
<tr>
<td>Ser195</td>
<td>0.482</td>
<td>0.250</td>
<td>0.666</td>
<td>195OG</td>
</tr>
</tbody>
</table>

**Table 3**

Benzamidine-inhibited trypsin: co-ordinates for benzamidine, serine 190, serine 195, glutamine 192, and associated solvent molecules
Gln192 in BA-trypsin, as shown in Figures 4 and 5. Gln192 acts as a polar flap covering the entrance to the binding pocket and must necessarily undergo a displacement on binding substrate side chains which interact with Asp189.

There are several other changes in the binding pocket when benzamidine is bound. There is a significant “squeeze” of the pocket, which leads to a close sandwiching of the inhibitor between the peptide planes of 190–191, and 215–216 (indicated in Figs 4 and 6). The movement is only visible in the right-hand side of the pocket seen in Figure 4. No movement is seen around Ser214 (which is hydrogen-bonded to Asp102). The disulfide bridge 191–220 moves in toward benzamidine by 0.7 ± 0.3 Å. The side chain of Ser190 rotates by 120° about its z–β bond and the y-oxygen forms a hydrogen bond to the hydroxyl group of Tyr228, and possibly to the carboxyl of Asp194. The position of Ser190 in DIP-trypsin is sterically incompatible with binding of benzamidine or an arginine side chain, so it presumably undergoes this type of burial on binding specific substrates or their analogues. The positions of inhibitor groups, and of groups in significantly different positions in BA-trypsin and DIP-trypsin are listed in Tables 3 and 4.

(b) The active site

Readjustments in the active site due to the BA-trypsin–DIP-trypsin difference include a shift in the imidazole ring of His57, binding of additional solvent molecules
Plate I. Five sections of the difference density map showing a block 25-2 Å (horizontal $dy = 0.04$ to 0.5) by 22.0 Å (vertical $dz = 0.54$ to 0.92) by 5.5 Å (thick $dz = 0.41$ to 0.51). Contours are at $\pm [0.10 (0.05)] e/Å^2$, and negative density is shown with dotted contours; P corresponds to the phosphorus atom position in DIP-trypsin. Similarly, DIP, Gin192 (in BA-trypsin), and BA (benzamidine) indicate their associated densities. Superimposed on the sections is a part of the skeletal model of trypsin showing the DIP group attached to Ser195. Benzamidine (BA) and solvent molecules (W1, W2, etc.) have been added to the model. The conformations of Gin192 and Ser190 are for DIP-trypsin, and the carboxyl of Asp189 is indicated. The negative region (SOL) corresponds to the displaced solvent in BA-trypsin.  

(facing p. 218)
and some rearrangements in the side chains of Leu99, Val227, Ser195 and Asp102. Two complementary positive and negative peaks in the difference map (see Table 2) indicate a rotation about the α-β bond of His57, a reorientation of the carboxyl group of Asp102, and the binding of a solvent molecule in the active site (W3). It is difficult to estimate the extent of movement of His57 from the difference peaks.

**Fig. 7.** The peaks associated with His57 on the difference map. The lower peak is negative density (−) while the other one is positive (+). The latter peak is a composite with a solvent molecule density (see text).

**Fig. 8.** Electron density for His57 in the DIP-trypsin Fourier map, computed for the plane parallel to the imidazole ring.
because both the movement of His57 and the solvent molecule W3 contribute to the positive peak and this solvent molecule shifts the peak away from the imidazole ring. The negative peak is also a hybrid due to rearrangements of the Asp102 and His57 side chains.

As a result, the vector moment of these two peaks comes to 36-4 electron Å, which would indicate a falsely large movement (1-04 Å) of the imidazole containing 35 electrons. The two peaks are shown in Figure 7, where the imidazole ring as it appears in DIP-trypsin is indicated by a solid line drawing. (It is clear from this Figure that the positive peak is too far from the imidazole to be due to histidine movement alone.) The movement of the imidazole was therefore measured by comparing a density map $\rho'(r)$ computed using terms $2|F_{BA}|-|F_{DIPT}|$, $\phi_0$, and the original $\rho(r)$ map for DIP-trypsin. The well-defined electron density for the imidazole in the latter map is shown in Figure 8 (and also in Plate IV of Stroud et al., 1974). The shift was evaluated from the concerted movements of the side-chain density between the two maps calculated from

