7

Chemical Modification: Side-Chain Specific Reagents

INTRODUCTION

Although chemical modification can be defined as any alteration of the covalent structure of a protein via introduction or removal of a chemical group, we use a more restrictive definition in the context of this chapter. To the enzymologist, chemical modification is the specific (at least attempted) modification of one or more amino acid side chains in a polypeptide chain, usually accompanied by the introduction of a substituent into that group. This type of protein modification has found a wide number of uses in protein chemistry, which we examine together with the practical requirements of the modification reagents that might be helpful in each context. The uses include:

1. The identification of residues involved in the catalytic mechanism of an enzyme or in a binding site on a protein.
2. The introduction of reporter groups, which can be used to report on their environment, to indicate conformational changes or to act as points of reference in distance measurements. Such introduced groups are often fluorescent or have magnetic properties suitable for ESR or NMR, although in theory they may contain any group with distinct spectral properties.
3. The direct detection of conformations or conformational changes in proteins. The reactivity of amino acid side chains is governed by their environment, and the nature of their reaction with chemical modification reagents can reflect this environment.

4. The chemical cross-linking, either of different polypeptide chains within an oligomer or of residues within the same polypeptide chain. Such modifications require bifunctional reagents.

5. The preparation of enzyme derivatives that can be used in hybridization studies to establish subunit composition or to examine subunit–subunit interactions. Such reagents must alter the charge of the protein to allow separation from native molecules.

The general characteristics of the modification reagents that might be required for these various uses can differ depending on the use. Before considering some of the theoretical and practical aspects of chemical modification we define six terms that will be used in this chapter and Chap. 8, on site-specific modification reagents.

Specificity: Refers to the chemical nature of the amino acid side chains modified by a particular reagent. In some cases it is advantageous to use a reagent with high specificity: that is, one that reacts only with, for example, cysteine residues. Such a reagent is particularly useful when attempting to modify a specific side chain in a protein to obtain evidence for its involvement in some aspect of the protein's function. In other instances a high specificity is detrimental—for example, with site-directed (see the following definition) or cross-linking reagents it makes it less likely that the reagent will be of use: They are designed for a particular purpose, not for reaction with specific residues.

Site Directed: Refers to a chemical modification reagent that has incorporated into the molecule some moiety which "directs" it toward a specific binding site on the protein such that any side chains modified have a high probability of residing in or near that site. When it is the active site of an enzyme the term "active-site-directed irreversible inhibitor" is often applied to such a reagent. Since their aim is often to label residues residing within the specific site for later identification, specificity and selectivity (see the following definition) are not highly sought-after characteristics of such reagents.

Selectivity: Refers to the sensitivity a particular reagent shows for the environment surrounding the group or groups it can modify. One with a high degree of specificity is of little use if it shows no selectivity. Little is gained concerning an understanding of the role of a particular type of side chain in a protein's function if all the residues react equally with the reagent. Ideally, it reacts with a high degree of selectivity, modifying either only one of a specific side chain or several of that side chain but with very different rates of modification. Such reagents are particularly useful for examining the environment of certain side chains in a protein.

Reversibility: Most chemical modifications are not readily reversible—in fact, it would be a serious handicap if they were as it would make identification of the modified residue or residues quite difficult. However, in certain circumstances it is advantageous to be able to reversibly modify a protein. This is particularly true in cases where chemical modifications are to be used in the construction of hybrid

molecules or where loss of biological activity as a result of a modification must be reacquired.

Photo-activatable: Refers to a group incorporated into a modification reagent which remains inactive until activated by light. Two types are commonly used: (1) so-called photo-affinity labels, where the photo-activatable moiety is the only chemical modification reagent in the molecule, the rest representing a site-directing moiety, and (2) hetero-bifunctional reagents, where the reagent contains two reactive moieties, one of which is reactive at all times and the other only after photo-activation. The advantages and uses of these types of reagents is discussed fully later in this chapter.

Suicide: Refers to reagents that *become* modification reagents after an enzymatic process has taken place that generates an active group on the reagent. They can be regarded as the ultimate in site-directed reagents.

Chemical modification experiments have taken many forms, but in general, reagent selection usually involves some form of trial and error. The pH dependence of kinetic parameters or ligand binding is often used to give a pK that can guide the design of a modification experiment. Once a modified protein is obtained it is often necessary to determine which amino acid residue (or residues) have been modified in terms of their location in the primary amino acid sequence of the protein. This involves the procedures discussed in Chap. 5.

As detailed at the start of this chapter, chemical modification experiments have been employed in a wide variety of ways. Because of this broad application and the detailed information that can be obtained, it is necessary to consider some of the pitfalls that await the experimentalist using chemical modification methods. The purpose is not to discourage the use of these approaches, but to give an awareness so that appropriate safeguards can be taken. Not all of the pitfalls listed here are always pitfalls—it depends on the nature of the information that the experiments are designed to provide.

1. Few reagents are specific to the extent that they react only with a particular type of side chain. Inferences regarding the functioning of specific side chains thus depend on the direct experimental demonstration that a particular type of side chain is being modified.

2. It is not possible to chemically modify an amino acid side chain without affecting the conformation of a protein. When chemical modification is used to introduce reporter groups, not only must the uniqueness of labeling be established but also the lack of *significant* (in terms of biological activity) effect on the protein's conformation must be confirmed before inferences concerning the potential role of measured conformational changes can be established. Reporter groups used to monitor the environment around a particular side chain *do not*: They reflect the environment around the reporter group, and since such groups are often bulky and hydrophobic, they can alter the environment of the side chain. Information concerning the side chain of a particular amino acid is best obtained by monitoring the

reactivity of that side chain, not the spectral properties of an introduced group after modification.

3. It is hard to predict the behavior of a particular reagent toward a protein—because it modifies a particular side chain in model compounds or in another protein does not mean that it will behave similarly in all proteins.

4. Interpretation of chemical modification data is not always straightforward. Except in the case of site-directed reagents it is difficult ruling out conformational effects as being the cause of loss of activity rather than the specific modification occurring in the appropriate specific site. Even ligand protection experiments do not allow an unequivocal determination to be made concerning the location of a specifically modified residue.

In this chapter we first examine various theoretical aspects of chemical modification experiments, followed by a discussion of different side-chain specific reagents. Finally, some applications of modification studies are considered.

THEORETICAL CONSIDERATIONS

There are two major factors affecting the reactivity of a specific amino acid side chain toward a reagent that under ideal conditions reacts with it. These are: (1) effects on the pK of the reacting group (which affects its nucleophilicity), and (2) steric effects.

p*K* *Values of Reacting Groups*

A number of factors influence the pK of an ionizable amino acid side chain in a protein, including the electrostatic field of the protein (especially the local charge distribution around the particular side chain), the solvation of the group and the hydrogen bonds, if any, that the group may be involved in. An examination of th, pK values commonly found in protein side chains shows a range from approximately 3.75 to >12. Table 7-1 shows pK values for amino acid side chains as determined in the free amino acids.

These pK values do not, however, necessarily represent what would be found in a protein for each side chain. Even in dipeptides the pK for histidine's imidazole

TABLE 7-1 Side-chain pK values in free amino acids

α-COOH	3.75
β,γ-COOH	4.6
Imidazole	7.0
α-NH_2	7.8
ε-NH_2	10.2
—SH	9.2
Phenolyic	9.6
Guanidyl	>12.0

TABLE 7-2 Imidazole pK values in small peptides

His-Gly	6.22
Gly-His	6.95
His-Lys	6.48
Gly-His-Gly	6.72

is quite variable. Table 7-2 shows imidazole pK values determined by NMR measurements for a series of peptides.

If we examine the experimentally determined pK values for histidines in several proteins, we find further evidence for the range of pK values the imidazole side chain can have as a result of its environment in a protein. Carbonic anhydrase and staphylococcal nuclease each have four histidine residues. The pK values of these residues (determined by NMR) are:

Carbonic anhydrase: 5.91, 6.04, 7.0, 7.23
Staphylococcal nuclease: 5.37, 5.71, 5.74, 6.5

Between these two proteins we have a range from 5.37 to 7.23, almost 2 full pH units, for the pK of histidine.

The effects of environment on pK can also be demonstrated by examining the pK of a single group under a number of different circumstances. The γ-COOH of glutamate-35 in lysozyme has a pK of 5.9 in the native enzyme, compared to a value of 4.6 for the free amino acid. When lysozyme is denatured the pK of glutamate-35 falls to a value of 4.6. When the native enzyme is allowed to bind the inhibitor tri-*N*-acetylglucosamine, the pK rises to 6.4. The native conformation, and conformational changes involving a particular residue, can affect the pK of a particular side chain.

Competitive Labeling Method. The pK values of reacting side chains can be determined via this method, which is based on the fact that in the presence of a trace amount of a radioactive modification reagent, the various groups in a protein that can be modified by the reagent will compete for the label. As a result, the amount of radioactivity incorporated into any group of the protein is determined by the pK of the group and the environment. Because only a trace amount of label is used, complexity as a result of one modification affecting subsequent modifications is avoided.

In practice, the protein and an internal reference standard are reacted with a limiting amount of radioactive label. The reaction is quenched by addition of an excess of unlabeled reagent. This not only stops the reaction, but generates a chemically homogeneous product which is heterogeneous with respect only to isotopic label, preventing problems due to different chemical modifications in subsequent peptide analysis. After quenching, the protein is subjected to this analysis using fragmentation methods which give peptides containing only one modified residue. The

amount of label in the various peptides depends on the reactivity of the individual groups, and the rate of labeling of the various peptides relative to the rate of labeling of the internal standard is found from a time-dependence profile. The relative rate of modification is then determined at a series of pH values, and from a plot of the relative rate versus pH the pK of the reacting group can been calculated. Once the pK of the reacting group has been found, the rate constant of reaction with the group can be expressed by the pH-independent or specific rate constant given in the equation

$$K_{\text{obs}} = K_n \frac{K_i}{K_i + [\text{H}^+]} \tag{7-1}$$

where K_{obs} is the rate constant at a particular pH, K_n the pH-dependent rate constant, K_i the ionization constant of the group, and H^+ the hydrogen-ion concentration at the particular pH. The ratio of K_n to the pH-independent rate constant of modification of a standard, free residue (K_s) gives $K_r = K_n/K_s$. In practice, the *amount* of modification of each modifiable residue is taken as a ratio to the amount of modification of an added standard under the same conditions, to give a value of K_r. A Brønsted plot (see Fig. 7-1) of log K_r versus pK for all similar groups in the protein can be plotted. Points that lie on the line indicate groups that show normal reactivity, points that lie below indicate those that are subject to steric hindrance, while points that lie above indicate residues that are especially reactive, either as the result of a uniquely reactive orientation or of specific reagent binding in the proximity of the reacting group giving local concentration effects.

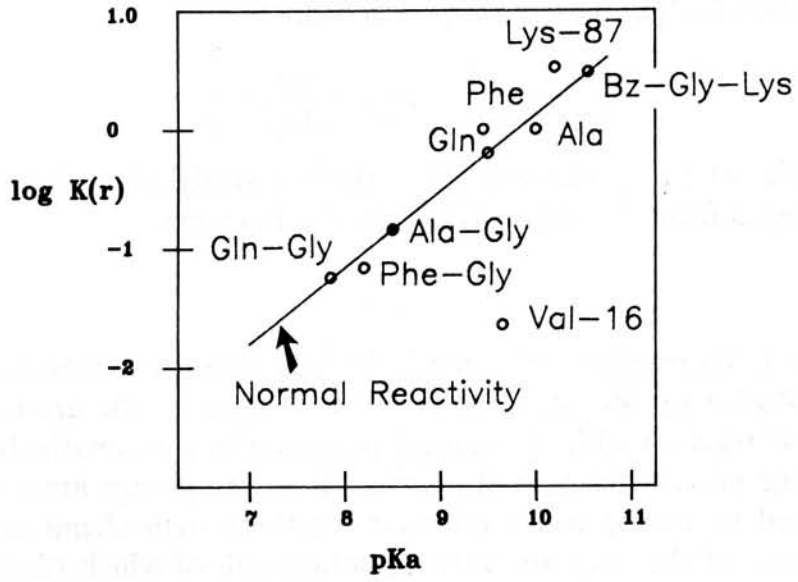


Figure 7-1 Brønsted plot for the reaction of acetic anhydride with various amines and with certain residues in native elastase. K_r is the ratio of the second-order velocity constant for the unprotonated amino group (K) to that for unprotonated phenylalanine (K_s).

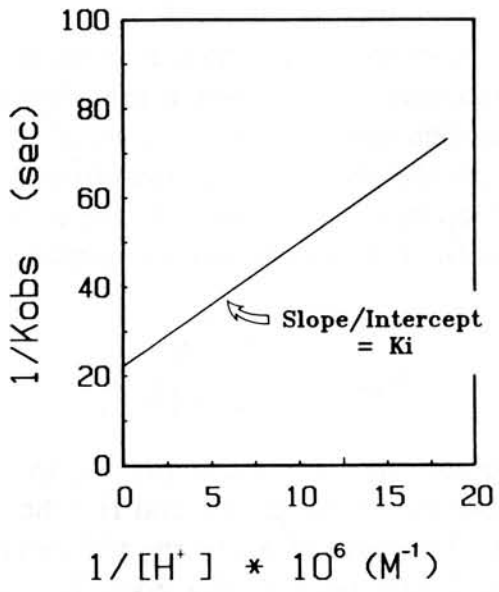


Figure 7-2 Plot of $1/K_{obs}$ versus $1/[H^+]$ to determine the ionization constant, K_i, for a reacting group.

Where a single group is known to react with a reagent, the pH dependence of the reaction can be expressed as

$$K_{obs} = \frac{K_{max}}{1 + K_i/[H^+])} \tag{7-2}$$

where K_{max} is the intrinsic pseudo-first-order rate constant when the group is protonated. Equation (7-2) in double reciprocal form,

$$\frac{1}{K_{obs}} = \frac{1}{K_{max}} + \frac{K_i}{K_{max}(1/[H^+])} \tag{7-3}$$

shows that a plot of $1/K_{obs}$ versus $1/[H^+]$ yields a straight line (Fig. 7-2) and that K_i can be obtained from the slope divided by the intercept.

Steric Effects

In a protein, the reactivity of a particular side chain to a modification reagent may well be affected by the accessibility of the reagent to the group. Such steric effects may slow reaction with the reagent (resulting in a below-the-line point in a Brønsted plot) or prevent reaction altogether. Reacting groups are sometimes classified as exposed or buried based on their reactivity with chemical modification reagents. Because of the reagents' varying natures, all of which react with similar specificity, some amino acid side chains may appear to be exposed to some reagents and buried to others. Classification of groups as exposed or buried on the basis of chemical reactivity alone may not be completely accurate. Generally, three classes are identified:

1. *Internal.* residues with completely buried side chains having no solvent accessibility
2. *Surface.* side chains partially buried, or accessible to solvent from one side only
3. *External.* side chains that project into the surrounding solvent and are freely accessible to modification reagents

QUANTITATION OF MODIFICATION AND DATA ANALYSIS

During a chemical modification experiment a number of experimental parameters can be determined and related to the number of residues being modified and their effects on activity. Here we examine some of the more usual methods of representing chemical modification data and discuss what information can be obtained from such analysis.

Time-Course Analysis

In most experimental situations there is a considerable excess of modification reagent to enzyme residues that can be modified, and the reaction can be regarded as pseudo-first order. The modification can be monitored either by following directly the modification of a particular type of side chain (which is quite easy if spectral changes result from the process), or by following the effect of modification on some enzymatic parameter of the protein (activity, regulation by an allosteric ligand, etc.).

The rate of inactivation, V_{inact}, is given by

$$V_{\text{inact}} = \frac{-d[\text{E}]}{dt} = K_{\text{inact}}[\text{E}] \tag{7-4}$$

Rearrangement and integration between the limits 0 and t for time and E_0 and E_t for [E] gives

$$\ln \frac{E_t}{E_0} = -K_{\text{inact}}t \tag{7-5}$$

where E_t is the activity at time t and E_0 is the initial activity (these are alternatively represented as A_t and A_0, respectively, by some workers).

Since $\ln (E_t/E_0)$ is the log of the residual activity at any time t, a semilog plot of log residual activity versus time, as shown in Fig. 7-3, gives a value for the rate constant of inactivation, K_{inact}.

In plots of log residual activity versus time, straight-line plots are obtained unless two or more residues per protein molecule, each contributing to the activity but reacting with quite different rate constants, exist. In such a case the plot of log residual activity versus time is multiphasic (if the rate constants are sufficiently separated), and rate constants can only be obtained by fitting the data to two or more exponentials. In plots of log percent residues modified versus time, complex plots are obtained if more than one residue reacts and the rate constants are sufficiently separated.

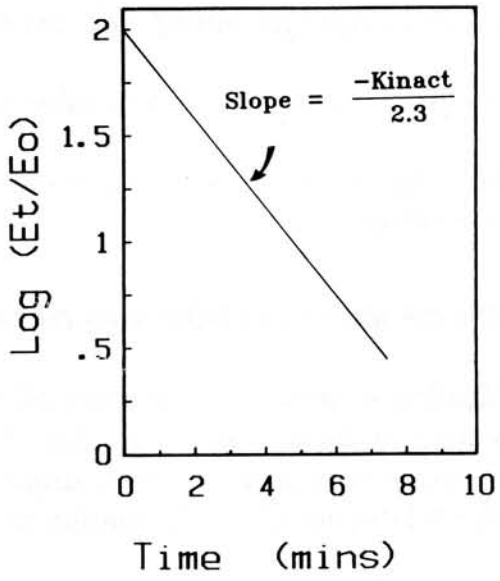


Figure 7-3 Semilog plot of log residual activity (E_t/E_0) versus time of incubation with reagent. The slope of this plot is $-K_{inact}$.

It is important to note that such analysis of the time dependence of inactivation assumes that the modification process is the only one utilizing the modification reagent. It is possible that the reagent may be subject to a competing process such as hydrolysis, and in such a case this process will contribute to a decrease in reagent concentration as a function of time. To obtain the true value of the apparent first-order rate constant for inactivation, a plot of log residual activity versus $(1 - e^{-k't})k'$ must be made, according to

$$\ln \frac{E_t}{E_0} = -\frac{k}{k'} I(1 - e^{-k't}) \tag{7-6}$$

where k' is the rate constant for hydrolysis of the reagent and I is the initial concentration of the reagent. k' must be experimentally determined independently. This problem is often encountered with reagents such as diethylpyrocarbonate, and Fig. 7-4 shows experimental data obtained with modification of *S*-adenosylhomocysteinase at a series of diethylpyrocarbonate concentrations. Also shown is a plot of K_{obs} versus [reagent].

The rate constant, K', for the decomposition of diethylpyrocarbonate is determined by incubating the reagent under the reaction conditions, withdrawing aliquots at appropriate time intervals, adding a large excess of imidazole and quantitating the formation of ethoxyformyl imidazole at 242 nm, using a molar extinction coefficient of 3200 cm^{-1}. In this way the quantity of reactive diethylpyrocarbonate at any time, t, is quantitated and the rate constant of decomposition is determined from a plot of log ($[DEP]_t/[DEP]_0$) versus time, as in Fig. 7-3.

