the well-known formol titration of amino groups (French and Edsäll, 1945; Kallen and Jencks, 1966).

\[ \text{H}_{\text{P}}\text{NH}_{2} + \text{C}=\text{O} \rightleftharpoons \text{H}_{\text{P}}\text{NHCH}_{2}\text{OH} \]

(1)

Formaldehyde is available from commercial sources in 37–40% aqueous solutions, containing small amounts of methanol and known as formalin. In such solutions, it exists primarily as a number of low-molecular-weight polymers of the type \( \text{H(OCH}_2\text{)}_n\text{OH} \). Formaldehyde may also be obtained as a stable solid known as paraformaldehyde, composed of high-molecular-weight polymers of the same type. Heating of paraformaldehyde can be used to generate pure gaseous formaldehyde.

Both sources of polymeric formaldehyde revert to monomer in dilute aqueous solutions. Under such conditions, formaldehyde is in very rapid equilibrium with a hydrated form (Equation 2), and more than 99.9% is in the form of the hydrate.

\[ \text{H}_{\text{C}=\text{O}} + \text{H}_2\text{O} \rightleftharpoons \text{HO-C-OH} \]

(2)

Compounds I through V (see below) have been shown to result from the reaction of formaldehyde with tyrosine, tryptophan, histidine, asparagine, and cysteine, respectively (French and Edsall, 1945). Presumably, their formation is preceded by the formation of very electrophilic immonium cations (Equation 3),

\[ \text{H}_{\text{N}}\text{CH}_{2}\text{OH} + \text{H}_2\text{O} \rightleftharpoons \text{H}_{\text{N}}\text{CH}_{2}\text{OH} \]

(3)

which react with the adjacent amino acid side chains. Similar reactions have been postulated to occur in proteins linking the \( \varepsilon \)-amino groups of lysine residues via methylene bridges to neighboring side chains (Fraenkel-Conrat et al., 1947; Fraenkel-Conrat and Olcott, 1948a, 1948b; Fraenkel-Conrat and Mecham, 1949). The postulated reactions are mechanistically similar to the widely studied Mannich reaction (Fernandez and Fowler, 1964; Alexander and Underhill, 1949). One
Both sources of polymeric formaldehyde revert to monomer in dilute aqueous solutions. Under such conditions, formaldehyde is in very rapid equilibrium with a hydrated form (Equation 2), and more than 99.9% is in the form of the hydrate.

\[
\begin{align*}
\text{H} & \quad \text{C} = \text{O} \quad + \quad \text{H}_2\text{O} \quad \rightleftharpoons \quad \text{HO} & \quad \text{C} & \quad \text{OH} \\
\text{H} & & & & \\
\end{align*}
\]  

(2)
\[
\text{H} \quad \text{C} = \text{O} \quad + \quad \text{H}_2\text{O} \quad \rightleftharpoons \quad \text{HO} - \text{C} - \text{OH}
\]
chemical modification of proteins

Gary E. Means
University of California, Davis

Robert E. Feeney
University of California, Davis

Holden-Day, Inc.
inactivation of certain proteases by diisopropylfluorophosphate involved reaction with specific serine residues presumably in their active sites.

Scientific developments during the following years led to increased interest in several physical-chemical techniques for studying proteins which, in turn, stimulated interest in protein modification. The use of X-ray diffraction for the study of proteins, for example, promoted interest in methods for isomorphic introduction of heavy atoms into proteins (see Section 3-2). For other studies, techniques for enhancing the fluorescence of proteins by introducing various fluorescent groups were developed. Improvements in many physical-chemical techniques have both contributed to, and made use of, protein-modification techniques.

The development of more accurate and more sensitive analytical techniques has similarly been a major contributor to, and beneficiary of, advances in protein modification. Development of the automated amino acid analyzer has been one of the most important improvements, having contributed immeasurably to all of protein chemistry.

Recently, great interest has been directed to the search for "affinity-labeling" or "active-site-directed" reagents. These reagents are designed to react preferentially with only those parts of protein molecules that are physically in the vicinity of particular biochemically active sites. Their use gives a highly selective way of chemically modifying a protein (see Section 2-2).

1-2 CHEMICAL MODIFICATIONS FOR ANALYTICAL AND INDUSTRIAL PURPOSES

For many years a motivating factor in the study of the chemical modification of proteins was the desire to determine quantitatively the amounts of proteins or their component amino acids. Many methods have been developed for such purposes. Because the intent was not to preserve the integrity of the protein, many of these methods are very harsh and in this way differ from most procedures described in this book. A few methods originally designed strictly for quantitative purposes have, however, also proved useful for selective chemical modification of proteins. The use of nitrous acid for the determination of amino groups in proteins, for example, is also of value for their selective modification.

Only a few methods are suitable for both modifying and determining amino acid side chains in proteins. Two such methods now in common usage are the reaction of sulfhydryl groups with \( p \)-mercuribenzoate and the reaction of amino groups with trinitrobenzenesulfonic acid. Both reagents can be used under relatively mild conditions which do not damage most proteins, and it is easy to measure the numbers of groups modified.

Commercial applications of the chemical modification of proteins have a long history related to the pharmaceutical, dyeing, and clothing industries. An early
application in the pharmaceutical industry was the use of formaldehyde to modify bacterial toxins and viruses; similar procedures are still important commercially. The purpose of this treatment is to kill, inactivate, or so change the virus or toxin as to render it incapable of eliciting its toxic or pathological response, while retaining its ability to elicit an immunogenic response when injected into an animal. The bacterial toxins, when so modified, are known as toxoids.