$$
\Delta x = \frac{\int \rho'(r) \cdot r \, dx}{\int \rho'(r) \cdot dx} - \frac{\int \rho(r) \cdot r \, dx}{\int \rho(r) \cdot dx},
$$

where $x$ is the line joining the two difference peaks in the direction of rotation. This expression is essentially a $\Delta x$ difference in centers of gravity of the peaks between $\rho'(r)$ and $\rho(r)$ and accounts for changes in thermal motion parameters. The imidazole ring of His57 has a somewhat larger thermal motion in BA-trypsin than in DIP-trypsin. In BA-trypsin the imidazole ring has moved 0-20±0-07 Å away from Asp102 by rotation about the C$_{a}$-C$_{b}$ bond.

This movement is insufficient to account for all of the negative peak between Asp102 and the ring. The remainder of this peak indicates that Asp102 has also moved away from the imidazole slightly in BA-trypsin. A corresponding positive region is not seen in the map and has presumably been relayed to the surface of the molecule by compensatory movements around Asp102. The side chains of Leu99 and Val227, for example, rotate about their C$_{a}$-C$_{b}$ bonds, such that their hydrophobic side chains move closer to the indole ring of Trp215. A solvent molecule not previously observed in DIP-trypsin is hydrogen-bonded to the nitrogen of the indole ring in BA-trypsin.

In BA-trypsin, as in DIP-trypsin (Stroud et al., 1974), one of the oxygen atoms on the carboxyl group of Asp102 is hydrogen-bonded to two backbone N–H groups (56 and 57), while the other oxygen forms a hydrogen bond to the hydroxyl of Ser214. The carboxyl group is shielded from solvent by the side chains of Ile99, Tyr94 and Leu57, such that there is no direct access to solvent. Although there is room for one, or perhaps two, water molecules to bind in a cavity behind Ile99, it would be impossible for a water molecule in this cavity to hydrogen-bond to Asp102 without disrupting the hydrogen bond to Ser214.

The γ-oxygen of Ser195, no longer tethered to the DIP group, is free to move; however, the large negative DIP density obscures any sign of the γ-oxygen position in the difference map. There are two almost equal and separate peaks for the hydroxyl group in the (2 BA-trypsin–DIP) electron density map. One of these lies close to the position found in stable acyl enzymes, indole acryloyl and carbamyl chymotrypsins (Henderson, 1970; Robillard et al., 1972). The second site involves an upward rotation of 120° about the $\alpha$–$\beta$ bond of the serine, making it close to the position found
for Ser195 in native α-chymotrypsin at pH 4.5 (Steitz et al., 1969). Based on our assigned position for His57, the hydroxyl group of Ser195 would form a more favorable hydrogen bond to the imidazole in the first (acyl enzyme) position, as shown in Figure 4. In spite of the clarity of the imidazole plane in density maps (Fig. 8 of this paper, and Plate IV of Stroud et al., 1974), exact assignment of the tilt of a ring within five or ten degrees from a 2.7 Å map is always open to some small unpredictable error, which should be kept in mind. However, this structure differs significantly from what we might expect to find for native trypsin, based on the structures of other native serine proteases. The uniquely high autolysis rate in trypsin has delayed preparation of native trypsin crystals; however, the crystallographic analysis of native trypsin at pH 7.5 is now under way.

It seems most probable that the side chains of Asp102, His57 and Ser195 in native trypsin are arranged as they are in native chymotrypsin. If transition to the acyl-like position is a direct consequence of binding the side-chain analogue benzamidine, it seems reasonable that it could also be a consequence of binding substrate side chains during formation of the Michaelis complex at pH 8.0.

(c) Other differences

There seems to be an overall increase in the number of solvent molecules specifically bound to the surface of BA-trypsin compared with DIP-trypsin, which may be due in part to slightly higher salt concentrations in BA-trypsin crystals. On the surface of the protein the side chains of Lys145, Asn25 and Val120 move slightly outward, and there are small shifts along the backbone chain between Tyr94 and Leu99. Solvent molecules that were bound between Lys204 and Cys202, between Lys109 and Ser84, and in a small crevice between Tyr172 and the cystine bridge 182-168 are no longer bound in BA-trypsin, while newly bound solvent appears near Ser96 and Ser110. Within the globular structure there is a small inward shift of the chain Ile47-Ser45, with the hydroxyl group of Ser45 rotating approximately 180° and Pro198 tilting slightly to accommodate the hydroxyl groups' new orientation.