An important property of K_{inact} is that except in cases of specific complex formation of the target protein with the modification reagent (MR), a plot of K_{inact} versus the concentration of the modification reagent should be linear and pass through the (0, 0) axis, as shown in Fig. 7-5.

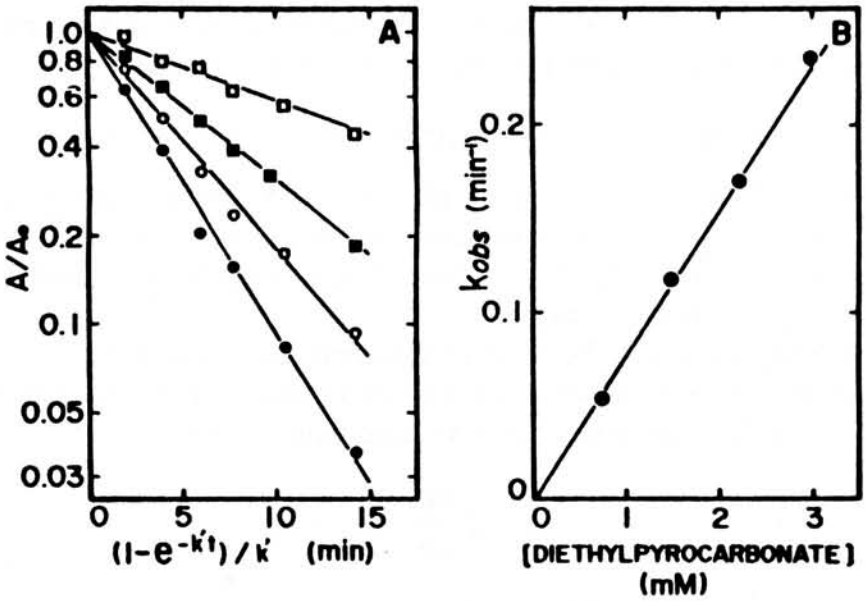


Figure 7-4 Inactivation of *S*-adenosylhomocysteinase by diethylpyrocarbonate. The enzyme (16 μM subunit) was incubated with 0.74 (□), 1.48 (■), 2.21 (○), or 2.98 (●) diethylpyrocarbonate in 0.1 *M* potassium phosphate buffer, pH 6.9, at 0°C. At time intervals, aliquots were removed for measurements of the residual enzyme activity. The first-order rate constant for decomposition of diethylpyrocarbonate (K') was determined separately as described in the text. Values on the abscissa are calculated with $K' = 7.5 \times 10^{-3}$ min^{-1}. *S*-Adenosylhomocysteinase incubated in the absence of diethylpyrocarbonate lost no activity under these conditions. (Reprinted with permission from: T. Gomi and M. Fujioka, *Biochemistry*, *22*, 137–143. Copyright 1983 American Chemical Society, Washington, D. C.)

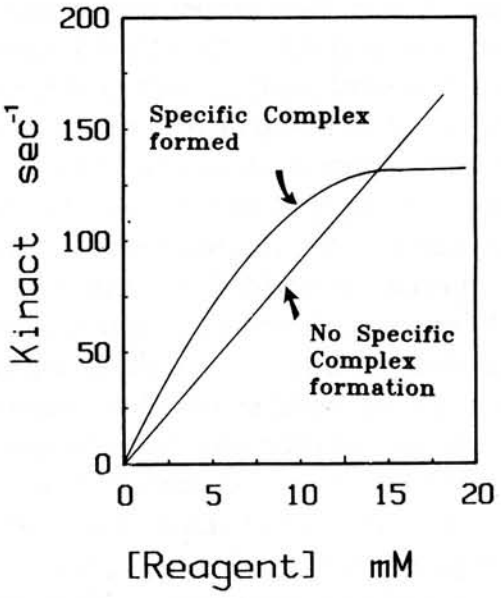


Figure 7-5 Plot of K_{inact} versus concentration of modification reagent.

In some instances a specific complex between the reagent (MR) and the protein may occur, and the modification process consists of two stages:

$$\text{protein} + \text{MR} \overset{K_d}{\rightleftharpoons} \text{protein} - \text{MR} \xrightarrow{k} \text{inactivation} \tag{7-7}$$

If such complex formation occurs, a plot of K_{obs} versus [reagent] concentration is hyperbolic and K_d, the dissociation constant for protein–reagent complex formation, can be obtained either by direct fitting of the hyperbola or from a double reciprocal plot of K_{obs} versus 1/[reagent].

Under conditions where the reagent concentration is not in large excess over the protein the assumption of pseudo-first-order kinetics is not valid, and the second-order rate constant for inactivation can be calculated from

$$\frac{2.303}{a-b} \log \frac{b(a-x)}{a(b-x)} = k_t \tag{7-8}$$

where a and b are the concentrations of reagent and enzyme at time $t = 0$, and x is the amount of enzyme modified at any time, t.

This analysis does not indicate how many of a particular residue must be modified for activity to be lost. The *minimal* order of the reaction is the slope obtained from a plot of log $1/t_{1/2}$ versus log [reagent]. This type of plot, although often used, has been criticized since it can lead to an underestimate of the number of residues necessary for activity in circumstances where a slow, rate-limiting reaction with one residue leads to a more rapid, kinetically unobservable modification of other residues that are essential for activity. *Estimates of the number of residues reacted from kinetic measurements are not a substitute for direct chemical evidence.*

Quantitation

As indicated in the last section, the number and types of groups that have been modified in a particular experiment must be directly determined by use of radioactive modification reagents, by amino acid analysis after modification, or by spectroscopic means when the modification results in the incorporation of a spectrally active group. There are two basic ways that quantitation can be achieved: The unreacted residues can be estimated by amino acid analysis and the number of residues reacted established by difference or alternatively, the reacted residues can be quantitated directly. In many ways the latter approach is preferable. With the first approach it is often a case of determining a difference between two quite large numbers, both of which have significant experimental uncertainty. Also, the modified residues may not separate sufficiently from the parent residue to allow unique determination of the unmodified residues. Direct quantitation can be achieved easily and accurately if radiolabeled reagents are available. Many reagents lead to the incorporation of chromophoric groups, which, after removal of unreacted reagent, may be quantitated spectrally. One particular problem with this appraoch is the environment of the group, which can affect the spectral properties. When environmentally sensitive

chromophores are used, it is advisable to denature the protein after modification and quantitate the incorporation in such a uniform state where tertiary structure does not influence the spectral properties. Where quantitation can be carried out via spectral measurements, the utility of this approach is indicated for individual reagents discussed. Assuming that this can be achieved, plots of residual activity versus number of groups modified per mole of enzyme can be made.

We consider several types of chemical modifications, depending on the number of amino acid residues modified and the relationship between the rate of chemical modification and the rate of inactivation.

In the most simple case, the rate of modification equals the rate of inactivation and there is a single residue involved. The rate of modification, V_{mod}, is given by

$$V_{\text{mod}} = \frac{dx}{dt} = K_{\text{mod}}(x_m - x) \tag{7-9}$$

where x is the number of modified groups and x_m is the maximum number for modification.

$$\int_0^{x_1} \frac{dx}{x_m - x} = \int_0^1 k_{\text{mod}}\, dt \tag{7-10}$$

Gives

$$\ln \frac{x_m - x_1}{x_m} = -k_{\text{mod}} t \tag{7-11}$$

Since $K_{\text{inact}} = K_{\text{mod}}$, we get

$$\ln \frac{x_m - x_1}{x_m} = -k_{\text{inact}} t = \ln \frac{E_t}{E_0} \tag{7-12}$$

or

$$\frac{E_t}{E_0} = 1 - \frac{x_1}{x_m} \tag{7-13}$$

and a plot of E_t/E_0 as a function of x_t, the number of residues modified per molecule of enzyme, is linear with a negative slope of $1/x_m$. When $E_t/E_0 = 0$, $x_t = x_m$. Figure 7-6 shows such a case.

When the rates of modification and inactivation are not equal, we can define a ratio of the rate constants, r.

$$r = \frac{K_{\text{inact}}}{K_{\text{mod}}} \tag{7-14}$$

As before,

$$\ln \frac{E_t}{E_0} = -k_{\text{inact}} t \tag{7-15}$$

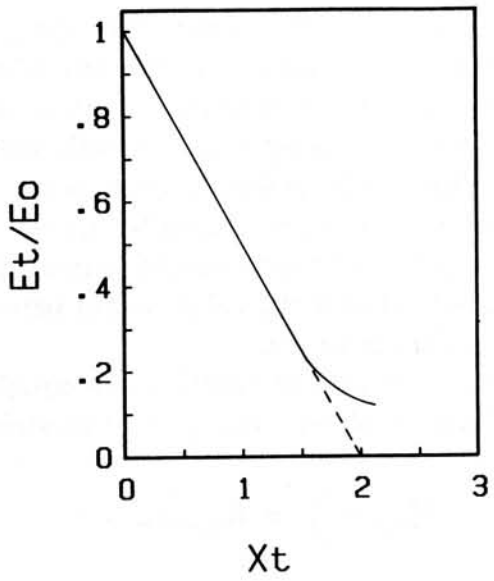


Figure 7-6 Plot of E_t/E_0 versus X_t for determining the maximum number of "essential" groups. E_0 is the initial enzyme activity at time t. The maximum number of "essential" groups in this case is interpreted to be two.

and

$$v_{\text{mod}} = \frac{dx}{dt} = k_{\text{mod}}(x_m - x) \tag{7-16}$$

which, in conjunction with Eq. (7-14), gives

$$r \ln \frac{x_m - x_1}{x_m} = -k_{\text{inact}} t \tag{7-17}$$

and we get

$$r \ln \frac{x_m - x_1}{x_m} = \ln \frac{E_t}{E_0} \tag{7-18}$$

which can be rearranged to give

$$\left(\frac{E_t}{E_0}\right)^{1/r} = 1 - \frac{x_t}{x_m} \tag{7-19}$$

In terms of E_t/E_0, we get

$$\left(\frac{E_t}{E_0}\right) = \left(1 - \frac{x_1}{x_m}\right)^r \tag{7-20}$$

which predicts the family of parabolic curves shown in Fig. 7-7, with the shape depending on the value of r. When $r = 1$ the plot resembles the previous one, but a variety of plots are possible.

Since it is inherently unlikely that $r > 1$, plots with $r < 1$ are most likely to be encountered. Mechanistically, such a plot indicates that inactivation is the result of a rate-limiting conformational change resulting from the modification event.

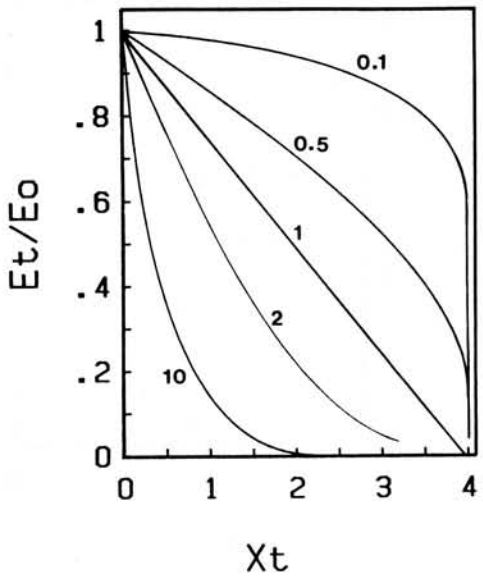


Figure 7-7 Effects of the value of r on the shape of plots of E_t/E_0 versus X_t.

A more likely situation is one in which there are several classes of a particular type of residue that can react and lead to loss of activity, each of which react with different rates, K_{mod_n}. The overall rate of modification, V_{mod}, is the sum of the rates of the individual classes:

$$v_{\text{mod}} = \sum^{n} \frac{d(x_n)}{dt} = \sum^{n} k_{\text{mod}_n}[(x_n)_m - (x_n)] \tag{7-21}$$

where $(x_n)_m$ is the maximum number of groups available in each class and x_n is the number of groups reacted. For each class of residue,

$$\ln \frac{(x_n)_m - (x_n)_t}{(x_n)_m} = -k_{\text{mod}_n} t \tag{7-22}$$

and the overall modification reaction is described by

$$\sum^{n} \ln \frac{(x_n)_m - (x_n)_t}{(x_n)_m} = -\sum^{n} k_{\text{mod}_n} t \tag{7-23}$$

$$\sum^{n} (x_n)_t = x_T \tag{7-24}$$

and is the quantity experimentally measured—that is, the sum of all modified residues at time t.

E_t/E_0 as a function of x is obtained as follows:

$$(x_1)_t = x_T - \sum_{n=2}^{n} (x_n)_t \tag{7-25}$$

and from previously,

$$(x_n)_t = (x_n)_m[1 - \exp(-k_{\text{mod}_n} t)] \tag{7-26}$$

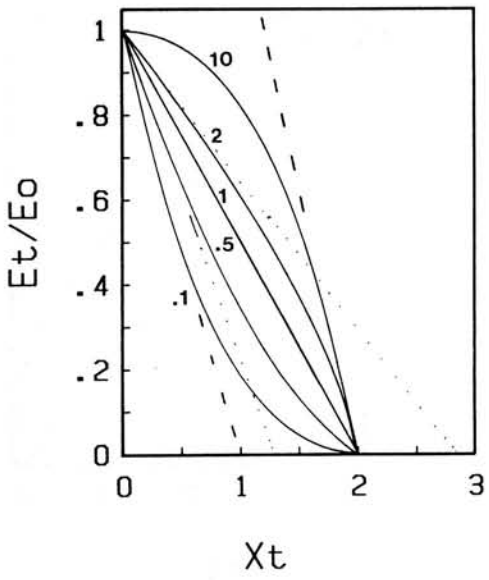


Figure 7-8 Plot of Eq. (7-27) for $n = 2$; $(X_1)_m = (X_2)_m = 1$; $K_{\text{mod}_1} = K_{\text{inact}}$; and $K_{\text{mod}_2}/K_{\text{mod}_1}$ as indicated for each solid curve. ---, lower and upper limiting curves; · · ·, extrapolation of curves to $E_t/E_0 = 1$ and $E_t/E_0 = 0$.

Combining Eqs. (7-25) and (7-26) and the equation for E_t/E_0 when there is a single inactivation and modification rate from previously [Eq. (7-20)], we get

$$\frac{E_t}{E_0} = 1 + \frac{\sum_{n=2}^{n} \{(x_n)_m[1 - \exp(-k_{\text{mod}_n}t)]\}}{(x_1)_m} - \frac{x_T}{(x_1)_m} \tag{7-27}$$

which predicts that a plot of E_t/E_0 versus x_t will generally be exponential since t is an independent variable. Plots of this equation are shown in Fig. 7-8 for a variety of combinations of $(x_n)_m$ and K_{mod_n} and n.

In the event that K_{mod} for one of the groups is equal to K_{inact} and the extent of modification can be determined independently of other modifications, we get

$$\ln \frac{(x_1)_m - (x_1)_t}{(x_1)_m} = -k_{\text{inact}}t = \ln \frac{E_t}{E_0} \tag{7-28}$$

Therefore,

$$\frac{E_t}{E_0} = 1 - \frac{(x_1)_t}{(x_1)_m} \tag{7-29}$$

which is analogous to the simplest case first considered.

Since this is the only situation that can give unique information, it is necessary, where simple E_t/E_0 versus extent of modification plots are not obtained, to experimentally determine the extent of modification of individual residues where more than one residue can be modified, and to correlate activity losses with such unique modifications.

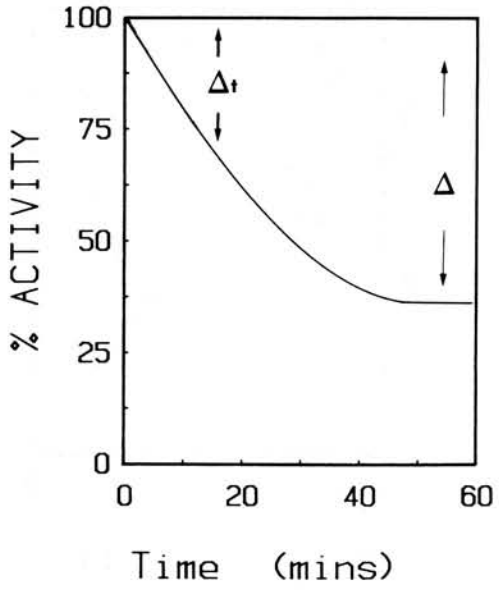


Figure 7-9 Plot of data indicating that modification of one or more residues leads to only *partial* loss of activity. The maximum loss of activity is Δ, and the activity loss at any intermediate time is Δt.

In the examples thus far considered we have assumed that maximal modification leads to *complete* loss of activity (whether it is catalytic activity or regulatory activity). It is quite possible (and often encountered) that modification leads to only partial loss of activity, as indicated in Fig. 7-9.

When such plots are constructed using varied time it is important that the same value of Δ be obtained at different concentrations of modification reagent. If increasing values of Δ are obtained as the concentration of the modification reagent is increased, two explanations are possible. In the first, the progress curves all behave as simple exponentials and the increased values of Δ indicate that an equilibrium between bound and unbound modification reagent is involved. In the second, the progress is multiphasic, indicating that more than one class of residues is being modified, each affecting the activity, possibly to different extents.

In the simple case where Δ does not change with increasing reagent concentration, the parameter $\Delta t/\Delta$ is used in place of E_t/E_0.

The Tsou Plot. When both the amount of modification per protein molecule *and* the percent residual activity can be determined as a function of time, information concerning the number of modified residues essential for activity can be obtained by application of an equation first developed by Tsou:

$$nx = pa^{1/i} + (n - p)a^{\alpha/i} \tag{7-30}$$

where n is the total number of residues in the protein of the type being modified, x the molar function of those residues remaining unmodified after a given time, a the residual activity, p the number of the residues that are modifiable, i the number

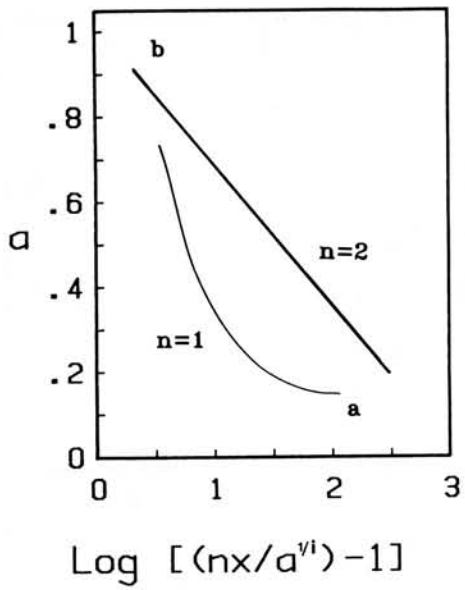


Figure 7-10 Tsou plot for a hypothetical situation with $n = 7$ and $p = 3$. Line (a) results from $i = 1$, line (b) results from $i = 2$. The linearity of line (b) indicates that $i = 2$ in this hypothetical situation.

of those modifiable residues essential for activity and α the ratio of the rates of modification of nonessential versus essential residues. Equation (7-30) can be rewritten as

$$\log\left(\frac{nx}{a^{1/i}} - p\right) = \log(n - p) + \frac{\alpha - 1}{i}\log a \tag{7-31}$$

A plot of $\log a$ versus $\log(nx/a^{1/i} - p)$ has a slope of $(a - 1)/i$.

