Kinetics and Mechanism of *N*-Substitution of Indoles and Carbazoles in Vilsmeier–Haack Acetylation

By **Antonio Cipiciani, Sergio Clementi, Paolo Linda,* Gianlorenzo Marino,** and **Gianfranco Savelli,**
Istituto di Chimica Organica, Università di Perugia, Via Elce di Sotto 10, 06100 Perugia, Italy

The rate constants of acetylation of several pyrroles, indoles, and carbazoles by the Vilsmeier–Haack reagent (*NN*-dimethylacetamide–carbonyl chloride) have been measured in 1,2-dichloroethane at 25 °C. Most substrates undergo *N*-substitution and the data strongly support a mechanism involving rate determining direct attack on the nitrogen atom. The order of susceptibility to the electrophile for positions 1—3 of the indole nucleus is C-3 > N-1 > C-2. Differences in behaviour towards *N*-substitution of pyrrole, indole, and carbazole are discussed in terms of loss of resonance energy in the transition states.

THE view that the unsubstituted indole nucleus is susceptible to electrophilic substitution, in mildly acidic or neutral media at C-3, although true in many cases,[1] has some noteworthy exceptions. Among these are the postulated *N*-nitrosation followed by rearrangement to a 3-nitroso-product;[2] the rapid electrophilic hydrogen exchange at the nitrogen atom;[3] the Mannich reaction leading to an *N*-dimethylaminomethyl derivative, readily isomerised to the C-3 substituted derivative;[4] and the *N*-trifluoroacetylation recently reported.[5] Even more striking is the tendency to substitution at N-1 rather than C-2 of 3-substituted indoles exhibited in Mannich[6] and Vilsmeier–Haack[7] reactions, in contrast with C-2 substitution under typical Friedel–Crafts conditions.[8] Similarly 2,3-dimethylindoles undergo ready *N*-substitution in the Mannich[6a,b] and Vilsmeier–Haack reactions,[7b,9] whereas C-6 substitution prevails in Friedel–Crafts reactions.[9,10] Similar behaviour is shown by carbazole, which undergoes *N*-formylation,[11a] and *N*-benzoylation[11b] when subjected to the Vilsmeier–Haack procedure, but C-3 substitution under Friedel–Crafts conditions.[12]

The mechanism of electrophilic substitution in 3-substituted indoles has received much attention, and the

[1] R. J. Sundberg, 'The Chemistry of Indoles,' Academic Press, New York and London, 1970, p. 1 ff.
[2] B. C. Challis and A. J. Lawson, *J.C.S. Perkin II*, 1973, 918.
[3] (*a*) M. Koizumni and T. Titani, *Bull. Chem. Soc. Japan*, 1939, **14**, 491 and earlier papers: (*b*) B. C. Challis and E. M. Millar, *J.C.S. Perkin II*, 1972, 1111, 1116, 1618. 1625; (*c*) D. M. Muir and M. C. Whiting, *ibid.*, 1975. 1316; 1976, 388.
[4] S. Swaminathan and K. Narasimhan, *Chem. Ber.*, 1966, **99**, 889.
[5] A. Cipiciani, S. Clementi, P. Linda. G. Savelli, and G. V. Sebastiani, *Tetrahedron*, 1976. **32**, 2595.
[6] (*a*) S. Swaminathan and S. Ranganathan, *J. Org. Chem.*, 1957, **22**. 70: (*b*) S. Swaminathan, S. Ranganathan. and S. Sulochana, *ibid.*, 1958, **23**. 707: (*c*) J. Wolinsky and J. E. Sundeen, *Tetrahedron*, 1970, **26**, 5427.
[7] (*a*) C. W. Whittle and R. N. Castle, *J. Pharm. Sci.*. 1963, **52**, 645: (*b*) S. Clementi, P. Linda, and G. Marino, *J.C.S. Chem. Comm.*. 1972, 427: (*c*) A. Chatterjee and K. M. Biswas, *J. Org. Chem.*, 1973, **38**, 4002.
[8] G. Magnanini, *Ber.*, 1888, **21**, 1936.
[9] N. F. Kucherova, V. P. Evdakov, and N. K. Kochetkov, *J. Gen. Chem.* (U.S.S.R.), 1957, **27**, 1131.
[10] W. Borsche and H. Groth, *Annalen*, 1941, **549**, 238.
[11] (*a*) B. S. Joshi, V. N. Kamat, D. H. Gawad, and T. R. Govindachari. *Phytochemistry*, 1972. **11**. 2065: (*b*) C. G. Raison, *J. Chem. Soc.*, 1949, 3319.
[12] E. Meitzner, *J. Amer. Chem. Soc.*, 1935, **57**, 2327.