There is a crevice in the surface of the molecule behind the specificity pocket and between Gly188 and Lys145 that is normally filled by a solvent molecule hydrogen-bonded to the carbonyl group of Cys220 (in one HgBr2 derivative of trypsin, a heavy-atom group fills this crevice, Stroud et al., 1974). When the lining of the specificity pocket moves in towards benzamidine, the molecule in the surface crevice is lost.

A solvent molecule W1 is hydrogen-bonded to the ε-nitrogen of His57, the amide moiety of Gln 192 and is at hydrogen bond distance from the hydroxyl of Ser195 in BA-trypsin (Fig. 5). Another solvent molecule, W2, lies 4.4 Å above W1.

4. Discussion

Benzamidine binds in the specificity pocket of trypsin, which normally binds the positively charged side chains of lysine or arginine during proteolysis. How much can we now infer about the binding of real substrates by trypsin? If BA-trypsin is a good model for side-chain binding, why is BA-trypsin an inhibitor even for non-specific substrates?
The binding pocket and specificity of trypsin: a model for substrate binding

The entrance to the binding pocket on the enzyme is a slit that provides enough room for the entry of a lysine or arginine side chain. Binding of benzamidine leads to induced changes in the specificity pocket that involve displacement of the γ-oxygen of Ser190, and the closure of the pocket to sandwich the benzamidine molecule between the peptide planes of 190–191 and 215–216. These close contacts are presumably made between the methylene groups of lysine and arginine and the enzyme in normal substrate binding. Benzamidine contains an amidinium cation, coplanar with the phenyl ring, which closely resembles the guanidinium cation of an arginine side chain (Fig. 1). Therefore, we built a model of an arginine amide bound to trypsin, such that the guanidinium cation superposed exactly onto the position of the amidinium moiety in BA-trypsin. The carbonyl carbon of arginine, which forms an acyl bond to the enzyme via Ser195 during catalysis, is brought to within 0.3 Å of the phosphorus position of DIP. The phosphorus is covalently bound to Ser195 in DIP-trypsin. This model allows the carbonyl group of arginine to adopt a position very close to that postulated for a “good” substrate (Henderson, 1970).

In a second model experiment, the model substrate was built from the serine-bound end, using bond lengths and angles compatible with an acyl enzyme in which the serine γ-oxygen is bonded to the carboxyl carbon of arginine. The arginine side-chain was constricted to lie in the benzamidine plane. In this case, the guanidinium cation ends up rotated slightly with respect to the amidinium group of benzamidine (see Fig. 9). Slight readjustments of the arginine side-chain confor-

![Diagram](image-url)

**Fig. 9.** A section through the difference map (BA-trypsin–DIP-trypsin). The position of benzamidine is shown in open line. An arginine side chain (A) of an acyl enzyme is placed so that the carbonyl carbon is at the same position as the phosphorus (P) of DIP-trypsin. The side chain is assumed to be planar, bringing the cation very close to the benzamidine cation. The scarp between benzamidine and DIP is clearly visible.
mation to preserve cation-anion interactions lead to a model in which N$_{(2)}$ and the carboxyl of Asp189 are about 2.9 Å apart.

In a third model-building experiment, the structure of the bovine pancreatic trypsin inhibitor (Huber et al., 1970) was used in conjunction with the BA-trypsin structure to generate a model for substrate binding. An arginine side-chain was built into the inhibitor structure replacing the side chain of Lys15. Models of trypsin and the modified inhibitor were brought together such that the “Arg 15” side-chain was superposed onto the benzamidine density in the difference map (Fig. 9). The side chain of Gln192 was moved to accommodate the inhibitor. Otherwise, the trypsin active-site configuration was that of BA-trypsin. The resulting model complex

