BIOORGANIC CHEMISTRY 4, 22–29 (1975)

Reaction of Formaldehyde with the Histidine Residues of Proteins

CHARLES J. MARTIN AND DAVID P. LAM

Department of Biochemistry, The Chicago Medical School, University of Health Sciences, Chicago, Illinois 60612

AND

MARIO A. MARINI

Department of Biochemistry, Northwestern University Medical School, Chicago, Illinois 60611

Received February 4, 1974

Imidazoles or imidazoles substituted in the 2 or 4(5) position but with the ring nitrogen free, give a positive Pauly reaction. *N*-Methylimidazole does not react. In the presence of formaldehyde, all imidazole compounds give a negative Pauly reaction. In agreement with nmr studies, it is concluded that formaldehyde reacts with the imidazole ring nitrogen in acid solution to form *N*-hydroxymethyl derivatives.

The Pauly color yield of various proteins (chymotrypsin, TPCK-chymotrypsin, lysozyme, ribonuclease, and reduced ribonuclease) is reduced 90–95% when the reaction is performed in the presence of formaldehyde. The color yield in water is essentially accounted for by the known reactive histidine and tyrosine content. In the presence of formaldehyde the color yield can be interpreted as arising from the known tyrosine content. It is therefore concluded that the histidine residues of the proteins examined have reacted with formaldehyde to form *N*-hydroxymethyl derivatives.

In contrast to the Pauly color yield of chymotrypsin (A_M, 52 720) which can be accounted for by the contribution of its two histidine and three reactive tyrosine residues, the color yield of TPCK-chymotrypsin (A_M, 47 685) is higher than would be expected on the basis of the reported site of reaction of TPCK with chymotrypsin. The experimental molar extinction coefficient should be close to that calculated (A_M, 31 300) for its presumed one reactive histidine and three tyrosine residues. That it is not is in agreement with a previous report from our laboratories suggesting that His-57 is not the only site of reaction of TPCK with chymotrypsin.

INTRODUCTION

The hydrolysis of acetyl-L-tyrosine ethyl ester by α-chymotrypsin in the presence of formaldehyde results in an approximate eightfold decrease in the specific rate constant and a 2.5-fold increase in K_m' at pH 8.0 but with no change in the active site concentration (*1*, *2*). From a pH-dependence study, the pK′ of the group on the free enzyme which can be equated with that involved in the rate-determining step for the decomposition of the ES-state is shifted down scale by about 0.3 unit in the presence of form-

aldehyde (*1*). Similar pK′ shifts and a reduction in catalytic efficiency have also been observed for the hydrolysis of *p*-nitrophenyl acetate by imidazole and its derivatives in formaldehyde solution (*3*).

From the effect of formaldehyde on the absorption spectra of imidazole compounds at 210–230 nm and absorption characteristics of chymotrypsin in the region of 230 nm, it has been shown that the data are compatible with the reaction of formaldehyde with the imidazole ring of both histidine residues of the enzyme (*4*). Apparently, however, only the reaction of formaldehyde with a single histidine residue is of kinetic importance (*2*). These and other results have led to the suggestion that formaldehyde is an active site label for chymotrypsin which affects catalytic efficiency through formation of a modified enzyme (*2*).

The implications of this proposal made it important to acquire additional supporting evidence for the occurrence of a reaction between formaldehyde and the histidine residues of chymotrypsin. The older literature is equivocal as to whether formaldehyde reacts with the histidine residues of proteins (*5*). In more recent publications it has been reported that formaldehyde reacts with the imidazole ring of model compounds, presumably to form an *N*-hydroxymethyl derivative (*4*, *6*, *7*).

More definitive evidence to support this conclusion has been obtained by an nmr study of the reaction (*8*). Imidazole reacts with formaldehyde to form an *N*-hydroxymethyl derivative in alkaline solution; in strong acid, the product is *N*,*N*′-dihydroxymethylimidazole. Data obtained from the binding of ^{14}C-formaldehyde to either α-chymotrypin, δ-chymotrypsin, or acetylated-δ-chymotrypsin indicate that a similar stoichiometry must prevail in the reaction with the imidazoyl groups of the histidine residues of these proteins (*9*).