Once the rate curves of the inactivation are determined, and assuming that n and p are known from independent experiments, the data are plotted according to Eq. (7-31) with a series of values of i. The correct value of i results in a linear plot, as shown in Fig. 7-10.

Protection Experiments

As will be discussed in Chaps. 11 and 17, chemical modification studies can be used in a variety of ways to study ligand binding or conformational changes induced by ligand binding. Further aspects of chemical modification studies in relation to these topics are dealt with in the appropriate chapters. Here we briefly discuss the concept of protection experiments and what information from them may mean with regard to the existence of certain amino acid residues in a particular binding site. One of the prime aims of chemical modification experiments is to attempt to elucidate what types of residues may be important in certain protein functions. Modification studies are often designed to establish important residues in substrate or ligand binding sites. Frequently, studies report that modification of a particular type of

residue results in loss of ligand binding ability by the protein, and that the presence of a particular ligand prevents the modification and loss of activity. The inference of such experiments is often that these results indicate the presence of the particular type of residue being studied in or near the ligand binding site. The "or near" is added to cover the possibility that the ligand in its binding site may sterically block access of the modification reagent to a nearby reactive residue, and vice versa. Such conclusions are unwarranted since there is another equally probable explanation. Modification of a residue some distance removed from a particular ligand binding site may lead to a conformational change in the protein which prevents ligand from binding. In the presence of ligand, there is a ligand-induced conformational change which prevents or hinders the modification reagent from attacking its target group. Chemical modification experiments with side-chain specific reagents are not particularly useful in defining what types of residues may be present in or near a ligand binding site. These comments do not of course apply to site-specific reagents, and with these, as is discussed later, ligand protection experiments can provide useful information.

SIDE-CHAIN SPECIFIC REAGENTS

This section of the chapter is not designed to give a comprehensive account of the many reagents available for the chemical modification of amino acid side chains in proteins. The reagents discussed here have been selected on the basis of (1) their widespread use, and (2) the purpose behind the modification reaction (i.e., the introduction of a reporter group, the alteration of the charge of a residue, or their reversibility). For ease of discussion the reagents described are dealt with in terms of their prime specificity, although, as emphasized earlier, many are not particularly specific. In some cases specificity is improved under certain conditions, and some of these are mentioned. During the descriptions of these various reagents, comments regarding the quantitation of modification by spectral methods are made where appropriate.

Modification of Amino Groups

Modifications of amino groups can be categorized in terms of how they effect the charge of the side chain: the formal charge at neutral pH can be retained, eliminated, or reversed. Table 7-3 shows the average amino acid composition and fractional exposure of residues in proteolytic enzymes whose crystal structures have been examined. Lysine is clearly the most exposed residue in these proteins and makes up a fairly large proportion of the composition of the "average" protein. In fact, of the residues that can be readily modified, lysine is present in much higher proportions than the others. Although Table 7-3 is based on one class of proteins, they are in general of "average" size and compact structure and can be considered as representative of proteins in general. As a result, it is likely that most proteins have lysine residues that can be modified.

TABLE 7-3 Exposure and average percent composition of amino acids in proteolytic enzymes of known crystal structure

Residue	Fractional exposure	Percent composition
Lysine	0.58	7.0
Glutamate	0.49	5.3
Glutamine	0.43	4.1
Arginine	0.43	3.9
Serine	0.43	7.8
Asparagine	0.43	4.8
Aspartate	0.42	5.2
Proline	0.40	5.5
Threonine	0.37	6.5
Glycine	0.33	7.6
Alanine	0.32	8.1
Histidine	0.27	2.2
Tyrosine	0.25	3.4
Phenylalanine	0.17	3.5
Valine	0.16	6.9
Isoleucine	0.14	4.6
Cysteine	0.13	3.4
Methionine	0.11	1.6
Tryptophan	0.11	1.2
Leucine	0.10	7.3

Modification with Retention of Charge. The following are examples of reagents that leave an ionizable group capable of being positively charged on the lysine side chain. The charge is, however, located at a different position than it was originally.

Reductive Alkylation: Reaction with an aldehyde gives a Schiff's base which is subsequently reduced with sodium borohydride (Fig. 7-11). The aldehyde used can have different R groups, such as formaldehyde, acetaldehyde, propionaldehyde, and so on, which allows information to be obtained regarding the steric environment of the side chain.

$$\mathrm{ENZ{-}NH_2} + \mathrm{RHCO} \underset{+H_2O}{\overset{-H_2O}{\rightleftharpoons}} \mathrm{ENZ{-}N{=}CHR} \xrightarrow{[H]} \mathrm{ENZ{-}NH{-}CH_2R}$$

Figure 7-11 Schiff's base formation and reduction with sodium borohydride.

Amidination with Methylacetimidate: The reaction (Fig. 7-12) gives a derivative that is quite stable at acid or neutral pH but labile at alkaline pH. The formal positive charge, which is retained, is delocalized over several atoms by resonance.

Modification with Loss of Charge. These reagents suppress the protonation of the lysine ε-amino group, and thus the derivatives can carry no charge.

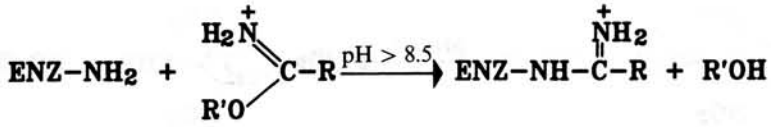


Figure 7-12 Amidination of lysine residue.

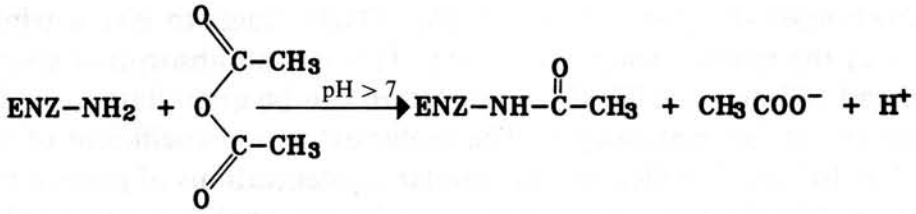


Figure 7-13 Reaction of amino group with acetic anhydride.

Acylation: Reaction of amino groups with acetic anhydride gives an acylated derivative (Fig. 7-13). Although acetic anhydride reacts readily with amino groups at neutral or slightly alkaline pH values, reaction with either sulfhydryl or imidazole groups can occur.

Carbamoylation: This reaction is analogous to that discussed in Chap. 5 for amino-terminal labeling. Reaction of amino groups with cyanate (Fig. 7-14) gives a quite stable derivative. As is also shown in Fig. 7-14, cyanate can react with sulfhydryl and phenolic groups, but at mildly alkaline pH values (approximately pH 8.0) derivatives formed by these reactions are not stable.

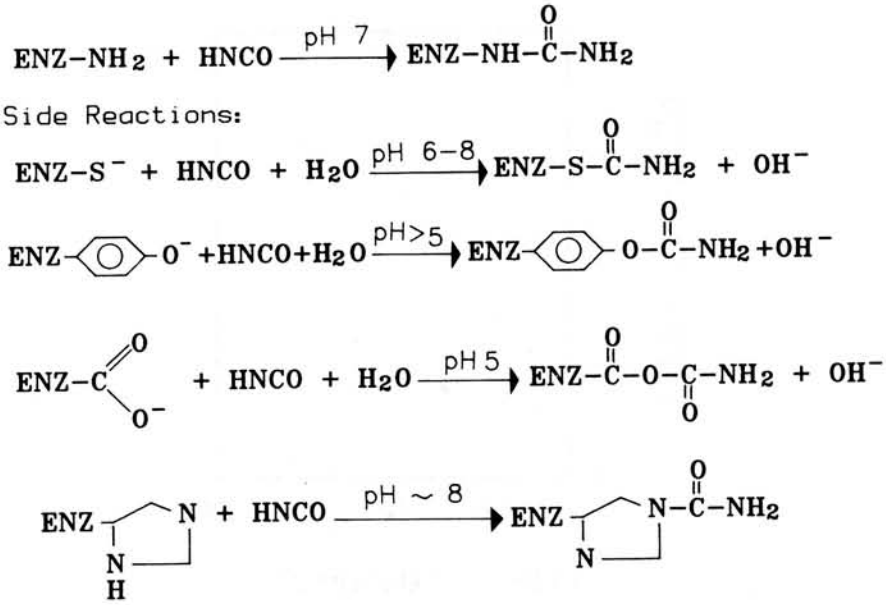


Figure 7-14 Carbamoylation of amino groups; also shown are various potential side reactions with other side chains.

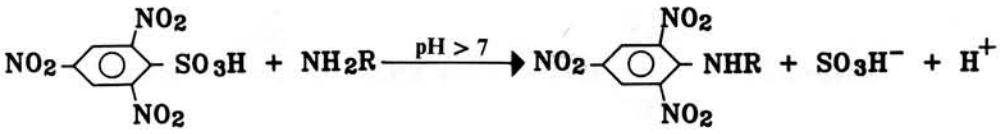


Figure 7-15 Trinitrobenzylation of amino groups by TNBS.

2,4,6-Trinitrobenzene-1-sulfonic Acid (TNBS): TNBS reacts to give a trinitrobenzyl derivative of the amino group (Fig. 7-15). This has an absorption spectrum centered around 367 nm, and TNBS modification can be quantitated via absorbance measurements at this wavelength. The molar extinction coefficient of the derivative is $1.1 \times 10^4\ cm^{-1}$, which at micromolar concentrations of protein means that even one modification per polypeptide can be reasonably quantitated. As with many of the reagents discussed here, reaction of TNBS with sulfhydryl groups can occur, although this problem has not been frequently reported.

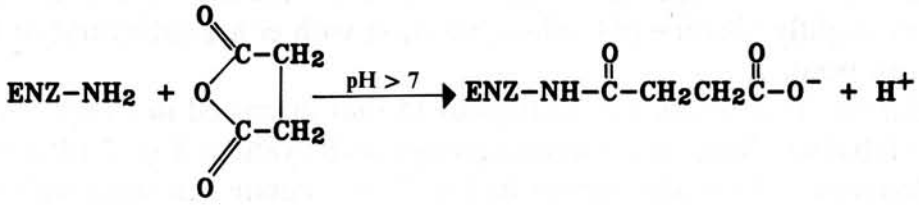


Figure 7-16 Acylation of amino group by succinic anhydride.

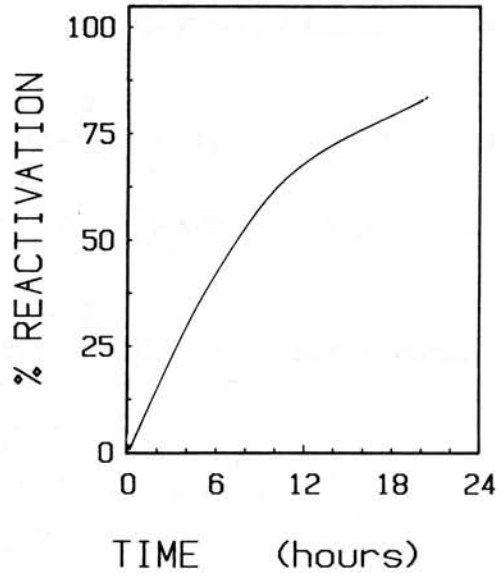


Figure 7-17 Reactivation of acylated protein by incubation at low pH.

Modification with Introduction of Negative Charge. As with modification reactions that retain a positive charge on the modified side chain, these reagents introduce a negative charge, but at a location spatially distinct from the original positive charge.

Acylation: This reaction is analogous to that discussed with acetic anhydride except that the anhydride carries a negative charge which results in the introduction of an overall negative charge into the side chain. Reagents such as succinic anhydride and tetrahydrophthalic anhydride have been successfully used for introducing a negative charge onto lysine side chains (Fig. 7-16). Succinic anhydride introduces a single negative charge per modified residue, while tetrahydrophthalic anhydride introduces a greater charge. Either modification can be completely reversed by incubation at low pH, and the rate of reversal is pH dependent (Fig. 7-17). As shown in Fig. 7-17, it may take up to 24 hours to achieve complete reversal. A potential problem is the stability of the protein at low pH for long periods. However, this approach to reversal has been used successfully in a number of cases, with full regain of native activity.

Pyridoxal-5'-Phosphate (PLP) Modification: Reaction of lysine residues with PLP at neutral pH results in a Schiff's base formation (Fig. 7-18). Subsequent reduction of the Schiff's base with sodium borohydride gives an irreversible derivative of

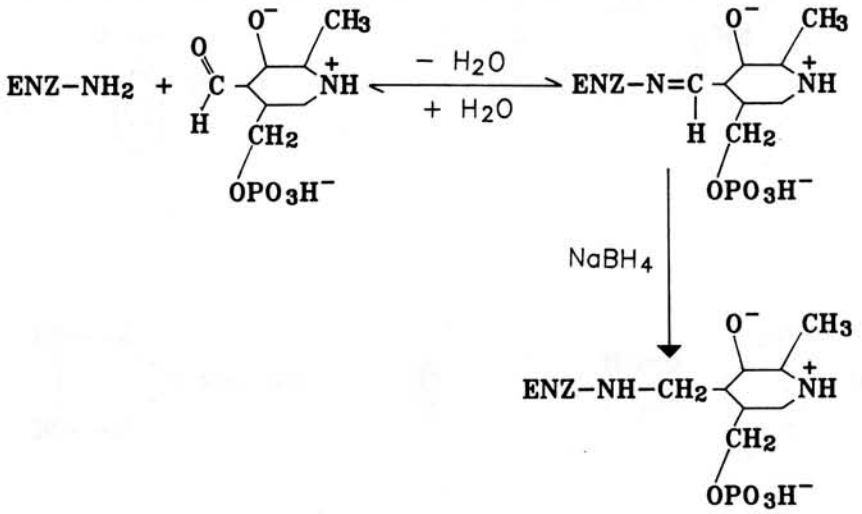


Figure 7-18 Pyridoxylation of amino groups.

the lysine. The introduced PLP group has a number of properties that make this an attractive modification for many purposes. The reaction product has an overall negative charge, as indicated in Fig. 7-18. The introduced group is also a chromophore, with a millimolar extinction coefficient at 316 nm of 8.5 cm^{-1}, allowing for quantitation. The chromophore is also fluorescent. The excitation and emission spectra of protein-bound PLP is shown in Fig. 7-19.

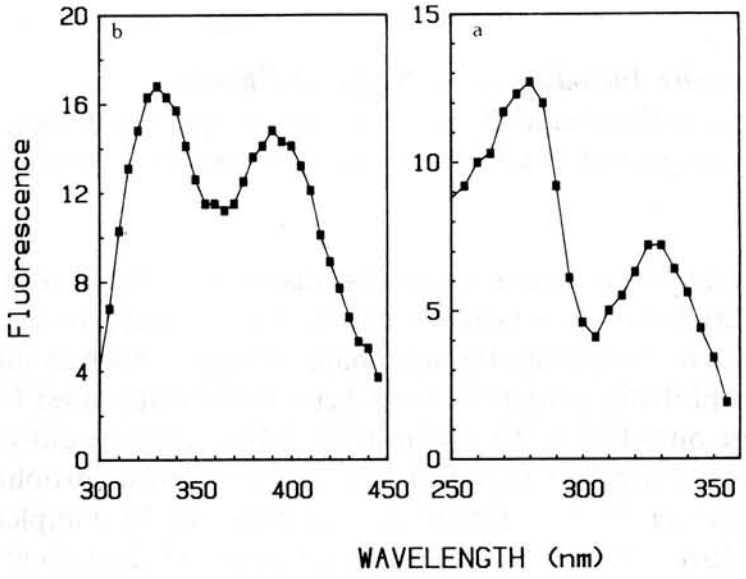


Figure 7-19 Fluorescence emission (A) and excitation (B) spectra of pyridoxylated glutamate dehydrogenase. In (A), excitation was at 280 nm. In (B), emission was at 400 nm.

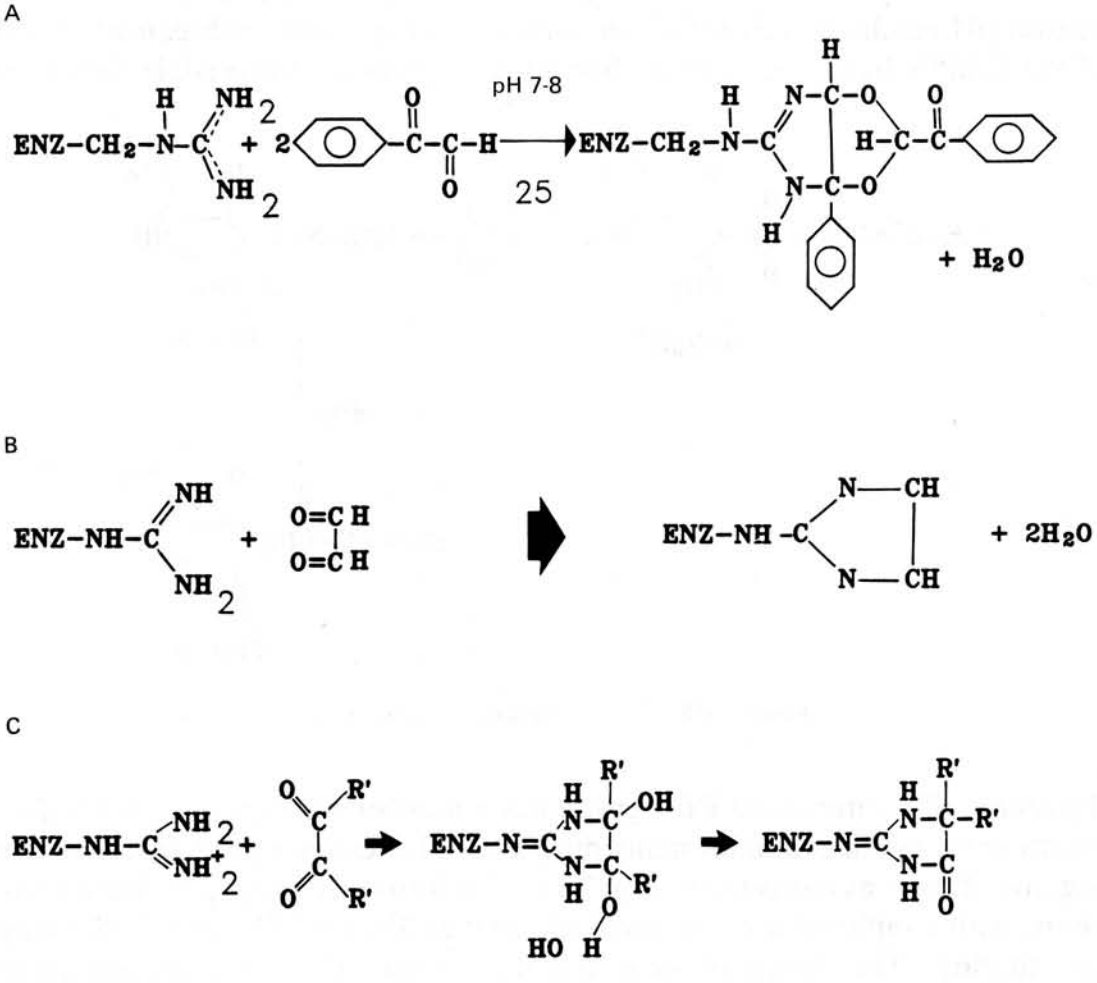


Figure 7-20 Reactions of arginine residues with (A) phenylglyoxal, (B) glyoxal, and (C) 2,3-butanedione.

