CHEMICAL TOPOGRAPHY OF SIDECHAINS IN ENZYMES

F. B. STRAUB

Institute of Biochemistry, Hungarian Academy of Sciences, Budapest

ABSTRACT

Information concerning the structure and the structural flexibility, the motility, of polypeptide chains in an enzyme may be obtained by use of the method of sidechain modification. Evaluation of such information requires a rigorous kinetic analysis of the chemical modification process. The need to restrict the study to a single residue is emphasized and illustrated with examples. It is shown that the reaction of the second reactive SH group of triosephosphatedehydrogenase with an appropriate reagent is the result of the fluctuation of the polypeptide chain in the native enzyme. A similar reaction due to the same cause is described for the reaction of the buried SH group of pancreatic amylase. The complex analysis of the carboxymethylation of aldolase, using correlated kinetic, enzymological and sequence studies, has established in the author's laboratory, the relation of a less reactive SH group to the active site. Similar considerations have been applied to the study of SH groups involved in subunit interaction within lactatedehydrogenase.

In the lifetime of a generation the chemistry of proteins has made a full turn from the ignoramus et ignorantibus of the early 'twenties to the complete amino acid sequences and detailed three-dimensional models deduced from x-ray crystallography. During the development of these modern methods, chemical topography or surface chemistry of enzymes has contributed a great deal to our knowledge of the biologically important functional groups of enzymes. The rapid progress in enzyme chemistry has led to the belief that knowledge of the nature and position of all atoms in the enzyme molecule is enough to explain its biological function. Such reductionism is, in my view, not warranted. In our laboratory in Budapest we have collected in the past few years some data on the location of sidechains and on their role in biological activity. These results I should like to use to illustrate the potentialities of the method of sidechain modification in order to show in which respects chemical topography may contribute to the understanding of the biological function.

Chemical modification of sidechains, and some physicochemical methods specific for one kind of sidechain, have long been used to determine which of these sidechains might be on the surface of the protein and again which of them might be involved in the mechanism of enzyme action. The scattered data obtained in this way have been the only evidence available for a long time. With
the advent of x-ray crystallography, some of these conclusions have been proved to be wrong, but most of them have found their proper place in understanding the mechanism of substrate binding and catalytic activity. The reasons for which x-ray crystallography will not dispose of the necessity for chemical topography, are the following.

First, any structural model has to account for the chemical modification results, for enhanced or reduced reactivity of the individual sidechains. Secondly, it appears that the harvest of enzyme crystals amenable to x-ray crystallography is not increasing very fast. It is most dangerous to make negative forecasts, but we do not yet see how the complicated structure of proteins in the high molecular weight classes could be deciphered. Whereas the difficulties caused by the size and higher organization of the molecule increase rapidly for the x-ray crystallographer, the results of chemical topography are not so sensitive to an increase in molecular weight, as shown by results achieved on such molecules as myosin or such organizations as the red cell membrane. Finally, chemical topography is now taking a lesson from crystallography: a new chapter of chemistry is evolving, the study of the reactivity of amino acid sidechains embedded in a known three-dimensional structure. Such studies will undoubtedly help in interpreting further data with new materials, the structure of which is yet unknown.

There is, in addition, one very important aspect of the structure which has to be understood and this is the fluctuation of part of the protein molecule. Structural changes caused by the binding of substrate or inhibitors may be studied by physical methods. But this is only part of the potential motility. In my belief, enzyme activity, subunit interaction, integration of the protein in the cell structure, its regulation by ligands is best understood if we assume its flexibility, the steady fluctuation of its polypeptide chains between such limits which include a few states having meaningful physiologically active structures. These fluctuations may be best studied by the chemical approach, using a combination of sidechain modification with the study of some physical parameters.

As an example of the use of chemical modification to determine the location of some residues in the structure of an enzyme, I should like to recall the recent experiments of Libor and Elödi in our laboratory, concerning the position of tyrosine residues in the coenzyme-free protein of glyceraldehyde-3-phosphate dehydrogenase from pig muscle.

Spectrophotometric titration of tyrosines gives information about the exposed and buried residues. Further refinement is possible through the method of Herskovits and Laskowski. Perturbation difference spectra obtained in the presence of 20 per cent ethylene glycol have indicated that in our enzyme, out of the nine tyrosines present in each subunit, four are accessible to the solvent and five are buried.

