The Crystal Structure of *N*(α)-t-Butoxycarbonyl-*N*(π)-benzyloxymethyl-L-histidine

Tom Brown, John H. Jones,* and John D. Wallis
The Dyson Perrins Laboratory, South Parks Road, Oxford, OX1 3QY

X-Ray analysis of *N*(α)-t-butoxycarbonyl-*N*(π)-benzyloxymethyl-L-histidine hydrate has confirmed that the *im*-protecting group is located on the π-nitrogen atom, in agreement with deductions from n.m.r. coupling constants.

We recently substantiated [1] Veber's suggestion [2] that the location of *im*-protection in peptide synthesis with histidine (which is often troublesome) might be a critical consideration, and have developed [3] a successful new approach in which *N*(α)-t-butoxycarbonyl-*N*(π)-benzyloxymethyl-L-histidine (1), prepared as shown in Scheme 1, is a key intermediate. According to our interpretation, the superiority of (1) over previously employed protected histidines is largely due to the obstruction of the π-nitrogen in the former in contrast to the uncertain but probable τ-location of *im*-protection in the latter.

The structure shown for (1) follows from that shown for (2) and (3), which follow in turn from presumed steric selectivities operative during their preparation. Since, however, the location of the *im*-protecting group is at the heart of our thesis it seemed desirable to establish the structure (1) beyond argument by crystallography. This we have now done. The structure of our protected acid is indeed (1) (Figures 1 and 2).

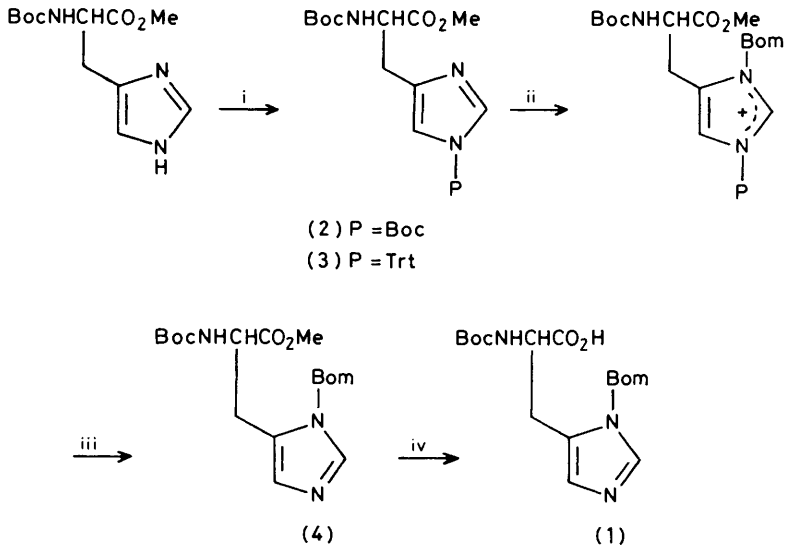
Bond lengths and angles are presented in Tables 1 and 2. There are two independent molecules of (1) in the unit cell (molecules 1 and 1A) and two molecules of water of crystallisation, one of which is equally disordered between two sites. The molecules of (1) are zwitterionic. The orientation of the amino-acid side-chains relative to their respective imidazolium rings are quite different in the two molecules. In molecule 1 the C(12)–C(13) bond lies nearly in the plane of the heterocyclic ring [C(3)–C(2)–C(12)–C(13): −12.8°],† while in molecule 1A the equivalent bond lies almost perpendicular to the corresponding plane [C(3A)–C(2A)–C(12A)–C(13A): 79.7°]. In both molecules the carboxy-group lies approximately *trans* to the C(2)–C(12) bond [C(2)–C(12)–C(13)–C(14): molecule 1, 178° molecule 1A, 160 °], so that the overall conformation of this side chain in molecule 1A resembles that found in DL-histidine hydrochloride dihydrate [4] and in one of the independent molecules of *N*(α)-acetyl-L-histidine monohydrate.[5] No conformation similar to that of molecule 1 has been observed previously in ring-protonated histidine derivatives. The bond lengths in the imidazolium ring are in good agreement with those of *N*(α)-acetyl-L-histidine monohydrate [5] and L-histidine hydrochloride monohydrate [6] with just one exception (Table 3). The N(2A)–C(2A) bond in molecule 1A appears to be somewhat too long [1.434 (8) Å] while the corresponding bond in molecule 1 is in good agreement with the comparison compounds. Even when the structure was refined with the ring dimensions constrained to be similar to those in Table 3, the large N(3A)–C(2A) bond length reappeared after refinement without the constraints.