Modification of the Guanidino Group of Arginine

The modification of arginine side chains is usually based on the formation of heterocyclic condensation products, with reagents having two adjacent or closely proximal carbonyl groups. Compounds with 1,2- or 1,3-dicarbonyl groups readily participate in such reactions because the spacing of the carbonyl groups closely matches that of the two unsubstituted nitrogens in the side chain. Although such compounds can react with amines or sulfhydryls, these reactions, at neutral pH, are very much slower than reaction with the guanidino group. Three reagents in particular have found widespread use: phenylglyoxal, glyoxal, and 2,3-butanedione. With phenylglyoxal and 2,3-butanedione the stoichiometry of reaction does not appear to be 1:1—products with two phenylglyoxal moieties are obtained and there is evidence that 2,3-butanedione reacts first to give a trimer which then reacts with the guanidino group. The modification reactions of these reagents are shown in Fig. 7-20. Modification of arginine residues with these reagents can be reversed, in the absence of excess reagent, at alkaline pH values, although the product obtained with phenylglyoxal modification has been shown to be slowly reversed at pH 7.0.

Modification of Carboxyl Groups

The primary means of modifying carboxyl groups in proteins involves carbodiimide activation via *O*-acylisourea formation at slightly acid pH (Fig. 7-21). The activated intermediate can either rearrange via an O → N acyl shift or react with an added nucleophile to give the corresponding amide. The nature of the added nucleophile can vary depending on the nature of the adduct desired, although glycine methyl ester is frequently used.

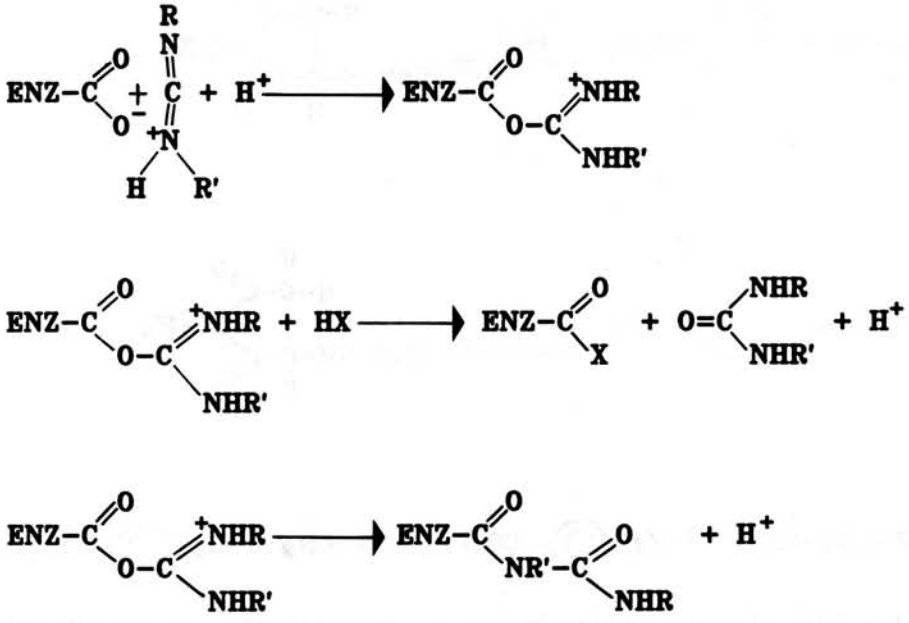


Figure 7-21 Carbodiimide activation of a carboxyl group and subsequent reactions.

Modification of Sulfhydryl Groups

A wide variety of reagents are available for the modification of sulfhydryl side chains in proteins. Many of the reagents are based on three reactions, involving modification by (1) iodoacetic acid or iodoacetamide, (2) maleimide or maleic anhydride, or (3) mercury derivatives such as *p*-chloromercuribenzoate (Fig. 7-22).

Iodoacetate derivatives are quite stable to various fragmentation or amino acid analysis conditions and are often used to alkylate cysteine residues during sequence

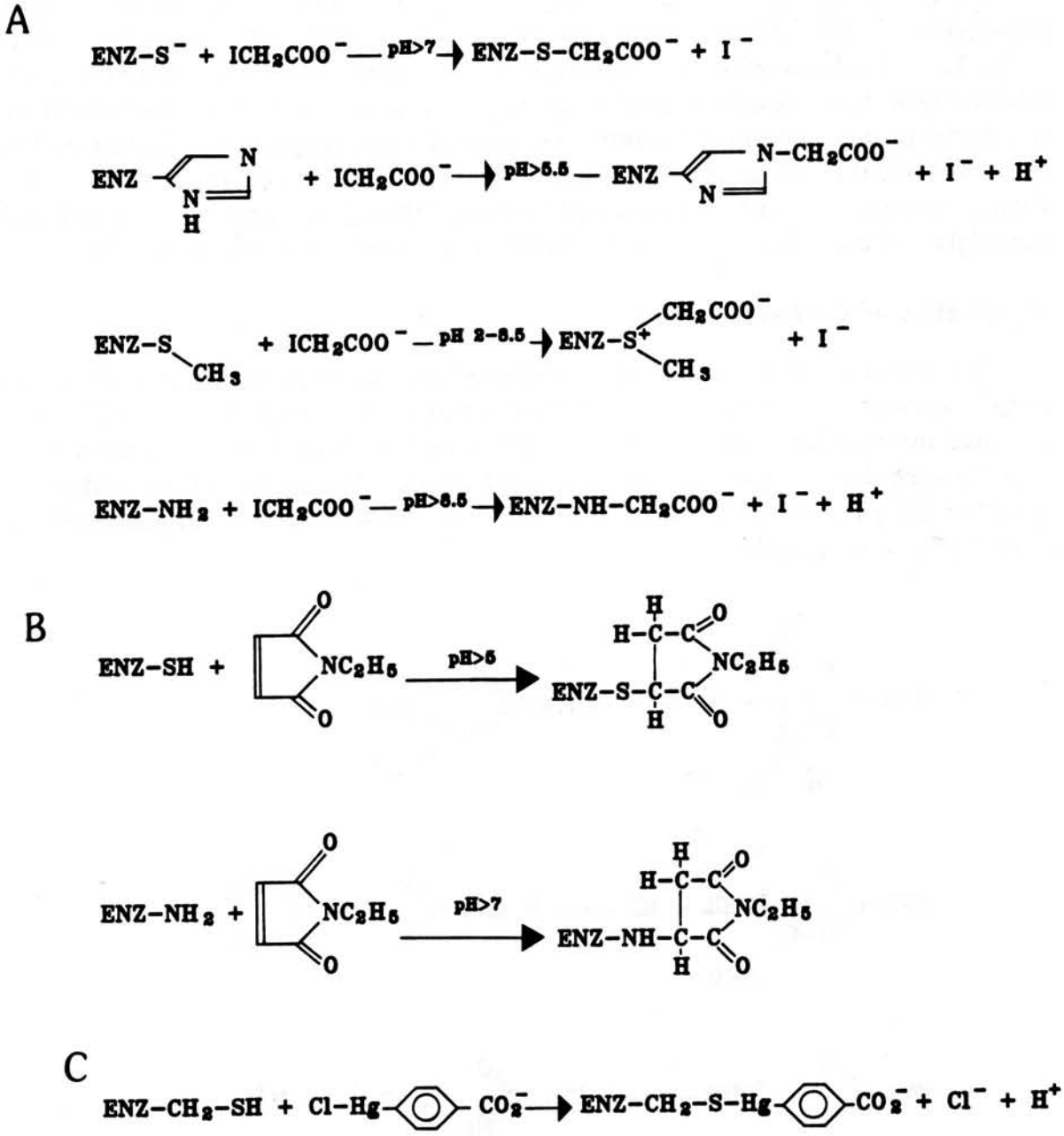


Figure 7-22 Modification of sulfhydryl residues by (A) iodoacetic acid, (B) maleic anhydride, and (C) *p*-chloromercuribenzoate. Also shown are potential side reactions in parts (A) and (B).

TABLE 7-4 Spectrally active reagents based on iodoacetamide

N-Iodoacetylaminoethyl-5-naphthylamine-1-sulfonic acid	
1,5,I-AEDANS	350F495
1,8,I-AEDANS	350F495
7-Diethylamino-3-[(4′-iodoacetylamino)phenyl]-4-methyl coumarin	390F460
2-Anthraceneiodoacetamide	320F420

TABLE 7-5 Spectrally active reagents based on maleimide

N-Pyrene maleimide	386F405
N-1-Anilinonaphthyl-4-maleimide	355F488
7-Diethylamino-3-(4′-maleimidylphenyl)-4-methyl coumarin	390F460
N-[*p*-(2-Benzoxazoyl)phenyl]maleimide	310F375
4-Dimethylamino-4′-maleimido-stilbene	345F480

analysis to prevent disulfide bond formation. A wide series of fluorescent reagents based on iodoacetate have been developed and some of them are listed, together with their fluorescence properties, in Table 7-4. All of these alkylating reagents can react with other amino acid side chains, such as methionine, lysine, or histidine.

Maleic anhydride- or maleimide-based reagents react with sulfhydryls to give acid-stable derivatives. Reactions with other groups may occur, but these derivatives tend to be acid labile. The derivative produced by modification with *N*-ethylmaleimide has an absorption maximum at 300 nm, with a molar extinction coefficient of 620 cm^{-1}, which is too low to allow reasonable quantitation except at high protein concentrations or high degrees of modification. As with the iodoacetate-based reagents, a variety of spectrally active reagents are available based on maleimide, some of which are given in Table 7-5.

The mercury-based reagents are the most specific for sulfhydryls. The *p*-mercuribenzoate derivative of a sulfhydryl has an extinction coefficient at 250 nm of 7500 M^{-1} cm^{-1} at pH 7.0, which allows reasonable quantitation spectrophotometrically even at low (i.e., μM) protein concentrations. A particularly useful derivative of pMB is *S*-mercuric-*N*-dansyl cysteine, whose reaction with sulfhydryls (Fig. 7-23)

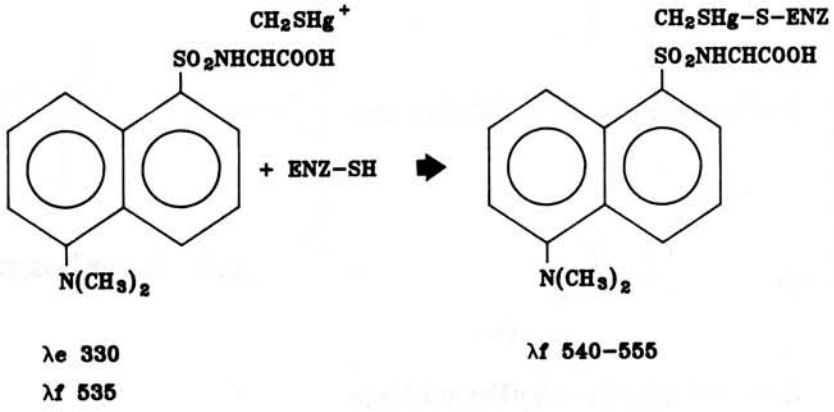


Figure 7-23 Reaction of *S*-mercuric-*N*-dansyl cysteine with sulfhydryl residues.

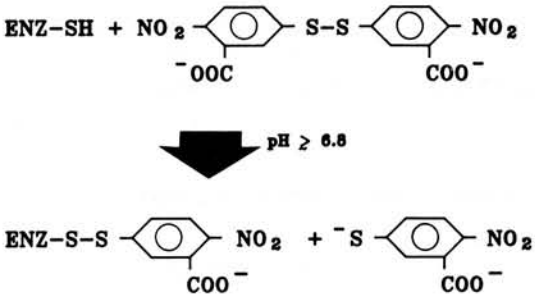


Figure 7-24 Reaction of sulfhydryl group with DTNB.

gives a derivative that can be used to estimate sulfhydryl groups by fluorescence titration, or to introduce a fluorophore onto a reactive sulfhydryl.

Iodoacetic acid and maleic anhydride can both react with other residues in addition to sulfhydryl moieties, and some of these potential side reactions are shown in Fig. 7-22.

The final reagent that is widely used with sulfhydryls is 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB). This forms a mixed disulfide with cysteine (Fig. 7-24) and releases the thionitrobenzoate anion, which can be quantitated spectrophotomerically by absorbance measurements at 412 nm. At pH 8 in aqueous buffers the thionitrobenzoate anion has a molar extinction coefficient of 13,600 cm^{-1}. This coefficient, which is quite dependent on other compounds in the buffer, allows easy and rapid titration of sulfhydryl groups in either native or denatured proteins. It is advisable to determine the extinction coefficient under individual conditions by titration of cysteine with DTNB in control experiments. The derivatives produced by DTNB modification are easily reversible by dithiothreitol or mercaptoethanol.

Modification of Histidine Residues

As was indicated, iodoacetate alkylates histidine residues and both mono- and disubstituted derivatives can be obtained (Fig. 7-25). A particularly useful reagent for histidine modification is diethylpyrocarbonate. The reaction proceeds (Fig. 7-26)

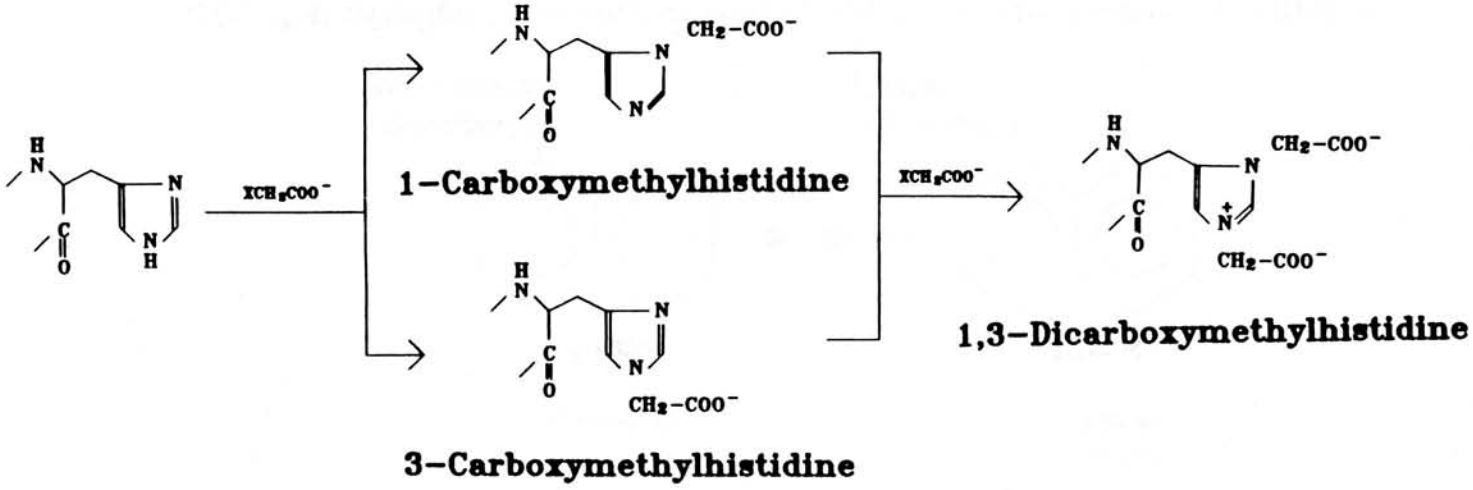


Figure 7-25 Modification of histidine with iodoacetic acid derivatives.

