BIOORGANIC CHEMISTRY, 2, 235–247 (1973)

NMR Studies of the Reaction of Formaldehyde with the Imidazole Side Chain of Histidine: A Reactive Amino Acid in Enzyme Catalysis[1]

PATRICIA DUNLOP AND MARIO A. MARINI

Northwestern University Medical School, Department of Biochemistry, 303 East Chicago Avenue, Chicago, Illinois 60611

HENRY M. FALES AND EDWARD SOKOLOSKI

National Heart and Lung Institute, National Institutes of Health, Bethesda, Maryland 20014

CHARLES J. MARTIN

The Chicago Medical School, Department of Biochemistry, 2020 West Ogden Avenue, Chicago, Illinois 60611

Received October 25, 1972

The reaction of imidazole and imidazole derivatives with formaldehyde can be demonstrated with NMR techniques. The results show that only one nitrogen of the imidazole ring reacts to form a *N*-hydroxymethyl derivative in alkaline solution. Under acidic conditions both nitrogen positions can support *N*-hydroxymethyl derivatives.

This reaction represents a useful tool for the further investigation of enzyme mechanisms involving the imidazole nucleus.

INTRODUCTION

There have been numerous studies on the reaction of formaldehyde with the imidazole and amino groups of histidine. Some workers found that formaldehyde does not react with the imidazole group (*1–3*); in other cases it does react (*4–8*). Some have concluded that the reaction occurs with the free amino group followed by a cyclization (*2, 9*) with the imidazole ring.

In an elegant potentiometric study, Kallen and Jencks (*6*) reported that formaldehyde reacts with the nitrogens of the ring. The nonprotonated imidazole forms a monohydroxymethyl adduct while the protonated ring forms a dihydroxymethyl derivative. Changes in pK' of a number of hydroxymethyl imidazole derivatives in water and guanidine solution also indicate that a definite reaction has occurred (*8*). The reaction of formaldehyde and imidazole has also been detected spectrally (*5, 7*). It has also been reported that both imidazoles of chymotrypsin react with formaldehyde (*7*). The

[1] This investigation was supported in part by Public Health Research Grant HE 14803 of the National Heart and Lung Institute and GM 10902 of the Institute of General Medical Sciences, National Institutes of Health.

catalytic efficiency of chymotrypsin and model imidazole compounds is reduced in formaldehyde but is not abolished (*10, 11*).

Currently there appears to be a reasonably large body of evidence to justify confidence that a hydroxymethyl derivative of imidazole can be formed. However, the nature of its formation and its structure are not clearly understood, since it is not experimentally feasible to isolate it because of the ease with which the reaction is reversed upon dilution. Because formaldehyde modified the active site of α-chymotrypsin with retention of activity (*11*), it is of some interest to know more of the structure so that mechanistic interpretations of catalysis might be possible.

Following the work of Wolff et al. (*12*) on the structure of ophidine, it appeared that an examination of the nuclear magnetic resonance (NMR) spectra would reveal the formation of a hydroxymethyl imidazole derivative by the resonance frequency characteristics of a methylene group located between a nitrogen and an oxygen. Furthermore, it was hoped that perturbations of the ring protons might also lead to an understanding of the nature of the bond(s) formed.

EXPERIMENTAL

Sixty megahertz NMR spectra were obtained on a Varian Associates A-60 analytical NMR spectrometer coupled with a V-6040 NMR variable temperature controller. Refined spectra were obtained with 90 megahertz, 100 megahertz, and 200 megahertz

TABLE 1

NUCLEAR MAGNETIC RESONANCE FREQUENCY PATTERNS FOR IMIDAZOLE DERIVATIVES[a]

Compound	pD	Proton bands in PPM				
		C_2	C_4C_5	α-C	β-C	Methyl
Imidazole	9	7.8(t)	7.18(d)			
	1	8.87(t)	7.62(d)			
N-methyl-imidazole	9	7.6(st)	7.02—7.1(d of q)			3.68(s)
	1	8.90(bs)	7.61(d of q)			4.10(s)
	1[b]	9.22(t)	7.83—7.73(d of q)			4.70(s)
1-methyl-imidazole-4-acetic acid	10	7.53(cs)	6.94(cs)		3.45(d)	3.69(s)
	2	8.72(cs)	7.43(cd)		3.94(cs)	3.94(cs)
N-acetyl-L-histidine	8.5	7.76(s)	6.98(s)	4.48(q)	3.12(2cd)	2.0(s)
	7.5	8.13(s)	7.1(cs)	4.5(q)	3.17—3.02(m and 2d)	1.98(s)
	5.0	8.63(d)	7.3(m)	4.4(q)	3.28—3.10(dt and 2d)	2.0(s)
	2.5	8.65(d)	7.35(m)		3.3—3.20(dt and 2d)	2.0(s)

[a] Legend to symbols: s—singlet; d—doublet; t—triplet; q—quartet; m—multiplet; bs—broad singlet; cs—complex singlet; cd—complex doublet; cm—complex multiplet; st—smeared triplet; dt—doublet of triplets; dm—doublet of multiplets.