As far as we are aware this is the first crystal structure to be reported for any *im*-modified histidine derivative: the only previous proof of structure for such a compound was by chemical degradation,[7] in the case of *N*(α)-benzyloxycarbonyl-*N*(τ)-2,4-dinitrophenyl-L-histidine. The present case also establishes the position of the *im*-substituent in (2) and (3) and it would seem that it is safe to generalise that the major product of *im*-protection of histidine side chains by electrophilic reagents is always that which is derivatised at the τ nitrogen. This is obviously not an unexpected result but is nonetheless a useful thing to have firmly established, since differentiation between *N*(τ)- and *N*(π)-isomers is not a trivial matter and there has been much confusion over it in the past. Barring crystallography or chemical degradation, which is only feasible in special cases, the only available criterion is the empirical rule enunciated by Matthews and Rapoport.[8] This states that the cross-ring coupling constant between the imidazole ring protons in 1,4-disubstituted imidazoles, of which *N*(τ)-blocked histidines are examples, lies in the range 1.1—1.5 Hz whereas that for the 1,5- or *N*(π)-isomers lies in the range 0.9—1.0 Hz. These coupling constants are often difficult or impossible to measure with precision; in complex structures, because of line broadening and/or overlapping, it is not possible to measure them at all. Structural assignments on this basis alone have so far not commanded our complete confidence since the coupling constants are small, the ranges

Table 1. Bond lengths for (1) in Å with e.s.d.s in parentheses

	Molecule 1	Molecule 1A
O(1)–C(4)	1.422(10)	1.392(9)
O(1)–C(5)	1.383(12)	1.422(10)
O(2)–C(14)	1.198(7)	1.252(7)
O(3)–C(14)	1.186(8)	1.237(7)
O(4)–C(15)	1.209(6)	1.208(7)
O(5)–C(15)	1.356(6)	1.331(7)
O(5)–C(16)	1.434(7)	1.445(6)
N(1)–C(1)	1.296(8)	1.305(9)
N(1)–C(3)	1.377(7)	1.370(9)
N(2)–C(1)	1.329(8)	1.324(8)
N(2)–C(2)	1.384(7)	1.434(8)
N(2)–C(4)	1.471(9)	1.443(9)
N(3)–C(13)	1.445(6)	1.451(6)
N(3)–C(15)	1.317(7)	1.343(7)
C(2)–C(3)	1.359(8)	1.355(9)
C(2)–C(12)	1.490(7)	1.498(7)
C(5)–C(6)	1.501(8)	1.496(8)
C(6)–C(7)	1.361(8)	1.379(8)
C(6)–C(11)	1.373(8)	1.382(8)
C(7)–C(8)	1.383(8)	1.381(8)
C(8)–C(9)	1.372(9)	1.369(9)
C(9)–C(10)	1.351(9)	1.380(9)
C(10)–C(11)	1.382(8)	1.385(9)
C(12)–C(13)	1.528(7)	1.533(7)
C(13)–C(14)	1.535(7)	1.526(7)
C(16)–C(17)	1.518(7)	1.532(7)
C(16)–C(18)	1.525(8)	1.521(7)
C(16)–C(19)	1.518(8)	1.528(7)

† A negative value corresponds to an anticlockwise rotation when viewed from atom 3 to atom 2.