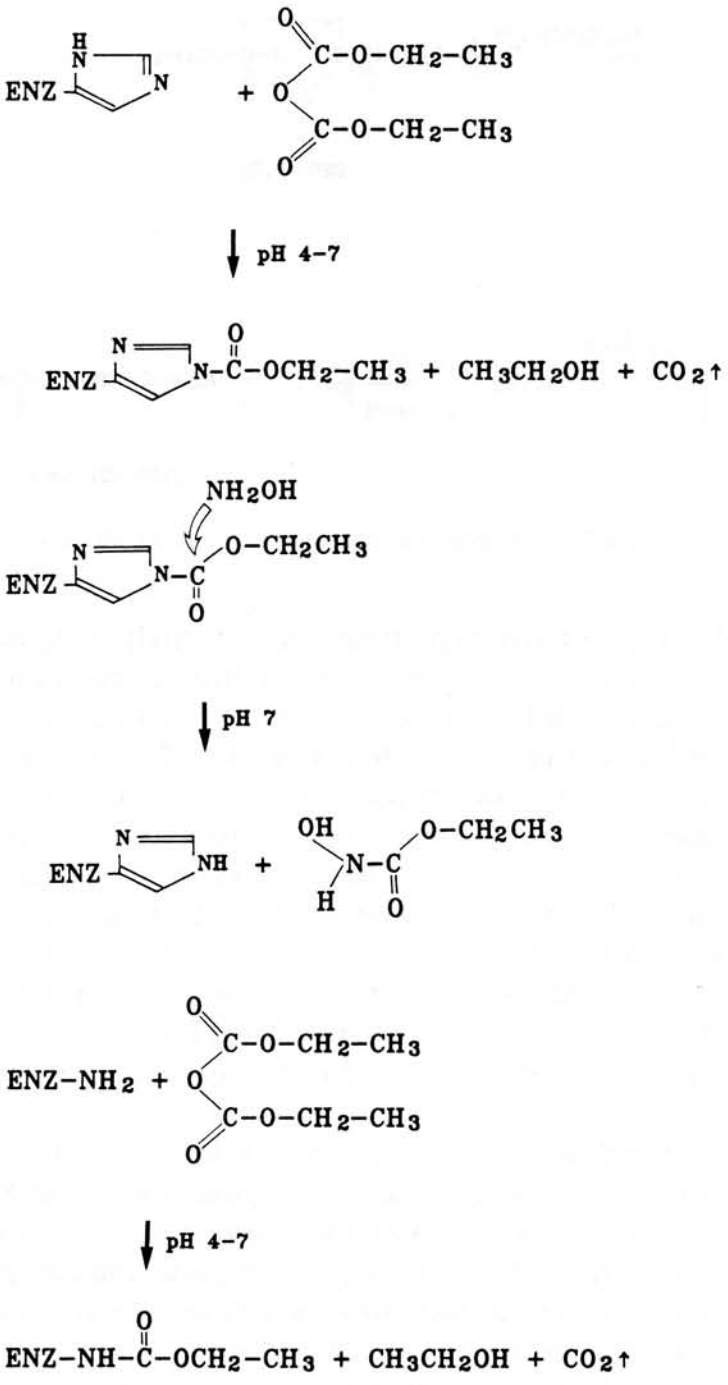


Figure 7-26 Modification of histidine by diethylpyrocarbonate and reversal by hydroxylamine. Also shown is the side reaction with lysine residues.

to produce an *N*-carbethoxy derivative, which has an absorbance band between 230 and 250 nm. At 240 nm, a molar extinction coefficient for this derivative of 3200 cm^{-1} allows the reaction to be quantitated and followed spectrophotometrically. Diethylpyrocarbonate is subject to a rapid hydrolysis, to carbon dioxide and ethanol, and as described earlier, time-course studies must take this into account. In addition to being easily monitored, modification of histidine by diethylpyrocarbonate is readily

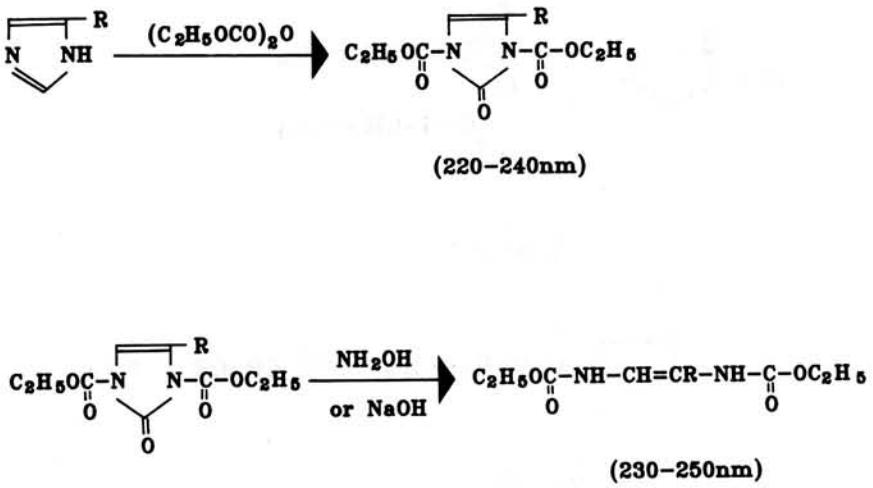


Figure 7-27 Formation of disubstituted imidazole ring.

reversed by incubation with hydroxylamine. Since diethylpyrocarbonate can ethoxyformylate tyrosine, lysine, or sulfhydryl residues (although reaction is minimal at pH 6, where histidine reacts readily), reversal of histidine modification by hydroxylamine is important, as it does not reverse modification of sulfhydryls or lysine side chains.

A point of caution concerning the use of diethylpyrocarbonate must be raised. At very high excesses of diethylpyrocarbonate to histidine it is possible to get modification at both imidazole nitrogens, giving a derivative with a higher extinction at 240 nm, which can lead to misquantitation of the number of histidine residues modified. The disubstituted histidyl derivative (Fig. 7-27) reacts with hydroxlamine, not to give reversal of the modification, but to give cleavage of the imidazole ring and a derivative that still absorbs at 230 to 250 nm. Thus if the disubstituted derivative is formed, it may appear on the basis of hydroxylamine reversal that lysine or cysteine has been modified.

Both iodoacetic acid and diethylpyrocarbonate have been shown to modify *either* N atom in the imidazole ring of histidine. In the case of iodoacetic acid it is possible to isolate 1- and 3-substituted carboxymethyl derivatives and relate activity effects to which atom is modified. Although diethylpyrocarbonate can react with both N atoms to give stable disubstituted histidines, in the monosubstituted product the N-1 substituent is much more stable than the N-3 substituent and diethylpyrocarbonate can effectively be considered as specific for N-1 in monosubstitution reactions.

N-1 modification by diethylpyrocarbonate of an N-3 carboxylmethyl histidine residue in bovine α-lactalbumin has been shown to produce complete inactivation, while the N-3 substituent retains activity.

Modification of Tryptophan Residues

In Chap. 5 we examined the use of brominating reagents such as *N*-bromosuccinimide to cleave polypeptide chains at tryptophan residues. Under mild conditions this reagent oxidizes tryptophan side chains, resulting in a loss of the absorbance

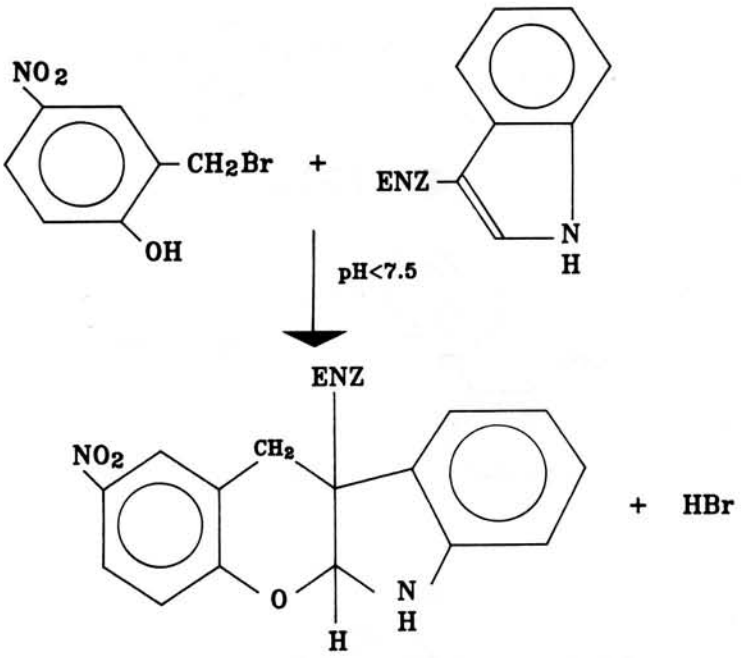


Figure 7-28 Modification of tryptophan with 2-hydroxy-5-nitrobenzylbromide.

properties in the region 270 to 295 nm. Although it has been used successfully to modify tryptophan in some proteins, the conditions must be carefully monitored since too high a reagent concentration leads to peptide-bond cleavage at the modified tryptophan.

Various benzyl halides have been used to alkylate the indole ring of tryptophan. The most popular is 2-hydroxy-5-nitrobenzyl bromide, first proposed as a tryptophan modification reagent by Koshland. Because of solubility problems the analogous dimethyl(2-hydroxy-5-nitrobenzyl) sulfonium salt is often used (Fig. 7-28). In the absence of cysteine residues these reagents are quite specific for tryptophan.

Because of product heterogeneity this reagent is often passed over in favor of various sulfenyl halides, which react with similar specificity but give a single product (Fig. 7-29). When sulfenyl halides with a nitrophenyl substituent are used the chromophoric properties of the nitrophenyl group allow quantitation of accessible tryptophan residues.

Modification of Tyrosine Residues

Tetranitromethane has been extensively used to modify tyrosine residues. The reaction, which can be conducted at mildly alkaline pH, proceeds via a proposed free-radical mechanism (Fig. 7-30) to give a 3-nitrotyrosine derivative that has a pK value of approximately 7.0.

The nitrophenoxide ion has an intense visible absorption. At low pH an absorption maximum of 360 nm with a molar extinction of 2790 cm^{-1} is found, while at higher pH the absorption maximum shifts to 428 nm and the extinction coefficient

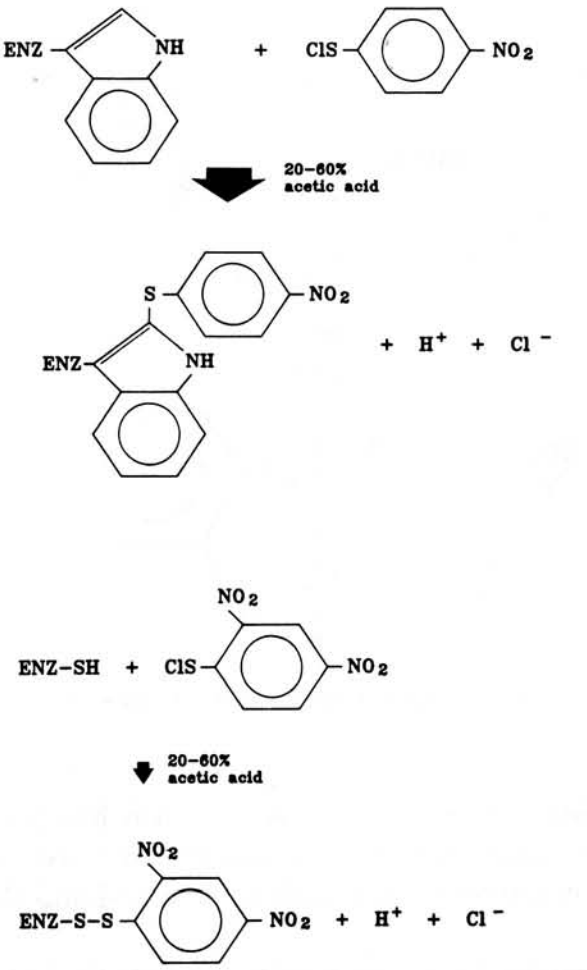


Figure 7-29 Modification of tryptophan with sulfenyl halides. Also shown is a competing reaction with sulfhydryl residues.

increases to 4200 M^{-1} cm^{-1}. This allows for a direct determination of the pK of the modified tyrosine residue. The nitrotyrosine is readily reduced to 3-aminotyrosine by sodium hydrosulfate. The pK value of this aromatic amino group is about 4.8, making it considerably lower than the other amino groups in the protein. This makes 3-aminoytrosine a particularly reactive target of amino-group-specific reagents such as 1-fluoro-2,4-dinitrobenzene (Chap. 5) at low pH, where the reactivity of other

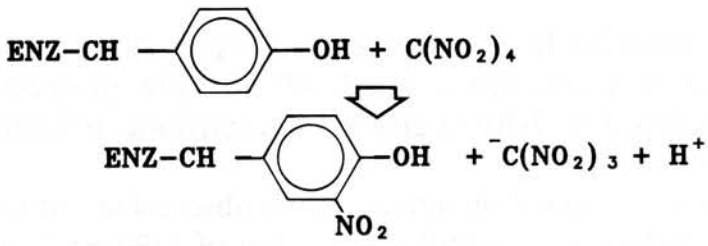


Figure 7-30 Nitration of tyrosine side chains by tetranitromethane.

TABLE 7-6 Specificity of commonly used modification reagents

Reagent	Side chain modified								
	Lys	Glu	Cys	Arg	Ser	His	Tyr	Trp	Met
Acetic anhydride	×		×		×		×		
Acyl anhydride	×		×				×		
Aldehydes	×		×	×		×			
N-Bromosuccinimide			×			×	×	×	
Carbodiimides		×							
Phenylisothiocyanate	×		×				×		
Tetranitromethane			×				×		
Trinitrobenzenesulfonate	×								
Diethylpyrocarbonate	×					×			
Iodoacetate	×		×			×			×
Maleimide	×		×						

amino groups is considerable lower. Where tyrosine can be specifically nitrated, subsequent reduction can allow for specific introduction of a modification with amino-group reagents.

Although tetranitromethane can modify cysteine, methionine, and tryptophan residues, the major problem encountered is the possibility of either intra- or intermolecular cross-linking as a result of side reactions of the proposed tyrosine free-radical intermediate. As discussed in a later section, tetranitromethane is sometimes used as a cross-linking reagent.

As has been emphasized in this section, many of the reagents discussed are not uniquely specific. Table 7-6 gives the specificities of many commonly used modification reagents.

CROSS-LINKING REAGENTS

In the recent past a tremendous amount of work has gone into the synthesis and use of a wide variety of chemical cross-linking reagents. These reagents, which contain two reactive moieties, can be used to cross-link residues within a polypeptide chain, between polypeptide chains in an oligomer, or between protein molecules which may, for whatever reason, associate with one another. The identification of residues (or proteins) that can be cross-linked together can give much important information in the areas of protein conformation and protein–protein interactions. A number of considerations go into the design or selection of a cross-linking reagent, and prior to considering the reagents themselves we examine briefly some of these considerations and their implications.

1. *Reaction Specificity.* As with most chemical modification reagents, reaction specificity and conditions of reaction are important. The appropriate groups the

reagent reacts with must be present and, if structural information is to be obtained, they must react under "native" conditions. Most of the cross-linking reagents available react with amino groups or sulfhydryl groups or either. Cross-linking reagents fall into two groups based on their types of reactivity. When both reactive moieties are the same, the reagent is *homo-bifunctional.* When one of the reactive moieties has a different specificity from the other, the reagent is *hetero-bifunctional.* Hetero-bifunctional reagents often react with amino groups through one of their reactive moieties and a different group through the other. A special class of hetero-bifunctional reagents are the ones that include a *photo-activable* moiety as one of their reactive groups. In many situations such reagents have a distinct advantage over other cross-linking reagents, as once reaction via the non-photo-activable moiety has occurred, cross-linking via the photo-generated free-radical reacting is almost assured and does not depend on the availability of a particular amino acid side chain to react with. This can in some circumstances lead to random cross-linking events; however, such an occurrence may be desirable.

The availability of groups to react to give cross-linking leads to a consideration of the second criterion in selecting a cross-linking reagent.

2. *Cross-Linking Distance.* From a number of standpoints cross-linking distance is an important parameter. As the length between the two functional groups increases, there is an increased probability that a second group on the protein exists which can react with the cross-linker to also give cross-linking. Cross-linking with a reagent that spans, for example, 10 Å can occur only if a second reactive group is within 10 Å of the first. Because a cross-linking reagent *can* span 10 Å does not mean that the second group is 10 Å from the first, as many of the reagents are quite flexible and can react with groups located at distances less than the maximum span length. If a series of cross-linking reagents can be used with similar reaction specificities but varying span length, considerable spatial information concerning groups can be obtained.

Depending on the type of cross-linking that is being attempted (i.e., intra-peptide or inter-peptide), one might select reagents with either short cross-linking distances or long ones. In many instances where inter-peptide cross-linking is attempted, none is found until reagents of a certain minimum length are used.

The third characteristic that must be considered is useful both in the identification of cross-linked peptides and in controlling that any effects observed after cross-linking are due to the cross-linking itself rather than to the chemical modification event.

3. *Cleavability.* The development of cleavable cross-linking reagents has greatly assisted the isolation and characterization of cross-linked fragments. These reagents contain a moiety between the two functional groups which can, after cross-linking, be cleaved by a reagent, optimally under conditions that are not deleterious to the protein itself. Combined with two-dimensional gel electrophoresis, cleavable cross-linking reagents considerably ease the task of identification and isolation of cross-linked components.

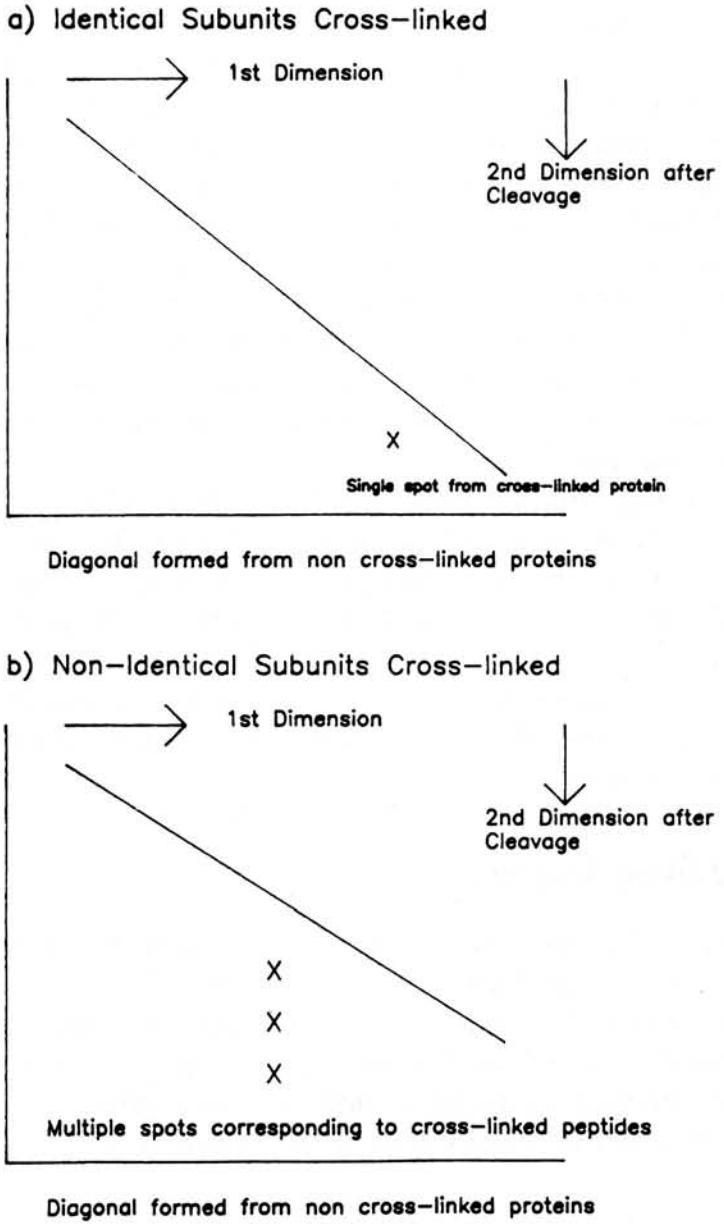


Figure 7-31 Schematic representation of two-dimensional PAGE results for homo- and hetero-polymers cross-linked with cleavable reagents.

Two situations can be envisaged involving multi-subunit proteins that give different types of experimental observations in two-dimensional gels with cleavage between the dimensions. They are summarized in Fig. 7-31.

In addition to this, cleavable cross-linking reagents can assist in control experiments to show whether or not any altered biological activity is the result of the cross-linking event or simply the chemical modification event. In the former case, the biological activity would be regained after cleavage of the cross-link (hence the importance of a cleavage procedure that does not harm the protein's integrity), whereas

if the activity change resulted from the modification event, cleavage of the cross-link would not regenerate activity. A second way to examine this question is to use monofunctional reagents with the same reaction specificity as the cross-linking reagent. In this case, cross-linking cannot occur and the question of whether the modification itself causes the loss of activity can be examined. It must be emphasized that simply because a cross-linking reagent is being used it cannot be inferred that cross-linking is the only type of chemical modification taking place. It is quite probable that, in addition, monofunctional modification may be occurring. This monofunctional modification can arise from two causes: (a) a lack of an available second amino acid side chain for cross-linking, and (b) degradation of one of the reactive moieties of the cross-linking reagent prior to cross-link formation, essentially converting the reagent into a monofunctional reagent.

