The "active site" is a concept equally applicable to enzymes and antibodies. Although it is not simple to define active site precisely, the term is extremely useful in indicating that fraction of the total protein which is most intimately involved in its function. The active site of enzymes is, therefore, pertinent to this symposium since there is reason to believe that the "specificity" part of the active site may be quite similar in antibodies and enzymes. Binding constants for substrates are frequently of the same order of magnitude as binding constants for haptens. Specificity patterns for substrates are similar to specificity patterns for antigens. Therefore there is reason to believe that studies on the properties of the active sites of enzymes may develop principles and techniques which will be helpful in understanding antibody action.

In this paper some studies on enzyme properties which are believed to be pertinent to antibodies will be discussed. The particular studies chosen are: a) the evidence that residues far removed from each other in the sequence play vital roles at the active site; b) indications on the size of
LEGAL NOTICE

This report was prepared as an account of Government-sponsored work. Neither the United States, nor the Contractor, nor any person acting on behalf of the Contractor:

1. Makes any warranties or representations, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use thereof will not infringe on any private rights.

2. Assumes any responsibility for loss to the user, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in this notice, "government or behalf of the United States" means the United States Government, its employees or contractors of the United States, or employees of such contractors, to the extent that such employees or contractors of the United States, or employees of such contractors possessed, communicated, or provided access to, any information pursuant to this employment or contract with the United States, or the employment with such contractor.
the active site; and c) evidence for flexibility at the active site. Since the relation between enzyme and antibody is of primary interest an attempt will be made to emphasize the fundamental conclusions derived from the observations and to evaluate critically the deduction on structure and function which derive from these conclusions.

**Roles of Methionine, Serine and Histidine in Chymotrypsin Action**

A number of years ago it was found that diisopropylfluorophosphate (DFP) inhibits a number of esterases of which chymotrypsin was an outstanding example (1). The group that was phosphorylated by the DFP was found to be serine (2) and its position in the amino acid chain of chymotrypsin has now been determined (3).

Also in these early studies preliminary indication of histidine involvement in chymotrypsin was obtained from pH studies in which activity could be correlated with a group having a pK of 7 (4). Since pH data is ambiguous for identifying groups (5) particularly in the pH 7 region, we sought more definitive data using photooxidation, a technique originally applied by Weil and co-workers (6).

In Fig. 1 the effect of photooxidation on the histidines residues of chymotrypsin is shown. A biphasic curve is obtained which can be resolved by appropriate mathematical methods (7) into two components, one of which oxidized with a rate constant of 0.49 min\(^{-1}\) and a second which is oxidized at a rate approximately 1/10 of this value. Since there are two histidine residues in the molecule of chymotrypsin, the differing reactivity of these
two amino acids can be explained by the assumption that one of the histidines is on the surface of the protein whereas the other is buried in its interior and thereby made inaccessible to the photooxidation reagent.

When methionine was followed during photooxidation the curve shown in Fig. 2 was obtained. There are two methionine residues in the molecule and again it is noticed that apparently one methionine residue is in the interior of the protein and the second methionine residue is exposed to the photooxidation reagent.

If the overall activity loss of the protein is now compared with the rate of loss of these amino acids (Fig. 3) the identification of the amino acid residues responsible for activity loss can be ascertained. Mathematical analyses (8,9) showed that in chymotrypsin modification of the "surface" histidine produces an "inert" protein, i.e. one with less than 1% of the initial activity and that oxidation of the surface methionine produces a partially active protein, i.e. one with an appreciable fraction of the initial activity. Thus we were able to identify the type of residues which were involved in the activity modification, their relative contributions to the activity, and their approximate positions in the tertiary structure of the protein.

To add further support to the conclusion reached from photooxidation, we next looked for a reagent that would attack methionine alone. On incubation for 1400 minutes with 0.001 M hydrogen peroxide, the activity of the protein decreases and finally levels off at a figure of approximately 37% of the initial activity (cf. Fig. 4). A parallel loss in methionine is observed
over this interval finally leveling out at 50% or one residue of methionine modified. Further additions of peroxide did not produce any further change in methionine loss or activity loss. No other amino acid was affected by the peroxide. Alkylation with iodoacetic acid gives similar results. The demonstration of a partially active species in which the surface methionine was modified were particularly pleasing because it indicated that the interpretation of the roles of histidine and methionine discussed above was indeed correct. Further support for these conclusions required the establishment of the identity of the two methionines modified. This was done by amino acid sequence analyses which established that the same methionine was oxidized by $\text{H}_2\text{O}_2$ and photooxidized (10).

