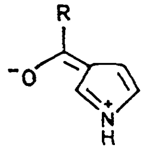
Crystallographic Studies in the Pyrrole Series. Part I. Crystal and Molecular Structure of Ethyl 4-Acetyl-3-ethyl-5-methylpyrrole-2-carboxylate

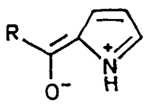
By **R. Bonnett**, **M. B. Hursthouse**, and **S. Neidle,*** Department of Chemistry, Queen Mary College, Mile End Road, London E1 4NS

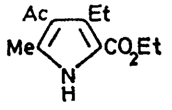
Crystals of the title compound are monoclinic, space group $P2_1/n$, $Z = 4$, with cell dimensions a = 6·772(3), b = 14·628(5), c = 12·187(5) Å, and β = 94·20(12)°. The structure was solved by direct methods, and refined by least-squares to R 0·0615 for the 1531 statistically significant reflections measured. The geometry of the molecule is compared with that observed for the pyrrole ring in other systems; in particular it has been observed that the normal symmetrical bonding arrangement in the ring is here considerably perturbed, owing to the electronic effects of the substituents.

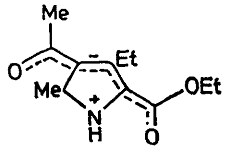
THE chemistry of the pyrrole series has received considerable attention, both on account of the biological importance of some of these compounds, and because of their inherent chemical and theoretical interest.[1] Nevertheless it is only the more complex of these substances, the phthalocyanins,[2] the corrinoids,[3] and, more recently, the porphyrins,[4] which have been investigated crystallographically. While some of these structural elucidations stand out very prominently in the history of X-ray crystallography, simpler members of the series have been virtually ignored. The present studies are intended to repair this omission, starting with monocyclic systems and moving by degree to more complex assemblies of pyrrole rings.

Since the molecular dimensions of pyrrole itself have been determined by microwave spectroscopy,[5] it was decided to investigate a structure in which the system was perturbed by an electron-withdrawing group. The carbonyl group was chosen since it is well known that in pyrrolyl ketones the reactivity of the carbonyl group to nucleophilic attack is depressed (the dipyrrolyl ketones, indeed, do not form normal ketonic derivatives),[6] this often being rationalised in terms of mesomeric structures such as (I) and (II). The importance of such canonical structures is also thought to be reflected in the i.r. (and Raman)[7] spectra of pyrrolyl ketones, where the stretching frequency of the carbonyl group is markedly lower than that found for simple aryl ketones [*e.g.* ν(CO): acetophenone in CCl_4, 1692; 2-acetylpyrrole in CCl_4, 1658; dipyrrol-2-yl ketone in $CHCl_3$, 1597 cm^{-1}].[8] If canonical structures such as (I) and (II), representing the 'amide-like' character of pyrrolyl carbonyl compounds, are important, molecular dimensions would be expected to differ significantly from those of pyrrole itself. To test this we have examined the crystal and molecular structure of ethyl 4-acetyl-3-ethyl-5-methylpyrrole-2-carboxylate (III), which was regarded as a particularly favourable case since in the zwitterionic representation the negative charge may be drawn as delocalised over two carbonyl oxygens [as in (IV)].


(I)


(II)


(III)


(IV)

EXPERIMENTAL

Crystals of the pyrrole (III) were obtained from ethanol.