4. *Reactive Moiety*: Finally, we can consider the reaction characteristics of the reactive moiety of the cross-linking reagent. As with many chemical modification reagents, the site of reaction can be controlled to a certain extent. It is possible to direct the reaction by considering the hydrophobicity, hydrophilicity, or charge of the reactive moiety. In addition, it may be possible to incorporate aspects of a site-directed irreversible inhibitor into one moiety of a hetero-bifunctional reagent, allowing reaction of one end of the cross-linker to be directed toward a particular binding site on the protein.

Types of Cross-Linking Reagents

A wide variety of cross-linking reagents have been developed in recent years with varying specificity, cleavability, and span lengths. They can be categorized into three basic classes, and within each class can be either cleavable or noncleavable. In many of the types of reagents used, a series of analogous reagents having differing span lengths are obtained by using a suitable spacer group between the reactive moieties. The three types of cross-linking reagents that we consider are summarized in Fig. 7-32.

Homo-bifunctional Reagents. This first class of reagents have the same reactive moiety at either reactive center, and their specificity is determined by this moiety. Typical of this class of reagent are the bisimidates, a series of bifunctional imido esters having $(CH_2)_n$ spacer groups with $n = 1$ (malonimidate) to $n = 6$ (suberimidate), spanning 5 to 11 Å and reacting with amino groups. Such reagents can be converted to cleavable homo-bifunctional reagents by inclusion of a disulfide bridge, as in dimethyl-3-3′-dithiobispropionimidate, which has a span length of 12 Å and is cleavable by mercaptans. Although imidoesters are often used in homo-bifunctional reagents, other amino-specific reagents such as *N*-hydroxysuccinimide esters or 1,5-difluoro-2,4-dinitrobenzene have also been employed.

Essentially all homo-bifunctional reagents are amino-group specific, with the exception of reagents such as glutaraldehyde and formaldehyde. Although these aldehyde reagents have been used in cross-linking studies, their uncharacterized and

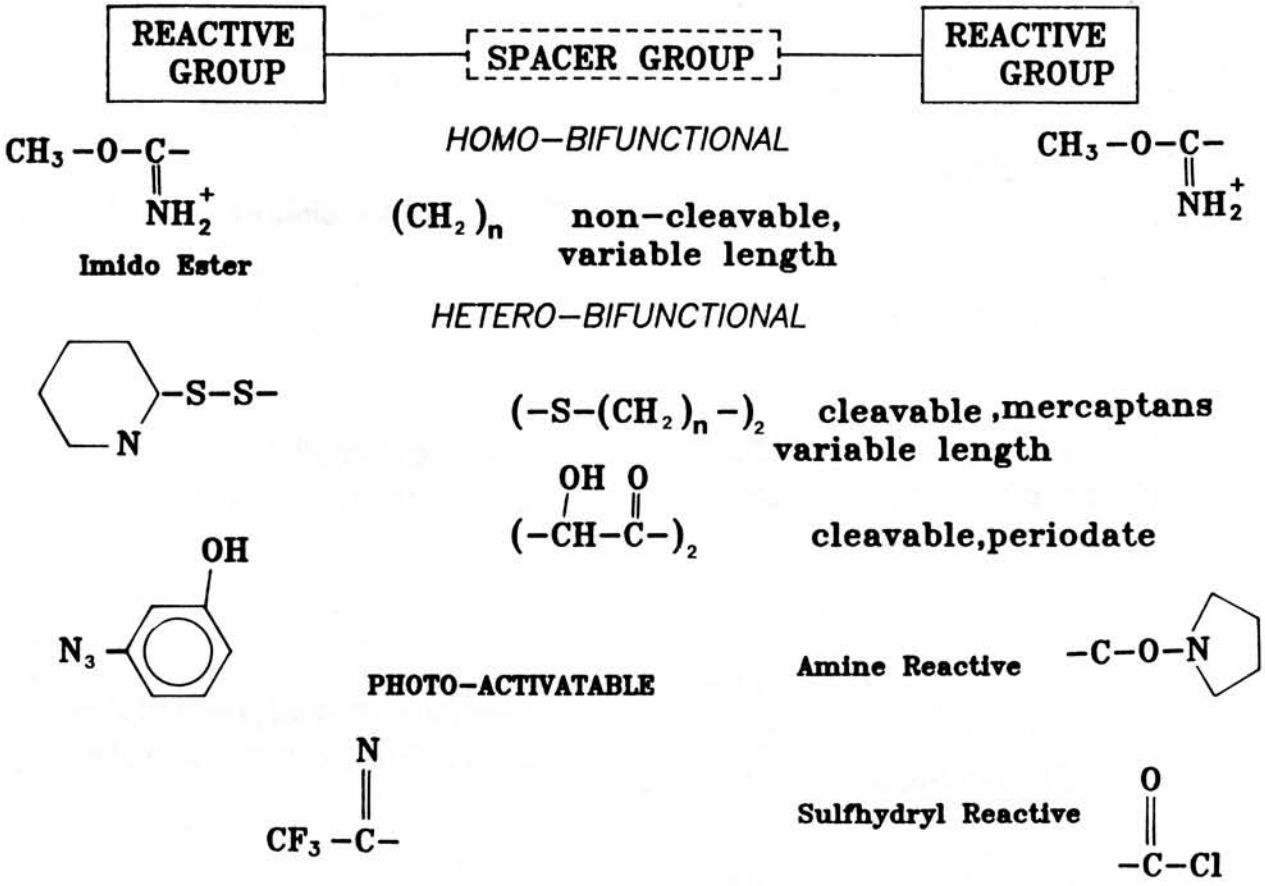
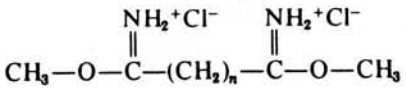


Figure 7-32 Types of cross-linking reagents.

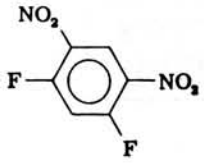
nonspecific reactivity makes them reagents of last resort, especially when limited extents of cross-linking are required.

Hetero-bifunctional Reagents. In all cases, one of the reactive moieties in this class of reagents reacts with an amino group, and the other reactive moiety usually reacts with a sulfhydryl side chain, although in the case of the carbodiimides the second reacting group is a carboxyl group. Except for the various carbodiimides these reagents usually contain the amino-reactive moiety *N*-hydroxysuccinimide and often a sulfhydryl-reactive maleimido or dithio moiety.

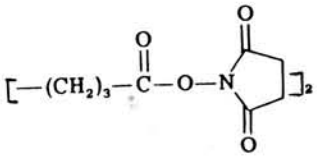
Photo-activable Reagents. This group of reagents is really an extension of the hetero-bifunctional reagents but has a more widespread utility. This comes from the fact that one of the reactive moieties is generated in situ after reaction of the first reactive moiety with the protein, which helps reduce the potential for monofunctional modification, and from the fact that the generated second reactive moiety is either a carbene (from diazo reagents) or a nitrene (from azido reagents) free radical that is highly reactive but quite nonspecific. As a result, there is little requirement for certain types of residues to be located close to one another to allow for cross-linking.



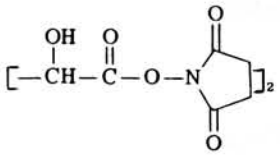
Bisimidates



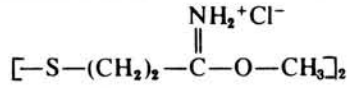
1,5–Difluoro–2,4–(dinitrobenzene)



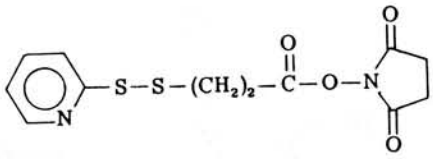
N–Hydroxysuccinimide ester of Suberic acid



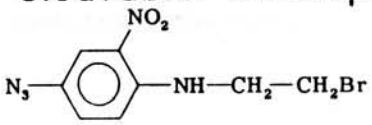
Disuccinimidyl Tartarate
Cleavable: Periodate



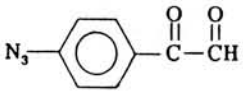
Dimethyl–3.3'–dithio–bispropionimidate
Cleavable: Mercaptans



N–Succinimidyl–3–(2–pyridyldithio)propionate



4–(Bromoaminoethyl)–2–nitrophenylazide



4–Azidoglyoxal

Figure 7-33 Structures of some commonly used cross-linking reagents.

The initial reactive moiety is either amino or sulfhydryl reactive as with other heterobifunctional reagents, and as with other reagents discussed, various span lengths or cleavability can be built into the reagents.

The structures of some of the more popular cross-linking reagents that have been used are shown in Fig. 7-33, together with the basis of their cleavability where appropriate.

Examples of the Uses of Cross-Linking Reagents

Determination of Interlysyl Distances in Glycogen Phosphorylase b

Reference: J. Hadju, et al., *Biochemistry, 18*, 4037–4041 (1979).

This enzyme undergoes a dimer–tetramer association reaction that appears to be related to the allosteric properties of the enzyme. Experiments involved cross-linking the enzyme with a series of imido esters whose cross-linking span length

ranges from 3.7 to 14.5 Å, electrophoresing the cross-linked products in an SDS–PAGE system, and quantitating the amount of each species, after staining, by densitometry. From such data two parameters are obtained: r_k, the rate-constant ratio of cross-linking, and C_d, the percent of cross-linked parameter, which are defined in Eqs. (7-32) and (7-33), respectively,

$$r_k = \frac{k_L}{k_0} \tag{7-32}$$

where k_L and k_0 are the apparent first-order rate constants for the disappearance of the monomer bound in the presence or absence of an added ligand, respectively.

$$C_d = \frac{[\text{trimer} + \text{tetramer}]}{[\text{total}]} \times 100 \tag{7-33}$$

Since all samples can be cross-linked under identical conditions except the presence of added ligand, both parameters can be calculated from densitometer scans

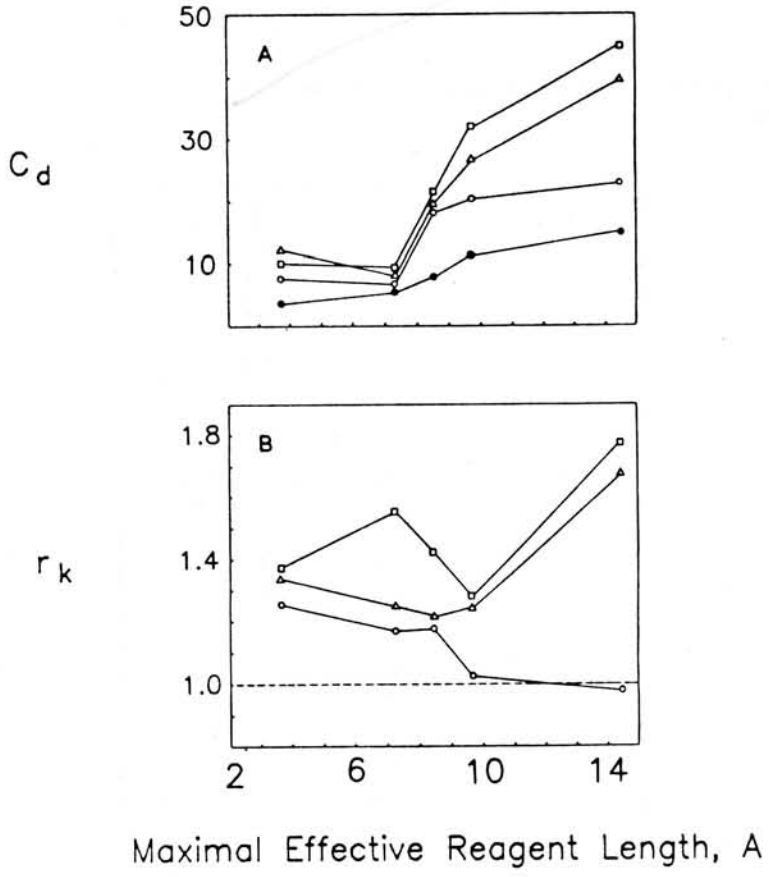


Figure 7-34 Effects of span length and AMP concentration on cross-linking parameters: cross-linking was performed with diimidates varying from 3.7 to 14.5 Å in span length in the absence of AMP (●) or in the presence of AMP: 0.1 m*M* (○); 0.3 m*M* (□); and 1 m*M* (△). (Reprinted with permission from: J. Hajdu, V. Dombradi, G. Bot, and P. Friedrich, *Biochemistry*, *18*, 4037–4041. Copyright 1979 American Chemical Society, Washington, D. C.)

of enzyme cross-linked in the absence or presence of ligand, and the parameter r_k is calculated from

$$r_k = \frac{\ln (\text{monomer/total})_L}{\ln (\text{monomer/total})_0} \tag{7-34}$$

The effects of either cross-linker length or of an added allosteric ligand, AMP, on these two parameters is shown in Fig. 7-34. The amount of trimer and tetramer formed, as shown by C_d, increases with a span length greater than 8 Å, and this increase is amplified by the presence of AMP (Fig. 7-34A), indicating that for effective cross-linking to give tetramer there must exist two lysine residues approximately 8 Å apart. AMP clearly functions to increase the amount of tetramer present rather than bringing the lysines closer together in the tetramer, since even in the presence of saturating AMP concentrations, no significant tetramer is formed with cross-linking reagents having a span of <8.5 Å.

Covalent Cross-Linking of the Active Sites of Vesicle-Bound Cytochrome b5 and NADH–Cytochrome b5 Reductase

Reference: C. S. Hackett and P. Strittmatter, *J. Biol. Chem.*, *259*, 3275–3282 (1984).

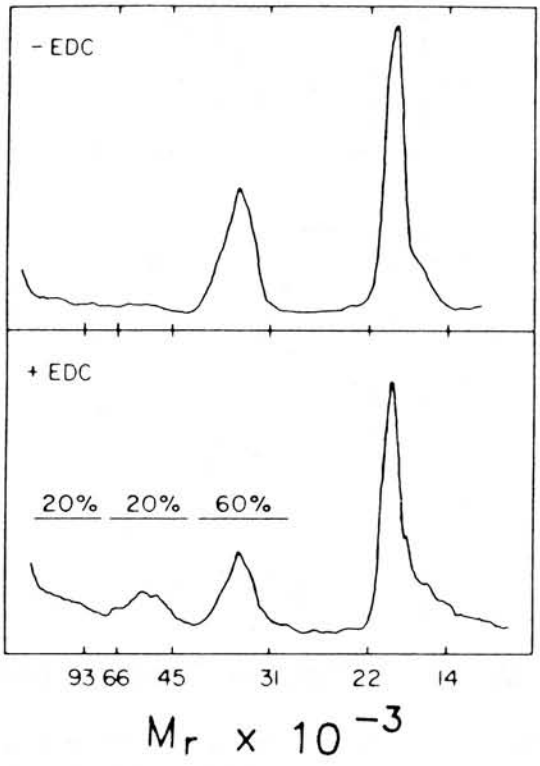


Figure 7-35 SDS-gel electrophoresis of cytochrome b_5–cytochrome b_5 reductase vesicles before and after cross-linking with EDC. [From *J. Biol. Chem.*, *259*, 3275–3282 (1984). Reprinted with permission of the copyright owner, The American Society of Biological Chemists, Inc., Bethesda, Md.]

The water-soluble carbodiimide 1-ethyl-3-(2-dimethylaminopropyl)carbodiimide hydrochloride (EDC) was used to cross-link cytochrome b_5 and NADH–cytochrome b_5 reductase. EDC is an interesting cross-linking reagent, being a so-called "zero-length" cross-linker. It acts by promoting the formation of an amide linkage between a carboxyl group on cytochrome b_5 and a lysyl residue on the reductase. Figure 7-35 shows a gel scan of cytochrome b_5 and the reductase electrophoresed with or without cross-linking by EDC. The cross-linked complex can then be incorporated into vesicles to study the effects of interaction of the two proteins during function.

Cross-Linked Galactosyltransferase and α-Lactalbumin: Use in Site–Site Distance Estimates by Fluorescence Resonance Energy Transfer Measurements

Reference: E. T. O'Keeffe, T. Mordick, and J. E. Bell, *Biochemistry*, *19*, 4962–4966 (1980).

A problem frequently encountered in physical measurements on the interaction of two proteins is the weak nature of such interactions. This is particularly difficult in fluorescence measurements where one component is fluorescent and the other can act as a resonance energy transfer acceptor of that fluorescence. This was the case in studies of the interaction of α-lactalbumin with galactosyltransferase, where dansylated α-lactalbumin was used to estimate a distance from the dansyl to cobalt bound to the metal-ion binding site of the transferase. Under normal circumstances an approximately 1000-fold excess of α-lactalbumin is required to bind the transferase completely. Since significant amounts of the 1:1 complex are required for the fluorescence measurements this would have led to an extremely unfavorable signal-to-background ratio. This problem was overcome by chemically cross-linking the dansyl α-lactalbumin to the transferase in a 1:1 complex and performing the fluorescence measurements on this complex. SDS-PAGE shows that cross-linking of α-lactalbumin to galactosyltransferase occurs. Since the cross-linked complex is not much larger than the un-cross-linked enzyme, it is difficult to separate the cross-linked from residual un-cross-linked activity by gel filtration. However, they can easily be separated using an α-lactalbumin affinity column. The resulting complex can be titrated with cobalt to estimate the quenching due to resonance energy transfer between the dansyl group on the α-lactalbumin (shown to be uniquely labeled at the amino terminal) and the cobalt metal binding site of the transferase. From the quenching a distance of 32 Å between the metal site and the dansyl group on the α-lactalbumin was calculated.

Possible Quaternary Structures for the Hexamer Glutamate Dehydrogenase Established by Cross-Linking Studies

Reference: T. J. Smith and J. E. Bell, *Arch. Biochem. Biophys.*, *239*, 63–73 (1985).

