*Tetrahedron* Vol 32, pp. 2595 to 2597. Pergamon Press 1976. Printed in Great Britain

# UNUSUAL SITE OF ELECTROPHILIC ATTACK ON INDOLE AND CARBAZOLE

A. CIPICIANI, S. CLEMENTI, P. LINDA,* G. SAVELLI and G. V. SEBASTIANI
Istituto di Chimica Organica, Università di Perugia, Perugia, Italy

(*Received in UK* 12 *April* 1976; *Accepted for publication* 11 *May* 1976)

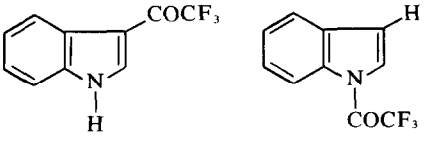
**Abstract**—Whereas trifluoroacetylation of pyrrole in 1,2-dichloroethane gives exclusively C-substitution and carbazole gives exclusively N-substitution, trifluoracetylation of indole gives a mixture of N- and C-derivatives. This behaviour is discussed in terms of different loss of resonance energy in the transition state leading to N-substitution. Formation of N-trifluoroacetyl-2-(3-indolyl)indoline in trifluoroacetylation of indole is also discussed.

We recently measured the relative reactivities of pyrrole, indole and of a number of indole derivatives in the Vilsmeier–Haack formylation by competitive methods.[1] To confirm these results we undertook a kinetic study on trifluoroacetylation by trifluoroacetic anhydride, the kinetics of which has been recently elucidated.[2]

Pyrrole is known to give only 2-trifluoroacetylation in very high yield.[3,4] Rate constants were determined at 0° in 1,2-dichloroethane and the expected third order kinetics were observed (the mean value on six runs is 36 l$^2$ mol$^{-2}$s$^{-1}$).

In contrast, reaction of indole under the same conditions gave, surprisingly, only 20% of the expected 3-trifluoroacetylindole (**1**), whereas 50% of the mixture was N-trifluoroacetylindole (**2**), and 15% was N-trifluoroacetyl-2-(3-indolyl)indoline (**3**), together with 15% of indole (either unreacted or, most probably, derived from hydrolysis of **2**). The product ratio was obtained by GLC analysis. Compound **1** was already known,[5,6] while the structures of **2** and **3** were elucidated by chemical and spectroscopic methods.

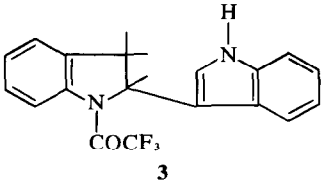
In fact the NH signal was absent in the IR and NMR spectra of compound **2**, which is quantitatively hydrolyzed to indole under basic conditions. The NMR spectrum of compound **3** is consistent with the presence of one aromatic NH only and with the typical pattern of an indoline nucleus. Under basic conditions **3** is readily hydrolyzed to 2-(3-indolyl)indoline.[7] Besides, the CO band frequency of both compounds is significantly higher than that of **1**, thus indicating a less important resonance interaction between the acyl group and the aromatic ring: a strong evidence for N-acyl substitution in contrast to C-substitution.



**1**



**2**

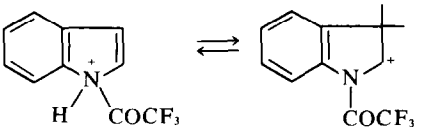


**3**

Furthermore the product composition is strongly dependent on experimental conditions (e.g. concentration of reactants, mixing order, solvent) as shown in Table 1.

Formation of N-trifluoroacetyl-2-(3-indolyl)indoline (**3**) is favoured by high concentration of reactants, whereas at lower concentrations N-trifluoroacetylindole (**2**) is the main product. Prevalence of **2** over **3** is also favoured by adding dropwise the indole to the solution of trifluoroacetic anhydride, rather than the usual addition of anhydride to the substrate. Both results are consistent with a mechanism involving the primary formation of the very reactive N-trifluoroacetyl carbocation (**4b**). This can subsequently react rapidly with the excess of indole, if present, giving **3**, or just lose a proton giving **2**: therefore the local concentration of indole is highly crucial.

A similar mechanism is probably involved in the reaction between indole and acetic-formic anhydride reported by Bergman.[8] The alternative mechanism, involving the primary formation of the so-called dimer of indole, 2-(3-indolyl)indoline, due to trifluoroacetic acid always present in the mixture because of the ready hydrolysis of the anhydride, and subsequent trifluoroacetylation of the strongly basic N atom of the indolenine, is less probable since, in analogous experiments, N-methylindole gives only N-methyl-3-trifluoroacetylindole in very high yield (90–95%) with the dimerization product present in traces.



