The Mechanism of Trifluoroacetylation of Indoles [1]

By **Antonio Cipiciani,* Sergio Clementi,* Gianfranco Giulietti, Gianlorenzo Marino,** and **Gianfranco Savelli,** Dipartimento di Chimica, Università di Perugia, Italy
Paolo Linda, Istituto Chimico, Università di Trieste, Italy

The mechanism of the reaction between substituted indoles and trifluoroacetic anhydride has been studied by product identification and n.m.r. spectroscopy. Experimental evidence is given for the stability of an ion pair intermediate.

THE chemistry of indoles has received much attention, mostly because of the biological interest of some derivatives, and there is a large literature on this subject. Although electrophilic substitution does not seem to take place in biological systems, its study affords valuable information on the charge distribution in the aromatic nucleus, the importance of steric requirements, and the relative stability of non-aromatic derivatives involved along the mechanistic pathway of reacting indoles.

In this connection, and in view of our interest in hetero-aromatic substitutions,[2] we have previously investigated quantitative aspects of the Vilsmeier–Haack acylation of alkylindoles,[3,4] as well as the reaction products formed in the trifluoroacetylation of indole itself [5] and its 2,3-dimethyl derivative.[6] The main aims of these studies were the determination of the reactivities of individual ring positions towards electrophiles and the elucidation of the mechanism of substitution in some substituted indoles.

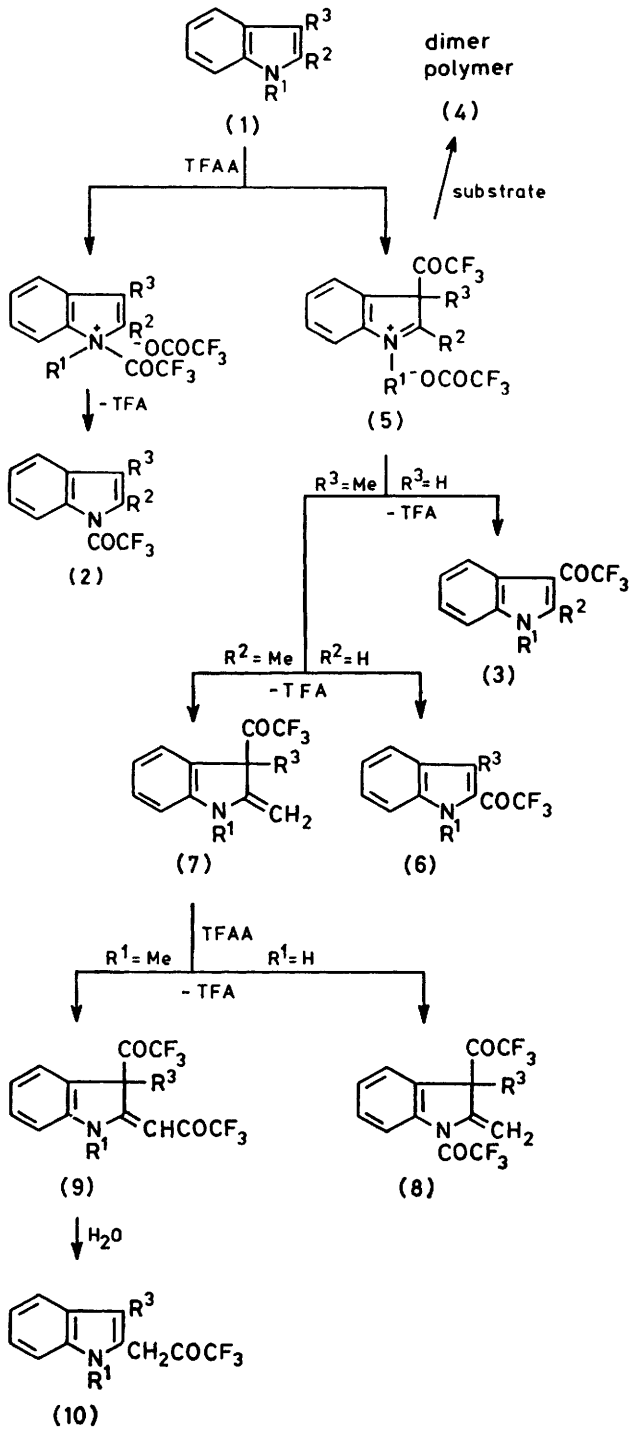
Our previous work,[3-6] suggesting that initial attack occurs predominantly at the 3- or *N*-position depending upon the substrate structure and the experimental conditions, and the valuable series of papers by Jackson and his co-workers,[7-9] pointing out the widespread occurrence of preliminary *ipso*-attack at the 3-position and subsequent rearrangement of the electrophile to the 2-position, led us to clarify the relative importance of *ipso*-attack at C-3, direct *N*-substitution, and rearrangement. The reaction investigated was trifluoroacetylation by trifluoroacetic anhydride (TFAA), whose mechanism was elucidated by kinetic study,[10] showing that the electrophile is the ion pair between the protonated anhydride and the anion of trifluoroacetic acid (TFA).

Following our preliminary account [6] we report here the systematic study of indoles substituted by alkyl groups in the various heterocyclic ring positions, carried out by product identification in 1,2-dichloroethane and by n.m.r. spectrometry in CCl_4.