Scheme 1. *Conditions*: i, $(Boc)_2O$ or TrtCl; ii, $PhCH_2OCH_2Cl$; iii, methanol; iv, aqueous NaOH

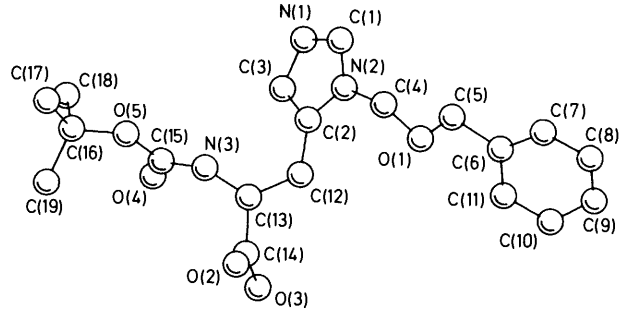


Figure 1. A ball and stick view of molecule 1, drawn with SNOOPI. (E. K. Davies, ' SNOOPI User Guide,' Chemical Crystallography Laboratory, University of Oxford, 1981.)

for the two substitution patterns abut each other, and the rule seemed to be based on only a modest number of examples. The present case, however, further substantiates the rule: the ester (4) exhibited a cross-ring coupling constant of 0.9 Hz but that of the isomer (5) prepared as in Scheme 2 was 1.5 Hz. We have yet to encounter a case which contradicts it, although we have had many instances where its application was frustrated by our inability to make the necessary measurements.

Experimental

N(α)-t-*Butoxycarbonyl*-N(π)-*benzyloxymethyl*-L-*histidine Hydrate* (1).—The anhydrous acid was prepared as described previously [3] and crystallised from aqueous methanol to give the *hydrated acid*, m.p. 108—110 °C, $[\alpha]_D^{20}$ +6.7° (*c* 0.5, in MeOH) (Found C, 58.4; H, 6.5; N, 10.75. $C_{19}H_{25}N_3O_5 \cdot H_2O$ requires C, 58.0; H, 6.4; N, 10.7%)

Crystal Data for (1).—$C_{19}H_{25}N_3O_5 \cdot H_2O$, M_r = 393.5, triclinic, *P*1, *Z* = 2, *a* = 8.750 (1), *b* = 9.689 (2), *c* = 13.466 (3) Å, α = 98.09 (2), β = 92.75 (2), γ = 108.51 (2)°, *U* = 1066.7 Å^3, μ = 0.99 cm^{-1} (Mo-K_α), final *R*-value 0.061.

X-*Ray Crystal Structure Determination of* (1).—A crystal of (1) was grown by slow evaporation of a solution in aqueous methanol. The crystal was mounted on a glass fibre and transferred to the goniometer of a computer-controlled Enraf Nonius CAD4-F *X*-ray diffractometer. Accurate cell parameters were determined from the setting angles of 25 strong reflections located by the SEARCH routine. The crystal

Table 2. Bond angles for (1) in ° with e.s.d.s in parentheses

C(4)-O(1)-C(5)	115.5(7)	114.0(6)
C(15)-O(5)-C(16)	122.6(4)	121.8(4)
C(1)-N(1)-C(3)	108.5(5)	110.3(5)
C(1)-N(2)-C(2)	108.6(5)	107.7(5)
C(1)-N(2)-C(4)	125.7(5)	124.8(6)
C(2)-N(2)-C(4)	125.5(5)	127.4(5)
C(13)-N(3)-C(15)	122.0(4)	120.5(4)
N(1)-C(1)-N(2)	109.7(5)	109.3(6)
N(2)-C(2)-C(3)	105.4(5)	105.3(5)
N(2)-C(2)-C(12)	122.2(5)	120.4(5)
C(3)-C(2)-C(12)	132.4(5)	134.2(6)
N(1)-C(3)-C(2)	107.8(5)	107.4(6)
O(1)-C(4)-N(2)	111.2(6)	110.3(6)
O(1)-C(5)-C(6)	107.1(8)	107.4(7)
C(5)-C(6)-C(7)	118.1(6)	120.3(6)
C(5)-C(6)-C(11)	123.7(6)	119.9(6)
C(7)-C(6)-C(11)	118.1(5)	119.7(5)
C(6)-C(7)-C(8)	120.4(6)	120.3(6)
C(7)-C(8)-C(9)	120.2(6)	120.1(6)
C(8)-C(9)-C(10)	118.7(6)	120.1(6)
C(9)-C(10)-C(11)	119.9(6)	120.1(6)
C(6)-C(11)-C(10)	121.0(6)	119.6(6)
C(2)-C(12)-C(13)	112.7(4)	111.7(4)
N(3)-C(13)-C(12)	110.0(4)	110.7(4)
N(3)-C(13)-C(14)	111.3(4)	113.2(4)
C(12)-C(13)-C(14)	108.2(4)	108.5(4)
O(2)-C(14)-O(3)	121.2(7)	124.9(5)
O(2)-C(14)-C(13)	119.0(5)	114.3(5)
O(3)-C(14)-C(13)	119.4(6)	120.8(5)
O(4)-C(15)-O(5)	123.9(5)	123.6(5)
O(4)-C(15)-N(3)	126.0(5)	123.3(5)
O(5)-C(15)-N(3)	110.1(4)	111.0(5)
O(5)-C(16)-C(17)	103.9(4)	111.0(5)
O(5)-C(16)-C(18)	110.7(5)	104.2(4)
O(5)-C(16)-C(19)	109.9(5)	110.1(4)
C(17)-C(16)-C(18)	110.3(6)	110.3(5)
C(17)-C(16)-C(19)	110.6(6)	110.8(5)
C(18)-C(16)-C(19)	111.2(6)	110.3(5)