[b] In DMSO, the acid added was the same as for pD 1 in D_2O.

instruments. All instruments were tuned with neat acetaldehyde. All samples were dissolved in deuterium oxide (D_2O) or dimethyl sulfoxide (DMSO), spectral grade. The internal reference for D_2O was 3-(trimethyl-silyl)-1-propanesulfonic acid sodium salt; for DMSO the reference was tetramethyl silane (TMS). All assignments of chemical shifts were in reference to these standards. The spectral results are reported in terms of

TABLE 2

NUCLEAR MAGNETIC RESONANCE FREQUENCY PATTERNS FOR IMIDAZOLE DERIVATIVES IN FORMALDEHYDE

Compound	pD	Proton bands in PPM[a]					
		C_2	C_4C_5	α-C	β-C	Methyl	Hydroxymethyl
Imidazole	9	7.81(t)	7.18(dt)				5.46(s)
	1	8.83(t)	7.75(dt)				5.75(s)
		9.00(t)	7.58(dt)				
		9.17(t)					
N-methyl-	9	7.63	7.10–7.03			3.70(s)	None
imidazole	1	8.78	7.53(q)				
		8.95	7.67(q)			4.05(s)	5.7(s)
	1[b]	9.10	7.96–7.77(m)				
		8.93				4.20(s)	5.83(s)
1-methyl-	10	7.53(cs)	6.94(cs)		3.45(d)	3.69(s)	None
imidazole-	5.5	8.63	7.36(cs)		3.73(d)	3.89(s)	5.58(s)
4-acetic		8.82					
acid	2	8.78(bs)	7.5(cs)		4.10(d)	4.03(s)	5.73(s)
		8.98(bs)	7.6(cs)				
N-acetyl-	8.5	7.93	7.08	In HOH	3.08	1.98	5.47
L-histidine	7.5	8.27	7.2	In HOH	3.13	1.97	5.55
	5	8.63	7.3	In HOH	3.24	2.0	5.6
		8.77	7.42	In HOH			5.68
		8.97					
	2.5	8.65	7.36	In HOH	3.3	2.02	5.63
		8.82	7.48				5.70
		9.07					

[a] Symbols are as in Table 1.

[b] In DMSO with an equivalent amount of HCl added as used for pD in D_2O.

δ (ppm). All spectra in these solvents were recorded at room temperature; the low temperature studies were run in deuterated chloroform as a two-phase system.

Imidazole, free base form, was obtained from J. T. Baker Chemical Company. *N*-Acetyl-L-histidine, *N*-methylimidazole, and 1-methylimidazole-4-acetic acid were obtained from CalBiochem Corporation. Formaldehyde, 12.4 *M* (37% w/v, 1% methanol) was donated by the Celanese Chemical Corporation.

For the determination of spectra, 50 mg samples of each derivative were dissolved in 0.3 ml of of solvent. The pH was recorded using full range pH paper and equated to pD. For adjustment to the basic range, pulverized dry sodium hydroxide was used. Adjustment to acid pH was done with concentrated hydrochloric acid. Unless otherwise stated, the formaldehyde concentration when present was 3.1 *M*. Just prior to use, the

formaldehyde was neutralized by passage through Dowex 1-OH until the pH was 6.5–7.0. No dilution was incurred in this procedure.

An NMR spectrum of formaldehyde thus treated shows the resonance frequency of the methanol residue at 4.95 ppm, slightly downfield from the 4.85 ppm peak for water. Under all the conditions used in these studies the chemical shift of formaldehyde never exceeded the downfield value of 4.98 ppm or the upfield value of 4.88 ppm. In all cases, the formaldehyde frequency merged with the water frequency to give one very broad band.

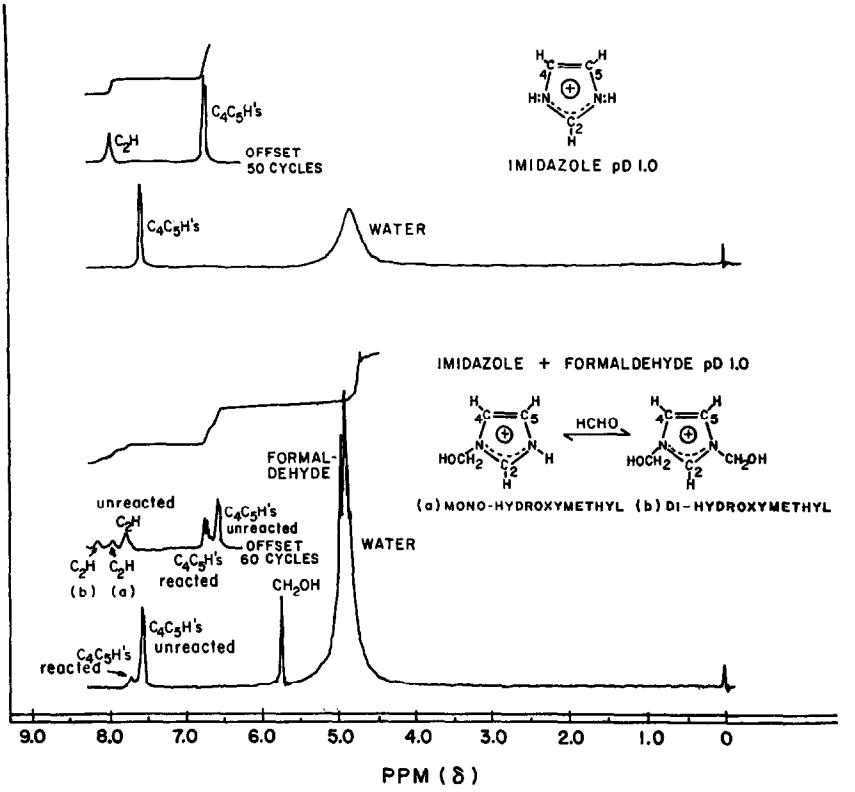


FIG. 1. Spectra for imidazole and imizadole in 3 *M* formaldehyde (pD 1.0) at room temperature (the lower inset trace shows the time-dependent formation of the derivative).