RESULTS AND DISCUSSION

Reactions in 1,2-dichloroethane (DCE) were carried out at 0 °C by adding the substrate to a slight excess of TFAA in solution, so that the reactants were present in equimolar amounts (*ca.* 0.5M). This procedure was needed in order to suppress as much as possible the formation of dimeric and/or polymeric products, which are favoured by a high local concentration of aromatic compound.[5]

The reaction mixtures were analysed by g.l.c. both before and after the usual work-up with $NaHCO_3$,



SCHEME

because in several cases the product distribution is changed dramatically by the intervention of water. The individual products were isolated by column chromatography or preparative g.l.c., and their structures usually assigned on the basis of n.m.r., i.r., and mass spectra. The results for each substrate are listed in Table 1. The relative amount of the products is given as a percentage yield, as quite often a significant amount of unchanged material is still present according to g.l.c. analysis.

Further information was obtained by carrying out the reactions in an n.m.r. tube placed in the probe of the spectrometer at 0 °C and recording the ^{1}H n.m.r. spectra of the reacting mixtures, so that the formation and disappearance of intermediates could be followed. The solvent used in this case was CCl_4, but appropriate experiments showed that no appreciable variation was

TABLE 1

Products formed in the trifluoroacetylation of indoles in DCE at 0 °C

Substrate	R^1	R^2	R^3	Products (yield %)
(1A)	H	H	H	(2A) (45), (3A) (20), (4A) (30)
(1B)	CH_3	H	H	(3B) (>97) [a]
(1C)	H	CH_3	H	(3C) (80), (2C) (15)
(1D)	H	H	CH_3	(2D) (>97), (4D) (trace)
(1E)	CH_3	CH_3	H	(3E) (> 97)
(1F)	CH_3	H	CH_3	(6F) (50), (4F) (trace)
(1G)	H	CH_3	CH_3	(2G) (25), (8G) (10) [b]
(1H)	H	$[CH_2]_4$		(2H) (>80)
(1J)	CH_3	CH_3	CH_3	(9J) (*Z*, *E*, 25) ⟶ (10J)

[a] Ref. 5. [b] Ref. 6.

TABLE 2

^{1}H N.m.r. spectra of indoles and their products formed on addition of TFAA in CCl_4: (I) starting substrate at 0 °C; (II) immediately after mixing at 0 °C; (III) after some time at room temperature

	NR^1	2-R^2	3-R^3	ArH	Others
(i) 1,2-Dimethylindole					
(I) (1E)	3.21	2.19	5.98	6.75—7.00 (3 H)	7.25 (1 H, m, 7-H)
(II) (3E)	3.65	2.68		7.10—7.40 (3 H)	7.95 (1 H, m, 7-H)
(III) (3E)	3.65	2.70		7.10—7.40 (3 H)	7.95 (1 H, m, 7-H)
(ii) 2-Methylindole					
(I) (1C)	6.40br (s)	2.12	5.98	6.75—7.10 (3 H)	7.35 (1 H, m, 7-H)
(II) (3C), (2C)		2.62		7.00—7.40 (4 H)	7.88 (1 H, m, 7-H)
(III) (3C), (2C)		2.65		7.05—7.40 (4 H)	7.90 (1 H, m, 7-H)
(iii) 1,3-Dimethylindole					
(I) (1F)	3.28	6.35	2.21	6.80—7.05 (3 H)	7.35 (1 H, m, 7-H)
(II) (5F)	3.65	6.58	2.21	6.90—7.10 (3 H)	7.30 (1 H, m, 7-H)
(III) (5F) (80%)	3.65	6.60	2.25	6.90—7.10 (3 H)	7.30 (1 H, m, 7-H)
(6F) (20%)	3.90		2.60	6.90—7.25 (3 H)	7.58 (1 H, m, 7-H)
(iv) 3-Methylindole					
(I) (1D)		6.52	2.25	6.60—7.15 (4 H)	7.25 (1 H, m, 7-H)
(II) (2D)		*ca.* 7.30	2.25	6.80—7.75 (3 H)	8.20 (1 H, m, 7-H)
(III) (2D)		*ca.* 7.30	2.25 (d)	7.00—7.40 (3 H)	8.20 (1 H, m, 7-H)
(4D) (trace)			1.45, 1.70		
(v) 2,3-Dimethylindole					
(I) (1G)		2.32	2.17	6.55—7.05 (4 H)	7.20 (1 H, m, 7-H)
(II) (2G) (25%)		2.46	2.17	7.05—7.70 (3 H)	7.75 (1 H, m, 7-H)
(5G) (75%)		2.50	1.68	7.05—7.60 (5 H)	
(III) (2G) (25%)		2.42	2.12	6.95—7.60 (3 H)	7.70 (1 H, m, 7-H)
(5G) (50%)		2.65	1.76	6.95—7.60 (5 H)	
(8G) (20%)		4.85 (d), 5.75 (d)	1.70	6.95—7.45 (3 H)	7.78 (1 H, m, 7-H)
(9G) (5%)		5.54	1.73	6.95—7.80 (3 H)	
(vi) 1,2,3-Trimethylindole					
(I) (1J)	2.95	2.10	1.92	6.70—7.00 (3 H)	7.20 (1 H, m, 7-H)
(II) (5J) (35%)	3.55	2.27	1.68	6.70—7.50 (4 H)	
(9J)-(*E*) (40%)	3.32	5.63 (1 H)	1.68	6.70—7.50 (4 H)	
(9J)-(*Z*) (20%)	3.55	*ca.* 7.20 (1 H)	2.20	6.70—7.50 (4 H)	
(III) (5J) (25%)	3.53	2.25	1.70	6.70—7.50 (4 H)	
(9J)-(*E*) (45%)	3.34	5.67 (1 H)	1.70	6.70—7.50 (4 H)	
(9J)-(*Z*) (20%)	3.57	*ca.* 7.20 (1 H)	2.22	6.70—7.50 (4 H)	
(vii) 1,2,3,4-Tetrahydrocarbazole					
(I) (1H)		2.57 (4 H, m, 1- and 4-H$_2$)		6.75—7.00 (4 H)	1.85 (4 H, m, 2- and 3-H$_2$) 7.20 (1 H, m, 8-H)
(II) (2H)	2.62 (2 H, m, 1-H$_2$)	2.91 (2 H, m, 4-H$_2$)		7.05—7.40 (3 H)	1.87 (4 H, m, 2- and 3-H$_2$) 7.80 (1 H, m, 8-H)
(III) (2H)	2.62	2.92		7.02—7.35 (3 H)	1.90 (4 H, m, 2- and 3-H$_2$) 7.82 (1 H, m, 8-H)