Crystal Data.—$C_{12}H_{17}NO_3$, M = 223·0, monoclinic a = 6·772(3), b = 14·628(5), c = 12·187(5) Å, β = 94·20(2)°, U = 1204·0 Å, D_m = 1·24(2) (by flotation), Z = 4, D_c = 1·230, $F(000)$ = 448·0. Space group $P2_1/n$ (No. 14, C^5_{2h}) from systematic absences: $h0l$, $0k0$ for $h + l$, $k = 2n$. Cu-K_α radiation, λ_{mean} = 1·54178 Å; μ(Cu-K_α) = 7·31 cm^{-1}. Cell dimensions were obtained from diffractometer measurements of 2θ values. Intensity data were collected for a crystal of dimensions *ca.* 0·60 × 0·30 × 0·30 mm, mounted about the b axis, on a General Electric XRD 6 diffractometer, equipped with a manual goniometer, pulse-height analyser, and scintillation counter. Nickel-filtered copper radiation was used. The stationary-crystal–stationary-counter method of intensity estimation was used throughout, with a 4° take-off angle, and a counting time of 10 s. Individual backgrounds at $(2\theta_{hkl} + 1)°$ were measured for all reflections. The 006 and 060 reflections were used as reference reflections to check on crystal stability; in neither case was there any significant decline in intensity during the course of the data collection.

---

[1] For a comprehensive review see H. H. Inhoffen, J. W. Buchler, and P. Jäger, *Forschr. Chem. org. Naturstoffe*, 1968, **26**, 284.

[2] J. M. Robertson, *J. Chem. Soc.*, 1935, 615.

[3] D. C. Hodgkin, J. Kamper, J. Lindsey, M. MacKay, J. Pickworth, J. H. Robertson, C. B. Shoemaker, J. G. White, R. J. Prosen, and K. N. Trueblood, *Proc. Roy. Soc.*, 1957, *A*, **242**, 228.

[4] For a summary see E. B. Fleischer, *Accounts Chem. Res.*, 1970, **3**, 105.

[5] L. Nygaard, J. T. Nielsen, J. Kirchheimer, G. Maltesen, J. Rastrup-Anderson, and G. O. Sorensen, *J. Mol. Structure*, 1969, **3**, 491.

[6] H. Fischer and H. Orth, 'Die Chemie des Pyrrols,' vol. 1, Leipzig, 1937, p. 364.

[7] G. B. Bonino, R. Manzoni-Ansidei, and P. Pratesi, *Z. phys. Chem.*, 1934, **25**, *B*, 348.

[8] M. K. A. Khan and K. J. Morgan, *J. Chem. Soc.*, 1964, 2579; H. Rappoport and C. D. Willson, *J. Amer. Chem. Soc.*, 1962, **84**, 630.

The intensities of 1792 reflections with $2\theta \leqslant 120°$ were measured, of which 1531 were considered to be statistically significant {net counts were $\geqslant 3\sigma(I)$, where $\sigma(I)$, the standard deviation in the intensity, was taken[9] as $[I + 2B + (0{\cdot}03I)^2]^{\frac{1}{2}}$, and B is the background count}. The 261 ' unobserved ' reflections were excluded from any subsequent least-squares calculations. No absorption correction was applied.

Structure Solution and Refinement.—The structure was solved by a direct statistical method using the Sayre equation, applied to centro-symmetric structures.[10] The data was put on an absolute scale and a set of normalised structure factors (E's) were calculated by means of a Wilson plot.[11]

The 128 highest E-values were then input to the program REL,[12] which automatically chose a set of origin-determining reflections, with a number of other, initially symbolic, phases (Table 1). The program then reiteratively applied the Sayre equation to each of the 16 starting sets. The set which converged with the least number of iterations also has the highest consistency index C, of 0·945, where C is defined [12] as: $C = \langle|E_h \sum_{h} E_{h'} E_{h-h'}|\rangle / \langle|E_h| \sum_{h} |E_{h'}||E_{h-h'}|\rangle$. 127 out of the 128 phases were determined with a probability of >0·99. The next best solution had C 0·896.