RESULTS AND DISCUSSION

A complete summary of the NMR data is contained in Tables 1 and 2 and the conclusions discussed in the text are based on these data. On occasion, references are made to the original figures to illustrate specific points.

The spectra of imidazole at pD 1 (Figs. 1 and 2) and at pD 9 (Fig. 3) show that the chemical shifts of the C_2 and C_4C_5 hydrogens are in agreement with those previously reported (*13*). Integration of these peaks gives a ratio of 1:2, which would confirm the assignment of the furthest downfield peak to the C_2 hydrogen and the upfield peak to the C_4C_5 hydrogens. The splitting patterns indicate spin-coupled hydrogens and an examination of the spectra with a 90 megahertz instrument showed a typical A_2X

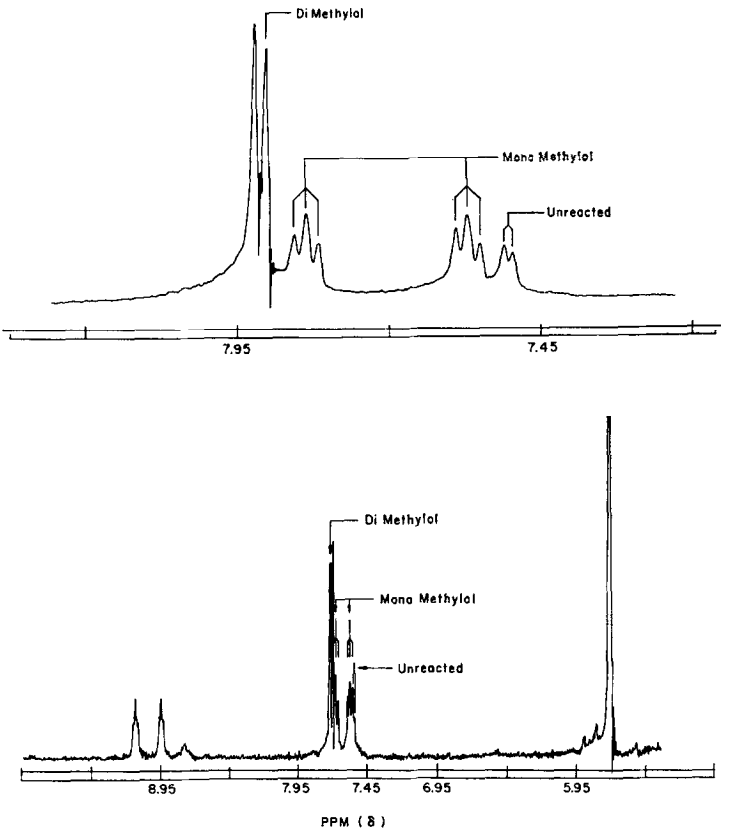
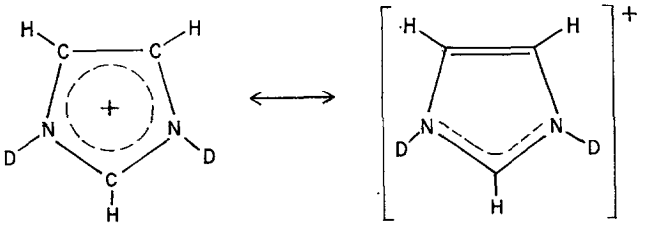


FIG. 2. Spectra of the C_2 and C_4C_5 hydrogens of imidazole in 12 *M* formaldehyde at pD 1.0 (lower, 100 MHz; upper, 200 MHz).

system with the downfield triplet broadened, perhaps from the effects of the nitrogen quadruple and/or coupling to the ^{14}N nucleus (spin = 1). A coupling constant of 1 Hz/sec between the C_2 and C_4C_5 hydrogens was derived from the 90-megahertz scan. The hydrogen on the nitrogen function is rapidly replaced by deuterium in deuterium oxide and is not observed.

It appears that the C_4 and C_5 hydrogen exert an equivalent effect on the C_2 hydrogen in both acid and base. At pD 1 this is understandable since the cation is fully symmetrical whether either resonance hybrid I or Ia is emphasized.