induced in the product distribution by the solvent change.

The ^{1}H n.m.r. spectra of substrates, intermediates, and products are listed in Table 2 and are discussed below. In general it is noteworthy that, in spite of the quite different products given by each substrate, a unique mechanism, illustrated in the Scheme, can be used to explain the preferred pathways in each case.

Indole (1A).—The behaviour of the parent substrate was previously described,[5] but the isomer distribution given in Table 1 is slightly different from that reported, since it depends markedly upon the absolute concentrations of reactants.[5]

The results show the close competition between *N*- and *C*-substitution, as also observed in chlorination,[11] the former prevailing in agreement with kinetic data on acid-catalysed hydrogen exchange in aqueous acetonitrile.[12]

The cation formed by the first TFAA attack has still enough electrophilic power to react further with the excess of indole, if present, giving a dimeric compound *N*-trifluoroacetyl-2-(3-indolyl)indoline (4A). The n.m.r. investigation of this reaction was not attempted since the identification of intermediates in the absence of alkyl substituents is not easy.

N-*Methylindole* (1B) *and* 1,2-*Dimethylindole* (1E).—These substrates, with a substituent linked to the nitrogen atom and an unsubstituted C-3, cannot exhibit any competition. Accordingly they give only 3-substitution in almost quantitative yield, leading to (3B and E), respectively, as indicated by their ^{1}H n.m.r. (disappearance of the 3-H peak and downfield shifts of all the other signals) and i.r. spectra (ν_{CO} *ca.* 1 655 cm^{-1}, typical of an aromatic ketone).

The reaction of (1B) has already been reported and was not investigated further. The ^{1}H n.m.r. study of (1E) shows (Table 2) that the spectrum of the final product appears immediately after mixing and does not change further. No evidence of the intermediate ion pair (5E) can therefore be presented. However, this is not surprising in view of the high reactivity of this substrate, greatly enhanced by the 2-methyl group, and loss of TFA from (5E) is clearly quite rapid at 0 °C.

2-*Methylindole* (1C).—This substrate is suitable for the examination of competition between *N*- and *C*-substitution: attack at C-3 should be favoured because of activation by the 2-methyl group, which has a larger effect at C-3 (operating by field and resonance effects) than at nitrogen (field effect only). Accordingly, the ^{1}H n.m.r. study indicates the predominant formation of the 3-substituted product (3C) by the same immediate change of the spectrum of the reacting mixture just described for the previous substrate (Table 2).

However, the spectrum is not clear enough to exclude the presence of some *N*-substituted product, and in fact the g.l.c. analysis shows (Table 1) that *ca.* 15% of (2C) is formed, a quantity higher than that found in the Vilsmeier–Haack acetylation (1.6%),[4] and in formyation.[3] This result is in keeping with the prevailing *N*-substitution found for indole itself in this reaction,[5] while only very small or negligible amounts of *N*-acyl derivatives were found in Vilsmeier–Haack acylations.[3,4] It is possible that the peculiar behaviour exhibited by trifluoroacetylation is related to the electronic demand of CF_3 in the electrophilic reagent, which facilitates a complex interaction with the nitrogen atom.[5]

The two products (2C) and (3C) were isolated by preparative g.l.c. and identified on spectral basis. Details are given in the Experimental section, but a simple though significant difference found for all substrates is to be stressed and this is in the carbonyl stretching frequencies: in this case ν_{CO} is 1 620 cm^{-1} for 2-methyl-3-trifluoroacetyl- and 1 720 cm^{-1} for 2-methyl-*N*-trifluoroacetylindole. On the other hand, the n.m.r. spectra of the two isolated products are quite similar. In fact the peaks of the 2-methyl group in (2C) and (3C) have almost identical chemical shifts and this explains why we could not obtain direct n.m.r. evidence of the 15% (2C) formed in the reaction mixture.

1,3-*Dimethylindole* (1F).—Compounds (1F and D) are particularly suitable for clarifying the importance of *ipso*-attack and subsequent rearrangements in the acylation reactions, but the results for the former are easier to interpret and are discussed first.

The widespread occurrence of *ipso*-attack, *i.e.* attack of the electrophile at the carbon atom bearing an alkyl substituent, is increasingly recognized in halogenation [13] and nitration.[14] Despite steric inhibition, this mechanism is favoured by the electron-donating power of the substituent, especially when the non-aromatic structure so formed is highly stabilized by resonance effects. The preliminary attack is usually followed by rearrangement to a side-chain position, giving what is called non-conventional substitution, as in polyalkylbenzenes and heteroaromatic analogues,[15] or to another nuclear position, as previously seen both in other heteroaromatic compounds,[16] and in the indole nucleus by Jackson and his co-workers, who claim that alkylation and acetylation of several 3-substituted indoles can lead to 2-substitution only *via* this pathway.[7,8]

The g.l.c. analysis of the reaction mixture in DCE shows (Table 1) that the only significant product, formed in *ca.* 50% yield, is the expected 1,3-dimethyl-2-trifluoroacetylindole (6F), besides the unchanged substrate and traces of a dimeric derivative. However, this result is not sufficient to clarify whether the product is obtained by direct attack at C-2,[4,17] or *via* the mechanism proposed by Jackson,[7] or even by rearrangement from a preliminary attack on nitrogen,[18] but the ^{1}H n.m.r. study can explain this matter.