TABLE 1

Starting set of phases for sign determination

Reflection	$\lvert E\rvert$ value	Starting phase	Final phase
2,6,3	2·82	+	Origin set
1,2,2	2·27	+	
1,3,1	2·14	+	
5,1,7	4·33	*a*	—
3,1,9	3·21	*b*	+
2,5,3	3·01	*c*	+
3,8,4	2·72	*d*	+

An E-map was then calculated with the 127 best-solution phases. This clearly revealed the complete structure, and three cycles of full-matrix least-squares refinement on the positions found, all assigned as carbon atoms, gave R 0·204. Correct assignment of oxygen and nitrogen atoms reduced R to 0·152. At this stage a number of reflections with poor agreement were remeasured, and three more cycles of isotropic refinement reduced R to 0·137. The temperature factors were then converted to their anisotropic (β_{ij}) equivalents, and four more cycles of refinement gave R 0·104. Examination of a difference Fourier map clearly revealed all the 17 hydrogen atom positions to be as expected. These were included in four subsequent cycles of full-matrix refinement, with their positional and isotropic thermal parameters being allowed to vary. The 021 and 111 reflections were omitted from the calculations at this stage because of suspected extinction. R was now 0·071. Examination of the variation of Δ^2F over ranges of $|F_o|$ and $\sin\theta$ suggested that a Hughes-type [13] weighting scheme was appropriate of the form $\sqrt{w} = 1$ for $F_o < F^*$ and $\sqrt{w} = F^*/F_o$ for $F_o > F^*$, with F^* set at 50·0. Inclusion of this weighting scheme in four final cycles of refinement gave R 0·0615 over the 1529 observed reflections. All parameter shifts were $<0{\cdot}02\sigma$, and refinement was judged to be complete. The final positional and thermal parameters for the atoms are listed in Tables 2—4. Supplementary Publication

TABLE 2

Final positional parameters for the non-hydrogen atoms, as fractions of the unit-cell edges, with standard deviations in parentheses

Atom	*x*	*y*	*z*
N(1)	0·2455(4)	0·2136(2)	0·1306(2)
C(2)	0·0892(5)	0·2546(3)	0·1688(3)
C(3)	0·0490(5)	0·2106(2)	0·2671(2)
C(4)	0·1895(5)	0·1395(2)	0·2860(2)
C(5)	0·3127(5)	0·1429(2)	0·2006(3)
C(6)	0·4804(5)	0·0879(2)	0·1776(3)
O(7)	0·5641(4)	0·1163(2)	0·0870(2)
C(8)	0·7376(5)	0·0661(3)	0·0580(3)
C(9)	0·7974(7)	0·1053(4)	−0·0477(4)
C(10)	−0·0082(7)	0·3321(3)	0·1045(4)
C(11)	−0·1061(5)	0·2362(2)	0·3397(3)
O(12)	−0·1002(5)	0·2065(2)	0·4345(2)
C(13)	−0·2680(7)	0·2993(3)	0·3009(5)
C(14)	0·2022(6)	0·0691(3)	0·3765(3)
C(15)	0·0645(7)	−0·0104(3)	0·3509(4)
O(16)	0·5471(4)	0·0241(2)	0·2327(2)

TABLE 3

Final anisotropic thermal parameters (× 10⁴) * for the non-hydrogen atoms, with standard deviations in parentheses

Atom	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
N(1)	233(7)	45(1)	58(2)	3(2)	51(3)	2(1)
C(2)	235(8)	41(2)	65(2)	2(3)	45(4)	−4(1)
C(3)	219(8)	43(2)	59(2)	−6(3)	44(3)	−7(1)
C(4)	227(8)	45(2)	54(2)	−1(3)	35(3)	−4(1)
C(5)	225(8)	44(2)	61(2)	1(3)	43(3)	1(1)
C(6)	234(8)	50(2)	73(2)	−2(3)	48(4)	1(2)
O(7)	269(6)	57(1)	80(2)	27(2)	84(3)	8(1)
C(8)	201(8)	65(2)	94(3)	21(4)	54(4)	−4(2)
C(9)	265(11)	79(3)	104(4)	7(4)	78(6)	−3(3)
C(10)	346(12)	54(2)	88(3)	31(4)	58(5)	11(2)
C(11)	281(9)	45(2)	83(3)	−10(3)	76(4)	−10(2)
O(12)	506(11)	70(2)	84(2)	29(3)	135(4)	2(2)
C(13)	323(12)	72(3)	137(5)	24(5)	123(7)	−2(3)
C(14)	298(11)	54(2)	60(3)	1(4)	45(4)	4(2)
C(15)	345(12)	61(2)	88(3)	−33(4)	25(5)	15(2)
O(16)	320(8)	70(2)	108(2)	46(3)	78(3)	28(2)