If we examine these results in the light of the primary structure of chymotrypsin as deduced from the sequence studies of Hartley (3), certain interesting facts emerge. In Fig. 5 the primary structure is shown schematically in which the sequences of amino acid residues are indicated by the heavy black lines, the lengths being approximately proportional to the number of residues in the chain. The uppermost chain with cystine N-terminal has 13 residues and is called the A chain. The middle chain with isoleucine N-terminal has 143 residues and is called the B chain. The remaining chain, the C chain, with alanine N-terminal, has 80 residues. The serine position is known quite accurately as indicated near the C-terminal end of the C chain. The position of the histidines are not known as accurately but they are in the B chain and somewhere near the middle of that chain. The kinetic analysis of Fig. 1 indicates that one of these two histidines is near the surface of the protein and the other more or less buried in the interior of
the molecule. The surface histidine is the residue which has been proven to be essential for the activity of the protein. and serine are

It might still be argued that histidine, not at the active site but at a distant position whose modification leads to unfolding of the protein. The recent experiments of Schoellman and Shaw (11) have helped to exclude this alternative. Using an alkylating agent which is partly analogous to a substrate of chymotrypsin, they have managed to alkylate a histidine residue. Since this reagent is "steered" to the active site by the portion which is analogous to the substrates of chymotrypsin, it is evident that the histidine which is alkylated is geometrically close to the specificity amino acids of the active site. This information can now be added to the previous observation that acetylation of the serine residue prevents absorption of the substrates of chymotrypsin to the active site (12). Hence both serine and histidine must be close to the specificity amino acids and to each other in three-dimensional space.

If we turn for a moment to the methionine residues, we see from Hartley's sequence

<table>
<thead>
<tr>
<th>17</th>
<th>16</th>
<th>15</th>
<th>14</th>
<th>13</th>
<th>12</th>
<th>11</th>
<th>10</th>
<th>9</th>
<th>8</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>Ala</td>
<td>Met</td>
<td>Ileu</td>
<td>Cys</td>
<td>Ala</td>
<td>Gly</td>
<td>Ala</td>
<td>Ser</td>
<td>Gly</td>
<td>Val</td>
<td>Ser</td>
<td>Ser</td>
<td>Cys</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>Gly</td>
<td>Asp</td>
<td>Ser</td>
<td>Gly</td>
<td>Gly</td>
<td>Pro</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

that one of these is only 3 residues away from the serine of the active site while the other is 15 residues away. One of these residues is followed by glycine and the other by isoleucine. This sequence information combined with
the cyanogen bromide procedure of Gross and Witkop (13) allowed us to identify which residue was being oxidized. It was found that both \(H_2O_2\) and photooxidation were oxidizing the methionine three residues from the active serine (10). Either oxidation produced a protein of 37% of the initial activity. The loss in activity could be accounted for entirely by a change in the \(K_M\) rather than the \(V_M\) of the enzyme. Thus, this conversion of a highly hydrophobic residue to a polar residue occurs with only modest change in the activity of the enzyme.

If one takes the maximum extension of the peptide chain, the distance between the sulphur of this methionine and the serine of the active site would be approximately \(13 \, \text{Å}\). Recently Lawson and Schramm (14) have been able to alkylate this methionine using a compound which acetylates serine prior to alkylation of methionine. Using the maximum extension of this compound a distance of \(7 \, \text{Å}\) between the serine and methionine is obtained. It indicates that the maximum distance apart of the serine oxygen and the methionine sulfur is \(7 \, \text{Å}\). The two residues could be considerably closer together, for example, as little as \(2 \, \text{Å}\) if they are in direct proximity. However, using these data we can see that in at least one direction from the serine residue at a distance of \(7 \, \text{Å}\) at a maximum, a very profound change in the nature of one of the chemical groups can be achieved without basic change in the reaction velocity of the enzyme.