In fact (Table 2) the substrate disappears completely immediately after mixing, as indicated by the absence of the *N*-methyl peak at δ 3.28, and a single compound is formed, whose NMe and 2-H signals are slightly shifted downfield. This structure is stable for several minutes at 0 °C. Only after that time or on increasing the probe temperature do two new singlets appear at δ 3.90 and 2.60, easily identified as the methyl groups of the 2-trifluoroacetyl derivative by comparison with the spectrum of the

isolated specimen of (6F), identified by the absence of the 2-H signal in the ^1H n.m.r. spectrum, and the carbonyl frequency (ν_{CO} 1 670 cm^{-1}). At the same time the reaction mixture darkens in the probe whilst other singlets begin to increase in the region of the aliphatic methyl peaks and the signals of the intermediate decrease (integration of 2-H over the homocyclic protons).

The mechanism shown by spectral and g.l.c. data is consistent with that depicted in the Scheme. In the presence of a substituent at the nitrogen atom attack is totally directed on C-3, in spite of the substituent it bears, to form the stable ion pair (5F) immediately. Its spectrum (Table 2) is compatible with that of the similar structure (5G) illustrated previously,[6] except that the 3-methyl peak is not shifted upfield with respect to (1F), but in this case we can easily assume that the shielding effect due to the loss of aromaticity is completely balanced by the deshielding effect due to the presence of the $COCF_3$ group, and we have previously seen that this balance does not always give the same result.[19]

The ion pair (5F) slowly rearranges to the final product (6F), although it remains the main species present in solution, while a small amount of a dimeric structure (4F) is formed owing to the electrophilic character of (5F) (see indole itself and other substrates with $R^2 = H$). No trace of direct attack on C-2 can be observed (the 3-methyl peak should have been shifted downfield), thus confirming that Jackson's hypothesis is correct. In addition, a rearrangement from N-1 to C-2, demonstrated for 3-alkyl-1-allylindoles,[18] can also be excluded. In all cases where the preliminary attack occurs on nitrogen the ^1H n.m.r. spectra show a change in the homocyclic proton region (Table 2) which is not observed here.

The g.l.c. results are also in agreement with this picture, as (5F) can readily revert to (1F) during work-up,[6] and this explains the large amount of 'unchanged' substrate still present. Furthermore, although the isolation of the dimeric or polymeric product (4F) was not attempted, the yield of (6F) steeply decreases in the presence of excess of substrate, that suggests that the ion pair does not take part in the slower rearrangement reaction.

On this basis it seems likely that the formation of the 2-substituted product obtained in the Vilsmeier–Haack formylation [3,20,21] and acetylation [4] of (1F) also occurs by the same mechanism.

3-Methylindole (1D).—The product analysis of the reaction of (1D) with TFAA in DCE shows (Table 1) almost exclusive formation of the *N*-trifluoroacetyl derivative (2D) (ν_{CO} 1 715 cm^{-1}), besides very small amounts of side products with a tendency to give polymer (4D), as already observed both with (1D and F).[20] The ^1H n.m.r. spectrum at the point of mixing in CCl_4 is in keeping with these observations, owing to the downfield shift of all peaks, especially those of 7- and 2-H, as occurs for all substrates substituted directly at the nitrogen atom. The spectrum does not change further: evidently, proton loss from the ion pair leading to (2) is as fast as it is from the ion pair (5) when $R^3 = H$, whereas when R^3 = alkyl structure (5) has a much longer life.

Nevertheless the spectrum of (2D) in the reaction solution is somehow peculiar, as the integration of the peaks indicates the presence of six protons and the 3-methyl group signal is a doublet. Clearly, possibly because of the higher basicity of this compound, (2D) is protonated, and protonation occurs at C-3, as demonstrated by the doublet observed for the conjugate acid of (1D).[22] This is the only case where the proton lost by the nitrogen is observed to be immediately taken up by C-3.

No evidence is found for the formation of (6D), *via* rearrangement from (5D), although *ca.* 5% of 2-substitution had been observed in Vilsmeier formylation [3] and acetylation.[4] Again this result should arise from the preference for TFAA to interact with the nitrogen atom, so that very little electrophile is directed to C-3. The predominance of *N*-substitution was already observed in Vilsmeier formylation under preparative conditions,[23] and the exclusive formation of the 2-substituted derivative found by other authors [20] can probably be attributed to the hydrolysis of the N-CHO derivative during the alkaline treatment of the reaction mixture.

2,3-Dimethylindole (1G) *and* *1,2,3,4-Tetrahydrocarbazole* (1H).—The results for (1G) were reported in the preliminary account [6] and can be briefly summarized as follows. Upon mixing, the substrate is completely consumed and a 1 : 3 mixture of (2G) and (5G) is immediately formed. The indoleninium ion pair (5G) is fairly stable for some minutes at 0 °C, then it slowly loses a proton to give the methyleneindoline (7G), which rapidly undergoes further substitution leading to (8G), whose appearance is observed in the ^1H n.m.r. spectrum.