two possible modes of C-2 substitution (direct attack on C-2 or primary attack at the 3-position with subsequent intramolecular rearrangement) have been thoroughly studied and discussed [13] for certain alkylation reactions. Jackson *et al.* conclude that the primary attack takes place at the 3-position, and is followed by rearrangement to C-2,[13a] but when a 6-methoxy-group is present the 2-position is activated in preference to the 3-position, and direct attack on C-2 is competitive.[13f] Other authors [6c,14] have shown that 2-substitution can arise from rearrangement of *N*-substituted 3-alkylindoles or from direct attack on C-2.[14b] Surprisingly, no detailed mechanistic study on the pattern of *N*-substitution appears to be available so far. Similarly the observation that carbazole readily undergoes *N*-substitution has received no attention.

In a preliminary communication on indole reactivity in Vilsmeier–Haack formylation,[7b] we suggested that there were at least five possibilities for *N*-substitution, *i.e.* reaction involving the conjugate base or the 3*H*-indole tautomer, direct attack on nitrogen of the neutral indole molecule, or attack on C-3 followed by a slow or a fast intra- or inter-molecular rearrangement. In view of our recent report on the unexpected *N*-trifluoroacetylation of unsubstituted indole and carbazole, and of the mechanistic interest arising from these results, we undertook a detailed kinetic investigation by measuring rate constants and isomer distribution in the Vilsmeier–Haack acetylation [4] of several pyrroles, indoles, and carbazoles. The results allow us to make a choice amongst the above-mentioned mechanisms.

RESULTS AND DISCUSSION

The acetylation rate constants of pyrrole, indole, carbazole, and a number of their derivatives with the *NN*-dimethylacetamide–carbonyl chloride complex * have been determined in 1,2-dichloroethane at 25 °C (Table 1). All the substrates exhibit second-order kinetics under the conditions examined.

Table 1 also records the rate constants for the individual ring positions in all cases where two products were obtained. These were calculated from the overall rates and the isomer distributions evaluated by g.l.c. (Table 2).

Whereas the three pyrroles give exclusive α-substitution, all the indoles and carbazoles with a free NH group give a detectable amount of *N*-substitution. The *N*-acetyl derivative is the main product from carbazoles and 3-alkylindoles, but, even when a free 3-position is present, indole, 2-methylindole, and 4-methylindole produce a significant amount of *N*-acetyl isomer. Moreover, a change of the 3-alkyl group from methyl to t-butyl leads to complete disappearance of *ortho*-substitution: *ca.* 5% of 2-acetyl derivative is found from 3-methylindole and exclusive *N*-substitution with 3-t-butylindole.