I Ia

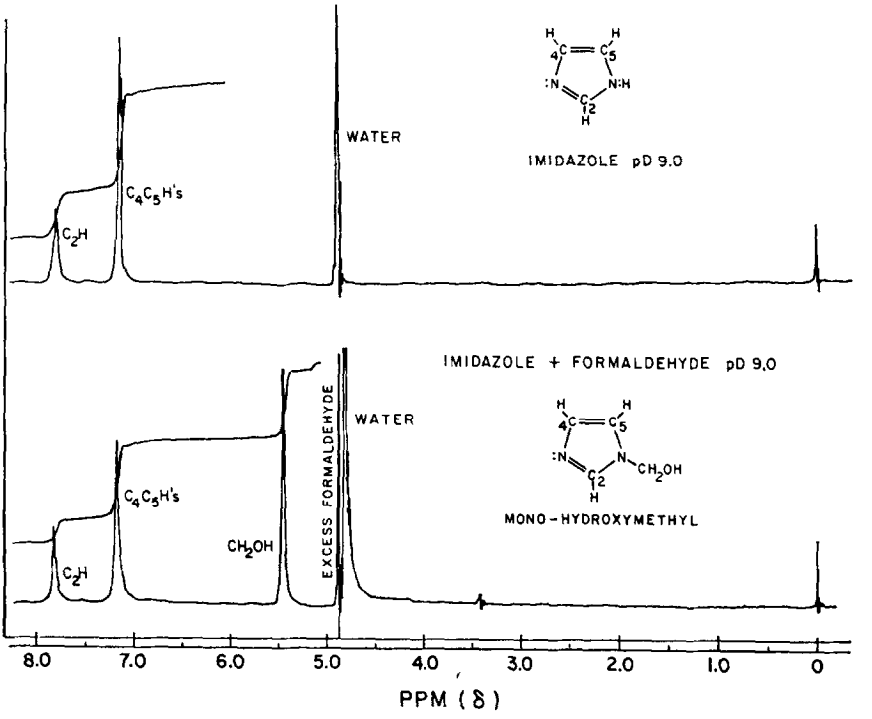
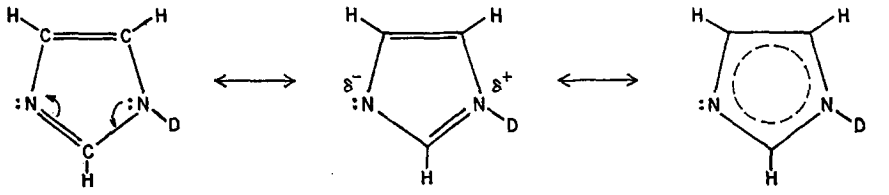


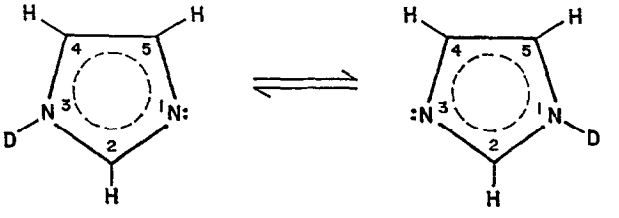
FIG. 3. Spectra of imidazole and imidazole in 3 *M* formaldehyde (pD 9.0) at room temperature.

But this is definitely not the case at pD 9 where the C_4C_5 hydrogens are separated from the C_2 hydrogen by nitrogens which are nonequivalent in the system II ↔ IIa ↔ IIb.



II IIa IIb

An obvious explanation for this phenomenon is that the exchange rate of the N-D group with solvent may be fast enough (compared to the NMR measuring time) that the identity of the C_4C_5 protons is lost (cf. Reaction 1).



REACTION 1

In the presence of slightly more than one molar equivalent of formaldehyde, unprotonated imidazole in water at pD 9 (Fig. 3) shows a new peak at 5.46 ppm in addition to the C_2 and C_4C_5 frequencies. The integration ratio for the three peaks observed in this experiment, assuming the farthest downfield frequency to be the C_2 hydrogen, is 1:3.2:3.2. The irregular value of 3.2 hydrogens obtained for the peak previously

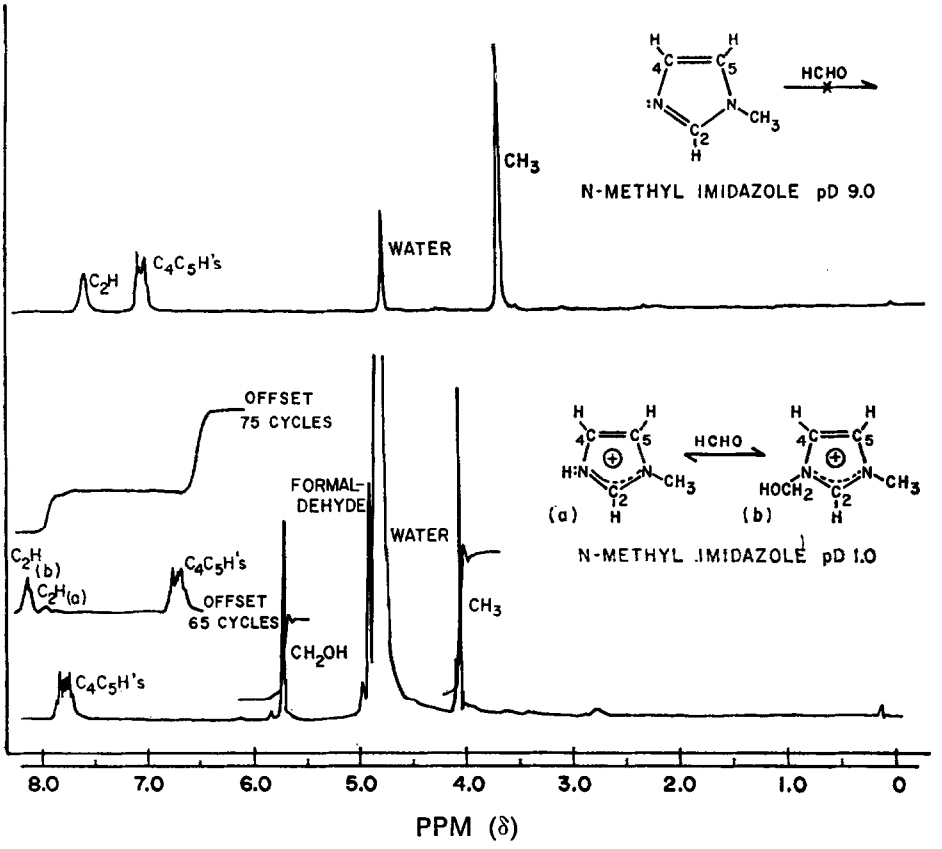


FIG. 4. Spectra of *N*-methylimidazole in 6 *M* formaldehyde at room temperature in DMSO at pD 1.0 and 9.0.

assigned to the C_4C_5 hydrogens is the result of slow exchange of the C_2 hydrogen for deuterium. Extrapolating back to zero time, the integration ratio becomes 1:2:2. The new peak at 5.46 ppm has the chemical shift and integration value expected for a methylene bonded between a nitrogen and an oxygen function (*14*).