To provide additional chemical evidence that would support the occurrence of a reaction between formaldehyde and the imidazoyl group of proteins, advantage has been taken of the specificity of the Pauly reaction. Imidazoles or imidazoles substituted in the 2 or 4(5) position, but with the ring nitrogens free, react with diazotized aromatic amines to give colored azo dyes (*10*). Imidazole compounds substituted on the ring nitrogen(s) give a *negative* Pauly reaction, and this property has been used, for example, to determine the peptide fragment containing the histidine residue (His-57) bearing the end product of the reaction between chymotrypsin and tosylamido-2-phenylethyl chloromethyl ketone (*11*). It therefore follows that, if formaldehyde reacts with the ring nitrogen(s) of the histidine residues of proteins, the Pauly color yield should be markedly reduced.

Tyrosine also gives a positive Pauly reaction although the color yield is less than one-fourth that generated by the imidazole ring chromophore (cf. Results and Discussion section). In the application of this test to proteins in aqueous solution, it has been reported (*12*) that the total color yield should be equal to the color contributions of the histidine and tyrosine content.

On the basis of the above considerations, we have therefore determined the Pauly color yield of model compounds in the absence and presence of formaldehyde. The experimentally determined molar extinction coefficients have then been compared with the Pauly color yield obtained from various proteins with and without formaldehyde present. Despite the lack of complete analytical correspondence between that calculated for each protein and that derived from model compound data, the marked

reduction in Pauly color that occurs upon reaction of the proteins with formaldehyde fits the interpretation that the aldehyde reacts with the imidazole ring nitrogens of the histidine residues. The residual Pauly color detected in the presence of formaldehyde suggests that it is derived, in whole or in part, from the available tyrosine residues.

EXPERIMENTAL

The sources of the enzyme preparations were as follows: α-chymotrypsin (lot No. 148) was a 3×-crystallized preparation from Armour Pharmaceutical Co. as used before (*1*); lysozyme (lot No. LY633), and ribonuclease (lot No. RAD6084) were crystalline preparations obtained from the Worthington Biochemical Corp. Formaldehyde (37% w/v, and containing 1% methanol) was a gift from the Celanese Corporation of America. Imidazole and its derivatives were obtained from the Aldrich Chemical Co. *N*-Methylimidazole was distilled twice before use (bp, 76°C at 6 mm Hg). *N*-Acetyl-L-tyrosine ethyl ester was obtained from Mann Research Laboratories, and Cleland's reagent (dithiothreitol) was a Calbiochem product.

The L-1-tosylamido-2-phenylethyl chloromethyl ketone (TPCK) derivative of chymotrypsin (*13*) was prepared as follows. Chymotrypsin (200 mg) was dissolved in 7.0 ml 0.05 *M* phosphate buffer, pH 7.0, and 3.0 ml of TPCK (32 mg in methanol), obtained from the Cyclo Chemical Co., were added. After incubation at 20°C for 20 hr, the solution was exhaustively dialyzed against 1 m*M* HCl and lyophilized. The TPCK-CT preparation contained less than 0.5% active enzyme.

Reduced ribonuclease was prepared according to the procedure of White (*14*) by incubation of the enzyme in 8 *M* urea at pH 8 for 1 hr but in the presence of 0.01 *M* dithiothreitol rather than mercaptoethanol. The solution was then dialyzed for 16 hr against cold deionized water.

For the determination of protein concentration by absorbance measurements at 280 nm, the following optical factors and molecular weights were used: chymotrypsin (0.500, 26 100) and TPCK-CT (0.500, 26 200); lysozyme (0.382, 14 800); and ribonuclease (1.4, 12 600).

The Pauly reaction was carried out according to the modification of Weiss-Szobolew (*15*). Absorbance measurements were made at 530 nm with a Beckman Model DU spectrophotometer and 10-mm cuvettes.

RESULTS AND DISCUSSION

The Pauly color yields (absorbance) of various model compounds are linearly related to their concentrations with imidazole or acetylhistidine having a molar extinction coefficient (A_M) of 18 700 M^{-1} cm^{-1} (Fig. 1). The A_M for 2-methylimidazole was 14 900 M^{-1} cm^{-1} and, as expected, *N*-methylimidazole gave essentially no reaction (A_M, 70 M^{-1} cm^{-1}). Acetyl-L-tyrosine ethyl ester produced only about 23% of the color yield of imidazole.