One of the oldest processes involving protein modification is the treatment of animal hides or hair for human use, as in the tanning of leather. Increased knowledge has led to recent improvements in these ancient procedures. For example, glutaraldehyde, used for cross-linking of proteins, is now also used for tanning leather. It apparently functions similarly by cross-linking collagen in the leather. Similarly, several different modifications are now used to give wool fibers superior performance for clothing. Chlorination or treatment with polyepoxides is being used commercially. The latter primarily react with amino groups. Use has also been made of reagents splitting disulfide bonds for the purpose of obtaining “permanent press” in finished clothing items.

1-3 CURRENT STATUS

The current interest in chemical modifications of proteins is indicated by the many reviews and books on its various aspects. Baker (1967), for example, has published a book dedicated to the organic chemistry of the active site (of enzymes), and Hirs (1967) has edited a detailed compilation of laboratory procedures for proteins. Other reviews, including general discussions of chemical reagents and their reactions, have been published by Cohen (1968), Glazer (1970), Vallee and Riordan (1969), Stark (1970), Shaw (1970), and Spande et al. (1970).

There is now a long series of chemical reagents and reactions used for: (a) synthesizing peptides and proteins (Stewart and Young, 1969), (b) sequential and stepwise degradation of proteins to determine their structures (Stark, 1970), and (c) preparation of derivatives of amino acids to increase their volatilities and detectabilities by vapor phase chromatography (Gehrke et al., 1968). The first type under (a) includes methods which must maintain the integrity of the polypeptide chain and the structures of the side groups of the amino acids, because the objective is to end with a normal polypeptide or protein. Special reagents are used to block or stabilize the partially synthesized protein. In general, however, the methods are milder than (b) and (c) types. This is not the case for the methods used for preparing derivatives for vapor phase chromatography. For this purpose, the only requirement is maintenance of the structure of the individual amino acids. Consequently, the methods used for these derivatives usually involve much harsher conditions than are used for intact proteins.

Many of the methods in use today are sufficiently simple and convenient to be
state including ordered but conformationally distinct forms produced by chemical modification.

Because a chemically modified protein differs chemically, then so must the relative stability of its various conformations also differ from those of the native protein. Differences in relative stability under different conditions is an additional complicating factor.

For most purposes, denaturation can usually be recognized by changes in rather obvious physical properties. Decreased solubility is a classic indicator of denaturation. The loss of biological activity is also frequently taken as a sign of denaturation. Such indications are relatively crude signs of conformational changes, and relatively more refined techniques are preferable following chemical modification. No single criterion is adequate to detect all structural changes, and the possibility exists that even a slight change may have important biological consequences. To distinguish between denaturation and those changes due to modification of functionally important residues is of utmost importance. A number of rather sensitive techniques are now available for detecting altered conformations. Some of the methods useful for detecting denaturation are mentioned in Chapter 4.

1-6 RELATIVE REACTIVITIES OF AMINO ACID SIDE CHAINS

The reactivity of one substance with another depends upon the respective properties of each and upon the environment in which they are placed. The reactivity of a group in a protein with a particular reagent depends, among other things, upon the effect of the environment on the group, and upon the ability of the reagent to enter that environment. Because proteins are quite large, many of their constituent amino acid residues are partially shielded from solvent, and hence are relatively unreactive with reagents dissolved therein. Conversely, some residues, perhaps as a result of their presence in a catalytic center, possess unusually high reactivities. Determination of relative reactivities toward a number of reagents can thus help to define the position of various residues in the three-dimensional structure of a protein. Such information is extremely valuable to the X-ray crystallographer for the interpretation of his results, and, until crystallographic information is available, affords the clearest source of information on three-dimensional structure. On the basis of such information, for example, a model of pancreatic ribonuclease A was proposed which proved remarkably similar to that later proposed on the basis of crystallographic data (Hammes and Scheraga, 1966). The determination of relative reactivities is currently one of the most active applications of protein modification.