* In the form: $f_i = f_i^0 \exp-(\beta_{11}h^2 + \beta_{22}k^2 + \beta_{33}l^2 + 2\beta_{12}hk + 2\beta_{13}hl + 2\beta_{23}kl)$.

TABLE 4

Final positional and isotropic temperature factors for the hydrogen atoms, with standard deviations in parentheses

Atom	*x*	*y*	*z*	B/Å²
H(0101)	0·289(6)	0·232(3)	0·066(4)	3·47(88)
H(0208)	0·707(5)	0·000(3)	0·053(3)	2·09(70)
H(0308)	0·865(7)	0·076(3)	0·121(4)	4·40(95)
H(0409)	0·801(7)	0·176(4)	−0·048(4)	5·63(119)
H(0509)	0·702(8)	0·096(4)	−0·102(5)	5·93(145)
H(0609)	0·961(8)	0·081(4)	−0·050(4)	6·82(128)
H(0710)	0·040(7)	0·347(3)	0·034(4)	4·83(105)
H(0810)	−0·038(7)	0·387(3)	0·143(4)	4·47(103)
H(0910)	−0·151(9)	0·318(4)	0·076(5)	6·63(138)
H(1013)	−0·199(7)	0·363(4)	0·299(4)	5·55(116)
H(1113)	−0·340(8)	0·278(4)	0·218(4)	6·20(129)
H(1213)	−0·387(7)	0·296(3)	0·340(4)	5·06(107)
H(1314)	0·161(6)	0·098(3)	0·441(4)	4·04(93)
H(1414)	0·334(6)	0·046(2)	0·384(3)	2·21(72)
H(1515)	0·060(6)	−0·058(3)	0·403(4)	3·54(86)
H(1615)	0·106(6)	−0·041(3)	0·285(4)	3·14(84)
H(1715)	−0·103(8)	0·009(3)	0·336(4)	6·52(117)

Mean C–H 1·02(12) Å. The last two digits in each identification number refer to the atom to which each hydrogen atom is bound.

---

[9] R. H. Eisenberg and J. A. Ibers, *Inorg. Chem.*, 1966, **5**, 411.

[10] D. Sayre, *Acta Cryst.*, 1952, **5**, 60.

[11] A. J. C. Wilson, *Acta Cryst.*, 1949, **2**, 318.

[12] R. E. Long, Ph.D. Thesis, University of California, Los Angeles, 1965.

[13] E. W. Hughes, *J. Amer. Chem. Soc.*, 1941, **63**, 1737.

No. SUP 20346 (4 pp., 1 microfiche), lists the final observed and calculated structure factors.*

All calculations were performed on the University of London CDC 6600 computer, using the '*X*-Ray '70' crystallographic computing system[14] except where otherwise stated. The scattering factors were taken from ref. 15.

DISCUSSION

The structure of the pyrrole (III) is shown in Figure 1.[16] Bond lengths and bond angles are shown in Figures 2 and 3. Table 5 compares the structural parameters observed for the ring system with those reported for pyrrole itself,[5] and for a typical pyrrole unit in the porphyrin nucleus.[4]

TABLE 5

Structural parameters and π-bond orders for the pyrrole ring in various systems

(*a*) Distances and π-bond orders

	Pyrrole unit in porphyrin [a]		Pyrrole [b]		Compound (III)	
	Bond length	π-Bond order *	Bond length	π-Bond order	Bond length	π-Bond order
N(1)–C(2)	1·37	0·27	1·370	0·27	1·330	0·53
N(1)–C(5)	1·37	0·27	1·370	0·27	1·395	0·13
C(2)–C(3)	1·45	0·16	1·382	0·58	1·404	0·44
C(4)–C(5)	1·45	0·16	1·382	0·58	1·382	0·58
C(3)–C(4)	1·34	0·91	1·417	0·35	1·417	0·35