Cross-linking studies can be used to distinguish between possible quaternary structural arrangements of subunits. This topic is dealt with in detail in Chap. 11, but a brief outline of the results obtained with glutamate dehydrogenase is given

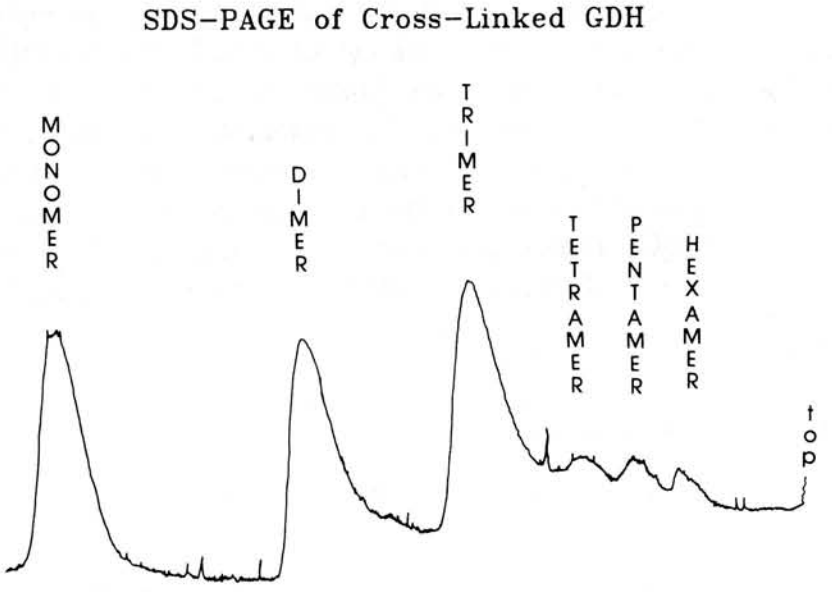


Figure 7-36 Densitometer scan of glutamate dehydrogenase cross-linked with dimethylpimilimidate.

here. The six subunits can be arranged either with cyclic symmetry (C-6) or with dihedral symmetry (D-3). They predict quite different cross-linking distributions of monomer, dimer, trimer, tetramer, pentamer, and hexamer obtained as a result of intersubunit cross-linking. Figure 7-36 shows a densitometer gel scan of glutamate dehydrogenase cross-linked with dimethylpimelimidate. From such data the amounts of each cross-linked species can be quantitated and compared with predicted stoichiometries. In this case the results eliminated cyclic symmetry from further consideration.

Identification of Intrapeptide Cross-Linking in the Monomer of Glutamate Dehydrogenase by Dimethylpimelimidate

Reference: T. J. Smith and J. E. Bell, unpublished results.

In attempts to predict tertiary structure on the basis of sequence information (see Chap. 10) the major problem encountered is the number of possible conformations a polypeptide may have. If this number can be limited by direct experimental information, it becomes feasible to attempt to predict tertiary structures. Chemical cross-linking of nearby residues and identification of the residues involved represents an attractive approach to obtaining information that can limit the number of conformations needing consideration. Clearly, experiments with cross-linking reagents having different span lengths can provide a wealth of information when intra-peptide cross-links can be identified. Experimentally, this is quite simple with a monomeric protein; however, with an oligomeric protein such as glutamate dehydrogenase the possibility of inter-polypeptide cross-links as well as intra-polypeptide cross-links presents difficulties. One approach is to use cross-linking reagents with short span lengths that cannot cross-link between subunits. This represents, however, an unnecessary limit on the possible information that can be obtained in such experiments.

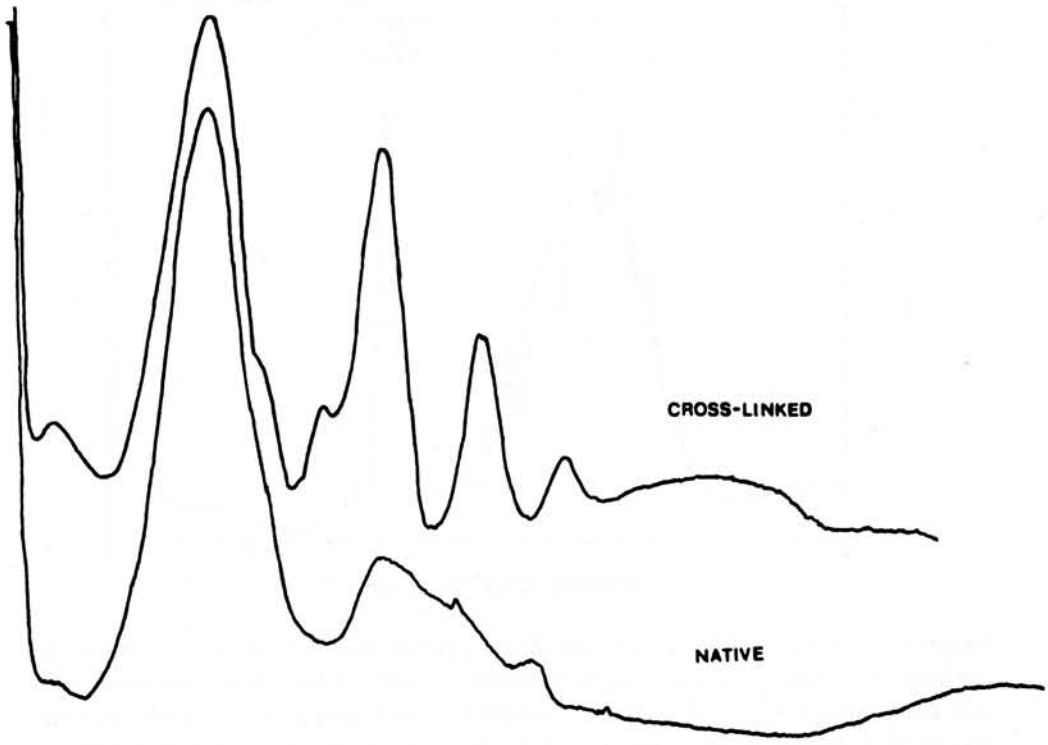


Figure 7-37 Densitometer scans of cyanogen bromide peptides obtained from native GDH and cross-linked trimer GDH.

In an alternative approach, cross-linked enzyme is electrophoretically separated as in the previous example, and the separated monomer, dimer, trimer, and so on, subjected to cleavage in situ in the gel. The cleaved monomer, dimer, trimer, and so on, are then electrophoresed in a second dimension together with a sample of un-cross-linked enzyme that has been electrophoresed and also cleaved in situ in the gel. Figure 7-37 shows a comparison of densitometer scans of native, un-cross-linked enzyme and cross-linked monomer. Clearly, a number of new peptides arising from intrapeptide cross-links are seen in the second dimension of the cross-linked monomer, which can be excised, further purified, and identified to give the desired information.

EXAMPLES OF MODIFICATION STUDIES

Uses of Chemical Modification in Hybridization Experiments

Reference: I. Gibbons and H. K. Schachman, *Biochemistry*, *15*, 52–60 (1976).

A particularly interesting example of some uses of chemical modification is provided by the work on aspartate trancarbamoylase. This enzyme consists of six catalytic peptides arranged in two trimers and six regulatory peptides arranged in three

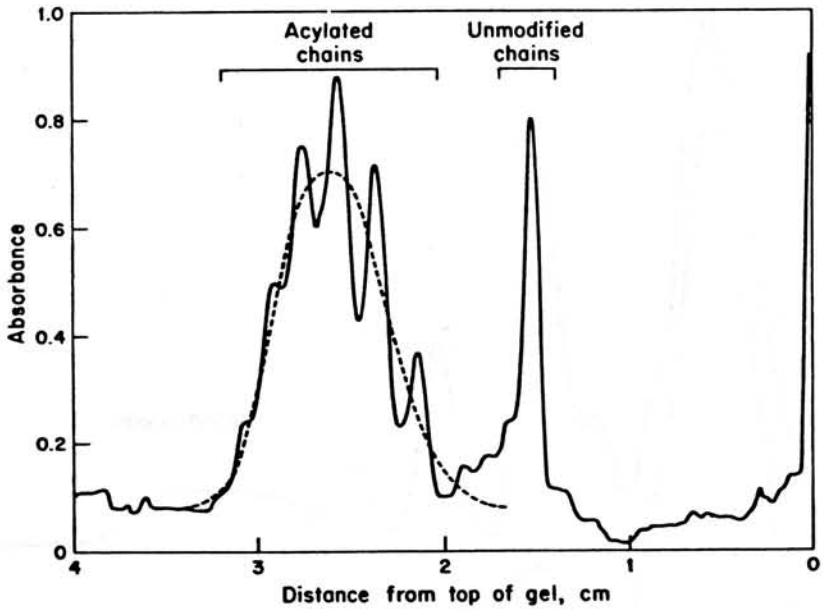


Figure 7-38 Electrophoresis patterns of polypeptide chains of catalytic subunit, the acylated derivative, and the deacylated product. Electrophoresis in polyacrylamide gels containing 8 m*M* urea was used. (Reprinted with permission from: I. Gibbons and H. K. Schachman, *Biochemistry, 15*, 52–60. Copyright 1976 American Chemical Society, Washington, D. C.)

dimers. The regulatory dimers appear to act as a "bridge" or "cross-link" between the two catalytic trimers, which have no direct contact. The catalytic trimers and the regulatory dimers can be separated and purified.

The catalytic subunits can be nitrated with tetranitromethane in the presence of the substrate carbamoyl phosphate and the analog succinate to give nitrated subunits with approximately 0.8 nitrotyrosine per polypeptide chain but 80 to 90% activity. These active, nitrated catalytic subunits can be pyridoxylated at one lysine per polypeptide chain with 90% loss of activity.

Nitrated, pyridoxylated catalytic subunits can be made more negative by modification with 3,4,5,6-tetrahydrophthalic anhydride. As described earlier, acylation can be reversed by incubation at low pH to give the nitrated, pyridoxylated catalytic subunits. Figure 7-38 shows the electrophoretic separation of acylated from nonacylated catalytic subunits.

Reconstituted enzyme can be made by mixing catalytic and regulatory subunits. If native catalytic subunits, C_n, and inactive, nitrated, pyridoxylated, acetylated (negatively charged) subunits, C_t, are mixed with an excess of regulatory subunits (R), a series of reconstituted molecules will be formed: C_n-R-C_n, C_n-R-C_t, and C_t-R-C_t. These hybrid molecules can be separated by ion-exchange chromatography, as shown in Fig. 7-39. Once separated the acylation can be reversed to give a variety of normally charged but chemically modified hybrids.

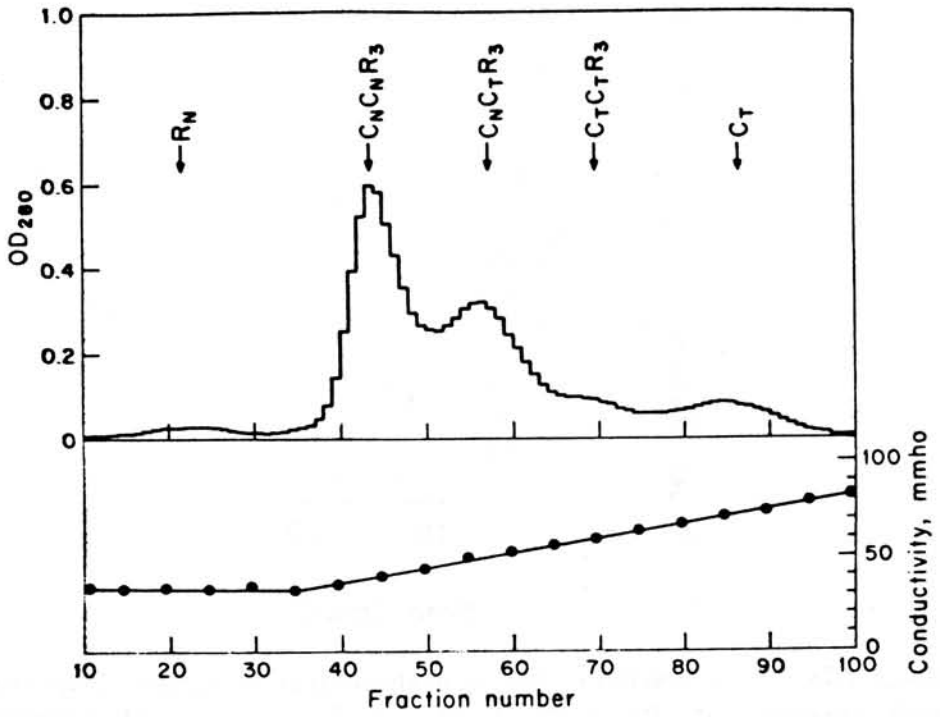


Figure 7-39 Chromatographic fractionation on DEAE-Sephadex of hybrid set formed by reconstitution of ATCase-like molecules from C_n, C_t, and R. (Reprinted with permission from: I. Gibbons and H. K. Schachman, *Biochemistry*, *15*, 52–60. Copyright 1976 American Chemical Society, Washington, D. C.)

Clearly, the possibilities of this type of manipulation with oligomeric proteins are almost unlimited, and some of the uses of this type of hetero-oligomer construction are examined in relation to allosteric proteins in a later chapter. Generation of differently charged subunits can also be of use in establishing quaternary structure, as detailed in Chap. 11.

Demonstration of Two Pyridoxal-5′-Phosphate (PLP) Reactive Lysines in Glutamate Dehydrogenase

Reference: J. C. Talbot et al., *Biochim. Biophys. Acta*, *494*, 19–32 (1977).

The time course of Schiff's base formation by PLP binding to glutamate dehydrogenase is biphasic (Fig. 7-40), indicating the probable existence of at least two classes of lysine residues per polypeptide chain.

Examination of the time course of inactivation of the enzyme by PLP showed approximately 95% loss of activity in the absence of protecting ligands and no loss of activity in the presence of NADPH and 2-oxoglutarate (Fig. 7-41).

Tryptic mapping of enzyme modification with PLP after reduction with sodium borohydride in the absence or presence of protection showed two modified peptides

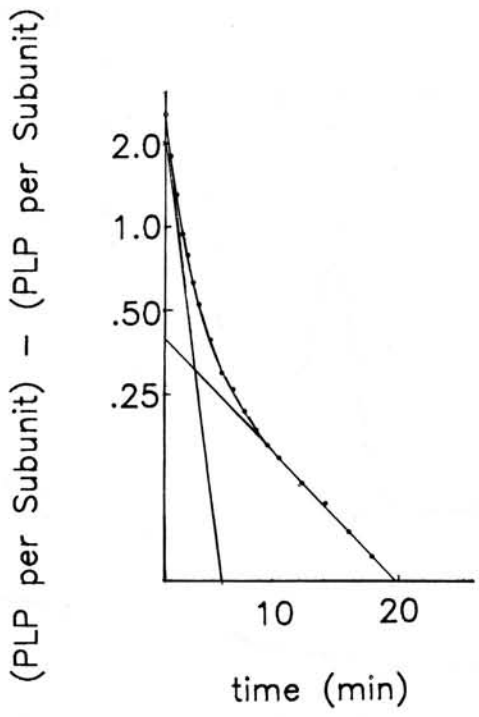


Figure 7-40 Time course of PLP binding to glutamate dehydrogenase. [Reprinted with permission from: J. C. Talbot, C. Gros, M. P. Cosson, and D. Pantaloni, *Biochim. Biophys. Acta*, *494*, 19–32 (1977).]

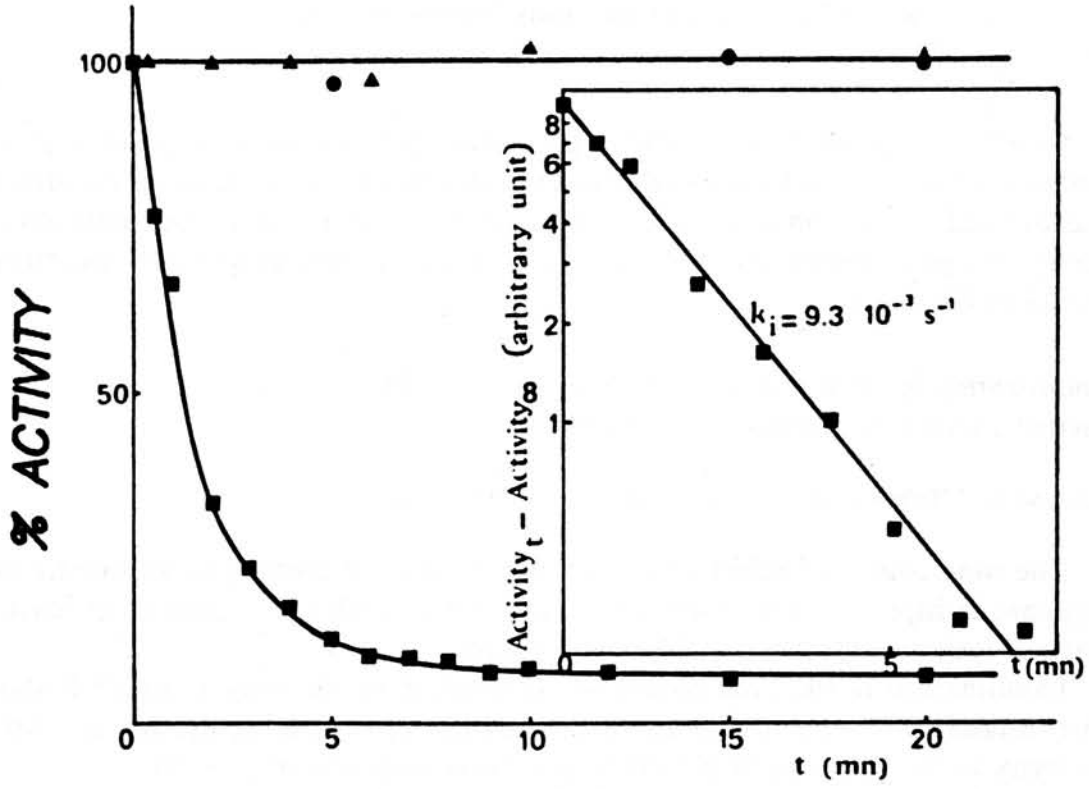


Figure 7-41 The inset shows a plot of log versus time and gives a rate constant of $9.3 \times 10^{-3}\ \text{sec}^{-1}$. [Reprinted with permission from: J. C. Talbot, C. Gros, M. P. Cosson, and D. Pantaloni, *Biochim. Biophys. Acta*, *494*, 19–32 (1977).]

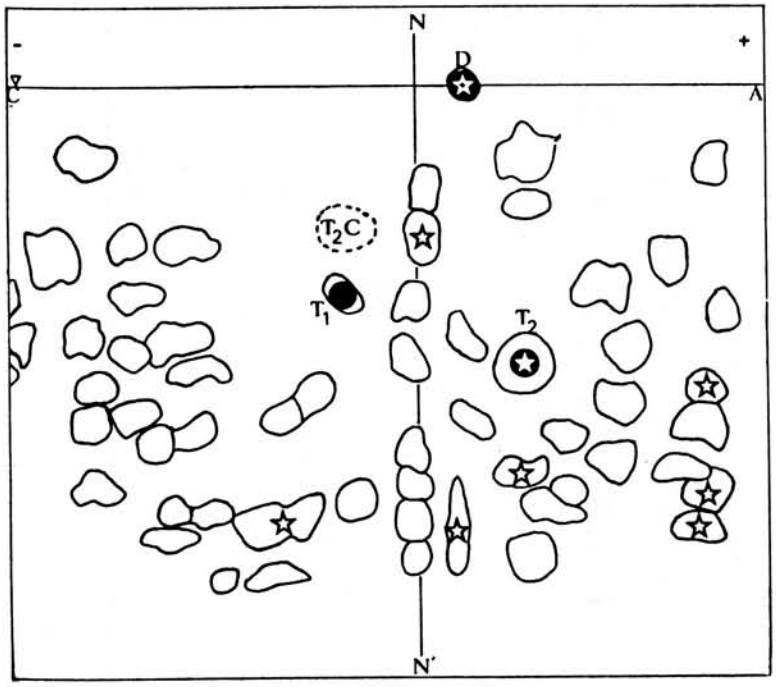


Figure 7-42 Peptide map of tryptic digest of [^{14}C]carboxymethylated reduced pyridoxal-P enzyme: T_1, peptide containing reduced pyridoxal-P-Lys (Lys-333); T_2, peptide containing reduced pyridoxal-P-Lys (Lys-126); T_2C, peptide containing reduced pyridoxal-P-peptide deriving from tryptic and chymotryptic digestion of pyridoxal-P enzyme, or from chymotryptic digestion of T_2. ●, Fluorescent-reduced pyridoxal-P-peptide spots under 300- to 350-nm ultraviolet light; ☆, radioactive spots; △, electrophoretic mobility of lysine. [Reprinted with permission from: J. C. Talbot, C. Gros, M. P. Cosson, and D. Pantaloni, *Biochim. Biophys. Acta*, *494*, 19–32 (1977).]

in the former case, corresponding to lysine-126 and lysine-333, but only a single modified peptide, corresponding to lysine-333, in the latter case (Fig. 7-42).