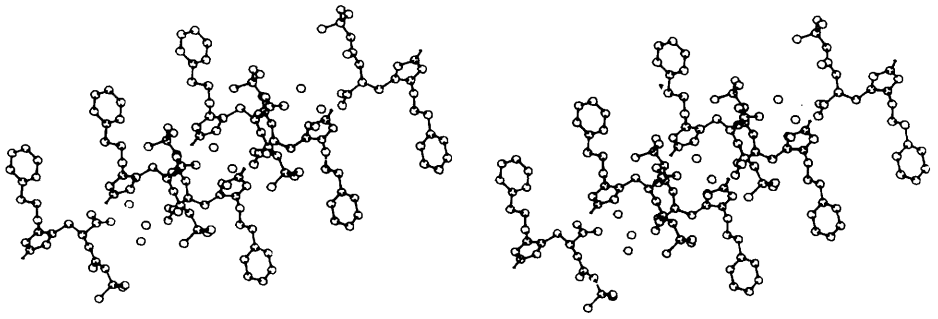
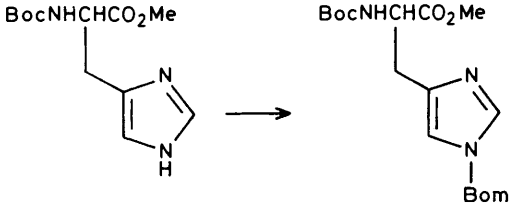


Figure 2. A stereo-view * of the molecular packing along the *a* axis in a crystal of (1). (* See caption to Figure 1.)

Table 3.

	(1)		*N*(α)-Acetyl-L-histidine monohydrate [5]		L-Histidine hydrochloride monohydrate [6]
	Molecule 1	Molecule 1A	Molecule A	Molecule B	
N(1)-C(1)	1.296(8)	1.305(9)	1.314	1.315	1.313(4)
N(1)-C(3)	1.377(7)	1.370(9)	1.370	1.370	1.377(4)
N(2)-C(1)	1.329(8)	1.324(8)	1.323	1.326	1.340(4)
N(2)-C(2)	1.384(7)	1.434(8)	1.371	1.374	1.383(4)
C(2)-C(3)	1.359(8)	1.355(9)	1.356	1.356	1.356(4)



(5)