This result is in agreement with the proposed formation of a *N*-hydroxymethylimidazole derivative.[2] However, the same problem of symmetry is found here as in free

[2] Both chemical shift and integration values militate against an alternative di-imidazolylmethane structure.

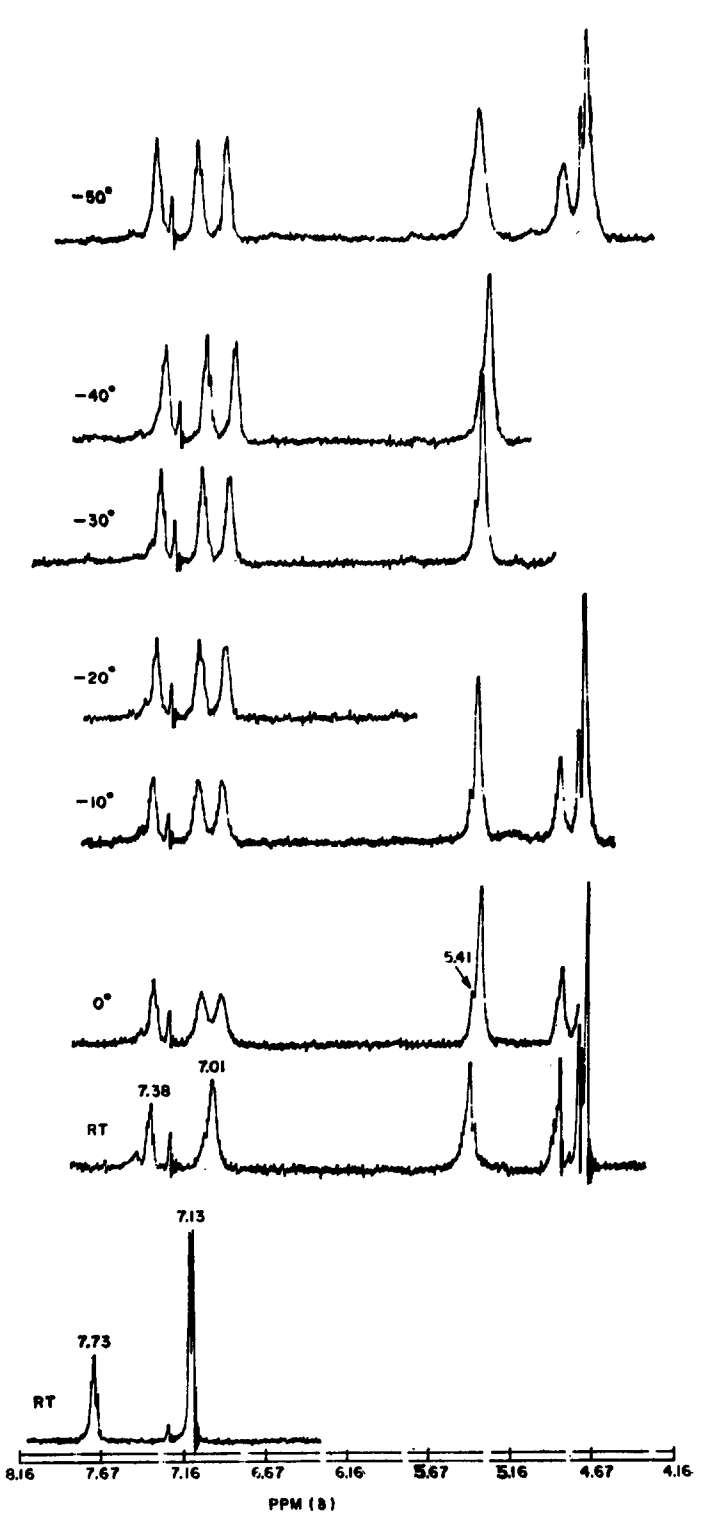
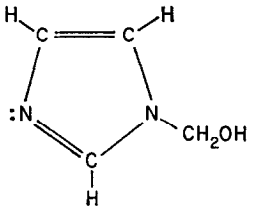


FIG. 5. Spectra of imidazole (bottom trace) at room temperature (RT) and hydroxymethyl imidazole from room temperature to −50°C in deuterated chloroform.

imidazole at pD 9.0, i.e., the coupling pattern (at 90 megahertz) shows a triplet for the C_2 hydrogen and a broad unresolved singlet for the C_4C_5 hydrogens. At pD 9 an unsymmetrical substitution, as shown by III,



III

similar to that found in *N*-methyl imidazole (Fig. 4) should have caused far more splitting than is actually observed, assuming the complex is a stable entity.