Subsequent studies indicated that enzyme with only lysine-333 pyridoxylated no longer underwent the concentration dependent aggregation the enzyme is known to show, but otherwise appeared normal.

Attachment of Metal Chelating Groups to Macromolecules Using "Bifunctional" Chelating Agents

Reference: C. Leung and C. F. Meares, *Biochem. Biophys. Res. Commun.*, *75*, 149–155 (1977).

In a novel approach to using chemical modification, diazonium chelating agents were used to derivatize proteins, mainly through reaction with lysine and histidine residues.

Once derivatized the proteins could be titrated with the fluorescent lanthanide europium (Eu), which indicated excellent agreement with the molar incorporation of EDTA groups (Fig. 7-43).

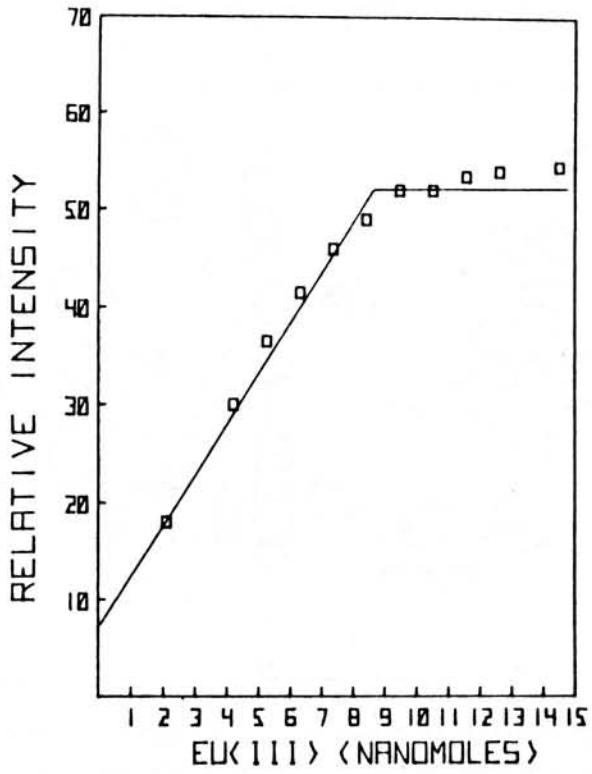


Figure 7-43 Fluorescence titration of albumin-bound EDTA groups (9 nmol) with Eu(III) in 0.1 *M* sodium citrate, pH 6.5. Fluorescence of Eu(III) excited at 310 nm and monitored at 618 nm. (Reprinted with permission from: C. S.-H. Leung and C. F. Meares, *Biochem. Biophys. Res. Commun.*, *75*, 149–155. Copyright 1977 Academic Press, Inc., New York.)

To be useful as a procedure for introducing fluorescent molecules onto a protein, such modifications must be shown to be unique if use in resonance energy transfer is anticipated.

Chemical Modification of Histidine Residues in p-Hydroxybenzoate Hydroxylase

Reference: R. A. Wijnands and F. Muller, *Biochemistry*, *21*, 6639–6646 (1982).

Figure 7-44 shows difference spectra of enzyme modified with diethylpyrocarbonate at pH 5.6 and at pH 8.0. The spectrum obtained at pH 5.6 is typical of histidine modification, and from the absorbance at 240 nm it was calculated that four histidine residues were modified, using an extinction coefficient of 3200 M^{-1} cm^{-1}. As shown in Table 7-7, essentially all the activity could be recovered after incubation of enzyme at pH 6, indicating the probability that only histidine residues are modified under these conditions.

The difference spectrum obtained by modification at pH 8.0 shows a distinctive region between 270 and 280 nm, suggesting that under these conditions tyrosine as well as histidine is being modified.

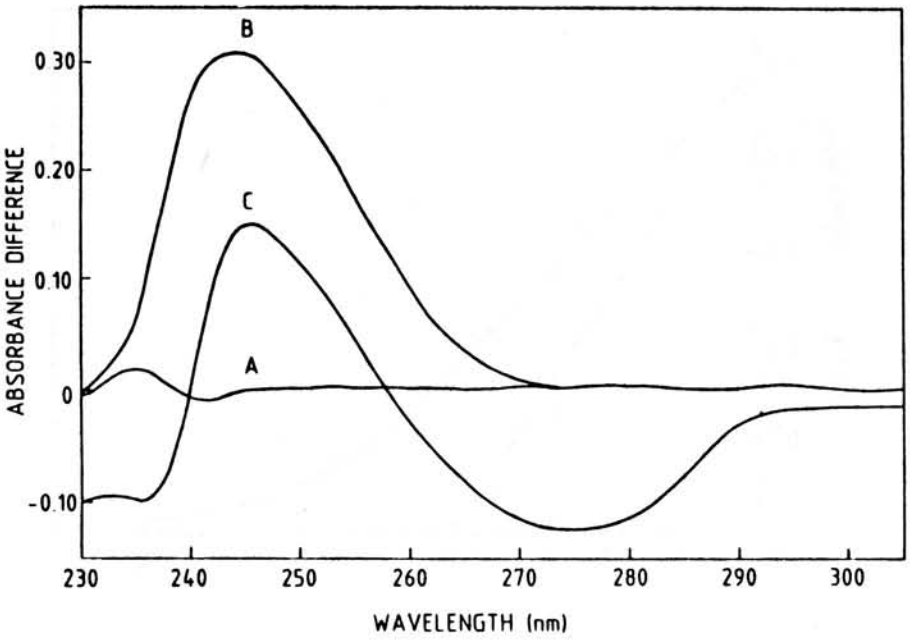


Figure 7-44 Ultraviolet difference spectra of hydroxybenzoate hydroxylase modified with DEP at pH 5.6 (B) or pH 8.0 (C). Line A is the baseline. (Reprinted with permission from: R. A. Wijnands and F. Muller, *Biochemistry*, *21*, 6639–6646. Copyright 1982 American Chemical Society, Washington, D. C.)

Figure 7-45 shows plots of percent activity remaining versus number of histidine residues modified in the absence of protecting ligand and in the presence of *p*-fluorobenzoate in part (A). Part (B) shows Tsou plots of the data, indicating that in the absence of *p*-fluorobenzoate, two of the four modifiable histidine residues are essential for activity, while in the presence of *p*-fluorobenzoate, one of the two modifiable residues is essential for activity.

Table 7-7 Time course of the reactivation of modified enzyme by hydroxylamine[a]

Time (hr)	Percent activity
0	2
0.5	9
1.0	14
1.5	75
2.0	86
2.5	92

[a] Enzyme was incubated with 116 m*M* hydroxylamine at pH 7.0.

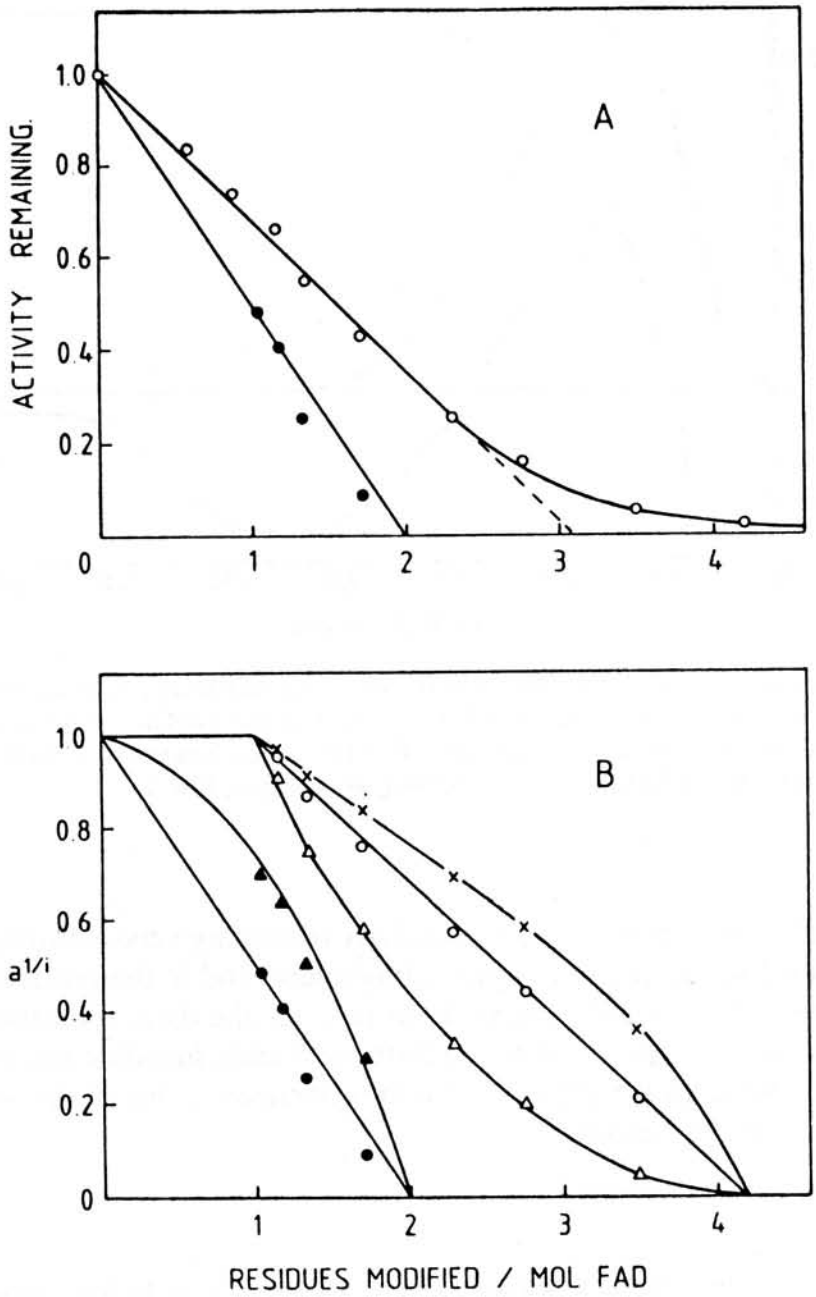


Figure 7-45 (A) Correlation between the number of histidine residues modified by diethylpyrocarbonate and the fractional activity remaining. The closed symbols represent the data for the enzyme–*p*-fluorobenzoate complex. (B) The data of part (A) are presented in the form of a Tsou plot for $i = 1$ (△), $i = 2$ (○), and $i = 3$ (X) (free enzyme), and for $i = 1$ (●) and $i = 2$ (▲) (*p*-fluorobenzoate-complexed enzyme). (Reprinted with permission from: R. A. Wijnands and F. Muller, *Biochemistry, 21*, 6639–6646. Copyright 1982 American Chemical Society, Washington, D. C.)

Modification of Arginine-148 in the β Subunit of Tryptophan Synthase

Reference: K. Tanizawa and E. W. Miles, *Biochemistry*, *22*, 3594–3603 (1983).

Reaction of apo-β-2 from tryptophan synthase with phenylglyoxal (Fig. 7-46) resulted in loss of the serine deaminase activity of the isolated subunit. Reciprocal pseudo-first-order rate constants calculated from the data in part A were plotted versus 1/[reagent] in part B, and the resulting plot indicated that the phenylglyoxal

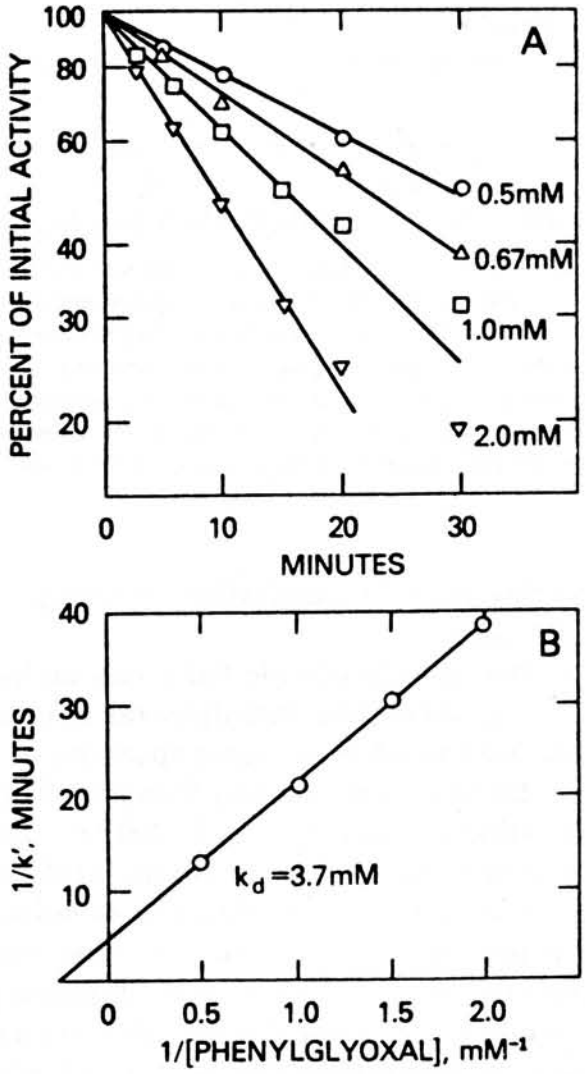


Figure 7-46 (A) Effect of phenylglyoxal concentration on the rate of inactivation of the apo-β-2 subunit of tryptophan synthase; (B) reciprocal first-order rate constants (k') calculated from part (A) are plotted versus the reciprocal concentration of phenylglyoxal. The dissociation constant (K_d) for phenylglyoxal calculated from the intercept on the abscissa ($1/K_d$) is 3.7 mM. (Reprinted with permission from: K. Tanizawa and E. W. Miles, *Biochemistry*, *22*, 3594–3606. Copyright 1983 American Chemical Society, Washington, D. C.)

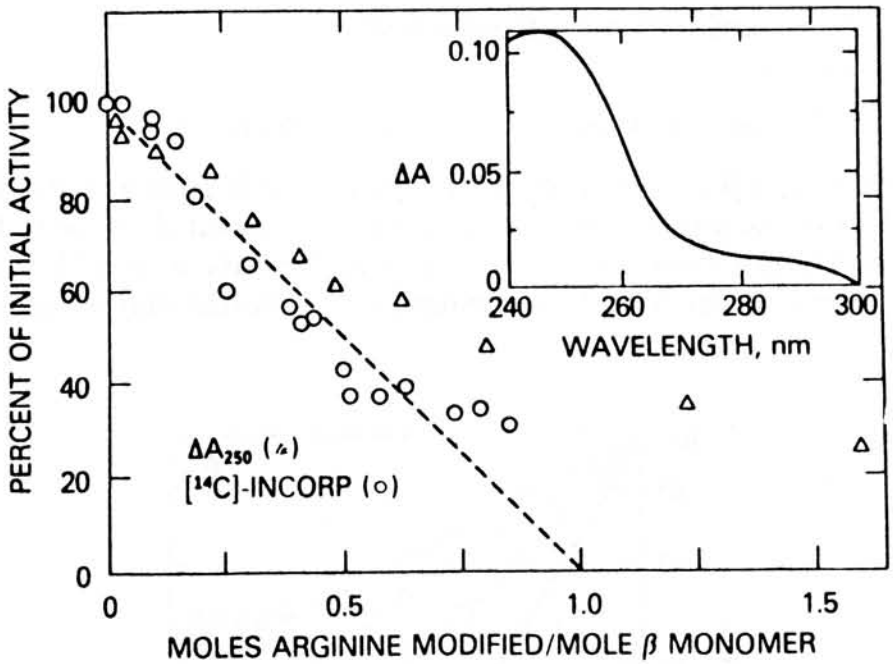


Figure 7-47 Effect of the extent of modification of the apo-β-2 subunit by phenylglyoxal upon activity. The extent of modification of arginyl residues determined from the change in absorbance at 250 nm after modification by phenylglyoxal in phosphate buffer (△) or from the incorporation of radioactivity after modification by phenyl-[2-^{14}C]glyoxal in borate buffer (○) is plotted versus the percent of initial activity. (Reprinted with permission from: K. Tanizawa and E. W. Miles, *Biochemistry*, *22*, 3594–3606. Copyright 1983 American Chemical Society, Washington, D. C.)

formed a reversible complex prior to inactivation occurring, with a calculated dissociation constant of 3.7 mM.

When percent of residual activity was plotted versus moles of arginine modified per subunit, determined using radioactive phenylglyoxal, a single site per subunit was modified (Fig. 7-47). The inset shows a difference spectrum of the modified enzyme. From the absorbance at 250 nm, approximately 0.96 modified arginine per subunit was calculated using an extinction coefficient of 11,000 M^{-1} cm^{-1} for the diphenylglyoxal adduct of arginine, in good agreement with the incorporation determined by radioactivity. Cyanogen bromide fragmentation and isolation of the radioactively labeled fragment gave a peptide containing two arginine residues on amino acid analysis. Pepsin digestion of this cyanogen bromide fragment gave a single labeled fragment, which from amino acid analysis was identified as containing arginine-148.

As can be seen from these examples, chemical modification experiments have been used in numerous ways to answer many important questions in protein chemistry. The reagents we have discussed are meant only to illustrate the types of considerations that can be addressed in the design and execution of a modification experiment. There are many more reagents available.

In the introduction to this chapter we mentioned potential uses of chemical modification experiments. First was the identification of residues involved in the

catalytic mechanism of an enzyme or in the binding site of a protein. Due to the possibility of a conformational change in the protein as a result of a chemical modification event, it is not possible to state unequivocally that because modification of a single amino acid residue in a protein blocks activity or binding that the residue is present in the catalytic or binding site. As discussed, even protection experiments do not allow this conclusion to be reached. The many other uses of chemical modification are not affected by this point, but where the purpose is to identify residues in or near particular binding sites of a protein, alternative approaches must be taken.