In the presence of 3.1 *M* formaldehyde, *all imidazole compounds examined were unreactive with the Pauly reagent.* That this result was not due to some unspecific interference with the Pauly reaction was shown by first permitting the diazotization step

with imidazole to occur, followed by the addition of formaldehyde. The pH was then raised by the addition of 10% sodium carbonate to develop the color. The color yield was essentially the same as that obtained without formaldehyde. This would indicate that the absence of a positive Pauly reaction when diazotized sulfanilic acid is added

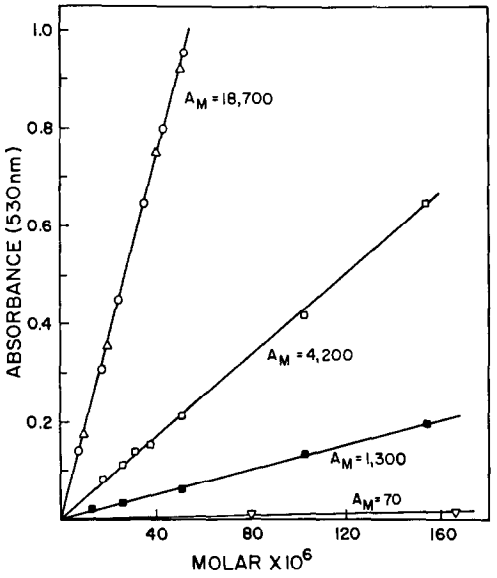

FIG. 1. The variation in Pauly color yields given by imidazole (△), acetylhistidine (○), *N*-methylimidazole (▽), and acetyl-L-tyrosine ethyl ester (□) in the absence of formaldehyde and by the tyrosine derivative in the presence of formaldehyde (■).

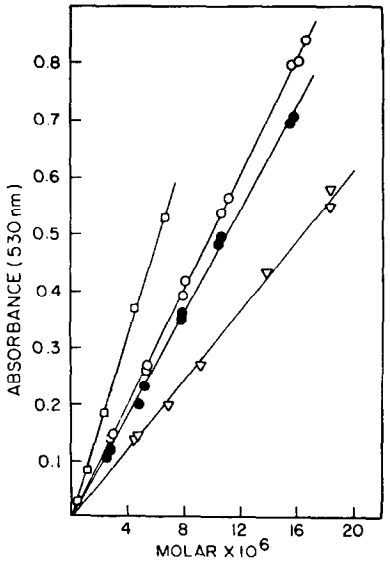

FIG. 2. The variation in Pauly color yields given by chymotrypsin (○), TPCK-CT (●), lysozyme (▽), and reduced ribonuclease (□), in the absence of formaldehyde.

to an imidazole compound in formaldehyde must reflect the prior formation of *N*-hydroxymethyl derivatives. As stated earlier (cf. Ref. *8*), these can exist under the strong acid conditions of diazotization. In contrast to the behavior of imidazole compounds, acetyl-L-tyrosine ethyl ester still gave a positive Pauly reaction in the presence of formaldehyde although the molar extinction coefficient was reduced about 70% (from 4200 to 1300; cf. Fig. 1).

The Pauly color yields obtained with proteins in aqueous solution are shown in Fig. 2. All proteins show a diminution in absorbance in the presence of formaldehyde (Fig. 3).

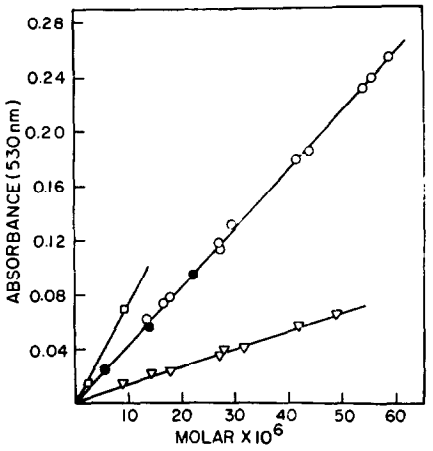

FIG. 3. The reduction in Pauly reaction absorbance for various proteins by the presence of 3.1 *M* formaldehyde. Symbols same as in legend to Fig. 2.