Side-Chain Specificity. The different chemical properties of the various amino acid side chains provide a basis for their differential modification. Most protein reagents react with more than one side chain (Table 1-1). Lack of specificity limits
<table>
<thead>
<tr>
<th>Reagent</th>
<th>Tyr</th>
<th>His</th>
<th>Arg</th>
<th>Trp</th>
<th>Cystine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetic anhydride (5-1, A-1)</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N-Acetyl imidazole (5-1, A-6)</td>
<td>±±</td>
<td>+++</td>
<td>+++</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Acrylonitrile (6-3)</td>
<td>±±</td>
<td>+++</td>
<td>+++</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Aldehyde/NaBH₄ (6-8, A-1)</td>
<td>+++</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N-Bromosuccinimide (8-9, A-8)</td>
<td>-</td>
<td>+++</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>N-Carboxyanhydrides (5-1)</td>
<td>+++</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cyanate (5-2, A-1)</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Cyanogen bromide (10-3)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1,2-Cyclohexanedione (10-1, A-5)</td>
<td>±</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Diacetyl trimer (10-1, A-5)</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Diazacacetates (7-1)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Diazonium salts (9-3, A-7)</td>
<td>+++</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Diketene (5-1)</td>
<td>+++</td>
<td>d</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dinitrofluorobenzene (6-5)</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5,5'-Dithiobis(2-nitrobenzoic acid) (8-3, A-2)</td>
<td>--</td>
<td>+</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Ethoxyformic anhydride (5-1, A-7)</td>
<td>+++</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ethylenimine (6-4)</td>
<td>±±</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>N-Ethylmaleimide (6-2, A-2)</td>
<td>±±</td>
<td>+++</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Ethyl thiocrofluoroacetate (5-1)</td>
<td>+++</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Formaldehyde (6-7, A-1)</td>
<td>+++</td>
<td>+++</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Glyoxal (10-2)</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Haloacetates (6-1, A-2)</td>
<td>+</td>
<td>+++</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hydrogen peroxide (8-7, A-9)</td>
<td>--</td>
<td>+++</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2-Hydroxy-5-nitrobenzyl bromide (6-6, A-8)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Iodine (9-1, A-6)</td>
<td>--</td>
<td>+++</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>α-Iodosobenzoate (8-4)</td>
<td>--</td>
<td>+++</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Compound</td>
<td>$-$</td>
<td>$+$</td>
<td>$++$</td>
<td>$+++\footnote{a}$</td>
<td>$+\pm\footnote{a}$</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Maleic anhydride (5-1, A-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$-Mercuribenzoate (10-2, A-2)</td>
<td>$-$</td>
<td>$+$</td>
<td>$++$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methanol/HCl (7-1, A-4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-Methoxy-5-nitrotropane (6-7)</td>
<td>$++\footnote{d}$</td>
<td>$-$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methyl acetimidate (5-3, A-1)</td>
<td>$++$</td>
<td>$-$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O-Methylisourea (5-3, A-1)</td>
<td>$++$</td>
<td>$-$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitrous acid (10-5)</td>
<td>$++\footnote{d}$</td>
<td>$++$</td>
<td>$+$</td>
<td></td>
<td>$+\pm\footnote{d}$</td>
</tr>
<tr>
<td>Performic acid (8-6)</td>
<td>$-$</td>
<td>$++$</td>
<td>$-$</td>
<td>$-$</td>
<td>$+$</td>
</tr>
<tr>
<td>Phenylglyoxal (10-1, A-5)</td>
<td>$+$</td>
<td>$-$</td>
<td></td>
<td>$+$</td>
<td>$++\footnote{d}$</td>
</tr>
<tr>
<td>Photooxidation (8-8, A-7)</td>
<td>$-$</td>
<td>$++$</td>
<td>$+\pm\footnote{d}$</td>
<td>$++$</td>
<td></td>
</tr>
<tr>
<td>Sodium borohydride (8-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Succinic anhydride (5-1, A-1)</td>
<td>$++\footnote{d}$</td>
<td>$++\footnote{c}$</td>
<td>$+$</td>
<td>$+$</td>
<td></td>
</tr>
<tr>
<td>Sulfenyl halides (10-4)</td>
<td>$-$</td>
<td>$+$</td>
<td>$+$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulfite (8-2, A-3)</td>
<td>$-$</td>
<td>$+\pm\footnote{d}$</td>
<td>$-$</td>
<td></td>
<td>$+$</td>
</tr>
<tr>
<td>Sulfonyl halides (5-4)</td>
<td>$++\footnote{d}$</td>
<td>$++$</td>
<td>$++$</td>
<td>$++$</td>
<td>$+$</td>
</tr>
<tr>
<td>Tetranitromethane (9-2, A-6)</td>
<td>$-$</td>
<td>$++$</td>
<td>$++$</td>
<td>$-$</td>
<td></td>
</tr>
<tr>
<td>Tetrathionate (8-5)</td>
<td>$-$</td>
<td>$+-\footnote{c}$</td>
<td>$-$</td>
<td></td>
<td>$+$</td>
</tr>
<tr>
<td>Thiols (8-1, A-3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$-$</td>
</tr>
<tr>
<td>Trinitrobenezesulfonic acid (6-5, A-1)</td>
<td>$++\footnote{d}$</td>
<td>$++\footnote{d}$</td>
<td>$-$</td>
<td>$-$</td>
<td>$+$</td>
</tr>
<tr>
<td>Water-soluble carbodiimide and nucleophile (7-2, A-4)</td>
<td>$+\pm\footnote{c}$</td>
<td>$+\pm\footnote{c}$</td>
<td>$-$</td>
<td>$-$</td>
<td>$+$</td>
</tr>
</tbody>
</table>

\footnote{a}{\textit{---}, $+\text{---}$, $++\text{---}$, and $+++\text{---}$ indicate relative reactivities. $\pm\text{---}$, $\pm\pm\text{---}$, and $\pm\pm\pm\text{---}$ likewise indicate relative reactivities which may or may not be attained depending upon the conditions employed.}

\footnote{b}{Numbers in parentheses are sections where reagent is discussed in most detail.}

\footnote{c}{Spontaneously reversible under the reaction conditions or upon dilution, regenerating original group.}

\footnote{d}{Easily reversible, regenerating original group.}
Reaction with HNBB can be used to quantitatively determine tryptophan in proteins (Barman and Koshland, 1967). The procedure is simple, involving measurement of the absorbance at 410 m\textmu\text{m} after treatment with HNBB and removal of hydrolyzed reagent. An extinction coefficient of 18,000 \textit{M}^{-1} \text{cm}^{-1} has been employed at pH values above 10. The possible formation of more than one product necessitates careful control of the reaction conditions. Cysteine residues have not been found to interfere in these determinations.