(*b*) Bond angles (°)

	Pyrrole unit in porphyrin [a]	Pyrrole [b]	Compound (III)
At N(1)	108	109·8	110·6
C(2)	109	107·7	107·6
C(5)	109	107·7	107·4
C(3)	107	107·4	107·8
C(4)	107	107·4	106·6

* Calculated following Cruikshank and Sparks' treatment,[17] and taking π-bond orders of 0 for σ bonds between sp^2 hybridised atoms, and 1 for double bonds.

[a] Ref. 4. [b] Ref. 5.

It is apparent that the presence of two carbonyl substituents has markedly altered the shape of the pyrrole ring. The C_{2v} symmetry of the parent system has disappeared, bond N(1)–C(2) being considerably shorter than N(1)–C(5), while C(2)–C(3) is longer than C(4)–C(5). Derivation of π-bond order, following the Cruikshank and Sparks treatment[17] (assuming 1·42 Å for the C–N bond involving trigonally hybridised atoms, and 1·27 Å for the C=N double bond),[18] gives 0·53 for the N(1)–C(2) bond in compound (III), to be compared with a π-bond order, similarly derived, of 0·27 for the C–N bonds of pyrrole itself. Other π-bond orders are given in Table 5, and overall they indicate the importance of the mesomeric dipolar structures (V) and (VI) summarised in (IV).

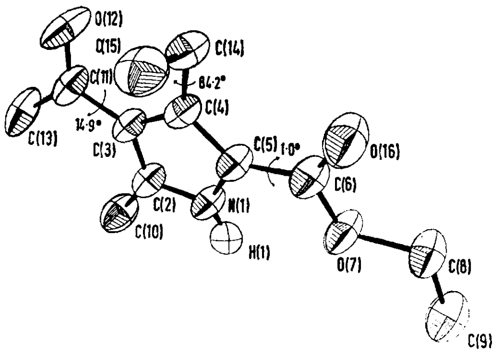


FIGURE 1 Perspective drawing of a molecule of ethyl 4-acetyl-3-ethyl-5-methylpyrrole-2-carboxylate. The thermal ellipsoids are scaled to include 50% probability. The angles shown denote dihedral angles with respect to the mean plane of the ring

Table 6 indicates that although the ring itself [plane (1)] is planar within the limits of experimental error, most of the substituent atoms in this tetrasubstituted molecule deviate significantly from this plane. Both C(10) and C(14) are out of plane by >5σ. The acetyl

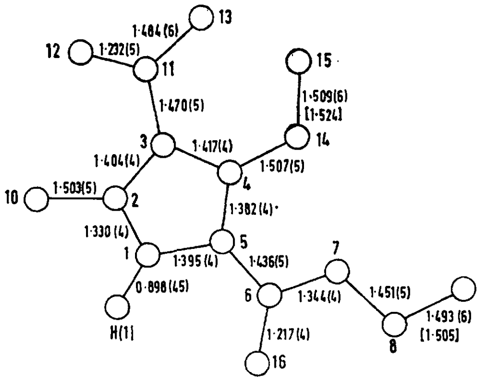


FIGURE 2 Bond lengths (Å); values in square brackets are corrected for thermal motion

---

* For details of Supplementary Publications see Notice to Authors No. 7 in *J. Chem. Soc.* (*A*), 1970, Issue No. 20.

[14] A revision of *X*-Ray '67, 'Program System for *X*-Ray Crystallography,' University of Maryland Technical Reports 67-58, 1967.

[15] 'International Tables for *X*-Ray Crystallography,' vol. III, 1962, Kynoch Press, Birmingham.