TABLE 1

Rate constants ($10^4 k_2$/l mol^{-1} s^{-1}) for Vilsmeier–Haack acetylation in 1,2-dichloroethane at 25 °C

	Rate constants [a]				
Substrate	Overall	N	α	β	Benzene ring (position)
Pyrrole	560		280 [b]		
N-Methylpyrrole	648		324 [b]		
2-Methylpyrrole	11 800		11 800		
Indole	150	3.30	0.056 [c]	147	
N-Methylindole	238			238	
2-Methylindole	569	9.10		560	
3-Methylindole	4.01	3.80	0.213		
4-Methylindole	121	5.69		115	
3-t-Butylindole	3.00	3.00			
1,2-Dimethylindole	506			506	
1,3-Dimethylindole	0.221		0.221		
2,3-Dimethylindole	3.43	3.32			0.106 (6)
1,2,3-Trimethylindole	0.113				0.113 (6)
Carbazole	26.1	24.3			0.90 (3) [b]
N-Methylcarbazole	2.80				1.40 (3) [b]
1,2,3,4-Tetrahydrocarbazole	1.89	1.79			0.102 (7)

[a] Rate constants for individual positions calculated from the isomer distributions. [b] Figure is divided by two to allow for the statistical factor. [c] Calculated value: see text.

TABLE 2

Isomer distributions (mole %) for the Vilsmeier–Haack acetylation of indole and carbazole derivatives

Substrate	*N*–Ac	*C*–Ac (position)
Indole	2.2	97.8 (3)
2-Methylindole	1.6	98.4 (3)
3-Methylindole	94.7	5.3 (2)
4-Methylindole	4.7	95.3 (3)
2,3-Dimethylindole	96.9	3.1 (6)
Carbazole	93.1	6.9 (3)
1,2,3,4-Tetrahydrocarbazole	94.6	5.4 (7)

The data of Table 1 make possible a discussion of individual reactivities in terms of electronic and steric effects.

Comparison of rate constants of *N*-methylated compounds with those of the unsubstituted compounds shows that *N*-methylation causes a slight increase in *C*-reactivity of all systems. This increment is very small in the α-positions of both pyrrole and indole (see figures for 3-methylindoles) and finite (k_{Me}/k_H *ca.* 1.5) in the β-position of indole and in the 3-position of carbazole, possibly owing to a primary steric effect exerted by the methyl group towards the bulky electrophile attacking the α-position.[15]

* This reagent was chosen instead of the more usual formylating agent *NN*-dimethylformamide–phosphoryl chloride because the reactions with the latter were too fast to be followed by usual methods.

[13] A. H. Jackson and A. E. Smith, *Tetrahedron*, (*a*) 1965, **21**, 989; (*b*) 1968. **24**. 403; (*c*) A. H. Jackson and P. Smith. *ibid.*, 1968. **24**. 2227; (*d*) A. H. Jackson. B. Naidoo. and P. Smith. *ibid.*, 1968, **24**, 6119; (*e*) K. M. Biswas and A. H. Jackson. *ibid.*, 1969, **25**, 227; (*f*) R. Iyer. A. H. Jackson. P. V. R. Shannon, and B. Naidoo, *J.C.S. Perkin II*, 1973, 872.

[14] (*a*) G. Casnati. R. Marchelli. and A. Pochini, *J.C.S. Perkin I*, 1974, 754; (*b*) G. Casnati, A. Dossena, and A. Pochini, *Tetrahedron Letters*, 1972. 5277.

[15] C. F. Candy. R. A. Jones, and P. H. Wright, *J. Chem. Soc.* (*C*), 1970, 2563.

In contrast, the activating effect exerted by a 2-methyl group on the reactivity at N-1 is larger, as shown by the difference between the results for 2-methylindole and those for 3-methylindole or indole itself (k_{Me}/k_H *ca.* 2.4).

Moreover, the unusually low activation factors due to a methyl group in the *ortho*-position (*i.e.* 2-methylindole is only 3.8 times more reactive than indole at C-3) are certainly connected with large steric requirements of the electrophile, as suggested in previous studies.[16] An estimate of the true electronic activation given by a methyl group from a conjugative position may be derived from the increase in reactivity found for 2-methylpyrrole in comparison with pyrrole (k_{Me}/k_H 35). Steric inhibition thus decreases reactivity by a factor of *ca.* 10.