The accessibility of tryptophan residues in \textalpha-chymotrypsin to reaction with HNBB has been examined as a function of pH (Oza and Martin, 1967). At pH 2, all eight tryptophans are reactive, whereas at pH 4 only one is available. This one does not appear to be essential for catalytic activity. In porcine pepsin, only two of the four tryptophans are reactive at pH 3.5, and modification of these has only a slight effect on catalytic activity (Dopide and Jones, 1968). Disruption by reduction and alkylation (Chapter 3 and Section 8-1) or alkali treatment, respectively, render the third and fourth tryptophans reactive. Modification of one tryptophan with HNBB has little effect on the properties of apomyoglobin whereas modification of the second residue is accompanied by drastic changes in its physical properties (Atassi and Caruso, 1968).

HNBB is useful for quantitatively determining tryptophan in proteins, for determining their relative accessibilities under various conditions, and for studying the effects of such modification of tryptophan on the properties of proteins. It may also be used as an environmental probe, its spectrum being sensitive to changes in local environment, especially in pH (Koshland et al., 1964; Horton and Koshland, 1965).

The related compound 2-methoxy-5-nitrobenzyl bromide is considerably less reactive, similar to unsubstituted benzyl bromide. It reacts with tryptophan, cysteine, and methionine in neutral or acidic solutions, although at a much slower rate than HNBB. Its spectral properties are very sensitive to changes in solvent polarity but not to pH. It has been used as a pH-insensitive environmental probe to follow and assess conformational changes of proteins (see Subsection 3-2.3) (Horton et al., 1965):

6-7 FORMALDEHYDE AND SOME OTHER CARBONYL COMPOUNDS

6-7.1 FORMALDEHYDE

In neutral or alkaline solutions, the amino groups of proteins can react in a readily reversible manner with 2 moles of formaldehyde (Equation 1) giving, with the addition of each, an apparent lowering in amino-group pK of 2 to 3 pH units. The resulting displacement of the acid-base equilibrium by formaldehyde is the basis of
the well-known formol titration of amino groups (French and Edsall, 1945; Kallen and Jencks, 1966).

\[
P\text{NH}_2 + C=O \rightleftharpoons P\text{NHCH}_2\text{OH} \rightleftharpoons P\text{NCH}_2\text{OH}
\]

Formaldehyde is available from commercial sources in 37–40% aqueous solutions, containing small amounts of methanol and known as formalin. In such solutions, it exists primarily as a number of low-molecular-weight polymers of the type \(H(O\text{CH}_2)_n\text{OH}\). Formaldehyde may also be obtained as a stable solid known as paraformaldehyde, composed of high-molecular-weight polymers of the same type. Heating of paraformaldehyde can be used to generate pure gaseous formaldehyde.

Both sources of polymeric formaldehyde revert to monomer in dilute aqueous solutions. Under such conditions, formaldehyde is in very rapid equilibrium with a hydrated form (Equation 2), and more than 99.9% is in the form of the hydrate.

\[
\text{H}C=O + \text{H}_2\text{O} \rightleftharpoons \text{HO-C-OH}
\]

Compounds I through V (see below) have been shown to result from the reaction of formaldehyde with tyrosine, tryptophan, histidine, asparagine, and cysteine, respectively (French and Edsall, 1945). Presumably, their formation is preceded by the formation of very electrophilic immonium cations (Equation 3),

\[
\text{H-N-CH}_2\text{OH} \xrightarrow{\text{H}_2\text{O}} \text{H-N-CH}_2 \rightleftharpoons \text{H-N}^+\text{CH}_2
\]

which react with the adjacent amino acid side chains. Similar reactions have been postulated to occur in proteins linking the \(\varepsilon\)-amino groups of lysine residues via methylene bridges to neighboring side chains (Fraenkel-Conrat et al., 1947; Fraenkel-Conrat and Olcott, 1948a, 1948b; Fraenkel-Conrat and Mechem, 1949). The postulated reactions are mechanistically similar to the widely studied Mannich reaction (Fernandez and Fowler, 1964; Alexander and Underhill, 1949). One
Both sources of polymeric formaldehyde revert to monomer in dilute aqueous solutions. Under such conditions, formaldehyde is in very rapid equilibrium with a hydrated form (Equation 2), and more than $99.9\%$ is in the form of the hydrate.

\[
\text{H} - \text{C} = \text{O} + \text{H}_2\text{O} \rightleftharpoons \text{HO} - \text{C} - \text{OH}
\] (2)
such product, the methylene-bridged lysine-tyrosine compound (VI), has been isolated from acid hydrolysates of formalin-treated tetanus and diphtheria toxins (Blass et al., 1965).