<table>
<thead>
<tr>
<th></th>
<th>z</th>
<th>y</th>
<th>z</th>
<th>Atom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro13</td>
<td>0.487</td>
<td>0.137</td>
<td>0.745</td>
<td>13CA</td>
</tr>
<tr>
<td></td>
<td>0.489</td>
<td>0.185</td>
<td>0.728</td>
<td>13C</td>
</tr>
<tr>
<td></td>
<td>0.488</td>
<td>0.180</td>
<td>0.753</td>
<td>13O</td>
</tr>
<tr>
<td>Cya14</td>
<td>0.469</td>
<td>0.167</td>
<td>0.720</td>
<td>14N</td>
</tr>
<tr>
<td></td>
<td>0.470</td>
<td>0.158</td>
<td>0.714</td>
<td>14NH</td>
</tr>
<tr>
<td></td>
<td>0.469</td>
<td>0.191</td>
<td>0.711</td>
<td>14CA</td>
</tr>
<tr>
<td></td>
<td>0.468</td>
<td>0.180</td>
<td>0.691</td>
<td>14CB</td>
</tr>
<tr>
<td></td>
<td>0.496</td>
<td>0.200</td>
<td>0.708</td>
<td>14C</td>
</tr>
<tr>
<td></td>
<td>0.513</td>
<td>0.186</td>
<td>0.706</td>
<td>14O</td>
</tr>
<tr>
<td>“Arg”15</td>
<td>0.495</td>
<td>0.223</td>
<td>0.709</td>
<td>15N</td>
</tr>
<tr>
<td></td>
<td>0.487</td>
<td>0.231</td>
<td>0.710</td>
<td>15NH</td>
</tr>
<tr>
<td></td>
<td>0.519</td>
<td>0.235</td>
<td>0.704</td>
<td>15CA</td>
</tr>
<tr>
<td></td>
<td>0.520</td>
<td>0.258</td>
<td>0.715</td>
<td>15CB</td>
</tr>
<tr>
<td></td>
<td>0.518</td>
<td>0.255</td>
<td>0.739</td>
<td>15CG</td>
</tr>
<tr>
<td></td>
<td>0.542</td>
<td>0.264</td>
<td>0.751</td>
<td>15CD</td>
</tr>
<tr>
<td></td>
<td>0.545</td>
<td>0.252</td>
<td>0.770</td>
<td>15NE</td>
</tr>
<tr>
<td></td>
<td>0.552</td>
<td>0.261</td>
<td>0.786</td>
<td>15CW</td>
</tr>
<tr>
<td></td>
<td>0.553</td>
<td>0.285</td>
<td>0.787</td>
<td>15N2</td>
</tr>
<tr>
<td></td>
<td>0.556</td>
<td>0.248</td>
<td>0.801</td>
<td>15N1</td>
</tr>
<tr>
<td></td>
<td>0.520</td>
<td>0.242</td>
<td>0.684</td>
<td>15C</td>
</tr>
<tr>
<td></td>
<td>0.531</td>
<td>0.256</td>
<td>0.678</td>
<td>15O</td>
</tr>
<tr>
<td>Ala16</td>
<td>0.509</td>
<td>0.227</td>
<td>0.672</td>
<td>16N</td>
</tr>
<tr>
<td></td>
<td>0.503</td>
<td>0.214</td>
<td>0.675</td>
<td>16NH</td>
</tr>
<tr>
<td></td>
<td>0.508</td>
<td>0.229</td>
<td>0.650</td>
<td>16CA</td>
</tr>
<tr>
<td></td>
<td>0.483</td>
<td>0.222</td>
<td>0.643</td>
<td>16CB</td>
</tr>
<tr>
<td></td>
<td>0.530</td>
<td>0.215</td>
<td>0.639</td>
<td>16C</td>
</tr>
<tr>
<td></td>
<td>0.547</td>
<td>0.201</td>
<td>0.648</td>
<td>16O</td>
</tr>
<tr>
<td>Arg17</td>
<td>0.529</td>
<td>0.215</td>
<td>0.618</td>
<td>17N</td>
</tr>
<tr>
<td></td>
<td>0.521</td>
<td>0.222</td>
<td>0.614</td>
<td>17NH</td>
</tr>
<tr>
<td>Gln192</td>
<td>0.601</td>
<td>0.204</td>
<td>0.699</td>
<td>192CB</td>
</tr>
<tr>
<td></td>
<td>0.612</td>
<td>0.194</td>
<td>0.680</td>
<td>192CG</td>
</tr>
<tr>
<td></td>
<td>0.590</td>
<td>0.170</td>
<td>0.682</td>
<td>192CD</td>
</tr>
<tr>
<td></td>
<td>0.569</td>
<td>0.170</td>
<td>0.671</td>
<td>192NE1</td>
</tr>
<tr>
<td></td>
<td>0.597</td>
<td>0.155</td>
<td>0.692</td>
<td>192OE2</td>
</tr>
</tbody>
</table>
was essentially the same as that proposed earlier by Stroud et al. (1971). Bond lengths of some of the previously predicted hydrogen bonds are listed in Table 6, and the structure of the complex around the active center of trypsin is shown in Figure 10.

