Methods of Peptide Sequencing. Part II.[1] Cyclisation of *N*-2-Amino-6-nitrophenyl and *N*-3-Amino-2-pyridyl Derivatives of Amino-acids and Peptides

By **Robert A. W. Johnstone*** and **T. Jeffrey Povall**, The Robert Robinson Laboratories, The University, Liverpool L69 3BX
Ian D. Entwistle, Shell Research Ltd., Sittingbourne, Kent

Peptides and amino-acids were treated with either 1-fluoro-2.6-dinitrobenzene or 2-fluoro-3-nitropyridine to give *N*-nitroaryl derivatives. The latter were reduced by transfer catalysis to *N*-aminoaryl derivatives. which cyclised with, in the peptide case, release of the peptide minus its *N*-terminal amino-acid residue. The cyclised materials were readily identified by mass spectrometry. This cycle of operations was used to obtain a partial sequence of amino-acids in a peptide.

THE Edman method for sequencing amino-acid residues in a peptide chain is widely used with various modifications.[2] In this method, the free amino-end of a peptide is treated with an isothiocyanate to release the *N*-terminal amino-acid residue as a thiohydantoin; the remaining peptide can be subjected repetitively to the same cycle to yield the sequence of amino-acid residues. 1-Fluoro-2,4-dinitrobenzene [3] and 2-fluoro-3-nitropyridine [4] have been used in determining the *N*-terminal amino-acid in a peptide but a major disadvantage of these reagents is that, in obtaining the derivative of the *N*-terminal amino-acid, the rest of the peptide is destroyed.

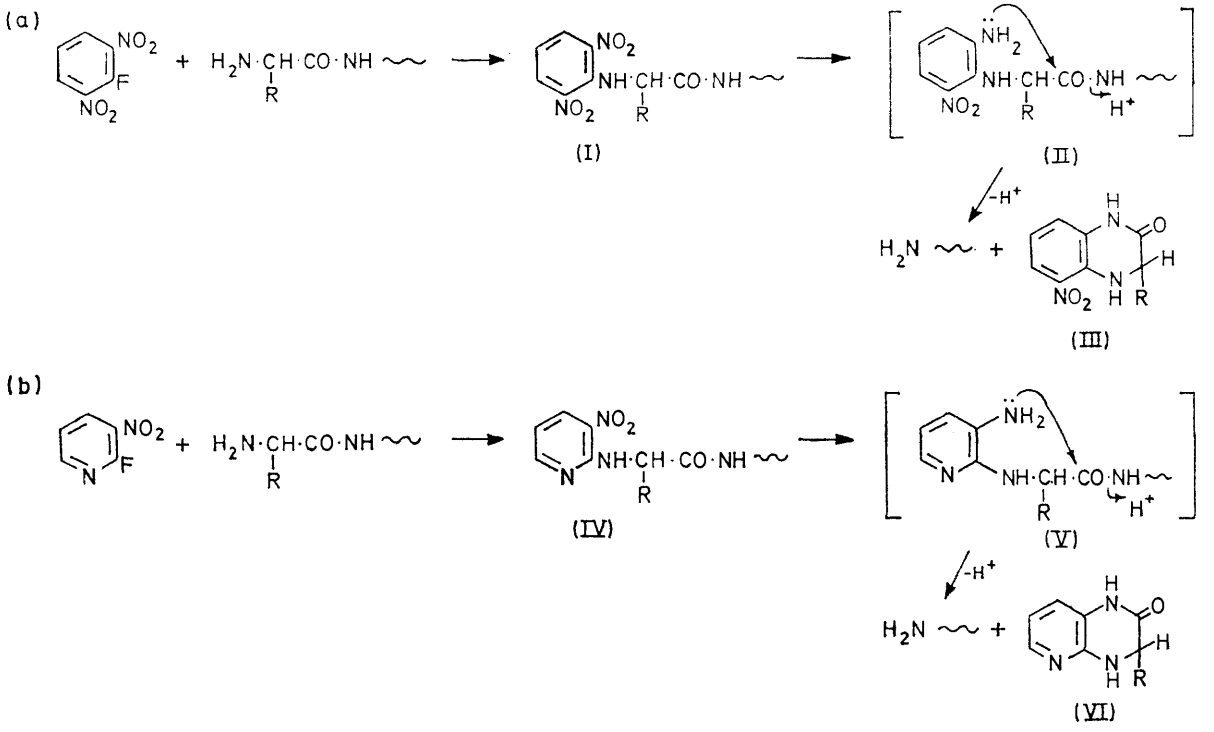
In a variation of this method, 2,4-dinitrophenyl derivatives of simple peptides were hydrogenated catalytically and treated with acid to form substituted aminoquinoxalines.[5] However the latter are readily oxidised in air. Other derivatives tried include 4-methoxycarbonyl-2-nitrophenyl [6] and 5,7-dinitrobenz-