An independent approach to the study of the distribution of tyrosines is the reaction with iodine in slightly alkaline milia. Exposed tyrosines are expected to be easily substituted. Moreover, in model experiments it has been shown that the formation of monooiodotyrosine and the consecutive formation of diiodotyrosine are proceeding at different rates according to the polarity of the solvent. It is therefore expected that the tyrosines which are surrounded by a non-polar region, would yield only diiodotyrosines, whereas a transient
CHEMICAL TOPOGRAPHY OF SIDECHAINS IN ENZYMES

appearance of monoiodotyrosine should be observed if the particular tyrosine is exposed to a polar environment.

When glyceraldehyde-3-phosphate dehydrogenase is treated with iodine, the first groups to react are the SH groups of cysteinyl sidechains. Actually, before any substitution occurs, these are oxidized to intra-subunit disulphides, as shown in our laboratory. If native enzyme is treated near 0°C at pH 9.2 with a limited amount of iodine (33 moles per subunit) all cystein residues will be rapidly oxidized, some tyrosines will be substituted, but no further substitutions will occur. Under these circumstances five out of the nine tyrosines become substituted.

If, at the same pH and temperature, the reaction is performed in the presence of 8 M urea, all tyrosines will be substituted. As the amino acid sequence of this enzyme is known, it is possible to identify the tyrosines which react without urea and those which react only in urea solution, by a study of the radioactive peptides obtained from the tryptic hydrolysate of the enzyme treated with labelled iodine.

Libor and Elödi have measured the formation of monoiodotyrosine and diiodotyrosine as a function of iodine concentration (Figure 1) and at high iodine concentration as a function of time (Figure 2). It is seen that the transient formation of monoiodotyrosines can be detected even if the enzyme is treated

![Figure 1](image-url)

**Figure 1.** Formation of iodo derivatives of tyrosine in glyceraldehyde-3-phosphate dehydrogenase as a function of iodine concentration. 0.50 mM enzyme in 0.15 M potassium iodide containing 0.2 M borate buffer, pH 9.2, was treated, at 0°C, for an hour with KI₃ in different concentrations. After removing excess of iodine the number of residues reacted was calculated from the absorption differences measured in 0.01 M hydrochloric acid, 0.01 M sodium hydroxide and 0.1 M borate, pH 9.2. □, unreacted tyrosine; ○, monoiodotyrosine; ×, diiodotyrosine; △, sum of tyrosyl derivatives.

P.A.C.—25/1—L
Table 1. Effect of iodine on the tyrosines of glyceraldehyde-3-phosphate dehydrogenase

<table>
<thead>
<tr>
<th>Tyrosine residues</th>
<th>Reacts with iodine before urea denaturation</th>
<th>Transient formation of monoiodotyrosine</th>
</tr>
</thead>
<tbody>
<tr>
<td>137, 252</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>39, 42, 46</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>91, 273, 311, 317</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

In this way, labelling with radioactive iodine combined with spectrophotometric and solvent perturbation methods allows one to assign the tyrosine residues of the known sequence to the outer, intermediate and core layers of the protein. However, one must be careful with the interpretation of these data. First, we have to remember that iodination in these experiments cannot be
CHEMICAL TOPOGRAPHY OF SIDECHAINS IN ENZYMES

performed without the previous oxidation of the sulphhydryl groups. In the case of glyceraldehyde-3-phosphate dehydrogenase this process leads to the formation of disulphide bonds within the subunit but to no serious change in the conformation; actually it is a process which is reversible to some extent. The second problem is that we have a number of substitutions which may take place later and the succeeding reactions may be the consequence of the conformational change brought about by the first substitution. Finally, it is conceivable that a ‘buried’ residue is exposed for part of the time to the solvent because of the fluctuation of the polypeptide chain. Such a mechanism has already been suggested for the iodination of some tyrosines⁸.

I have cited the work on the topography of tyrosines in glyceraldehyde-3-phosphate dehydrogenase because it illustrates the power of the method in yielding information and at the same time it illustrates the problems of evaluation. In order to substantiate the conclusions, the method of sidechain modification must be combined with kinetic analysis of the reaction and should be restricted to the study of one single residue at a time. As an example I should like to quote the work of Vas and Boross⁹ from our laboratory, dealing with the same enzyme.