In Table 1 the experimentally determined Pauly A_M values given by the proteins with or without formaldehyde present are tabulated. Also included in the table are calculated Pauly yields for the various proteins in aqueous solution. These values are

TABLE 1

PAULY COLOR YIELDS OF VARIOUS PROTEINS IN WATER AND IN FORMALDEHYDE

	A_M		A_M (calcd)[a]
	In water	In formaldehyde	In water
Chymotrypsin	52 720	4530	50 000 (2 His + 3 Tyr)
Lysozyme	30 700	1340	31 300 (1 His + 3 Tyr)
RNase	65 000	4600	87 400 (4 His + 3 Tyr)
			68 700 (3 His + 3 Tyr)
Reduced RNase	82 000	7300	87 400 (4 His + 3 Tyr)
			83 200 (4 His + 2 Tyr)
TPCK-chymotrypsin	47 685	4550	31 300 (1 His + 3 Tyr)
			46 260 (1.8 His + 3 Tyr)

[a] On the basis of the indicated number of residues and using A_M of 18 700 for His and 4200 for Tyr.

based on the assumptions (cf. Refs. *12* and *15*) that the color yield in water is due to the reactive tyrosine plus histidine content of each protein and that the molar extinction coefficient of model compounds is applicable to similar groups in a protein.

The content of reactive tyrosine residues has primarily been inferred from the titration behavior of the phenolic hydroxyl groups. Some tyrosines titrate abnormally; i.e., they exhibit a time-dependent ionization that requires a pH of 12–13 for completion. Chymotrypsin has 2 normally ionizing tyrosines out of 4; ribonuclease, 3 out of 6; and lysozyme, 2 or 3 out of 3 (*16*). These values are in reasonable agreement with tyrosine reactivity studies using a variety of reagents (*17*). However, chymotrypsin has been reported to have 3 tyrosine residues that react with cyanuric fluoride (*18*). From a compilation of the available literature (*17*), all of the histidine residues of these proteins appear to be "exposed" on the basis of chemical reactivity studies. It is reasonable then to calculate the expected color yield of the Pauly reaction on the basis of the reported reactive histidine and tyrosine content of the proteins. This is not, however, a totally defensible position since execution of the Pauly reaction requires exposure of each protein to strong acid conditions during the diazotization step. This treatment may normalize "buried" tyrosine residues with a resulting increase in color yield.

Comparison of the data in the first two columns of Table 1 shows a marked decrease in the A_M value of a protein when the Pauly reaction is performed with formaldehyde present. For example, with chymotrypsin there is a 92% decrease of the color yield. Essentially similar reductions were obtained with the other proteins. Thus, for the proteins examined, the A_M values obtained by applying the Pauly procedure to a solution of the protein in formaldehyde are only 4–10% of the values obtained in water. Since model imidazole compounds in formaldehyde solution give no Pauly color and since the color yield of the proteins with formaldehyde present are considerably less than could be accounted for by even one reactive imidazoyl group (A_M, 18 700), the conclusion that formaldehyde has reacted with all the histidine residues on the N position of the imidazole ring seems inescapable.

With the exception of ribonuclease and TPCK-chymotrypsin (see below) the calculated (column 3, Table 1) and experimental (column 1, Table 1) values approximate those expected from the known histidine and reactive tyrosine content. The A_M values determined in the presence of formaldehyde are less amenable for comparison with attempted calculated values due to uncertainties of the state of the tyrosine residues. However, one can calculate that the color yield could arise from 3.5 tyrosine equivalents for chymotrypsin, 1 tyrosine equivalent for lysozyme, 3.5 tyrosine equivalents for ribonuclease, 5.6 tyrosine equivalents for reduced ribonuclease, and 3.5 tyrosine equivalents for TPCK-chymotrypsin using A_M, 1300 for the phenolic group in formaldehyde. It is to be noted that in no case does the color yield in the presence of formaldehyde exceed the limit of the analytical tyrosine content. It may therefore be concluded that the color yield of a protein in formaldehyde is related to its tyrosine but not its histidine content.

In the case of ribonuclease in aqueous solution, the calculated A_M value based on 4 His + 3 Tyr is higher than found by experiment. This may indicate that some of the postulated reactive chromophores are inaccessible to the reagent (the data are in reasonable agreement with 3 His + 3 Tyr) and is supported by comparison with the

result for reduced ribonuclease. Treatment of the enzyme with dithiothreitol in a denaturing environment increased A_M from 65 000 to 82 000, in fair agreement with a calculated value based on 4 His + 2 Tyr. The color yield in formaldehyde was also increased.