The reaction of formaldehyde with thiols (Equation 4) is analogous to, and very much more rapid than, its reaction with amines (Lewin, 1956; Barnett and Jencks, 1967). Histidine side chains undergo a similar reaction (Equation 5) (Levy, 1935; Kallen and Jencks, 1966), which is thought to be responsible for the decrease in catalytic activity of α-chymotrypsin in the presence of formaldehyde (Martin and Marini, 1967). The formation of such adducts is accompanied by an increase in absorbance near 230 mμ (Saidel and Carino, 1966; Martin and Marini, 1967). The reaction of formaldehyde with peptide bonds brings about an increase in absorbance near 220 mμ (Saidel et al., 1965) and is presumably similar to its reaction with urea (Smythe, 1953). Arginine residues are also thought to react with formaldehyde, but the nature of the reaction is not known (Fraenkel-Conrat and Olcott, 1948a).

\[
P - \text{SH} + \overset{\text{H}}{\overset{\text{H}}{\overset{\text{C} &= \text{O}}{\overset{\text{H}}{\overset{\text{P}}{\overset{\text{SCH}_2\text{OH}}{}}}}}
\]

\[
P - \overset{\text{H}}{\overset{\text{H}}{\overset{\text{C} &= \text{O}}{\overset{\text{H}}{\overset{\text{N}}{\overset{\text{H}}{\overset{\text{P}}{\overset{\text{NCH}_2\text{OH}}{}}}}}}}
\]
Although the reactions of formaldehyde have been used industrially for many years, little is really known about the many underlying reactions. Its potential reaction with so many different amino acid side chains and the difficulties surrounding the detection of such have largely precluded its use as a specific protein reagent. Under carefully controlled conditions, it can be safely used as the carbonyl component for the reductive alkylation of proteins (Means and Feeney, 1968) (see Section 6-8).

### 6.7.2 Other Aldehydes and Ketones

The reactions of other aldehydes and ketones with proteins have not been widely studied but are generally considered to be much less extensive than those of formaldehyde. The isolation of ε-N-isopropyllysine from proteins exposed to low concentrations of acetone and reducing agents, however, affords clear evidence that the reaction of acetone with primary amino groups is much more extensive than generally believed (Means and Feeney, 1968). The same can presumably be said for a large number of aldehydes and ketones. Oxytocin reacts with acetone to give what is thought to be a 2,2-dimethyl-4-imidazolidone derivative (Ferrier et al., 1965; Yamashiro et al., 1965; Yamashiro and du Vigneaud, 1968). This is a specific reaction of oxytocin and certain closely related peptides with acetone, but it illustrates the kind of unusual reactions that may occur under certain circumstances.

Salicylaldehyde reacts readily with amino groups of proteins at moderately alkaline pH giving salicylidene adducts (Equation 6) resembling those formed from amino groups and pyridoxal phosphate (Section 6-9) (Williams and Jacobs, 1966, 1968). The reaction is easily reversed by dilution or by removing excess reagent. Upon incubation with salicylaldehyde at pH 9.6, horse heart cytochrome c, for example, has been shown to lose its ability to reduce DPN in a rat heart cytochrome c reductase system. If high ratios of salicylaldehyde are used, all amino groups can be modified. Dialysis to remove the reagent gives a 20% return of DPN reductase activity and, if followed by brief exposure to either pH 3 or 11, nearly complete recovery occurs.

\[
\begin{align*}
\text{O} & \quad \text{C} \\
\text{HO} & \quad \text{H} \\
\text{P} & \quad \text{NH}_2 + \\
\leftrightarrow & \\
\text{P} & \quad \text{N} \quad \text{C} \\
\text{H} & \quad \text{O} \\
\end{align*}
\]

(6)

2,4,6-Trinitrobenzaldehyde reacts at neutral pH with bovine serum albumin to give an unstable colored complex with absorption maxima at 440 and 520 mµ. Amino groups of the protein are necessary for the reaction (Whitehouse and Skidmore, 1966).
Benzoquinone reacts with amino groups of proteins to give disubstituted quinones. Optimum reaction is obtained at pH 8.0. These substituted quinones can be reduced by ascorbate and reoxidized by oxygen to produce hydrogen peroxide. Under appropriate conditions the rate of ascorbate oxidation can be used to determine the amount of 1,4-addition product. Cytochrome c at pH 7.4 reacts with benzoquinone to give a derivative containing 2 to 4 moles of benzoquinone. The product was oxidizable by cytochrome oxidase but at a much reduced rate as compared to unmodified cytochrome c. In the presence of phenazine methosulfate, benzoquinone-treated cytochrome c was reduced but required greater amounts of NADH than untreated cytochrome c (Morrison et al., 1969).

\[
\text{O} \quad \text{NH}_2 + \text{CH}_3\text{O} \quad \text{NO}_2 \xrightarrow{\text{pH 8.5}} \quad \text{O} \quad \text{NH} \quad \text{NO}_2 + \text{CH}_3\text{OH}
\]

(7)

Amino groups of proteins react with 2-methoxy-5-nitrotropane at room temperature in slightly alkaline solutions (pH 7.0 to 8.5) to give an adduct absorbing maximally at 420 m\(\mu\) (\(\varepsilon = 2.07 \times 10^4\) M\(^{-1}\) cm\(^{-1}\)) (Tamaoki et al., 1967) (see Equation 7). No other amino acid side chains are believed to react under these mild conditions. With taka-amylose A, 9 out of 19 amino groups were readily modified, whereas the remainder were relatively resistant. Such modification brought about a progressive decrease in amylase activity but slightly increased the maltosidase activity. When the modified enzyme was incubated for several hours at 38°C in 1 to 2 M hydrazine at pH 8.5 to 9.0, the 420-m\(\mu\) absorption band was replaced by a new maximum at 354 m\(\mu\), due to release of the blocking group upon reaction with 2 moles of hydrazine. This treatment substantially restored the enzyme's amylase activity.