In an attempt to clarify the exact nature of this reaction, spectra were obtained on a similar mixture of imidazole (free base) and formaldehyde dissolved in chloroform (two phases were present) to allow observation at low temperature. The free base form

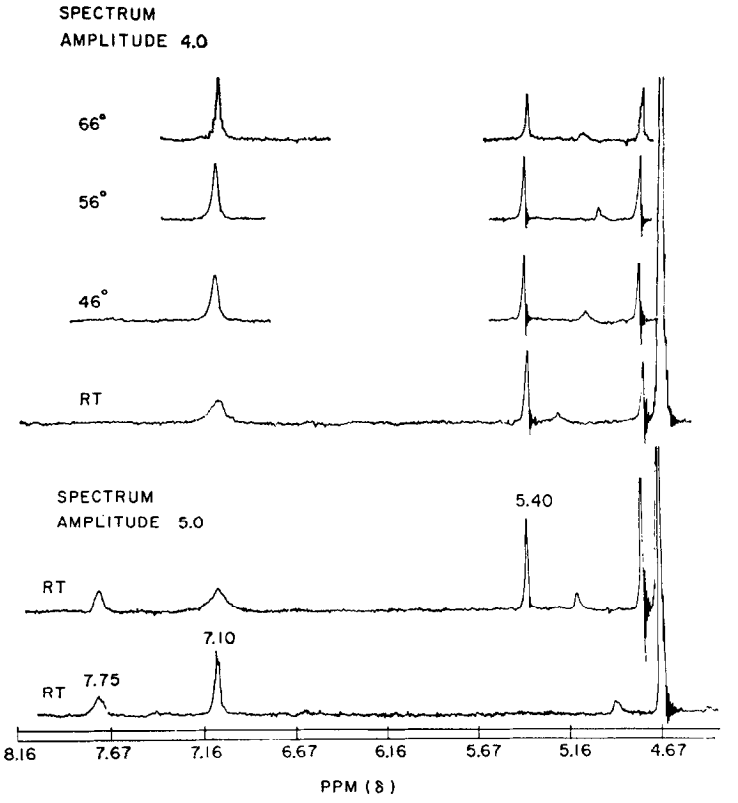


FIG. 6. Spectra of imidazole at room temperature (bottom trace) and hydroxymethyl imidazole from room temperature to 66°C in D_2O. The upper four traces are at spectrum amplitude 4; the other trace of hydroxymethyl imidazole (RT) is at spectrum amplitude 5.

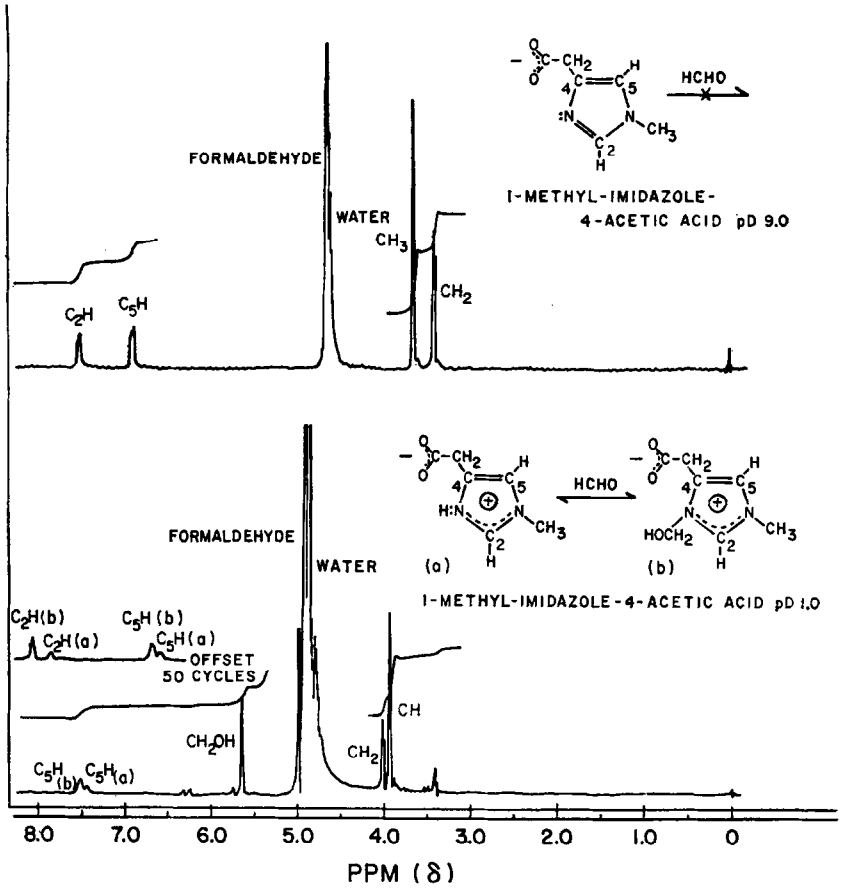
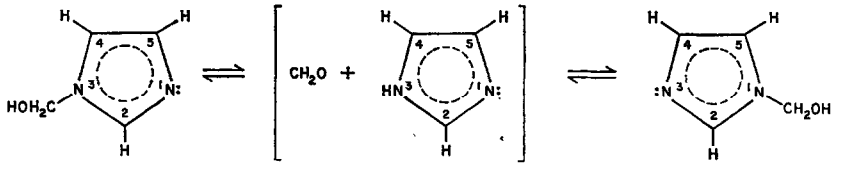


FIG. 7. Spectra of 1-methyl-imidazole-4-acetic acid in 3 *M* formaldehyde at room temperature.

of imidazole produced an alkaline medium without the addition of base. The resulting spectrum (Fig. 5), at room temperature, resembled that observed in water. The broad unresolved peak at 7.01 ppm, presumably due to the C_4C_5 hydrogens, splits into a pair of poorly resolved triplets as the temperature is lowered to $\sim -10°C$. We interpret this observation to suggest that at room temperature the exchange rate between structurally identical molecules (Reaction 2) which contain the methylol group on different nitrogens is sufficiently fast so that the NMR instrument sees an average of the two and distinctions between the C_4 and C_5 hydrogens are lost as in the case of proton exchange in imidazole itself.