**4a** **4b**

The competition between N- and C-substitution is more difficult to rationalize, especially because the behaviour of the indole nucleus towards trifluoroacetylation seems unusual in respect to most of the other electrophilic substitutions, giving mainly C-substitution.[9] At the moment we are unable to give a complete explanation of this phenomenon; nevertheless we can suggest a complex mechanism for the N- vs C-trifluoroacetylation involving the peculiar electronic demand of $CF_3$ in the electrophilic agent. Moreover, the solvent effect discussed below could be a cue for a different acid assistance on the bond breaking of the intermediates at the two positions. We should point out that the whole picture seems to be consistent with the proposed mechanism for dimerization of indoles.

Table 1. Product composition (% mole) in trifluoroacetylation of indole derivatives

| Substrate (M)[a] | Reagent (M)[a] | Solvent[b] | N-$COCF_3$ | 3-$COCF_3$ | III | Unreacted | Total yield |
|---|---|---|---|---|---|---|---|
| TFA (0.5) | Indole (0.1) | DCE | 73 | 15 | — | 12 | 74 |
| Indole (0.1) | TFA (0.5) | DCE | 50 | 20 | 15 | 15 | 70 |
| Indole (2.5) | TFA (2.5) | DCE | — | — | 90 | 10 | 90 |
| Indole (2.5) | TFA (2.5) | DMF | — | 90 | — | 10 | 88 |
| N-Methylindole (2.5) | TFA (2.5) | DCE | — | 100 | | — | 95 |
| Carbazole | TFA (0.5–2.5) | DCE | 100 | — | | — | 100 |

[a] Reagent indicates the reactant which is dropped into the solution of the other reactant (substrate); the molarity is reported in brackets.

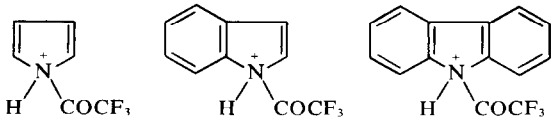
[b] TFA = trifluoroacetic anhydride; DCE = 1,2-dichloroethane; DMF = N,N-dimethylformamide.

Table 2. Observed third-order rate constants for the trifluoroacetylation of pyrrole in 1,2-dichloroethane at 0°

| Run | $10^3$[ArH]/M | $10^3$[$(CF_3CO)_2O$]/M | $10^3$[$CF_3COOH$]/M | $k^{III}_{obs}$ l² $mol^{-2}s^{-1}$ |
|---|---|---|---|---|
| 1 | 19.74 | 4.90 | 4.37 | 37 |
| 2 | 18.69 | 5.36 | 4.39 | 41 |
| 3 | 8.00 | 3.65 | 6.15 | 34 |
| 4 | 8.40 | 4.28 | 6.80 | 35 |
| 5 | 7.62 | 3.61 | 7.47 | 35 |
| 6 | 5.75 | 5.36 | 8.09 | 35 |

Most interesting is the dramatic solvent effect: in N,N-dimethylformamide only 3-trifluoroacetylindole was obtained. This result could be tentatively explained in terms of different basicity of the solvent, i.e. DMF could behave as acid scavenger in a rate determining step involving assistance by acid. This hypothesis is in keeping with the suggested mechanism of aromatic trifluoroacetylation, where the assistance by an acid molecule on the bond breaking of the intermediate could not be excluded.[2] Since the kinetic order for the trifluoroacetylation of indole is less clear than that of pyrrole, we conclude that this step could really become rate determining under certain conditions or for one of the two positions of attack.

Finally, carbazole, under various conditions, undergoes exclusive N-trifluoroacetylation in quantitative yield. We suggest that the observed order of 'facile' N-trifluoroacetylation: pyrrole (0%) < indole (50%) < carbazole (100%) must be related to the different degree of the overall loss of aromaticity in the transition state for the examined compounds as approximated by the Wheland intermediate structures **5–7**. The nitrogen quaternization resultant by direct attack on the heteroatom leads to a larger degree of bond fixation in pyrrole than in indole (one benzene ring unchanged), than in carbazole (two benzene rings).




**5** **6** **7**

## EXPERIMENTAL

NMR spectra were obtained using a JEOL C-60HL spectrometer using TMS as internal standard and DMSO d-6 as solvent. The chemical shifts were given in ppm. GLC analyses were carried out with a C. Erba fractometer model GI equipped with a flame ionization detector using a 2 mt × 2 mm column packed with GSE 301 10% at 140°. The IR spectra were obtained on film with a Perkin–Elmer spectrophotometer.