2-Methoxy-5-nitrotropane is a commercially available, stable crystalline compound. It decomposes slowly in neutral aqueous solutions and quite rapidly in solutions more alkaline than pH 8.5. Although it has not been widely used, it appears promising as an easily reversible amino-group reagent.

Certain 1,2- and 1,3-dicarbonyl compounds have been used to modify arginine residues of proteins (see Section 10-1). Glutaraldehyde, a 1,5-dialdehyde, has been used as a bifunctional reagent to form cross-links between various side chains in proteins. It has been thus used to insolubilize crystals of carboxypeptidase A, in which case only a slight loss of catalytic activity was observed (Quiocio and Richards, 1966). The chemical bases of the reactions remain obscure, but they appear to involve the side chains of lysine, cysteine, histidine, and tyrosine residues (Habeeb and Hiramoto, 1968).
6-8 REDUCTIVE ALKYLLATION

\[ \text{P--NH}_2 + \text{RHCO} \xrightarrow{H_2O/\text{H}_4\text{O}} \text{P--N=CHR} \xrightarrow{(\text{H})} \text{P--NH--CH}_2\text{R} \]  \hspace{1cm} (1)

Aliphatic aldehydes and ketones react very rapidly and in a highly reversible manner with amino groups of proteins (see Section 6-7). The adducts so formed can be reduced by mild reducing agents to stable alkylamino groups (Equation 1) (Means and Feeney, 1968a). Under the mild, slightly alkaline conditions required for extensive alkylation of protein amino groups, other side-chain groups do not give stable derivatives. Both \( \alpha \)- and \( \varepsilon \)-amino groups react. The procedure is applicable to a wide range of proteins.

Extensive modification is easily obtainable under mild conditions with sodium borohydride as the reductant. The required concentration of sodium borohydride is low so as not to cleave disulfide bonds (Section 8-1). The reaction is strongly pH-dependent, best results being obtained near pH 9. At pH 7, similar modification can be obtained, but with rather inefficient utilization of the reagents. Low temperatures (0°–4°C) favor more extensive reaction.

Using formaldehyde as the carbonyl component, the reaction proceeds rapidly giving \( \varepsilon \)-N,N-dimethyllysine residues as the principal products. Monomethyllysine is formed initially, but its conversion to dimethyllysine is very fast. It predominates only at very early or incomplete stages of reaction. With other aldehydes and ketones, reaction with a second carbonyl molecule is greatly retarded such that predominantly monoalkyllysine residues are obtained. When acetone is the carbonyl component, for example, only the monoalkyl derivative, \( \varepsilon \)-N-isopropyllysine, is obtained. Mixed \( \varepsilon \)-N-methylalkyllysines, such as \( \varepsilon \)-N-methyl-N-isopropyllysine, can be prepared by treating first with the higher carbonyl compound and then, in a second step, with formaldehyde.

\[ \text{P--NH}_2 + (\text{CH}_3)_2\text{CO} + \frac{1}{2} \text{NaBH}_4 \xrightarrow{0^\circ \text{C, pH 9}} \text{P--NHCH(CH}_3)_2 + \frac{1}{2} \text{NaH}_2\text{BO}_3 + \frac{1}{2} \text{H}_2\text{O} \]  \hspace{1cm} (2)

\[ \text{P--NH}_2 + 2\text{H}_2\text{CO} + \frac{1}{2} \text{NaBH}_4 \xrightarrow{0^\circ \text{C, pH 9}} \text{P--N(CH}_3)_2 + \frac{1}{2} \text{NaH}_2\text{BO}_3 + \frac{1}{2} \text{H}_2\text{O} \]  \hspace{1cm} (3)

Alkylation of amino groups in this way does not greatly affect their basicity. The \( pK \) values of dimethylamino groups (i.e., tertiary amines) are about 0.4 to 0.6 pH
units below those of primary amino groups. Monoalkylamino groups (i.e., secondary amines) usually have pK’s from 0.1 to 0.7 pH units higher than primary amino groups. These differences result in slight alterations in the titration curves and electrophoretic mobility of modified proteins at alkaline pH. Because of the relatively small change in basicity and the relatively small space occupied by the methyl groups, ε-N,N-dimethyllysine residues are more similar to unmodified lysines than most other derivatives of lysine.

When bovine pancreatic ribonuclease is reductively methylated, catalytic activity is lost with little change in the protein’s physical properties (Means and Feeney, 1968b). Its ability to bind 3’-cytidylic acid is impaired slightly and its high reactivity with iodoacetate is increased. The loss of catalytic activity appears to result from minor alterations of its active site. A similar loss of catalytic activity occurs with reductive isopropylation (i.e., using acetone as the carbonyl component) and is accompanied by slightly greater impairment of the ability to bind 3’-cytidylic acid. A somewhat greater enhancement of its reactivity with iodoacetate is also obtained. In this instance, small changes in optical rotation show it to have experienced some change in tertiary structure.

Turkey ovomucoid, in which a lysine residue is essential for trypsin-inhibitory activity, loses its antitrypsin activity upon reductive methylation or reductive isopropylation. Both derivatives retain their ability to inhibit α-chymotrypsin, showing the two inhibitory sites to be independent (Means and Feeney, 1968a).
When reductively methylated bovine chymotrypsinogen A is activated by trypsin, chymotryptic activity is obtained against benzoyltyrosine ethyl ester equivalent to that obtained by activation of the unmodified zymogen. The terminal amino group formed upon activation is resistant to subsequent alkylation (Means and Feehey, 1970).