**Flexibility Studies**

Even after the amino acids of the active site have been identified, the nature of the specificity process is not solved. It still remains to determine in what way the amino acid residues act to discriminate between
closely similar chemical structures. Although the details of these interactions are not understood, certain generalizations about the mechanism of this process are beginning to emerge. One of the most important of these concerns the flexibility of the active site. The classical template theory is based on a relatively rigid active site in which substrate and enzyme fit like a "key and lock" (15). In chemical terms this does not mean that all vibrations were excluded and absolute rigidity exists. Rather, molecular vibrations are presumed to occur but they are minor and not a significant feature of the specificity process. The rigidity however is essential since it provides the force which is postulated to prevent the access of some compounds to the active site.

Because of experiments which could not be explained readily by the template hypothesis, we proposed an alternative mechanism based on a flexible active site (16). In this mechanism a significant change in the conformation of the protein was a requirement for activity. Many substances might produce changes but substrates would produce a proper fit whereas nonsubstrates would not. This "induced fit" theory was able to explain many phenomenon not explainable by the template hypothesis and further experiments to develop even more definitive examples compatible with a flexible site were pursued (17).

In addition to the specificity studies there was another way of obtaining evidence for this hypothesis. The flexible site theory predicted that conformation changes would be induced by the substrate. Whether these conformation changes would be large enough to be detectable by the currently available means was of course not predictable a priori but nevertheless such studies were undertaken in the hope that sufficiently large changes in atomic positions would occur. A fairly substantial list of enzymes are now available
for which conformation changes have been reported as a result of the treatment of an enzyme with a substrate (18). Not all of these studies can be considered definitive for a variety of reasons but the common explanation of the wide variety of phenomena observed supports the flexible site explanation.

To give an example in which specific experiments were designed (19) to distinguish conformation changes from general environmental effects the muscle enzyme, phosphoglucomutase, was found: 1) to have a specificity pattern conforming to an induced fit type enzyme, 2) to show a conformation change in the presence of substrate by the most unequivocal method for demonstrating such changes, i.e. by the enhanced reactivity of an amino acid residue in the presence of substrate, 3) to show a characteristic difference spectrum in the tyrosine-tryptophan region showing a change in environment of these side chains on exposure of the enzyme to substrate.

Some typical data are shown in Table 1 and Figure 6. Treatment of phosphoglucomutase with iodoacetamide caused increased inactivation and increased reactivity of an SH group (Table 1) in the presence of glucose-6-phosphate when compared to a control using glycerophosphate. Glycerophosphate is a phosphate ester, having the same pK as glucose phosphate, but it is not a substrate for the enzyme. Since it will create an ionic environment essentially identical to that of the glucose phosphate, it is used as control to obviate the possibility that some nonspecific ionic binding effect is responsible for the increased reactivity in the presence of glucose phosphate.

A question might now arise concerning the magnitude of the change. Differences in reactivity of 20 per cent are indeed quite large, but one might prefer an all-or-none difference in rates. Although the latter is preferable,
it would certainly seem to be neither necessary nor likely. For an enzyme of molecular weight ca. 70,000 and a substrate of molecular weight of 300, it is readily seen that the number of amino acids in actual contact with the substrate must be few. In other words, an active site comprising only those amino acids in contact with the substrate must be only a very small fraction of the total enzyme surface. Moreover, the extent of the shape change necessary to bring the catalytic groups into alignment is not known. If, as seems probable, enzymes operate by very fine adjustments, changes of a fraction of an angstrom may be sufficient. Moreover, the position of the SH group is not known; it might be quite distant from the site of substrate absorption and hence a larger quantitative change at active site might result in a diminished effect at a more distant point in the tertiary structure. A true difference of 20 per cent, therefore, would seem sufficient. The agreement between the amino acid analyses for the unreactive residues of better than 2 per cent, and the repeated differences of ca. 20 per cent on replication of the glucose-phosphate-glycerophosphate experiments, indicates a truly enhanced reactivity of that magnitude.