The influence of steric factors in these systems is also shown by the behaviour of 4-methylindole, which is slightly less reactive at C-3 than indole itself, owing to *peri*-interactions.

The benzene ring (C-6 in 2,3-dimethylindole and 1,2,3-trimethylindole, and C-7 in 1,2,3,4-tetrahydrocarbazole) is only slightly less reactive than the heterocyclic position 2 in 3-methyl- and in 1,3-dimethyl-indole, owing to the absence of steric inhibition by adjacent methyl groups and the conjugative activating effect of the C-2 methyl groups. The reactivity of the benzene ring in the foregoing three compounds appears to be little affected by structural modification of the heterocyclic moiety. Furthermore, the ratio of *N*- to homocyclic *C*-substitution is *ca.* 30 in 2,3-dimethylindole, in good agreement with the value obtained for carbazole.

Regarding the effect of annelation on pyrrole, the reactivity of the β-position of indole is not markedly different from that of the α-position of pyrrole (k_P/k_I = 1.90). To a first approximation, assumption of the same *ortho*-activation by the methyl substituent in 2- and 3-methylindole enables us to calculate a value of 5.6×10^{-6} l mol^{-1} s^{-1} for the rate constant for reaction at the 2-position of unsubstituted indole. Therefore k_P/k_I is 5.000 for the α-position, and less than unity for the β-position. The whole picture is in keeping with the analogous annelation of thiophen and furan.[17]

Use of the calculated *k*-value for C-2 substitution gives the reactivity order of the heterocycle in the indole nucleus as follows: 3-position (2.600) > 1-position (59) > 2-position (1). Thus the 1-position is more reactive than the 2-position, and this behaviour is in accord with the well established greater ease of proton exchange at position 1 than position 2, even under acidic conditions.[3c] Although the overall reactivity order is pyrrole > indole > carbazole, the rate constants of individual positions (Table 1) give completely different orders of reactivity for the various positions of substitution for NH, carbazole > indole > pyrrole, for C-2, pyrrole > indole, for C-3, indole > pyrrole.

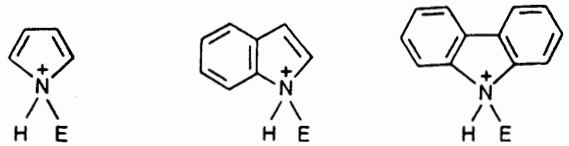
Our final point concerns the mechanistic features of attack on nitrogen in indole and its derivatives. Carbazole undergoes *N*-acetylation seven times as fast as indole. This eliminates the hypothesis that the main path is attack on position 3 of indole (or 3-substituted indole) followed by intra- or inter-molecular rearrangement, since in carbazole this path is very unlikely. A further consequence of this observation is that the mechanism involving a 3*H*-indole tautomer, which cannot exist in carbazole, can be disregarded.

Rate-determining attack on position 3 also seems unlikely on the basis of the kinetic results for various 3-substituted indoles. In fact a 3-t-butyl group should cause a marked decrease in rate constants in the case of a mechanism involving a C-3 ⟶ N rearrangement of this kind.

The possibility of a reaction through the conjugate base can be discarded on the basis of the similarity of the rate constants of *N*-methyl derivatives to those of the unsubstituted compounds.

We conclude that the most likely mechanism of *N*-substitution of indole and 3-substituted indoles is the direct attack on the heteroatom by the electrophile. A similar direct *N*-attack has been suggested in a recent kinetic study on acid-catalysed exchange of pyrrole and indole in aqueous acetonitrile,[3c] not supporting the view that exchange at nitrogen in indole takes place *via* 3-protonation and deprotonation of the resulting 3*H*-indolium.[18] In these processes the nature and the concentration of the acid and the solvent mixture play a fundamental role.[3c]

Finally the observed order of reactivity in *N*-acetylation by the Vilsmeier–Haack reagent (pyrrole ≪ indole < carbazole) already found in trifluoroacetylation,[5] should be related, as previously suggested,[5] to differences in degree of overall loss of aromaticity in the transition state as approximated by the Wheland intermediate structures (1)—(3). The nitrogen quaternisation resulting from direct attack of the electrophile on the heteroatom leads to a larger degree of bond fixation in pyrrole than in indole (one benzene ring unchanged) and still larger than in carbazole (two benzene rings).