This model for enzyme-substrate interaction embodies a substrate conformation that evolved to bind tightly to the enzyme, and an enzyme conformation which is presumably like that induced by binding of specific substrate side-chains. Benzamidine binding is, therefore, a good model for substrate side-chain binding, showing a number of interactions that we expect to be present during the specific interaction between real substrates and trypsin.

This important correspondence helps to explain why the length of the basic side-chain, as well as its charge, is important in catalysis. When the side chains of arginine or lysine esters are shortened or lengthened by a methylene group (Baines et al., 1964; Baird et al., 1965), the rate of tryptic hydrolysis drops markedly. The BA-trypsin structure explains this sharp specificity. The β-carboxyl group of Asp189 is tied to a fixed position by the orientation of the peptide chain and by hydrogen bonding. There is no movement of this carboxyl group on binding benzamidine; thus, even though there are changes in the binding pocket region, the position of the critical charge on Asp189 is stabilized in the specificity pocket and the orientation of this carboxyl group is unperturbed on binding. If an ester or amide of a basic aliphatic amino acid with a side chain one methylene group shorter, or one longer, than normal arginine or lysine side-chains is attacked by trypsin, the coulombic attraction between the positive side-chain and the negative carboxyl group of Asp189 will either pull or push the susceptible bond of the substrate away from the position and orientation required for optimal catalysis.

Data for the hydrolysis of various ester analogues of benzamidine and phenyl guanidine (Mares-Guia et al., 1967) shed more light on the relation between the length of the side chain and catalytic activity, and lead to the questions of how benzamidine inhibits trypsin, and how the structure of BA-trypsin relates to the
free enzyme and the enzyme substrate complex with specific substrates. Mares-Guia et al. (1967) report that trypsin catalyzed the hydrolysis of ethyl and methyl esters of \( p \)-guanidino and \( p \)-amidinobenzoic and phenyl acetic acids. The distance between the charged groups and the ester moiety in true substrates is more closely approximated in the phenyl acetates than in the benzoates (see Fig. 1), and the phenyl acetates are in fact hydrolyzed more rapidly. From the model of BA-trypsin (Fig. 5) the ester group of trypsin-ethyl \( m \)-amidinobenzoate would be forced to project away from the enzyme outward into the solvent, and this explains the observation that even though this ester binds more tightly than its \( p \)-para analogue, it is not hydrolyzed by trypsin. If trypsin can still hydrolyze esters of compounds resembling benzamidine, does benzamidine competitively inhibit trypsin by merely filling up the binding pocket and thereby sterically interfere with substrate binding, or do induced structural changes in the enzyme itself contribute to inhibition?

(b) Benzamidine inhibitor action

Model building shows that the top of the phenyl ring in benzamidine (\( C_{(3)} - C_{(4)} - C_{(5)} \)) lies too close to the \( \gamma \)-oxygen of Ser195 to permit substrates to approach and form an acyl-serine. For example, a prohibitively close contact (\(< 2 \ \text{Å}\)) would occur between \( C_{(3)} \) benzamidine and the \( \alpha \)-carbon of a glycyrl substrate bound in the manner shown in Figure 10.

While large alkyl amines, guanidines and amidines competitively inhibit trypsin\(^\dagger\), smaller amines (methyl, ethyl and \( n \)-propyl) and guanidines (methyl and ethyl), which are also competitive for specific substrates, promote the slow hydrolysis of the non-specific substrate acetyl glycine ethyl ester up to 11 times of that without these side-chain analogues. Far from contributing to inhibition, as is the case with some of the larger cations, these small positively charged molecules enhance the apparent reactivity of trypsin (Inagami & Murachi, 1964; Inagami & York, 1968; Erlanger & Castleman, 1964; Mares-Guia & Shaw, 1965). Thus, the BA-trypsin structure and these data indicate that benzamidine and other large amines, amidines

\(^\dagger\) Erlanger & Castleman (1964) have observed the tryptic hydrolysis of acetyl glycine ethyl ester activated by 2-aminoheptane and other large alkyl amines. It is possible that these large amines no longer bind in the specificity pocket, but rather at some second site (Sanborn & Bain, 1968), and they might affect catalysis in the same way that excesses of specific substrates act (substrate activation) (Trowbridge et al., 1963).
and guanidines competitively inhibit trypsin by sterically hindering substrate accessibility, not only to the specificity pocket but also to the active site itself.