On the other hand, g.l.c.–m.s. product analysis shows a large amount of unchanged substrate, with (2G) in 25% and (8G) in 10% yield. However, the reaction mixture before work-up contains a lower amount of substrate and *ca.* 20% of a more volatile monosubstituted product which could not be isolated. Since (2G) does not react with TFAA to give (8G) we concluded that this product was the enamine (7G), which ought to be very reactive, as its n.m.r. signals could not be identified. In spite of the apparent difference, the g.l.c. results are in agreement with the n.m.r. data, provided, as seems likely, that (5G) readily reverts to (1G) in the presence of water and loses a proton under the g.l.c. conditions to give (7G).

Besides the formation from (5G) of the two products (2G) and (8G), the n.m.r. spectrum of the reaction mixture in CCl_4 showed the formation of a smaller amount of a third product (5%), which was not isolated, but whose characteristic peaks are a singlet at δ 5.54 and a methyl signal very close to that of (8G). In view of the close similarity of this spectrum with that of the main product of the reaction with 1,2,3-trimethylindole (Table 2), we assumed structure (9G) for this product. Therefore it appears that the reaction of the second TFAA molecule at the enamine group of (7G) is just a bit slower than direct substitution at the nitrogen atom leading to (8G).

The behaviour of this substrate is particularly suitable for pointing out some specific mechanistic patterns of indole substitutions. First, it shows the close competition between direct *N*-attack and *ipso*-attack at C-3. Clearly, *ortho*-activation by the 2-methyl group on the 3-position overwhelms the attraction of the reactant for the nitrogen atom. Secondly, it indicates the importance of the presence of a 2-methyl group in the subsequent steps of reaction. The low acidity of the 2-methyl hydrogens allows the formation of the enamine (7G), the reaction of which with a second molecule of TFAA is so fast that the rearrangement of the electrophile from C-3 to the side-chain position (2-CH_2), observed in the systems of benzofuran and benzothiophen,[15,19] cannot occur.

When the two methyl groups linked to C-2 and -3 are substituted by a four-membered aliphatic ring as in (1H), the reaction takes a much simpler pathway. Both g.l.c. and n.m.r. data indicate that the only product immediately formed in high yield (80%) is the *N*-substituted derivative (2H), as again indicated by the amide carbonyl frequency (ν_{CO} 1 717 cm^{-1}) and by the large downfield shift of 7- and 1-H. Clearly, in this case steric hindrance due to the presence of the methylene chain, which cannot adjust its position as freely as a methyl group, depresses the rate of *ipso*-attack and the competing *N*-substitution is again favoured.

1,2,3-*Trimethylindole* (1J).—Trifluoroacetylation of this substrate has already been studied in benzene.[9] The main product was 1-(1,3-dimethylindol-2-yl)-3,3,3-trifluoropropanone (10J), in the presence of *ca.* 30% of the two geometrical isomers of the unstable 1,3-dimethyl-3-trifluoroacetyl-2-trifluoroacetylmethyleneindoline (9J). The formation of the latter is favoured by increasing the temperature and by an excess of TFAA. Our systematic investigation is in agreement with these results and the mechanistic hypothesis suggested therein, but it also gives experimental support to new relevant features of the reaction.

In fact, g.l.c.–m.s. analysis of the reaction mixture confirms (Table 1), besides a large amount of unchanged substrate, the presence of two main products, one of which (10J) (*m/e* 255) prevails over the other (9J) (*m/e* 351). However, analysis of the crude mixture before work-up shows that (9J) is the almost exclusive product and only a very small amount of (10J) is formed. The two products were isolated by preparative g.l.c. and it could be observed that (9J) is unstable and slowly changes to (10J), probably because of traces of moisture.

The ^1H n.m.r. spectra of the reacting mixture, although difficult to interpret, gives valuable information, and therefore the reaction was carried out in CCl_4, $CDCl_3$, and C_6D_6. Immediately after mixing the reactants in CCl_4 at 0 °C the absence of the peaks of (1J) indicates no starting material. Instead (Table 2), outside the aromatic region the spectrum of the mixture shows a singlet at δ 5.63, five more singlets in the methyl region (at δ 3.55, 3.32, 2.27, 2.20, and 1.68), and two small broad singlets at δ 3.97 and 2.67.

Comparison with an authentic specimen of (10J) permits the exclusion of this product in the organic medium. Therefore, in agreement with the g.l.c. results, it is concluded that the side-chain-substituted derivative (10J), which is eventually obtained as the main reaction product after work-up, is formed only after the intervention of water, by hydrolysis of (9J). Accordingly, the possible rearrangement of (7J) to (10J) does not occur in the organic solvent, as already observed in the reaction of (1G).

The spectrum in CCl_4 of the isolated specimen of the unseparated geometrical isomers *Z* and *E* (called *cis-s-cis* and *trans-s-cis* in ref. 9*b*) of (9J) indicates that it is responsible for the presence in the reaction mixture of the peaks at δ 5.63, 3.55, 3.32, 2.20, and 1.68. A careful analysis of the peak integrations of spectra relative to diverse experiments at various times shows that the singlet at δ 5.63 (methyne) is related to the methyl groups at δ 3.32 and 1.68, whereas the methyne signal related to the methyl groups at δ 3.55 and 2.20 is hidden behind the aromatic peaks (δ *ca.* 7.20).