There are four cysteinyl sidechains in each subunit of glyceraldehyde-3-phosphate dehydrogenase and it has previously been shown in our laboratory by Szabolcsi and her co-workers¹⁰,¹¹ that these SH groups react with p-hydroxymercuribenzoate with different kinetics. At pH 8.5 two of them react very rapidly, whereas the other two react with first-order kinetics independent of the concentration of p-hydroxymercuribenzoate. They have established that the denaturation blue shift measured at 295 nm follows first-order kinetics with the same rate constant as the parallel formation of these two mercaptide bonds (Figure 3). Obviously, the third and fourth cysteinyl residues become exposed and react only as a consequence of the structural change induced by blocking the first two SH groups.

\[
\text{Figure 3. Formation of mercaptide and denaturation blue shift in glyceraldehyde-3-phosphate dehydrogenase as a function of time. The enzyme was treated at 5° with 4 to 20 moles of p-OH-MB per subunit in 0.05 M tris buffer, pH 8.5. (a) mercaptide formation, measured according to Boyer}^{25}. K=0.131 \text{ min}^{-1}. (b) \text{denaturation blue shift measured at 295 nm}; k=0.136 \text{ min}^{-1}.
\]
Vas and Boross have established the conditions under which the reaction of the first two SH groups may be resolved into two phases. The mercaptide bond formation is measured at pH 7.0, using stoichiometric amounts of p-hydroxymercuribenzoate [Figure 4 (a)]. The reaction of the first SH is instantaneous, the second is rapid, the third and fourth—in agreement with former results—slower, and these latter reactions are due to the denaturation of the enzyme. If the reaction of the latter two cysteinyl residues is subtracted, the rate of reaction of the second SH group is obtained [Figure 4 (b)]. Although the rate is influenced by the concentration of p-hydroxymercuribenzoate, mercaptide formation is not a simple second-order reaction.

Actually it can be shown that the reaction of the second SH group is not connected with a conformational change caused by the reaction of the first one. One can block the rapidly reacting cysteinyl residue (this is Cys 149 in the sequence, part of the active site) with a carboxymethyl group or with one equivalent of p-hydroxymercuribenzoate. If the second SH is reacted now by the addition of p-hydroxymercuribenzoate, it will react at the same rate, even if this addition is done one hour after blocking the first SH group. The authors have advanced the explanation that the second SH, which is supposed to be Cys 153 in the sequence, is in fact a buried SH group, but one which is exposed for a fraction of the time to the solvent and solutes, due to fluctuation of the polypeptide chain.

$$E \xrightarrow{k_{-1}} E^* + p\text{-OH-MB} \xrightarrow{k_2} E\cdot p\cdot MB \xrightarrow{k_2} \text{denaturation}$$

where $E^*$ is the fraction of the enzyme [$E$] which reacts with the reagent, $k_{+1}/k_{-1}$ is the equilibrium constant, $k_2$ the second-order rate constant of mercaptide formation and $k_3$ the first-order rate constant of unfolding, which leads to the
unmasking of the third and fourth cysteinyl residues. The equation describing the rate of the reaction of the second SH has been calculated

\[ v = [E] \left( k_2 k_{41} [p-OH-MB] \right) \left( k_{-1} + k_2 [p-OH-MB] \right) \] (1)

and it has been found that the calculated times agree with the experimental data in Figure 4 (b).

A similar phenomenon was observed by us in the reaction of the masked SH groups of pancreatic amylase. It is known that pancreatic amylase contains four disulphide bonds and two masked cysteinyl residues, which would react with SH reagents only after denaturation by urea or by removal of the stabilizing Ca ions. These experiments are complicated by the fact that amylase deprived of Ca ions is very unstable in the presence of proteolytic impurities. Even after several recrystallizations amylase contains a sufficient amount of proteolytic impurities to be degraded rapidly after addition of ethylenediaminetetra-acetate. However, this proteolytic degradation may be suppressed for some time by the addition of diisopropylphosphofluoridate. Taking these precautions we have found that Ellman's reagent (2,4-dithiobis-nitrobenzoate) will react with the masked SH groups at pH 7.5 in the presence of 0.01 M ethylenediaminetetra-acetate, whereas at the same time the enzyme activity will not decrease significantly (Figure 5). When the reaction of the SH groups was nearly complete, we stopped the reaction by adding an excess of calcium ions. From this mixture we isolated an amylase which has practically full enzymatic activity but has its two SH groups in the form of a mixed disulphide with thionitrobenzoate. It is chromatographically distinct from native amylase and can be reverted to the original enzyme by the addition of mercaptoethanol.
If the enzyme alone was kept in ethylenediaminetetra-acetate solution the SH groups were exposed at a much slower rate and this was proportional to the rate of denaturation. It appears, therefore, that the removal of calcium has facilitated the fluctuation of that part of the polypeptide chain which shielded the SH groups.

The last feature of this type of investigation, concerning the role of the cysteinyl residues in rabbit muscle aldolase, should illustrate the point that an appropriate study of the environmental conditions and a study of the kinetics of sidechain modification make it possible to study the modification of the individual groups within a series of reactions.