(c) Histidine 57- aspartic acid 102

It is clear from the density maps that the δ-nitrogen of His57 in BA-trypsin or DIP-trypsin points directly towards the center of the carboxyl group of Asp102 (see Plate IV of Stroud et al., 1974) rather than to the lower carboxyl oxygen (as found in the α-chymotrypsin structure of Birktoft & Blow (1972) at pH 4-5). Our analysis was done at pH 8-0, where the native enzyme would be active. M. Hunkapiller, et al. (1973) have shown that the pKₐ of the equivalent residue to Asp102 in α-lytic protease (a bacterial serine protease related to the trypsin family, showing similar kinetics to elastase) is 6-7; therefore, the enzyme is inactive until Asp102 becomes charged (above pH 6-7). Consequently, the interaction we observe between the neutral His57 and the ionized Asp102 is one in which the protonated δ-nitrogen interacts with the delocalized negative charge on the carboxylic acid.

Structure analysis carried out in the acid pH range (where the enzyme is inactive) show interaction between a neutral Asp102 and neutral imidazole. (The imidazole is neutral above pH 4-5, Hunkapiller et al. (1973.) Thus Asp102 is the base during catalysis and its pKₐ, rather than that of His57, is responsible for the low limb of the pH activity profile in the serine proteases. The imidazole presumably acts as a proton relay between Asp102 and the substrate or solvent, and remains neutral throughout catalysis. Therefore, the effective pKₐ of the Asp102-His57 system as seen in kinetics is the pKₐ of the better base, Asp102 (pKₐ = 6-7). The mechanistic importance of this assignment is that no unfavorable charge separation is required during catalysis. In a proposed mechanism incorporating this assignment of pKₐ values (Stroud & Krieger, 9th Int. Congr. Biochem., 1973), negative charge is developed on the tetrahedral intermediate as negative charge is neutralized on the carboxyl group of Asp102. At no stage is the imidazole required to be positively charged, so smoothing the barriers between intermediate states during hydrolysis.

If, on the other hand, the pKₐ of Asp102 were low, or even normal for aspartic acid in solution, protonation of Asp102 would be unlikely even during the hydrolysis reaction. Then the role of Asp102 in enhancing enzymatic rate would be difficult to explain in other than purely structural terms. For example, its role might be to maintain and stabilize the orientation of His57 during catalysis. However, most other assignments of pKₐ values from kinetic data are inherently incapable of distinguishing between the pKₐ values of Asp102 and His57. Likewise, studies of the net proton release upon denaturation (Ferscht & Sperling, 1973) or competitive labeling techniques (Cruickshank & Kaplan, 1973; Beeley & Neurath, 1968) can only conclude that the Asp-His system has an apparent pKₐ near neutrality and, therefore, that one of these groups loses a proton as the pH is raised through neutrality.

Thus the presence of a carboxyl group of high pKₐ and a neutral side-chain of His57 with a low pKₐ suggests two compelling new evolutionary reasons why the Asp-His-Ser arrangement should be universal to serine proteases. First, by neutralizing a negative charge on Asp102, rather than generating a positive charge on His57 during formation of the tetrahedral intermediate, there is no unfavorable charge separation. This contributes to reducing transition-state internal energies, and so to rate enhancement. Second, if the charged Asp102 is to be a proton acceptor at physiological pH values, its pKₐ must be raised and it must have access to the proton
donor. The imidazole of His57 is ideally suited to both insulate Asp102 from solvent (so raising the $pK_a$ of the buried carboxyl group) and to serve as a proton conductor, transferring charge from the carboxyl group to the substrate. It is also important to note that both the reverse separation of the $pK_a$ values of Asp102 and His57 and the structure which shows a symmetric interaction between the charge on Asp102 and His57, are unlike the situation in aqueous solution and more like interactions and proton affinities expected for groups in the gas phase or in hydrophobic solvents. The protein environment of Asp102 is responsible for these effects.