To add further support to the chemical evidence, a physical probe of protein shape, the difference spectrum, was studied. The results of a dual beam experiment in which enzyme plus substrate was compared to enzyme plus glycerophosphate can be seen in the solid line of Fig. 6. There is a small but reproducible difference. Since the tubes contain the same concentration of phosphate ester, it is improbable that these are general media effects. To check this point further, 0.001 M glucose phosphate was added to both cuvettes.
Since the initial concentration of phosphate ester was 0.01 M in each cuvette, this addition would mean only a small change in the general environment. Hence, if the previous difference spectra were a general environment effect, little change would be expected. On the other hand, if the difference spectrum were an active site effect, the addition of glucose phosphate, which is more strongly absorbed than the glycerophosphate, would lead to a "washing out" of the difference spectrum. The dotted line of Fig. 6 shows this washing out, and indicates that the difference being observed is a consequence of an active site change.

These and related data led to the conclusion that protein conformation changes are indeed occurring and that they are occurring in a manner directly connected with the specificity process of the enzyme.

Conclusions

In closing it may be desirable to review the major conclusions obtained for enzyme sites and to extrapolate these implications to antibody action. Let us first examine the conclusion that histidine and serine are close together in three dimensional space. The fact that groups far from each other in the peptide sequence can be near each other in space is not new or surprising. What is significant is that these two groups are now known to play vital catalytic roles, i.e. their nearness is required for the function of the protein. Hence the coding of the gene must contain the information necessary to place these two groups together in a precise juxtaposition. Although we know that the whole protein structure is precise and is probably determined by the primary sequence, the possibility that active sites were always found in a limited region of a single sequence was a definite possibility. It still is such a possibility in the case of many
enzymes and may be the case in antibodies. The fact, however, that two groups so far apart in the sequence are indeed vital parts of the active site of at least one enzyme means that this type of site must be considered the most probable until further examples are obtained. It means, for example, that the active site cannot be defined by a single labeling. Independent evidence for the presence or absence of each region of the peptide chain at the active site will have to be obtained.

The evidence regarding methionine adds support to the generally accepted concept that the active site is a limited region of the large protein molecule. That this residue can have its properties changed drastically with only minor effect on the enzyme action indicates that it is at best at the periphery of the active site and that the active site in this direction extends at the most to $7\,\AA$. This at once presents hope and difficulty. Hopefully it shows that the active site will only contain a few amino acids and hence the job of identifying these will demand less than would be the case if many amino acids with gradually decreasing quantitative contribution were part of the site. Difficulties arise however since it is clear that each residue will have to be ascertained individually since even a residue adjacent to an active site residue may no longer be part of that vital region. Again it emphasizes that "labeling the active site" will require a labeling of each residue involved at the active site.

Finally the flexibility of the active site changes the mechanism of the specificity process itself. This is helpful in one sense since it allows us to explain otherwise contradictory behavior in specificity patterns.
It causes difficulties in another category, however, since it suggests that a still picture, e.g. as obtained by X-ray studies of the enzyme before interaction with substrate, will not per se explain specificity. The dynamic phases of the interaction will also require description before a full understanding of the specificity process is obtained.

An interesting illustration of the significance of these conclusions arises in interpretation of labeling studies. If it is true that 2 residues lying far apart must ultimately be aligned in close juxtaposition, then some of the intermediate residues, perhaps far removed from the two residues at the active site, must play important roles in orienting the tertiary structure of the molecule. If it is further true that substrate (or hapten) can cause conformation changes in the protein leading to exposure of residues, it is also true that the same process can cause the "burying" of other residues. The classical interpretation of a "protection" experiment is then in some doubt. A residue labeled in the absence of hapten and not labeled in the presence of hapten is conventionally presumed to be at the active site. It is possible however that the protection occurs because of a refolding of the molecule and the labeled group is actually a group made less reactive by refolding at some distant position of the tertiary structure.

Although none of these conclusions has yet been established in antibodies the specificity patterns in the two systems are so similar that some of it would be surprising if the structure-function relationship were not similar also. Hence it seems probable not only that the techniques of the enzyme studies may be helpful in antibody studies, but also that the conclusions regarding enzymes may be equally applicable to both systems.
References


Reactivity of SH Groups of Phosphoglucomutase

Conditions: 0.01 M tris or veronal buffer, pH 8.0, 10 minutes, 0.067 M iodoacetamide, 0.01 M phosphate ester, 4.6 x 10^{-5} M enzyme, amino acid residues corrected to leucine = 46.0. Column load of 1 equivalent to 3 μmoles of aspartic acid. Experimental error for unchanged residues such as phenylalanine ± 1-3%.