This spectrum should be due to the *Z*-isomer (the higher priority groups lie on the same side of the double bond, so that the $COCF_3$ group is near to the NMe group), since the methyne proton is largely deshielded by the effect of both $COCF_3$ groups, and also the *N*-methyl signal is shifted downfield with respect to the other isomer, because of the nearness of the α-trifluoroacetyl group. The relative shift of the 3-methyl groups (unattributed in ref. 9*b*) is less obvious: it seems that when both the $COCF_3$ groups are close to 3-methyl (*E*-isomer) the molecular geometry requires that their inductive effects are balanced by their anisotropic effects.

The experimental *E* : *Z* ratio is always between 2 : 1 and 3 : 1 and does not change appreciably with time, although it varies slightly in each experiment, at variance with the ratio 8 : 1 reported in ref. 9*b*. As in this work no other evidence of the existence of the two isomers could be obtained (the i.r. spectrum shows only two carbonyl bands at 1 750 and 1 650 cm^{-1}) and any attempt at separating them by chromatographic methods failed. Only the g.l.c. analysis on a LAC 728 column indicated that the specimen of (9J) was not unique, but the separation was insufficient for preparative purposes.

In the ^1H n.m.r. spectrum of the reaction mixture in CCl_4 (Table 2), after the identification of the two isomers of (9J), one should expect that the major species present is the ion pair (5J), both by analogy with other substrates and to justify the presence of a large amount of unchanged substrate in the g.l.c. analysis. Since the signals left unattributed cannot support the structure (5J), we assumed that some peaks of the ion pair could have the same chemical shift as those of (9J). In order to clarify this point we studied the solvent effect on chemical shifts by also carrying out the reaction in [$^2H_6$]benzene and in $CDCl_3$ at room temperature. The results, in terms of products and mechanism, are almost the same as in CCl_4, although the reaction is much slower in benzene and slightly faster in chloroform. Therefore, even in these solvents, and without excess of TFAA,[9]

(10J) is not formed in the organic medium, but only after work-up by hydrolysis of (9J).

In fact the spectrum of the reacting mixture in $CDCl_3$ shows the expected seven methyl singlets and the signals belonging to each structure are attributed by comparison with an authentic specimen and by peak integration. (*E*)-(9J) is responsible for the singlets at δ 5.84 (CH), 3.51 (N-Me), and 1.78 (3-Me), while (*Z*)-(9J) generates signals at δ 3.67 (N-Me) and 2.27 (3-Me). The peaks of the indoleninium ion pair (5J) resonate at δ 4.19 (N-Me), 2.89 (2-Me), and 1.94 (3-Me), the relative amounts of the three components being similar to those in CCl_4.

The spectrum in C_6D_6 shows, besides the methyne singlet at δ 5.67, nine typical methyl peaks. Again from comparison with authentic specimens and peak integrations we can elucidate the structure of (*E*)-(9J) (δ 5.67, 2.71, 1.60), (*Z*)-(9J) (δ 3.16, 2.05), and (5J) (δ 3.63, 2.35, 1.55 as being present in the mixture).

Accordingly, the methyl signals of (5J) in CCl_4 should be at δ 3.55, 2.27, and 1.68, in agreement with the results for (5G), *i.e.* the 3-methyl peak is shifted slightly upfield with respect to the substrate, whereas the 2- and *N*-methyl peaks are shifted downfield. This attribution of the peaks was also confirmed on adding benzene (or chloroform) to the CCl_4 mixture, and considering the specific shielding (or deshielding) effect induced by the added solvent.

Two more signals in C_6D_6 at δ 3.05 and 2.15 belong to a product present in much lower amount, and can be related to the two broad singlets present in CCl_4 at δ 3.97 and 2.67. Since in these media the reaction mixture darkens much more than in $CDCl_3$, we might tentatively attribute these peaks to the remaining methyl groups, present in some polymeric material.

The whole mechanistic framework of trifluoroacetylation of (1J) is in keeping with the general Scheme. The preliminary *ipso*-attack of TFAA at C-3 leads to the stable ion pair (5J), which loses TFA to yield (7J). Further attack of TFAA gives (9J) (*Z*- and *E*-isomers), which yield (10J) on contact with water.

Our results are in good agreement with those reported in ref. 9, except that we demonstrated that (10J) is formed only after work-up. It seems likely that the reaction of other substrates, and especially *N*-methyl-1,2,3,4-tetrahydrocarbazole,[9] proceeds in the same way. Moreover, we could clarify the spectra of the two isomers of (9J), as the 3-methyl peaks were unattributed and the methyne proton of the *Z*-isomer was not quoted at all in the cited paper.

The results on the trifluoroacetylation of (1J) seem to be in contrast with the previous studies on the Vilsmeier–Haack formylation [3] and acetylation,[4] where the products were reported to be the 6-acyl derivatives. These two reactions were therefore re-examined and the results published in a preliminary account.[21] It was shown that, whereas the acetylation reaction gives the product substituted at the homocyclic ring, the formyl group substitutes a proton of the 2-methyl group giving a structure similar to (10J). Therefore it seems reasonable that the mechanism illustrated in the Scheme has general validity for electrophilic substitution of 1,2,3-trimethylindole, provided that the steric requirements of the electrophile does not render *ipso*-attack much slower, thus making 6-substitution predominant. Work is in progress to obtain further evidence for other electrophilic reactions.

Conclusions.—This work indicates that, although different products are given by each substrate, the mechanism of trifluoroacetylation for all the indoles studied can be described by a single mechanism.