Simple mono- and dialkyllysines are stable under conditions for acid hydrolysis of proteins and are easily quantitated by slight modifications of currently used amino-acid-analysis systems. The optical absorption factors at 570 mµ for mono- and dimethyllysine after reaction with ninhydrin are 0.881 and 0.822, respectively, as compared to a value of 1.00 for lysine (Means and Feehey, 1968a). Mono- and dialkyllysine residues do not react with trinitrobenzenesulfonic acid (TNBS) (Habeb, 1966), and their sum can be obtained colorimetrically from the decrease in TNBS-reactive amino groups (Means and Feehey, 1968a). Borohydride concentrations can be easily determined by a simple colorimetric procedure which also uses TNBS (Means and Feehey, 1968a) (Sections 4-1, 6-5). Trypsin does not attack internal e-N,N-dimethyllysine residues (Gorecki and Shalitin, 1967; Lin et al., 1969a) and carboxypeptidase B does not cleave carboxy-terminal dimethyllysine (Lin et al., 1969b).

6-9 PYRIDOXAL PHOSPHATE

\[
\begin{align*}
P - \text{NH}_2 + \text{O}^- \text{C}^- \text{H} \text{CH}_2 \text{H} \text{CH}_2 \text{OPO}_3 \text{H}^- & \xrightarrow{\text{H}_2\text{O} + \text{H}_2\text{O}} \text{P} - \text{N}=\text{C}^- \text{H} \text{CH}_2 \text{H} \text{CH}_2 \text{OPO}_3 \text{H}^- \xrightarrow{\text{NaBH}_4} \\
\end{align*}
\]

Reduction with sodium borohydride has been used to stabilize linkages between pyridoxal phosphate (PLP) and many enzymes in which it occurs. The procedure has been used to label the PLP binding sites of these enzymes (Schirich and Mason, 1963; Fischer et al., 1958; Hughes et al., 1962; Wilson and Kornberg, 1963; Anderson and Chang, 1965; Phillips and Wood, 1965; Wilson and Crawford, 1965; Matsuo and Greenberg, 1959). As employed for rabbit muscle glycogen
phosphorylase (Fischer et al., 1958), urea, detergents, or extremes of pH (pH < 4.5 or > 9.5) are required to induce the enzyme-PLP complex into a reactive state. This process, presumably, involves its transformation from an unreactive aldamine form, I or II (λ_{max} = 330 m\mu), into a Schiff-base form, III (λ_{max} = 415 m\mu), which is then reduced (Equation 2). Borohydride-reduced glycogen phosphorylases a and b have absorption maxima at 325 m\mu and retain most of their catalytic activity. The effects of various conditions upon the reduction of glycogen phosphorylase have recently been described (Strausbaugh et al., 1967). Many PLP-containing enzymes are found primarily in the form of a Schiff-base-type complex and are easily reduced by sodium borohydride at neutral pH without deforming agents (Anderson and Song, 1967; Dempsey and Snell, 1963; Klein and Sagers, 1967).

\[
\begin{align*}
\text{CH}_3\text{–} & \text{N} & \text{–} & \text{CH}_2\text{PO}_3^- \\
& & & \\
\text{OH} & & & \\
\text{CH} & & & \\
\text{N} & & & \\
& & & \\
\text{XH} & & & \\
\end{align*}
\]

\[\text{I}\]

\[
\begin{align*}
\text{CH}_3\text{–} & \text{N} & \text{–} & \text{CH}_2\text{PO}_3^- \\
& & & \\
\text{OH} & & & \\
\text{CH} & & & \\
\text{N} & & & \\
& & & \\
\text{XH} & & & \\
\end{align*}
\]

\[\text{II}\]

\[
\begin{align*}
\text{CH}_3\text{–} & \text{N} & \text{–} & \text{CH}_2\text{PO}_3^- \\
& & & \\
\text{OH} & & & \\
\text{CH} & & & \\
\text{N} & & & \\
& & & \\
\text{XH} & & & \\
\end{align*}
\]

\[\text{III}\]

\[
\begin{align*}
\text{CH}_3\text{–} & \text{N} & \text{–} & \text{CH}_2\text{PO}_3^- \\
& & & \\
\text{OH} & & & \\
\text{CH} & & & \\
\text{N} & & & \\
& & & \\
\text{XH} & & & \\
\end{align*}
\]

\[\text{IV}\]

\[
\begin{align*}
\text{CH}_3\text{–} & \text{N} & \text{–} & \text{CH}_2\text{PO}_3^- \\
& & & \\
\text{OH} & & & \\
\text{CH}_2 & & & \\
\text{N} & & & \\
& & & \\
\text{XH} & & & \\
\end{align*}
\]

\[\text{II}\]

\[
\begin{align*}
\text{CH}_3\text{–} & \text{N} & \text{–} & \text{CH}_2\text{PO}_3^- \\
& & & \\
\text{OH} & & & \\
\text{CH} & & & \\
\text{N} & & & \\
& & & \\
\text{XH} & & & \\
\end{align*}
\]

\[\text{III}\]

\[
\begin{align*}
\text{CH}_3\text{–} & \text{N} & \text{–} & \text{CH}_2\text{PO}_3^- \\
& & & \\
\text{OH} & & & \\
\text{CH} & & & \\
\text{N} & & & \\
& & & \\
\text{XH} & & & \\
\end{align*}
\]

\[\text{IV}\]