REACTION 2

As the temperature is lowered further, the rate of the reaction is decreased and the C_4 and C_5 hydrogens can now be observed as distinct entities. It seems logical to suppose that formaldehyde and free imidazole are intermediates in this reaction and our failure to observe appropriate NMR peaks, indicating their presence, would mean that their concentration is very low. Water solutions freeze before exhibiting similar spectra, but proof that the same process is occurring is provided by Fig. 6 where a solution of

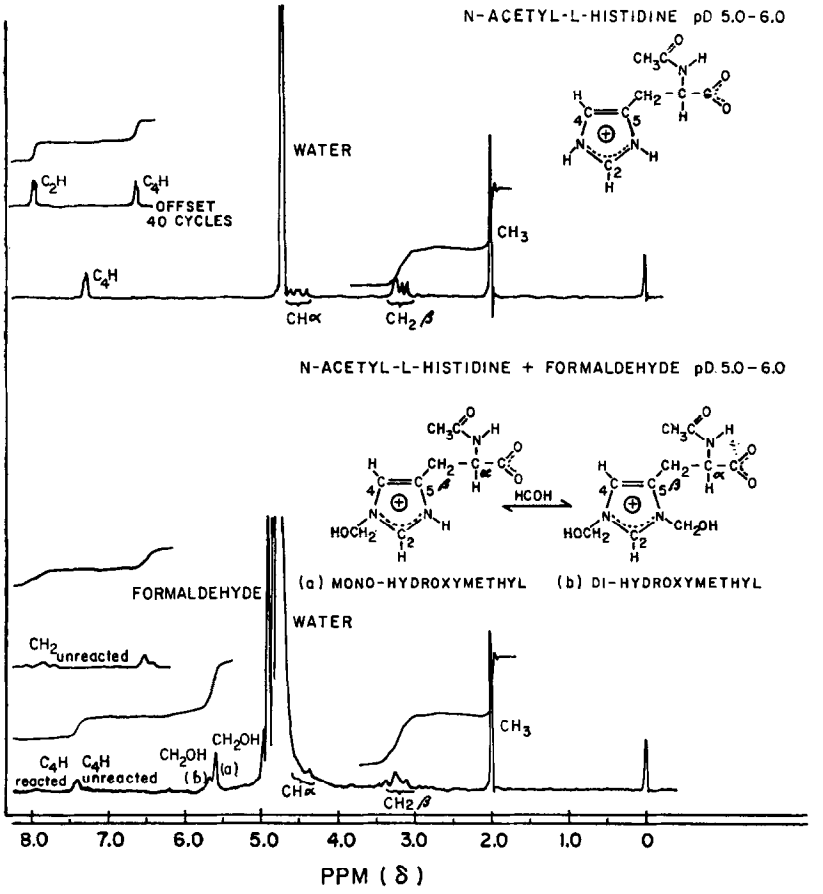
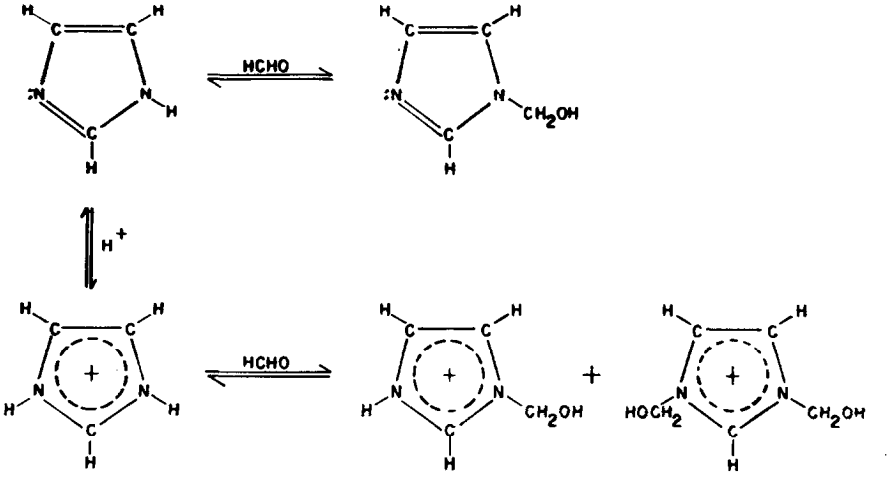


FIG. 8. Spectra of *N*-acetyl-L-histidine and *N*-acetyl-L-histidine in 3 *M* formaldehyde at room temperature.

imidazole and formaldehyde at pH 9.0 has been heated to 66°C resulting in considerable sharpening of this peak as the exchange rate is increased. In this case the C_2-hydrogen had exchanged completely so that the broadening observed at room temperature could not have resulted from spin coupling from this proton.