<table>
<thead>
<tr>
<th>Exp. No.</th>
<th>Buffer V=Veronal T=Tris</th>
<th>Treatment for Metal Ion</th>
<th>Column Load</th>
<th>Residues of SH Alkylated Per Mole of Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Glycero-P</td>
<td>Glucose-P</td>
</tr>
<tr>
<td>1</td>
<td>V</td>
<td>None</td>
<td>1</td>
<td>1.29</td>
</tr>
<tr>
<td>2</td>
<td>V</td>
<td>None</td>
<td>1</td>
<td>1.08</td>
</tr>
<tr>
<td>2</td>
<td>V</td>
<td>None</td>
<td>2</td>
<td>1.04</td>
</tr>
<tr>
<td>2</td>
<td>V</td>
<td>None</td>
<td>3</td>
<td>0.97</td>
</tr>
<tr>
<td>3</td>
<td>T</td>
<td>Dowex 50</td>
<td>1</td>
<td>0.76</td>
</tr>
<tr>
<td>3</td>
<td>T</td>
<td>Dowex 50</td>
<td>3</td>
<td>0.78</td>
</tr>
<tr>
<td>4</td>
<td>T</td>
<td>Dowex 50</td>
<td>1</td>
<td>0.93</td>
</tr>
<tr>
<td>4</td>
<td>T</td>
<td>Dowex 50</td>
<td>3</td>
<td>0.80</td>
</tr>
</tbody>
</table>
Figure 1. Loss of Histidine on Photooxidation of Chymotrypsin. Upper points show experimental determination of % histidine remaining as determined by amino acid analyzer (2 histidine residues per molecule of chymotrypsin). Time.

Final first order slope can be extrapolated to zero as shown by dotted line to give number of residues (1 in this case) which are disappearing at that rate. Subtracting the extrapolated values from the observed values gives new values (lower points) representing disappearance of remaining residue. Straightness of this line supports analysis method and allows calculation of rate constant for loss of the "surface" residue.

Figure 2. Loss of Methionine on Photooxidation of Chymotrypsin. Same type analysis as for histidine (cf. Fig. 1). There are two methionine residues per molecule of chymotrypsin.

Figure 3. Activity and Amino Acid Residue Losses of Chymotrypsin on Photooxidation. Dotted lines represent rates of loss of "surface" histidine and methionine residues under these conditions as deduced by data of Figures 1 and 2 respectively. Open circles represents a "rate" assay, i.e. a conventional assay which measures decreases caused both by "inert" enzyme and "partially active" enzyme. Filled circles represent "all-or-none" assay which measures decreases caused by production of inert enzyme only. Rate assay line is curved and initial slope has rate constant equal to sum of methionine and histidine constants. "All-or-none" assay line is straight and rate constant is equal to histidine constant.
Figure 4. Time curve of methionine and activity loss on treatment of chymotrypsin with 0.001 M H₂O₂.

Figure 5. A Schematic Version of Alpha Chymotrypsin. Only two of the 5 disulfide bonds are shown to indicate the chains are held together by these bonds. The lengths of the heavy black lines are proportional to the number of residues in the chain. The positions of the serine and methionine residues are quite precise since the sequence of the C chain is almost complete. The histidine position is only approximate since the B chain is only partially complete.

Figure 6. Difference Spectra of Phosphoglucomutase in Presence of Phosphate Esters after Pretreatment with Glucose-6-phosphate. Solid line shows enzyme plus 0.01 M glucose-6-phosphate versus enzyme plus glycerophosphate. Dotted line shows same solutions except 0.001 M glucose-6-phosphate has been added to each cuvette.
FIGURE 1
FIGURE 3
FIGURE 4

% REMAINING

TIME IN MINUTES

METHIONINE

ACTIVITY
FIGURE 5
FIGURE 6

(a) 0.001 M G-6-P vs 0.001 M G-6-P
(b) 0.01 M β-GLY-P vs 0.001 M β-GLY-P

OPTICAL DENSITY

MILLIMICRONS

210 220 230 240 250 260 270 280 290 300 310