*3-Trifluoroacetylindole* (**1**) was prepared according to Katner[5] by adding dropwise trifluoroacetic anhydride (6.45 g, 30.7 mmol) to indole (3 g, 25.6 mmol) in 10 ml of N,N-dimethylformamide. After usual work-up 5.63 g (90%) of **1** were obtained m.p. 209–11°, lit. 209–12°,[5] $\nu_{C=O}$ = 1650 $cm^{-1}$, $\delta$ 7.15–7.65 (3H, m, Ar–H), 8.13 (1H, m, 4-H), 8.39 (1H, m (coupling with fluorine), 2-H), 12.40 (1H, bs, N-H).

*N-Trifluoroacetylindole* (**2**). A soln of indole (0.47 g, 4 mmol) in anhyd 1,2-dichloroethane was added dropwise to trifluoroacetic anhydride (4.2 g, 20 mmol) in 1,2-dichloroethane (20 ml) at 0°. The mixture was stirred for 20 min, poured into sat. $NaHCO_3$ aq, the organic layer separated, dried ($Na_2SO_4$), filtered and concentrated. Fractional distillation of the residue yielded 0.43 g (51%) of pure **2**, b.p. 118–20°/20 mm, $\nu_{C=O}$ = 1750 $cm^{-1}$, $\delta$ 6.92 (1H, dd, $J_{3,2}$ = 3.8 Hz, $J_{3,7}$ = 0.7 Hz, 3-H), 7.25–7.75 (4H, m, Ar-H and 2-H), 8.30 (1H, m, 7-H).

*N-Trifluoroacetyl-2-(3-indolyl)indoline* (**3**). 3.81 g (90%) of dense yellow oil **3** were obtained by the same conditions as for **2**. The concentrated organic layer was washed repeatedly with water to remove indole and chromatographed on silica gel; $\nu_{C=O}$ = 1720 $cm^{-1}$, $\delta$ 3.08 (1H, d, $J_{gem}$ = 16 Hz, indoline 3-H), 3.70 (1H, dd, $J_{gem}$ = 16 Hz, $J_{vic}$ = 8.5 Hz, indoline 3-H), 6.16 (1H, d, $J_{vic}$ = 8.5 Hz, indoline 2-H), 6.85–7.45 (8H, m, Ar-H and indole 2-H), 8.20 (1H, bd, $J_{ortho}$ = 7 Hz, indoline 7-H), 10.50 (1H, bs, indole NH).

*N-Methyl-3-trifluoroacetylindole* was obtained under various conditions of concentration, mixing order and solvent; m.p. 104°, lit[8] 105°, $\delta$ 3.90 (3H, s, N–$CH_3$), 7.20–7.70 (3H, m, Ar–H), 8.20 (1H, m, 4-H), 8.45 (1H, m. (coupling with fluorine), 2-H).

*N-Trifluoroacetylcarbazole.* The standard procedure was repeated for carbazole in 1,2-dichloroethane soln to obtain N-trifluoroacetylcarbazole in 90% yield, m.p. 61°, $\delta$ 7.15–7.60 (4H, m, 2-3-6-7-H), 7.90–8.15 (4H, m, 1-4-5-8-H).

*Kinetic procedure.* The kinetic runs of trifluoroacetylation of pyrrole were carried out at 0° as previously described.[2] The third-order equation followed is

$$\frac{dx}{dt} = k^{III}(a-x)(b-x)(c+x)$$

were a, b, and c are the initial concentration of the substrate, anhydride and acid, respectively. The individual rate constants are reported in Table 2.

*Acknowledgements*—We thank the Consiglio Nazionale delle Ricerche (ROMA) for financial support and prof. G. Marino for discussion.

## REFERENCES

[1]S. Clementi, P. Linda and G. Marino, *J.C.S. Chem. Commun.* 427 (1972).
[2]S. Alunni and S. Clementi, *J. Chem. Soc.* Perkin II, 1521 (1972).
[3]W. D. Cooper, *J. Org. Chem.* **23**, 1382 (1958).
[4]S. Clementi and G. Marino, *Tetrahedron* **25**, 4599 (1969).
[5]A. S. Katner, *Org. Preparat. Procedures*, **2**, 297 (1970).
[6]W. B. Whalley, *J. Chem. Soc.* 1651 (1954).
[7]B. L. VanDuuren, *J. Org. Chem.* **26**, 2954 (1961).
[8]J. Bergman, *J. Heterocyclic Chem.* **8**, 329 (1971).
[9]R. J. Sundberg, *The Chemistry of Indoles.* Academic Press, New York (1970).