Muscle aldolase has disappointed the simple reductionists. Aldolase is an enzyme containing 28 cysteinyl residues in four subunits. It is inactivated by an excess of p-hydroxymercuribenzoate, but it has full activity when only eight out of the 28 SH groups are blocked. Moreover, it has been shown by a number of methods that such an enzyme is damaged in its structure, whereas in the presence of the substrate or inorganic phosphate this structure is again stabilized. On the ground of such experimental evidence aldolase has been regarded in turn as an SH-enzyme, and has again been excluded from that group.

Szabolcsi and her co-workers have recently studied the carboxymethylation of the cysteinyl groups of aldolase and this programme is connected with the study of the primary structure of aldolase carried out by Sajgó in our Institute.

Incorporation of radioactive carboxymethyl residues when 14C-labelled bromoacetate reacts with muscle aldolase has been investigated at pH 7.5 and 8.5 in tris buffer and this was correlated with loss of enzyme activity (Figure 6). It has been established that under these conditions three cysteinyl residues would react and nothing else. When loss of activity and carboxymethylation are plotted against each other it becomes evident that there is no linear relation.

![Figure 6](image-url)
CHEMICAL TOPOGRAPHY OF SIDECHAINS IN ENZYMES

(Figure 7). Kinetic analysis of the time course of inactivation and carboxymethylation shows that loss of activity is a single apparently first-order reaction, its rate depending on the concentration of bromoacetate. On the other hand, carboxymethylation is a complex reaction, the initial phase of which can be resolved into two reactions, the carboxymethylation of two SH groups (Cys I and Cys II) at different rates. It could be shown that the rate of loss of activity is always identical with the rate of the slower of these two carboxymethylations. The third cysteinyl residue (Cys III) is only labelled after extensive inactivation, therefore, blocking of Cys III does not enter into the inactivation process\textsuperscript{18}.

Especially revealing are the results obtained when carboxymethylation is studied in the presence of inorganic phosphate. It should be remembered that the substrate of the enzyme is a phosphorylated compound and the enzyme has two specific phosphate binding sites\textsuperscript{21}. Kinetic analysis of these experiments shows that the rate of the first carboxymethylation reaction is similar to that observed in the absence of phosphate, whereas the second one and the loss of enzyme activity are much inhibited (Table 2).

Table 2. Carboxymethylation and inactivation of aldolase in the absence and presence of inorganic phosphate: 37°, pH 8.5

<table>
<thead>
<tr>
<th></th>
<th>Rate constants in the absence of inorganic phosphate M(^{-1}) h(^{-1})</th>
<th>Rate constants in the presence of inorganic phosphate M(^{-1}) h(^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cys I</td>
<td>120</td>
<td>130</td>
</tr>
<tr>
<td>Cys II</td>
<td>60</td>
<td>20</td>
</tr>
<tr>
<td>Inactivation</td>
<td>60</td>
<td>20</td>
</tr>
</tbody>
</table>
After carboxymethylation of aldolase in the presence or absence of phosphate, using radioactive bromoacetate, the tryptic peptides containing labelled material have been isolated\(^9\). From these studies the following conclusions have been drawn.

Along the peptide chain of aldolase the cysteinyl groups are located according to the scheme of Figure 8. (The full sequence being as yet incompletely known, the numbers are approximations\(^20\).)

The cysteinyl residue which reacts with bromoacetate at the same rate both in the presence or absence of phosphate, is Cys-230. The cysteinyl residue which reacts more rapidly in the absence of phosphate and which is protected by phosphate (Cys II), is in the C-terminal part of the molecule, in the tryptic peptide: Ala.Leu.Ala.Cys.Ser. (Asn\(_1\)Gln\(_1\)Gly\(_1\)Ala\(_1\)Leu\(_1\)) Lys. The residue which reacts in tris buffer being the third in the time course of the reaction (Cys III) is similarly in the C-terminal part of the molecule, in the tryptic peptide: Cys.Pro.Leu. (Pro\(_1\)Leu\(_1\)Trp\(_1\)) Lys.

We can summarize these findings by underlining the unusual property of aldolase: its most reactive sulphhydryl group may be carboxymethylated with no obvious loss of activity and apparently this reaction has not much influence on the steric structure either. The second most reactive SH group is also on the surface of the molecule, but it is protected after inorganic phosphate has been bound to the enzyme. Further, it was shown that blocking of this cysteinyl residue damages the specific phosphate binding sites of aldolase\(^9\). Therefore it is assumed that this particular cysteinyl sidechain is in the vicinity of the substrate binding site. Its carboxymethylation brings about an unfolding which in turn exposes a neighbouring SH group. Perhaps it is worth mentioning that according to previous data\(^22\) the C-terminal tyrosine is important for full activity and specificity, and a lysine, which is approximately in the middle of the polypeptide chain, in the proximity of Cys I, forms a Schiff base with the substrate according to Lai et al.\(^23\). The SH group, which Szabolcsi and her co-workers have found to be connected with the substrate binding site, is rather far from both these residues.