[16] C. K. Johnson, ORTEP, ORNL-3794, Oak Ridge National Laboratory, 1965.

[17] D. W. J. Cruikshank and R. A. Sparks, *Proc. Roy. Soc.*, 1960, *A*, **258**, 270; D. W. J. Cruikshank, *Tetrahedron*, 1962, **17**, 155.

[18] T. A. Hamor, W. S. Caughey, and J. L. Hoard, *J. Amer. Chem. Soc.*, 1965, **87**, 2305.

group, though planar itself [plane (2)] is twisted by *ca.* 15° out of the ring plane whereas the ester carbonyl [plane (3)] is twisted by only one degree from this plane. This suggests that (VI) is more important than (V), a suggestion which is supported by the observation that C(5)–C(6) is somewhat shorter (by 7σ) than C(3)–C(11). Other bond lengths and angles do not differ appreciably from accepted values,[19] apart from C(14)–C(15) and C(8)–C(9) both of which are expected to be normal C_{sp^3}–C_{sp^3} bonds. Both C(9) and C(15) have larger thermal motion than the other atoms in the molecule, and the discrepancies are probably due to neglect of this effect. A correction utilising the rider model of Busing and Levy[20] has been calculated (Figure 2), which has resulted in some improvement.

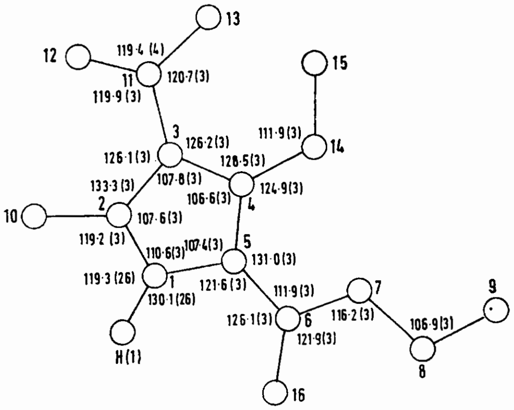
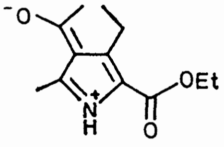
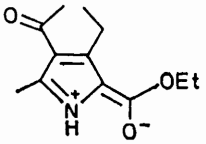


FIGURE 3 Valence angles (deg.)



(V)



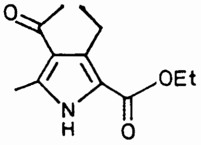
(VI)

TABLE 6

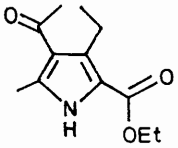
Equations of mean planes in the form $px + qy + rz = s$ in direct space, and in square brackets displacements (Å) from the mean planes through various parts of the molecule

	p	q	r	s
Plane (1): N(1), C(2)—(5).	3·859	9·313	5·808	3·697
[N(1) −0·003, C(2) −0·001, C(3) 0·005, C(4) −0·006, C(5) 0·005, C(6) 0·006, O(7) 0·068, C(10) −0·029, C(11) 0·066, O(12) 0·361, C(13) −0·197, C(14) −0·087, O(16) −0·011]				
Plane (2): C(3), C(11), O(12), C(13).	3·579	11·472	3·479	3·519
[C(3) 0·002, C(11) −0·007, O(12) 0·003, C(13) 0·002]				
Plane (3): C(5), C(6), O(7), O(16).	3·767	9·097	6·204	3·720
[C(5) 0·003, C(6) −0·009, O(7) 0·003, O(16) 0·004]				

The conformation of the ethoxycarbonyl group with respect to the ring is of some interest.[21,22] I.r. and other studies[23] have suggested that rotational isomerism is possible, and, taking pyrrole (III) as example, two isomers (VII) and (VIII) would be expected in solution. Such isomerism bears a vinylogous relationship to the rotational isomerism observed in amides and could also occur, in principle, at the C(3) acetyl group. Our study has shown that in the solid state only the form (VIII) is present, and that the only hydrogen bonding in the crystal is of the intermolecular type. In accordance with this the i.r. spectrum (KBr disc) shows a broad N–H stretching band at 3200 and carbonyl stretching frequencies at 1705 and 1630 cm^{-1} (*cf.* refs. 21 and 22).