(1) (2) (3)

The observation that C-3 in carbazole is more reactive than C-6 in 2-methylindoles by a factor of about 10 (and by yet a larger factor in indole) is probably again because attack at C-3 in carbazole involves less overall loss in resonance energy in the transition state than attack at C-6 in indoles.

[16] P. Linda, A. Lucarelli, G. Marino, and G. Savelli, *J.C.S. Perkin II*, 1974, 1610.

[17] S. Clementi, P. Linda, and G. Marino, *J.C.S. Perkin II*, 1971, 79.

[18] Ref. 1, p. 6.

EXPERIMENTAL

Materials.—3-t-Butylindole,[19] 1,3-dimethylindole,[20] 1,2,-3-trimethylindole,[20] 4-methylindole,[21] and *N*-methylcarbazole [22] were prepared according to reported methods. The other substrates and some products were available from previous studies or were commercial samples.

The following compounds were prepared as reported: 3-acetylindole,[23] 3-acetyl-*N*-methylindole,[24] 3-acetyl-2-methylindole,[10] 3-acetyl-1,2-dimethylindole,[10] 2-acetyl-3-methylindole,[25] 2-acetyl-1,3-dimethylindole,[25] *N*-acetyl-2,3-dimethylindole,[26] 6-acetyl-2,3-dimethylindole,[27] 6-acetyl-1,2,3-trimethylindole,[10,28] 3-acetylcarbazole,[29] 3-acetyl-*N*-methylcarbazole,[30] *N*-acetyl-1,2,3,4-tetrahydrocarbazole,[31] and 7-acetyl-1,2,3,4-tetrahydrocarbazole.[32]

Most of the *N*-acetyl derivatives were prepared by treatment of the substrate with equimolecular amounts of triethylamine and acetic anhydride under reflux for 16 h; [33] *N*-acetylindole had b.p. 153° at 15 mmHg (lit.,[34] 152—153° at 14 mmHg); *N*-acetyl-3-methylindole had m.p. 65° (lit.,[35] 68°); *N*-acetylcarbazole had m.p. 70° (lit.,[36] 68—69°); the n.m.r. spectra of these compounds agree with reported data.

Under the same conditions the following *N*-acetyl derivatives were prepared: *N*-acetyl-2-methylindole, b.p. 250°, δ (CCl_4) 2.45 (6 H, s, 2-CH_3 and NAc), 6.07br (1 H, s, 3-H), 6.80—7.30 (3 H, m, ArH), and 7.64 (1 H, m, 7-H); *N*-acetyl-4-methylindole, b.p. 172° at 15 mmHg, δ ($CDCl_3$) 2.46 (3 H, s, 4-CH_3), 2.58 (3 H, s, NAc), 6.47 (1 H, d, $J_{2,3}$ 3.8 Hz, 3-H), 6.82—7.30 (3 H, m, ArH + 2-H), and 8.20 (1 H, d, $J_{6,7}$ 8.0 Hz, 7-H); 3-acetyl-4-methylindole, m.p. 185—187°, δ ($CDCl_3$) 2.45 (3 H, s, 3-Ac), 2.75 (3 H, s, 4-CH_3), 7.0—7.5 (3 H, m, ArH), 8.25 (1 H, d, $J_{1,2}$ 2.5 Hz, 2-H), and 8.70br (1 H, s, NH); *N*-acetyl-3-t-butylindole, m.p. 118°, δ [$(CD_3)_2SO$] 1.35 (9 H, s. 3-CMe_3), 2.55 (3 H, s, NAc), 7.12 (2 H, m, 5- + 6-H), 7.26 (1 H, s, 2-H), 7.62 (1 H, m, 4-H), and 8.28 (1 H, m, 7-H).