Finally, I should like to point out that there is a more general problem in enzymology which has recently attracted the attention of an increasing number of investigators. This is the subunit structure of enzymes and the possible role thereof in the steric structure and enzymatic function of oligomeric proteins. Here again, information can be gained by chemical modification which can be controlled but might hardly be replaced by x-ray crystallography, as will be shown by Südi at the International Biochemical Congress at Interlaken.
fact is that there is a number of homo-oligomeric or quasi-homo-oligomeric
proteins like lactate dehydrogenase, glyceraldehyde-3-phosphate dehydrogen-
ase, aldolase, or even haemoglobin which, as a rule, do dissociate to some extent
in solution. I should like to remind you of the example of haemoglobin in
support of my view that neither the extent, nor the manner of this dissociation,
can be studied directly, or at least studied with reasonable economy, by x-ray
crystallography. On the other hand, an association—dissociation equilibrium
can be regarded as a special case of the fluctuation of protein structure and
therefore easily studied by making use of the differential chemical reactivity of
dissociated versus associated subunits. Indeed, by the appropriate choice
of reagents it is certainly possible to label covalently the contact surfaces of
oligomeric proteins, by analogy of active site labelling where protein—protein
interactions are replaced by protein—substrate interactions.

The above different lines of investigation carried out in our laboratory at
Budapest serve to illustrate our view on the use of chemical topography. I
believe that sidechain modification as a tool for the understanding of the
functioning of enzymes is still in its early development. Its great potentialities
have already underlined through the discovery of important groups in the
active site of several enzymes. As our general understanding of the detailed
structure of biologically active proteins is increasing through the application of
a variety of methods, this knowledge will in turn help us to evolve more
sophisticated methods of sidechain modification.

Obviously, we need more and more specific reagents, a variety of reagents
having different bulk, charge and steric properties and differing specificities.
Using a specific reagent to detect one kind of amino acid residue, we have to
expect to find more than one such residue on the surface of the protein. As most
of these are in different environments, we may expect that they will react
somewhat differently from each other, at least at a different rate. Conditions
should be looked for, which allow the separate reaction of each of them.
Perhaps it should be the maxim of sidechain modification that each sidechain
should be studied separately. In our experience the proper analysis of the
kinetics can usually help to achieve this.

It may be expected that the comparative study of the reaction of the same
amino acid sidechain in different micro-environments would lead to the general
understanding of the phenomenon of enhanced reactivity. It appears from
already existing data that enhanced reactivity of an amino acid sidechain in the
active site is not due to its covalent attachments but rather to the micro-
environment formed by surrounding amino acid sidechains coming from
distant segments of the same polypeptide chain. The biological phenomenon of
enzymatic catalysis appears to reside in the modified reactivity of surface
groups surrounded by the non-aqueous protein medium. As first shown by
Crestfield et al., the specificity and the increased rate of an enzymatic reaction
is (accordingly) in part due to the forced orientation of the reagent. This
concept has led to a search for bifunctional reagents and it is expected that their
use may enable us to open a new field in sidechain modification studies.

In a number of cases, sidechain modification causes a change in the confor-
mation, and some further modifications are a consequence of these changes. On
the other hand, as I have shown and as others have also observed, some primary
modifications are due to the physiological change of conformation. to fluctua-
tions of the polypeptide chain. Here is a challenging problem to solve: is fluctuation just a nuisance, is it an irrelevant fact or is it a pre-requisite of specific reactivity and biological activity?

Our choice of method will be influenced by our belief. Sidechain modification of individual groups, with thorough study of its kinetics, if possible, with the help of other analytical and physicochemical methods should be used when looking for the possible correlation of biological activity and fluctuation of the polypeptide chain.

May I close by pointing out the similarity between an enzyme activity and a smile on a face. There are many ways to study the basis and the mechanism of the smile of Mona Lisa Gioconda. You may take the x-ray picture of her head and you obtain the most definite knowledge of the basic structures. On the other hand, you may study the coherent movements of some surface structures. There are many possible approaches if you want to know a woman, and you have to choose the best combination.

REFERENCES

1 J. Herskovits and M. J. Laskowski, J. Biol. Chem. 235 PC, 56 (1960); 237, 2481 (1962).
12 F. B. Straub, unpublished.
13 M. Schramm, Biochemistry, 3, 1231 (1964).