Attack at C-3 and direct *N*-substitution compete closely. The latter prevails when R^1 = H, whereas the fate of the intermediate indoleninium ion pair depends on the presence of alkyl groups at C-2 and -3. When R^3 = H almost exclusive 3-substitution is observed, while if R^3 = alkyl and R^2 = H the rearrangement of the electrophile from C-3 to -2 becomes important. When both R^2 and R^3 = alkyl a reactive methyleneindoline is formed, which rapidly undergoes further substitution either at nitrogen (R^1 = H) or at the enamine group to give the Fisher base. A side-chain-monosubstituted aromatic product is obtained from this base only after the usual work-up, so that the non-conventional substitution can be ruled out.

It appears possible that the same mechanism can also apply to other electrophilic substitutions of indoles. In particular a close similarity exists between the trifluoroacetylation and Vilsmeier–Haack reactions. However, somewhat different results obtained for the latter point to the importance of the steric requirements of the electrophile and perhaps also the relevance of the acidity of the organic medium.

Quite an important feature of this study is the first spectral evidence for preliminary *ipso*-attack at C-3 and subsequent rearrangement of the $COCF_3$ group to C-2 in the reaction of 1,3-dimethylindole, thus confirming Jackson's mechanistic hypothesis.[7-9] Accordingly, those parts of the discussion on partial rate factors (derived from rate constants and product distribution) in the acetylation of indoles reported in our previous work [4] involving the reactivity at C-2 are probably in error, since the rate-determining step might also be the rearrangement reaction in that case. Nevertheless the rest of that work does not appear to be affected by the results of this study and is, therefore, to be considered still valid.

EXPERIMENTAL

Materials.—Compounds (1A—E, G, and H) were commercial samples purified by distillation or crystallization, while (1F) [24] and (1J) [25] were prepared according to literature methods. The ^1H n.m.r. spectra of all substrates in CCl_4 at room temperature are recorded in Table 3. The solvents were purified by standard procedures. Commercial TFAA was used without further purification.

Reaction Conditions.—A solution of the substrate was added dropwise at 0 °C (ice-bath) to a stirred solution con-

TABLE 3

^{1}H N.m.r. spectra of the substrates in CCl_4 at room temperature (δ values)

Compound	NR[1]	2-R[2]	3-R[3]	ArH	7-H
(1C)	6.41br (s)	1.89	5.92	6.50—7.10	7.30
(1D)		6.58	2.26	6.85—7.10	7.35
(1E)	3.10	2.08	5.95	6.70—7.05	7.26
(1F)	3.11	6.22	2.19	6.75—7.10	7.58
(1G)		2.05	2.02	6.52—7.05	7.20
(1H)		2.58 [a]	1.83 [b]	6.75—7.40	
(1J)	3.20	2.12	2.05	6.75—7.02	7.25

[a] 4 H, m, 1- and 4-H_2. [b] 4 H, m, 2- and 3-H_2.

final concentration of both reactants was *ca.* 0.5M in 0.5 ml, was put into an n.m.r. tube and placed in the probe of the instrument at 0 °C. An equimolar amount of substrate in CCl_4 at the same temperature was added to the cold tube and the spectrum taken several times, at intervals, at 0 °C. The temperature of the probe was then raised to room temperature and the spectrum again taken immediately and after some time. Chemical shifts are accurate to within 0.3 p.p.m.

Thanks are due to the Italian C.N.R. for financial support, to Dr. G. V. Sebastiani for helpful discussions, and to Messrs. C. Bastianelli and G. Gambini for recording the n.m.r. spectra.

[1/894 *Received*, 3rd *June*, 1981]

TABLE 4

Spectral data of the trifluoroacetylated products of indoles

Compound	N.m.r. chemical shifts (δ) [a] NR[1]	2-R[2]	3-R[3]	ArH	7-H	*m/e* (70 eV)	ν_{CO}/cm^{-1}
(2C) [b]		2.68	6.42	7.10—7.45	7.90	227	1 730
(3C) [c]		2.75		7.10—7.50	7.95	227	1 620
(2D)			2.25	7.00—7.40	8.23	227	1 715
(3E)	3.65	2.68		7.10—7.40	7.95	227	1 655
(6F)	3.90		2.60	6.90—7.25	7.58	241	1 670
(2G) [d]		2.42	2.12	6.95—7.35	7.71	241	1 715
(8G) [d]		4.85, 5.75 [e]	1.70	6.95—7.45	7.78	337	1 700, 1 750
(2H)		2.92, 2.62, 1.90 [f]		7.02—7.30	7.82 [g]	267	1 717
(9J)-(*E*) [b]	3.42	5.75 [h]	1.75	6.85—7.65	*i*	351	1 650, 1 750
(9J)-(*Z*) [b]	3.65	*ca.* 7.20 [h]	2.25	6.85—7.65	*i*	351	1 650, 1 750
(10J) [b]	3.52	4.12 [j]	2.25	6.90—7.25	7.40	255	1 770

[a] CCl_4 unless otherwise stated. [b] $CDCl_3$. [c] [$^2H_6$]Acetone. [d] Ref. 6. [e] Doublets, 1 H each, 2-H_2. [f] 2.92 (2 H, m, 1-H_2), 2.62 (2 H, m, 4-H_2), 1.90 (4 H, m, 2- and 3-H_2). [g] 8-H. [h] 1 H, s, CH. [i] Within ArH. [j] 2 H, s, CH_2.

taining a slight excess of TFAA, in order to allow for its high volatility and ready hydrolysis, so that the final concentration of the reactants was *ca.* 0.5M. After completing the addition the mixture was allowed to stand for *ca.* 0.5 h at room temperature. A portion of the mixture was then analysed as such, whereas the rest was worked up as usual with $NaHCO_3$. It should be pointed out that the reaction conditions were kept the same for all the substrates, to make possible the comparison of the results, and therefore no attempt to optimize the yield of trifluoroacetylated products was made.