(d) The role of glutamine 192

The side chain of Gln192 covers the entrance to the specific binding pocket in BA-trypsin and forms a hydrogen bond to W1 (Fig. 4), which in turn is hydrogen-bonded to the $\varepsilon$-nitrogen of His57 or to the $\gamma$-oxygen of Ser195 in the “down” position. This orientation is completely different from that found in DIP-trypsin. If Gln192 adopts the BA-trypsin conformation in native trypsin (a lid to the binding pocket) then it must move away during substrate binding. It is also possible that Gln192 might “pack” against a substrate once it has formed an enzyme–substrate complex. This residue (192) is glutamine in five different trypsin species sequenced so far (listed by Stroud et al., 1971), whereas in bovine chymotrypsin it is a methionine residue. Perhaps this substitution serves to offer a polar environment to exchange with polar side-chains in trypsin substrates, and a non-polar, hydrophobic environment for side chains of chymotrypsin substrates. A reorientation similar to that found for Gln192 has been observed in chymotrypsin (Met192) by Steitz et al. (1969).

The side chain of residue 192 might also have an important role in stabilizing the conformation of the specificity pocket region in the zymogen precursors of trypsin and chymotrypsin. In the crystal structure of chymotrypsinogen (Freer et al., 1970), the polypeptide chain lining the binding pocket is rotated away from the chymotrypsin orientation, and the side chain of Met192 fills the pocket. In the zymogen Met192 could be stabilizing the zymogen pocket through “hydrophobic” interactions. Gln192 could play a similar role in trypsinogen by rotating downward, and possibly hydrogen-bonding to Asp189. (The structure analysis of bovine trypsinogen is nearing completion in our laboratory.)

(e) Side-chain activation of trypsin

The rate of hydrolysis of specific substrates by trypsin is much faster than the rate for non-specific amides or esters (Inagami & Mitsuda, 1964); however, a marked increase in the rate of hydrolysis of acetyl glycine ethyl ester, a non-specific substrate, accompanies the binding of small alkyl amines or guanidines to the trypsin specificity pocket. Furthermore, this effect is primarily an effect on $k_{cat}$, with a negligible effect on $K_m$ (Inagami & Murachi, 1964). Inagami & York (1968) also showed that the binding of these small cations shifts the lower limb of the pH profile of acetyl glycine ethyl ester hydrolysis to lower pH values. For example, on binding methyl guanidine, the observed $pK_a$ shifts by 0-4 unit. On the basis of the recent determination of $pK_a$ values by Hunkapiller et al. (1973), we assume that this group with a $pK_a$ around neutrality is Asp102. The $pK_a$ change contributes to a rise in $V_{max}$; however, even after correction for this effect, methyl guanidine or ethyl amine still raise $V_{max}$ by factors of approximately 4 and 11, respectively, clearly implicating

18
other induced effects. Ester substrates of chymotrypsin show a similar relation between binding affinity and the \( pK_a \) of Asp102, indicating that charge interactions are insufficient to explain the \( pK_a \) change. If the binding of benzamidine to trypsin induces a similar activation state, which is masked by its action as a steric inhibitor, the BA-trypsin structure should resemble the activated state, and might suggest how the alkyl cations activate hydrolysis of non-specific substrates. When the structure of native trypsin (work in progress) is completed, the changes induced by benzamidine (or the small cations) can be assessed directly; however, there is a good deal of evidence indicating that induced structural changes are involved. The masking of tyrosine and tryptophan residues in BA-trypsin, or \( n \)-butylamine-trypsin are examples (Villanueva & Herskovits, 1971; D’Albis & Béchet, 1967). Lysine residues are also affected by benzamidine binding (Beaven & Gratzer, personal communication). Following this evidence, it may be that one function of side-chain binding on specific substrates is to reposition parts of the enzyme structure so as to enhance the catalytic efficiency. In any event “induced fit” (Koshland, 1968) is the most reasonable explanation of the small alkyl amine or guanidine effect on hydrolysis of non-specific substrates by trypsin.

The rate-limiting step for hydrolysis of esters by trypsin is deacylation. Therefore, the primary effect detected by Inagami & York (1968) is one on the deacylation rate-constant. In order to assess the effect of small cations on the acylation rate-constant, Chambers & Stroud (unpublished results) have studied the effect of alkyl cations on hydrolysis of other non-specific substrates. We have shown that guanidinium also activates hydrolysis of \( p \)-nitrophenyl acetate by trypsin, and that a major part of this activation is due to an increase in the deacylation rate-constant. The preliminary indication is that changes in internal energy of the acyl enzyme could explain the major contribution to side-chain activation.