The spectrum of 3 *M* formaldehyde and imidazole at pD 1 (Fig. 1) shows a new singlet peak at 5.75 ppm and multiple peaks for the resonance species of the C_2 and C_4C_5 hydrogens. These multiple peaks are due to reaction products and not to spin-coupling interactions. This evaluation is based upon close examination of the C_4C_5

patterns using 100 and 220 megahertz spectra (Fig. 2). These spectra show the di-species as a large doublet downfield, the mono-species as two doublets which appear as triplets and a triplet of lesser intensity at highest field. The upfield member of each set is unreacted imidazole cation. The downfield members are the resonance frequencies of the C_4C_5 and C_2 hydrogens under the influence of the newly reacted adjoining nitrogen functions. The 9.0 ppm peak is the C_2 hydrogen shift due to a single hydroxymethyl derivative involving one of the two nitrogen residues bonded to the C_2 hydrogen. The 9.17 ppm peak is caused by a dihydroxymethyl derivative, in which both nitrogen functions are involved. This observation is based upon a mathematical relationship observed between the integration values of the new species peaks and the total integration value observed for the product peak. The 5.75 ppm singlet is the chemical shift of both mono- and di-substituted hydroxymethyl derivatives. Integration of the area under this peak yields a value of two plus hydrogens which would support the presence of both derivatives. Assuming a direct substitution reaction, these data can be summarized by the following equilibria:



REACTION 3

N-methylimidazole (Fig. 4) and 1-methylimidazole-4-acetic acid (Fig. 7) at basic pD show no reaction product with excess formaldehyde. The acidic spectra of *N*-methylimidazole and formaldehyde shows an hydroxymethyl reaction product peak after standing several hours. 1-Methylimidazole-4-acetic acid in D_2O and formaldehyde shows a reaction peak for 5.7 ppm for the hydroxymethyl group. The absence of a reaction product for both compounds with formaldehyde at pD 9 is in agreement with the interpretation of kinetic measurements (*10*). It would thus appear that the imidazole ring can accommodate a disubstitution of the nitrogens in acid but only a single substitution in base.

The reaction of *N*-methylimidazole with formaldehyde in DMSO is a much more rapid reaction and within an hour the nonequivalent C_4C_5 hydrogens seen as a pair of closely spaced quadruplets at 7.73 and 7.83 ppm are accompanied by a new, unresolved multiplet at 7.77–7.93 ppm, a singlet at 5.83 ppm, and a poorly resolved multiplet for the C_2H at 9.22 ppm. The integrals of the aforementioned peaks are in the ratio of

2:2:1 and on this basis we suggest the *N*-methyl-*N'*-methylolimidazole cation structure although, as in the case of imidazole itself, we know of no precedent for such a reaction and confirmation and isolation of the compound is most desirable. The evidence for the existence of the *N*-hydroxymethylated cation is nonetheless compelling since the appearance of the band at 5.83 ppm has the correct ratio of hydrogen atoms.

The spectra of *N*-acetyl-L-histidine (Fig. 8) also show mono- and dihydroxymethyl reaction products in the presence of formaldehyde under the appropriate conditions. In acid solution, the C_2 and C_4 hydrogens show the expected multiplicities of chemical shift and the mono- and dihydroxymethyl derivatives are seen as two independent resonance peaks. Integration of the area under the peaks show that the downfield peak (5.70 ppm) is the chemical shift of the dihydroxymethyl derivative. The upfield member (5.63 ppm) is the mono-hydroxymethyl derivative. In the basic range, only the monohydroxymethyl derivative exists.

These results may be applied to histidine and its derivatives although the reaction with formaldehyde is much more complex since many more species are present. Work on this reaction is currently in progress.

ACKNOWLEDGMENT

The authors thank Daniel A. Netzel, Northwestern University, for his expert technical assistance.

REFERENCES

1. M. LEVY, *J. Biol. Chem.* **109**, 365 (1935).
2. D. FRENCH AND J. T. EDSALL, *Advan. Protein Chem.* **2**, 277 (1945).
3. E. H. FRIEDEN, M. S. DUNN AND D. C. CORYELL, *J. Phys. Chem.* **43**, 85 (1943).
4. T. W. BIRCH AND L. J. HARRIS, *Biochem. J.* **24**, 564 (1930).
5. L. J. SAIDEL AND R. L. CARINO, *Fed. Proc.* **25**, 796 (1966).
6. R. C. KALLEN AND W. P. JENCKS, *J. Biol. Chem.* **241**, 5864 (1966).
7. C. J. MARTIN AND M. A. MARINI, *J. Biol. Chem.* **242**, 5736 (1967).
8. M. A. MARINI AND C. J. MARTIN, *Anal. Biochem.* **26**, 231 (1968).
9. A. NEUBERGER, *Biochem. J.* **38**, 309 (1944).
10. C. J. MARTIN, N. B. OZA, AND M. A. MARINI, *Eur. J. Biochem.* **20**, 276 (1971).
11. M. A. MARINI AND C. J. MARTIN, *Eur. J. Biochem.* **19**, 153 (1971); P. C. DUNLOP, M. A. MARINI, AND C. J. MARTIN (manuscript in prep.).
12. J. WOLFF, K. HORISAKA, AND H. M. FALES, *Biochemistry* 7, 2455 (1968).
13. A. MANNSCHRECK, W. SEITZ, AND H. A. STAAB, *Ber. Bunsenges Physik. Chem.* **65**, 470 (1963).
14. J. R. DYER, "Application of Absorption Spectroscopy of Organic Compounds" (K. L. Rinehart. Jr., Ed.). Prentice-Hall, New Jersey, 1965.