Scheme 2. *Conditions:* $PhCH_2OCH_2Cl$

system was triclinic and the space group assumed to be *P*1 since the compound was known to be chiral. Diffraction intensities were measured by ω/2θ scans out to a maximum of $2\theta = 49°$ to give 3 744 unique reflections after the application of Lorentz and polarisation corrections and the merging of equivalent reflections. Initial attempts to solve the structure with MULTAN 80 [9] were unsuccessful, including the variation of the molecular fragments used in the calculation of the modified structure factors (*E*'s) and the variation of the choice of origin-defining reflections. *N*(α)-Acetyl-L-histidine monohydrate [5] also crystallises in the space group *P*1 with 2 independent molecules in the unit cell. The co-ordinates of all the non-hydrogen atoms common to (1) from this structure (excluding the acetyl methyl carbon), were used to form a bi-molecular fragment for which MULTAN 80 molecular scattering factors were calculated and used along with those for two benzene rings to calculate the *E*'s. The resulting best solution gave two promising 5 atom fragments. These were input as correctly positioned fragments in a re-run of MULTAN 80 and 5 further possible atoms were located. After a further round of MULTAN 80, a 17 atom fragment corresponding to one molecule of (1) without the t-butyl group or the benzene ring was used in a difference Fourier synthesis to locate most of the other molecule. After initial full-matrix refinement of the positional co-ordinates all the remaining non-hydrogen atoms in the molecules, barring 2 phenyl carbons, were located from a difference Fourier map. The refinement proceeded by blocked-matrix least squares with the parameters for each molecule blocked separately, and including isotropic temperature factors. Subsequent difference Fourier syntheses revealed the positions of the missing carbons and also of three oxygen atoms belonging to molecules of water of crystallisation. Two of the latter were later assigned fractional occupation numbers whose sum was constrained to unity. Waser constraints [10] were used in the refinement but only those on the t-butyl groups and benzyl groups were retained until convergence. The refinement was continued with anisotropic temperature factors, and then hydrogen atoms were placed geometrically and assigned an isotropic temperature factor of 0.085 except those on the benzyl groups which were assigned a higher value of 0.15. Hydrogens were also placed on N(1) and N(1A) but not on any of the carboxy oxygens since a difference Fourier map clearly suggested this to be correct for molecule 1 and since for each carboxy-group the C–O bond lengths were equivalent. Weights for the final rounds of refinement were computed from the Chebyshev series [11] $w = [274.2t_0(X) + 345.3t_1(X) + 81.2t_2(X)]^{-1}$ where $X = F_o/F_{max.}$ The structure converged with an *R*-value of 0.061. All calculations were performed with the CRYSTALS [12] package on a VAX 11/750 computer in the Chemical Crystallography Laboratory. Final fractional atomic co-ordinates, temperature factors and structure factor tables are given in Supplementary Publication No. 23436 (34 pp.).†

Measurement of the Cross-ring Coupling Constants of the Esters (4) *and* (5).—The esters were prepared as outlined in Schemes 1 and 2 and described in detail previously.[3] The cross-ring coupling constants were measured on a Bruker WH 300 instrument operating at 300 MHz, in $CDCl_3$ at room temperature, using the imidazole 2-H resonance, which was clear of other bands in both cases and which could be simplified to a doublet by irradiation of the resonance due to the side-chain methylene group.

† For details of the Supplementary publications scheme, see Notice to Authors No. 7, *J. Chem. Soc., Perkin Trans. 1*, 1981, Index issue.

References

1 A. R. Fletcher, J. H. Jones, W. I. Ramage, and A. V. Stachulski, *J. Chem. Soc., Perkin Trans. 1*, 1979, 2261.

2 D. F. Veber in ' Peptides: Chemistry, Structure, and Biology,' (Proc. 4th American Peptide Symp.), eds. R. Walter and J. Meienhofer, Ann Arbor Science Publishers, Ann Arbor, 1975, p. 307.

3 T. Brown, J. H. Jones, and J. D. Richards, *J. Chem. Soc., Perkin Trans. 1*, 1982, 1553.

4 I. Bennett, A. G. H. Davidson, M. M. Harding, and I. Morelle, *Acta Crystallogr., Sect. B*, 1970, **26**, 1722.

5 T. J. Kistenmacher, D. J. Hunt, and R. E. Marsh, *Acta Crystallogr., Sect. B*, 1972, **28**, 3352.

6 K. Oda and H. Koyama, *Acta Crystallogr., Sect. B*, 1972, **28**, 639.

7 J. R. Bell and J. H. Jones, *J. Chem. Soc., Perkin Trans. 1*, 1974, 2336.

8 H. R. Matthews and H. Rapoport, *J. Am. Chem. Soc.*, 1973, **95**, 2297.

9 P. Main *et al.*, ' MULTAN 80. A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data,' Department of Physics, University of York, 1980.

10 J. Waser, *Acta Crystallogr.*, 1963, **16**, 1091.

11 J. R. Carruthers and D. J. Watkin, *Acta Crystallogr., Sect. A*, 1979, **35**, 698.

12 D. J. Watkin and J. R. Carruthers, ' CRYSTALS User Guide ,' Chemical Crystallography Laboratory, University of Oxford, 1981.

Received 16*th June* 1982; *Paper* 2/1012