(VII)



(VIII)

The crystal structure consists of chains of hydrogen-bonded molecules, the chains running roughly parallel to the *c* axis (Figure 4). The hydrogen bonding is between H(1) of one molecule and O(12) of a screw-related molecule. Figure 5 shows the situation in more detail. The distances and angles involved correspond closely to those found in other N–H · · · O=C systems.[24,25] There are no other intermolecular contacts within van der Waals radii.

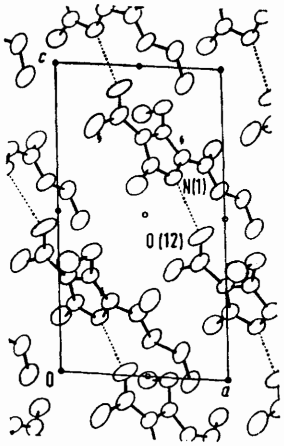


FIGURE 4 A view of the molecular packing, projected along the *b* axis direction. Dashed lines represent hydrogen bonds. N(1) and O(12) are related by the symmetry operation ($\frac{1}{2} + x$, $\frac{1}{2} - y$, $z - \frac{1}{2}$)

---

[19] *Chem. Soc. Special Publ.*, No. 18, 1965.

[20] W. R. Busing and H. A. Levy, *Acta Cryst.*, 1964, **17**, 142.

[21] R. Grigg, *J. Chem. Soc.*, 1965, 5149.

[22] R. A. Jones and A. G. Moritz, *Spectrochim. Acta*, 1965, **21**, 295.

[23] R. A. Jones, *Adv. Heterocyclic Chem.*, 1970, **11**, 456.

[24] See, for example, F. H. Allen, and J. Trotter, *J. Chem. Soc.* (*B*), 1971, 1073.

[25] J. Donohue in 'Structural Chemistry and Molecular Biology,' eds. A. Rich and N. Davidson, Freeman, San Francisco, 1968.

It is interesting to note the conclusion[4] that ring deformation of the sort observed here has not been detected in the porphyrin series. Because of the considerable aromatic stabilisation of the porphyrin system such effects would be expected to be small, and would require powerful electron-withdrawing substituents to reveal them. Of the systems listed by Fleischer[4] only one (nickel 2,4-diacetyldeuteroporphyrin-IX dimethyl ester[18]) fits into this category. A comparison of average bond dimensions with those of nickel aetioporphyrin[26] [N(1)–C(2) longer, C(2)–C(3) shorter in the latter] suggests that a mesomeric effect of the carbonyl group is being detected. However, the bond length differences, although in the right direction, are small and possibly within experimental error. Evidently it is desirable that further detailed structural studies on porphyrins with strongly polarising substituents (some of which, *e.g.* cytohaem-a, chlorophyll-c have important biochemical functions) should be undertaken.

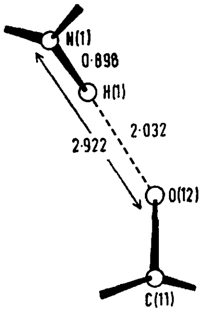


FIGURE 5 Detail of the hydrogen bonding. Distances in Å; angles (deg.): N(1)–H(1)–O(12) 70·7 (38)°, N(1)–O(12)–C(11) 129·6(3)°, and H(1)–O(12)–C(11) 126·7(27)°

The award of an I.C.I. Research Fellowship (to S. N.) is gratefully acknowledged.

[1/2129 *Received*, 11*th November*, 1971]

---

[26] E. B. Fleischer, *J. Amer. Chem. Soc.*, 1963, **85**, 146.