The data for TPCK-chymotrypsin are at variance with what one would expect on the basis of prior literature knowledge. The site of reaction of TPCK with the enzyme is reported to be the $N^{\varepsilon 2}$ position of His-57 (*19*). The peptide fragment bearing the carboxymethyl group as the end product of its isolation gives a negative Pauly reaction (*11*). Assuming this to be true, i.e., that only one histidine residue should be available for reaction with the Pauly reagent in TPCK-chymotrypsin, the calculated A_M value for 1 His + 3 Tyr gives 31 300 (cf. Table 1). The experimentally determined value is 47 685. Even if one calculates the A_M on the basis of one histidine residue and the total tyrosine content (4 residues), the value of A_M is increased to only 35 500. Thus, irrespective of the assumption concerning the number of reactive tyrosines, the experimental A_M value reflects a color contribution by 1.8 histidine residues rather than one residue as would be expected. In marked contradistinction to this result, chymotrypsin that has been methylated on the N position of the imidazole ring of His-57 (*20*) gives a Pauly color yield consistent with only one reactive histidine residue (*21*). A higher than expected histidine reactivity (1.6 groups) has also been detected in the reaction between TPCK-chymotrypsin and the histidine reagent, diazonium-1-H-tetrazole (*21*). Furthermore, when TPCK-chymotrypsin is first denatured and *S*-aminoethylated prior to hydrolysis, amino acid hydrolysis yields 1.6 histidine residues (*21*). These results suggest that perhaps the imidazole ring of His-57 in chymotrypsin is not the only site of reaction with TPCK.

ACKNOWLEDGMENTS

This work was supported in part by a research grant from the National Science Foundation (GB-33473) and by research grants from the Institute of General Medical Sciences (GM-10902) and the National Heart and Lung Institute (HE-14803), National Institutes of Health.

REFERENCES

1. M. A. MARINI AND C. J. MARTIN, *Eur. J. Biochem.*, **19**, 153 (1971).
2. C. J. MARTIN, N. B. OZA, AND M. A. MARINI. *Can. J. Biochem.*, **10**, 1114 (1972).
3. C. J. MARTIN, N. B. OZA, AND M. A. MARINI, *Eur. J. Biochem.*, **20**, 276 (1971).
4. C. J. MARTIN AND M. A. MARINI, *J. Biol. Chem.*, **242**, 5736 (1967).
5. D. FRENCH AND J. T. EDSALL, *Advan. Protein Chem.*, **2**, 227 (1945).
6. R. G. KALLEN AND W. P. JENCKS, *J. Biol. Chem.*, **241**, 5864 (1966).
7. M. A. MARINI AND C. J. MARTIN, *Anal. Biochem.*, **26**, 231 (1968).
8. P. DUNLOP, M. A. MARINI, H. M. FALES, E. SOKOLOSKI, AND C. J. MARTIN, *Bioorganic Chem.*, **2**, 235 (1973).
9. P. DUNLOP, M. A. MARINI, AND C. J. MARTIN, *Biochim. Biophys. Acta*, **243**, 320 (1971).
10. K. HOFMANN, *in* "The Chemistry of Heterocyclic Compounds: Imidazole and Its Derivatives" (A. Weissberger, ed.), Part 1, p. 136. Interscience, New York, 1953.
11. K. J. STEVENSON AND L. B. SMILLIE, *J. Mol. Biol.*, **12**, 937 (1965).
12. S. F. HITCH, Ph.D. thesis, University of London (1956).

13. G. SCHOELLMANN AND E. SHAW, *Biochemistry*, **2**, 252 (1963).
14. F. H. WHITE, JR., *in* "Methods in Enzymology" (C. H. W. Hirs, Ed.), Vol. XI, p. 482. Academic Press, New York, 1967.
15. P. ALEXANDER AND R. J. BLOCK, Eds., "The Composition, Structure and Reactivity of Proteins," Vol. 2, p. 92. Pergamon Press, New York, 1960.
16. C. TANFORD, *Advan. Protein Chem.*, **17**, 69 (1962).
17. M. J. KRONMAN AND F. M. ROBBINS, *in* "Fine Structure of Proteins and Nucleic Acids" (G. D. Fasman and S. N. Timasheff, Eds.), p. 271. Marcel Dekker, New York, 1970.
18. M. J. GORBUNOFF, *Biochemistry*, **10**, 250 (1971).
19. E. B. ONG, E. SHAW, AND G. SCHOELLMANN, *J. Amer. Chem. Soc.*, **86**, 1271 (1964).
20. Y. NAKAGAWA AND M. L. BENDER, *Biochemistry*, **9**, 259 (1970).
21. T. T. BLAIR, M. A. MARINI, AND C. J. MARTIN, *FEBS Lett.*, **20**, 41 (1972).