Product Analysis.—The concentrated residues of the organic layers of each reaction were analysed by g.l.c. on a C. Erba G.I. fractometer (f.i.d.; stainless steel column 2 m × 4 mm packed with 10% silicone SE-30; temperature 130—180 °C). Benzophenone was used as internal standard. It is noteworthy that the introduction of a $COCF_3$ group linked to an aromatic carbon decreases the volatility over the substrate, whereas the *N*-substituted derivatives and, even more, the indoline structures exhibit a much higher volatility.

G.l.c.–m.s. analyses were made on a Varian 311A instrument operating at 70 eV.

I.r. spectra were recorded on a Perkin–Elmer model 257 spectrophotomer.

Trifluoroacetylated Products.—The products formed from (1A and B) were described in ref. 5. All the other trifluoroacetylated products, synthesised here for the first time, were prepared under the conditions described above and isolated, usually, by preparative g.l.c. Their characteristic spectra data used for identification are listed in Table 4.

N.m.r. Studies.—N.m.r. spectra were recorded on a JEOL C-60 HL spectrometer with tetramethylsilane as internal standard. A solution of TFAA in CCl_4, so that the

REFERENCES

[1] Presented at the Second European Symposium on Organic Chemistry (ESOC II), Stresa, 1981.
[2] G. Marino, *Adv. Heterocycl. Chem.*, 1971, **13**, 235, and references quoted therein.
[3] S. Clementi, P. Linda, and G. Marino, *J. Chem. Soc., Chem. Commun.*, 1972, 427.
[4] A. Cipiciani, S. Clementi, P. Linda, G. Marino, and G. Savelli, *J. Chem. Soc., Perkin Trans. 2*, 1977, 1284.
[5] A. Cipiciani, S. Clementi, P. Linda, G. Savelli, and G. V. Sebastiani, *Tetrahedron*, 1976, **32**, 2595.
[6] A. Cipiciani, S. Clementi, G. Marino, G. Savelli, and P. Linda, *J. Chem. Soc., Chem. Commun.*, 1980, 784.
[7] (*a*) A. H. Jackson, B. Naidoo, and P. Smith, *Tetrahedron*, 1968, **24**, 6119; (*b*) A. H. Jackson and B. Naidoo, *ibid.*, 1969, **25**, 4843; (*c*) A. H. Jackson and B. Naidoo, *J. Chem. Soc., Perkin Trans. 2*, 1973, 548; (*d*) R. Iyer, A. H. Jackson, and P. V. R. Shannon, *ibid.*, 1973, 878, and previous papers of the series.
[8] A. H. Jackson, B. Naidoo, A. E. Smith, A. S. Bailey, and M. H. Vandrevala, *J. Chem. Soc., Chem. Commun.*, 1978, 779.
[9] (*a*) A. S. Bailey, J. M. Peach, and M. H. Vandrevala, *J. Chem. Soc., Chem. Commun.*, 1978, 845; (*b*) A. S. Bailey, J. B. Haxby, A. N. Hilton, J. M. Peach, and M. H. Vandrevala, *J. Chem. Soc., Perkin Trans. 1*, 1981, 382.
[10] S. Alunni and S. Clementi, *J. Chem. Soc., Perkin Trans. 2*, 1972, 1521.
[11] M. De Rosa and J. L. T. Alonso, *J. Org. Chem.*, 1978, **43**, 2639.
[12] D. M. Muir and M. C. Whiting, *J. Chem. Soc., Perkin Trans. 2*, 1976, 388.
[13] P. B. D. de la Mare, 'Electrophilic Halogenation,' Cambridge University Press, Cambridge, 1976.
[14] R. B. Moodie, and K. Schofield, *Acc. Chem. Res.*, 1976, **9**, 287.

[15] (*a*) R. S. Hartshorn, *Chem. Soc. Rev.*, 1974, **3**, 167; (*b*) E. Baciocchi, S. Clementi, and G. V. Sebastiani, *J. Chem. Soc., Perkin Trans. 2*, 1974, 1882, and previous papers of the series.
[16] C. Bastianelli, S. Clementi, and G. V. Sebastiani, *Chim. Ind. (Milan)*, 1980, **62**, 262.
[17] G. Casnati, A. Dossena, and A. Pochini, *Tetrahedron Lett.*, 1972, 5277.
[18] G. Casnati, R. Marchelli, and A. Pochini, *J. Chem. Soc., Perkin Trans. 1*, 1974, 754.
[19] (*a*) E. Baciocchi, S. Clementi, and G. V. Sebastiani, *J. Chem. Soc., Perkin Trans. 2*, 1976, 266; (*b*) E. Baciocchi, S. Clementi, and G. V. Sebastiani, *J. Heterocycl. Chem.*, 1977, **14**, 359.
[20] H. von Dobeneck, D. Wolkenstein, and G. Blankestein, *Chem. Ber.*, 1969, **102**, 1347.
[21] C. Bastianelli, A. Cipiciani, S. Clementi, and G. Giulietti, *J. Heterocycl. Chem.*, 1981, **18**, 1275.
[22] R. L. Hinman and J. Lang, *J. Am. Chem. Soc.*, 1964, **86**, 3796.
[23] A. Chatterjee and K. M. Biswas, *J. Org. Chem.*, 1973, **38**, 4002.
[24] M. Julia and J. Lenzi, *Bull. Soc. Chim. Fr.*, 1962, 2263.
[25] J. T. Fitzpatrick and R. D. Hiser, *J. Org. Chem.*, 1957, **22**, 1703.