Many proteins which do not normally contain PLP nevertheless can, under appropriate conditions, form specific complexes with it, in some cases, perhaps reflecting an enzyme–allosteric effector relationship. These complexes may be either
aldamines, Schiff bases, or mixed complexes. Bovine serum albumin, for example, binds 2 moles of PLP very tightly, one as a Schiff base and one as an aldamine. At higher PLP concentrations, additional PLP is bound (Dempsey and Christensen, 1962). The high affinity binding appears dependent upon specific interactions between the anionic phosphate ester moiety of PLP and certain cationic groups of the protein. The same can be said for most protein-PLP complexes.

Reduction of pig kidney fructose-1,6-diphosphatase in the presence of PLP gives an enzyme which is still active but no longer sensitive to allosteric AMP inhibition or to high substrate inhibition (Marcus and Hubert, 1968). Similar treatments inactivate *Candida utilis* phosphogluconate dehydrogenase (Rippa et al., 1967), rabbit muscle aldolase (Shapiro et al., 1968), bovine pancreatic ribonuclease (Means and Feeney, 1970), and bovine liver glutamic dehydrogenase (Anderson et al., 1966). PLP alone reversibly inhibits the catalytic activity of each of these enzymes. PLP lowers the oxygen affinity of hemoglobin, as do a number of other phosphate esters. Reduction with sodium borohydride has been used to determine the site of its interactions with hemoglobin (Benesch et al., 1969). In each of these cases, binding appears to be facilitated by, and the site largely determined by, the phosphate moiety.

Equimolar PLP and pancreatic ribonuclease in neutral solution react slowly, giving a complex which, after a few minutes, has a composite absorption maximum at a wavelength between that of PLP (\( \sim 388 \, \text{m\textmu} \)) and its Schiff-base derivatives (415 to 425 m\textmu) under the same conditions. Addition of sodium borohydride decolorizes the yellow solution, shifting its absorption maximum to shorter wavelength and stabilizing the linkage between enzyme and PLP. After purification, the pyridoxamine ribonuclease (\( \lambda_{\text{max}} = 324 \, \text{m\textmu} \), assumed \( \varepsilon = 0.97 \times 10^4 \, M^{-1} \, \text{cm}^{-1} \)) can be oxidized with performic acid and digested with trypsin to give one major fluorescent peptide, the position of which in the linear sequence of ribonuclease has been determined. The PLP label is at lysine residue 7, a position known from other chemical and X-ray-diffraction studies to be in or immediately adjacent to the active site.

PLP labeling with sodium borohydride appears to have general utility as a means to specifically label lysine residues in or near phosphate binding sites of proteins. These labels are readily detected by their characteristic absorption and fluorescent properties (Klein and Sagers, 1967).

REFERENCES


<table>
<thead>
<tr>
<th>Chemical Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimethyl adipimidate</td>
<td>Water soluble, specific for amino groups, product is positively charged; reacts similarly to ethyl acetimidate (Section 5-5). Has been used to determine inter-residue distances in pancreatic ribonuclease (Hartman and Wold, 1967). Diethyl malonimidate has been used similarly with bovine serum albumin and γ-globulin (Dutton et al., 1966).</td>
</tr>
<tr>
<td>Phenol-2,4-disulfonyl chloride</td>
<td>Water soluble, reacts principally with amino groups, forms linkages of 6.7–7 Å; cross-linked derivative of egg white lysozyme was catalytically active (Moore and Day, 1968).</td>
</tr>
<tr>
<td>Hexamethylenediisocyanate</td>
<td>Insoluble in water, reacts principally with amino groups; has been used with pancreatic ribonuclease and α-chymotrypsin (Ozawa, 1967b); disothiocyanate has also been used (Ozawa, 1967b); The similar reagent azophenyl-β-diisocyanate has been used with whale myoglobin (Fasold, 1965a, 1965b).</td>
</tr>
<tr>
<td>Glutaraldehyde</td>
<td>Complex reaction with proteins apparently involving many different side chains. Confers great structural rigidity. Reaction with crystalline carboxypeptidase A increased the mechanical stability without eliminating catalytic activity (Quirocho and Richards, 1966). Treatment of trypsin gives an insoluble product retaining much of its catalytic activity (Habeeb, 1967).</td>
</tr>
<tr>
<td>Formalin (formaldehyde)</td>
<td>Highly reactive water-soluble reagent, undergoes a variety of unknown cross-linking reactions (Fraenkel-Conrat and Oclott, 1948; Blass et al., 1965).</td>
</tr>
<tr>
<td>Woodward's reagent K</td>
<td>Water-soluble reagent; links existing carboxyl and amino groups. Has been used to prepare oligomeric α-chymotrypsin (Patel and Price, 1967).</td>
</tr>
<tr>
<td>Bisdiazobenzidine</td>
<td>Highly reactive water-soluble reagent, reacts primarily with tyrosine and histidine residues. Used to prepare water-insoluble catalytically active derivatives of papain (Süllman et al., 1966).</td>
</tr>
</tbody>
</table>