Chloroform and *NN*-dimethylacetamide were purified by standard procedures. Commercial carbonyl chloride (20% in toluene) was used without further purification.

Product Analysis.—All the substrates examined gave, under kinetic conditions, only the acetyl derivatives, as checked by g.l.c. analysis. Isomer distributions were measured by g.l.c. comparison with authentic specimens [C. Erba G.I. fractometer, equipped with flame ionization detector; stainless steel column (2 m × 4 mm) packed with 10% silicone SE-30 temperature 180—215 °C]. The results are in Table 2.

Kinetics.—Preparation of the *NN*-dimethylacetamide–carbonyl chloride complex and the kinetic method have been described previously.[16,37] Owing to the high reactivity of many substrates and to the precipitation of solid (reaction intermediate) during the early stages of the reaction, a batch-wise procedure was used.

The Consiglio Nazionale delle Ricerche is thanked for financial support, and Dr. G. V. Sebastiani for help in n.m.r. spectral measurements.

[6/1878 *Received*, 6*th October*, 1976]

[19] H. C. Bettembourg and S. David, *Bull. Soc. chim. France*, 1962, 772.
[20] M. Julia and H. J. Suizi, *Bull. Soc. chim. France*, 1962, 2266.
[21] J. A. Elvidge and R. G. Foster, *J. Chem. Soc.*, 1964, 981.
[22] V. P. Lopatinskii, E. E. Sirotkina, and M. M. Sukhoroslova, *Metody polucheniya Khim. Reaktivov i Preparatov*, 1964. 69 (*Chem. Abs.*. 1966. **64**. 19542).
[23] J. E. Saxton, *J. Chem. Soc.*, 1952, 3592.
[24] C. B. Barett, R. J. S. Beer, G. M. Dodd, and A. Robertson. *J. Chem. Soc.*. 1957, 4810.
[25] K. Ishizumi, T. Shioiri, and S. Yamada, *Chem. and Pharm. Bull.* (*Japan*). 1967. **15**, 863.
[26] C. M. Atkinson, J. C. E. Simpson. and A. Taylor, *J. Chem. Soc.*. 1954. 165.
[27] W. J. Gaudion, W. H. Hoow, and S. G. P. Plant, *J. Chem. Soc.*, 1947, 1631.
[28] N. N. Suvorov and N. P. Sorokina. *Zhur. Obshchei Khim.*, 1960, **30**, 2055 (*Chem. Abs.*, 1961, **55**, 6466).
[29] Y. Nagai and C. C. Huang, *Bull. Soc. Chem. Japan*, 1965, **38**, 951.
[30] N. P. Buu-Hoi and R. Royer, *Rec. Trav. chim.*, 1947, **66**. 533.
[31] W. H. Perkin and S. G. P. Plant. *J. Chem. Soc.*, 1921. 1825.
[32] S. G. P. Plant and K. M. Rogers, *J. Chem. Soc.*. 1936, 40.
[33] P. Linda and G. Marino, *Ricerca sci.*, 1967. **37**, 424.
[34] H. Plieninger and G. Werst. *Chem. Ber.*, 1956, **89**, 2783.
[35] (*a*) B. Oddo, *Gazzetta*, 1913, **43**, 385: (*b*) A. Borsodi and T. Toth. *Magyar Kém. Folyóirat*, 1969, **75**, 251 (*Chem. Abs.*, 1969, **71**, 49691).
[36] A. A. Berlin, *J. Gen. Chem.* (*U.S.S.R.*), 1944, **14**, 438 (*Chem. Abs.*, 1945, **39**, 4606).
[37] S. Alunni, P. Linda, G. Marino, S. Santini. and G. Savelli, *J.C.S. Perkin II*, 1972. 2070.