Other explanations are less likely. For example, small alkyl amines or guanidines might stabilize an active form of the enzyme versus an inactive one, thus apparently raising the catalytic rate-constant. An equilibrium such as this has been demonstrated for chymotrypsin (Ferscht & Requena, 1971), where the transition from inactive to active enzyme is correlated directly with the formation of a buried salt bridge between the \( \alpha \)-amino group of Ile16 and the carboxyl group of Asp194. At neutral pH, about 10% of the enzyme exists in the inactive form, which does not bind specific substrates and in which the amino-terminal is assumed to be outside of the molecule (Ferscht, 1972). Trypsin has a similar internal salt bridge between Asp194 and the \( \alpha \)-amino group of Ile16, although a similar conformational equilibrium has not yet been established. However, we expect the salt bridge to be much stronger in trypsin than it is in \( \delta \) or \( \alpha \)-chymotrypsin (thus leading to a much smaller proportion of “inactive” enzyme), because the apparent \( pK_a \) defining the high pH limb of the pH-activity profile associated with the \( \alpha \)-amino group of Ile16 is more than 1.5 \( pK_a \) units above that found for chymotrypsin. Typical values are about \( pK_{app} \) 8.5 to 8.9 for chymotrypsin, and \( pK_{app} \) of 10.1 for hydrolysis of \( \alpha \)-N-benzoyl-L-argininamide by trypsin (Spomer & Wootton, 1971). The indication is that the salt bridge is much stronger in trypsin, and, therefore, that the proportion of inactive enzyme, where the \( \alpha \)-amino terminus is no longer constrained in the salt bridge, is much smaller.

The binding of indole and DIP to chymotrypsin (Hess et al., 1970) and the binding of \( n \)-butylamine (D’Albis & Béchet, 1967) or DIP (Ghelis et al., 1967) to trypsin appear
SPECIFIC BINDING OF TRYPsin

to stabilize the enzymes in a conformation similar to the active, or "in" form. By
pushing the conformational equilibrium toward the active enzyme, the specific side
chain analogue activators could produce an increase in $V_{\text{max}}$ for non-specific sub-
strate reactions, but could not in any case explain a differential effect on deacyla-
tion or acylation rate-constants as we observe for hydrolysis of nitrophenyl acetate.
Furthermore, this effect cannot be responsible for most of the observed increase in
$V_{\text{max}}$, for even if the active-inactive trypsin equilibrium were similar to that of
chymotrypsin, only 10% of the enzyme would be affected, clearly less than required
for a factor of 11 in $k_{\text{cat}}$. Also, one would expect those alkyl cation effectors that
bind most strongly to trypsin's specificity pocket to have the greatest effect on the
conformational equilibrium; however, alkyl amines do not bind as tightly to trypsin
as do their corresponding guanidines, yet the amines of corresponding size are better
activators.

In summary, it seems inescapable that the small alkyl amines or guanidines induce
small changes in the catalytic site of trypsin, and that these changes are responsible
for activating the enzyme and affecting $k_{\text{cat}}$. The specific side chains of real substrates
may be responsible for similar rearrangements. It seems likely that the structure of
BA-trypsin resembles an "activated" species, although the enzyme is inhibited
by steric hindrance of productive substrate binding. If this is indeed true, we might
expect the side chains of real specific substrates to induce similar effects in the
enzyme.

We thank Robert Huber for sending us co-ordinates of the trypsin inhibitor, and a
preprint of some of his results. We are grateful to Lalit Samson for preparation of the
ORTEP drawings in Figs 5, 6 and 7, and to Michael Hunkapiller, Steve Smalcombe and
John Richards for permission to cite some of their results. We thank Sten Samson for
his help and advice in rebuilding our diffractometer system.

This work is contribution no. 4553 from the Norman W. Church Laboratory of Chemical
Biology of the California Institute of Technology. This work has been carried out with
the support of the United States Public Health Service (grant nos GM12121 and GM19984),
whose help is gratefully acknowledged. One of us (M. K.) is the recipient of a Danforth
Graduate Fellowship, and another (R. M. S.) is the recipient of a National Institutes
of Health Career Award (GM70469).

REFERENCES