[1] Part I, R. A. W. Johnstone and T. J. Povall, *J.C.S. Perkin I*, 1975, 1297.
[2] See for example, W. A. Schroeder, *Methods Enzymol.*, eds. C. H. W. Hirs and S. N. Timasheff, 1972, **25**, 298.
[3] F. Sanger, *Biochem. J.*, 1945, **39**, 507; *Bull. Soc. chim. biol.*, 1955, **37**, 23; A. Niederwiesser, *J. Chromatog.*, 1971, **54**, 215.
[4] A. Signor, L. Biondi, A. M. Tamburro, and E. Bordignon, *European J. Biochem.*, 1969, **7**, 328; E. D. Bergmann and M. Bentov, *J. Org. Chem.*, 1961, **26**, 1480.
[5] M. Jutisz and W. Ritschard, *Biochim. Biophys. Acta*, 1955, **17**, 548; V. M. Ingram, *ibid.*, 1956, **20**, 577.
[6] R. W. Holley and A. D. Holley, *J. Amer. Chem. Soc.*, 1952, **74**, 5445.

imidazol-4-yl;[7] by these methods, the *N*-terminal amino-acid residue can be removed and the remaining peptide subjected again to the cycle of operations.

The reduction of dinitro-compounds to the corresponding monoaminomononitro-compounds can be effected easily and in high yield through the use of a palladium catalyst and cyclohexene;[8] this type of reduction in which hydrogen is transferred to or from cyclohexene from or to a substrate *via* a metal catalyst has been used previously in other situations.[9]

The preparation of a nitroaryl derivative [(I) or (IV)] of a peptide followed by reductive cyclisation to form a tetrahydroquinoxaline (III) or a tetrahydropyridopyrazine (VI) with release of the remaining peptide forms the basis of the method described here (see Scheme). In the following discussion, results for dinitrophenyl derivatives (I) are dealt with and any distinct differences from the nitropyridyl derivatives (IV) are considered afterwards.



SCHEME

With 1-fluoro-2,6-dinitrobenzene, amino-acids and peptides with free amino-groups gave 2,6-dinitrophenyl derivatives (I) which, in these small-scale experiments, were purified by t.l.c. The dinitro-compounds (I) were 'half-reduced' to the corresponding monoaminomononitro-compounds (II) with a slight excess by weight of palladium–charcoal catalyst in refluxing ethanol containing cyclohexene. Under these conditions, the monoamino-derivatives (II) cyclised to the tetrahydroquinoxalines (III) and could not be isolated as intermediates; at the same time, the remaining peptide was released. The deep orange tetrahydroquinoxalines (III) were easily detectable by t.l.c. and afforded simple mass spectra by which they could be identified. The normal amino-acids (Tables 1—4) formed cyclic products (III) readily but one or two small variations were noted. Methionine gave the expected product (III; R = $CH_2 \cdot CH_2 \cdot SMe$) but the reaction was much slower than with the other amino-acids, probably owing to sulphur poisoning the catalyst.

The nitropyridyl derivatives (IV) behaved similarly but the reduction step was significantly faster and the cyclisation to pyridopyrazines (VI) much slower, so that the intermediates (V) could be isolated. Addition of acetic acid was necessary to make the cyclisation proceed at a similar rate to that of the dinitrophenyl compounds (I). The nitropyridyl derivative from dimethyl glutamate on reductive cyclisation gave the pyroglutamyl compound (VII) as well as the expected product (VI; R = $CH_2 \cdot CH_2 \cdot CO_2Me$) but the derivative from glutamine gave exclusively the pyroglutamyl compound (VII). The derivative from cysteine was reduced to the cyclic compound corresponding to alanine (VI; R = Me); in MeOD, reductive cyclisation gave a deuterioalanine derivative (VI) containing 15%

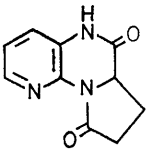
[7] K. L. Kirk and L. A. Cohen, *J. Org. Chem.*, 1969, **34**, 384, 390, 395.

[8] I. D. Entwistle, R. A. W. Johnstone, and T. J. Povall, *J.C.S. Perkin I*, 1975, 1300.

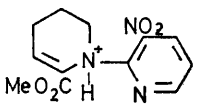
[9] E. A. Braude, R. P. Linstead, and K. R. H. Wooldridge, *J. Chem. Soc.*, 1954, 3586.

of deuterium (mass spectrometry). The tetrahydropyridopyrazines (VI) were colourless but readily detectable as blue compounds under u.v. light.

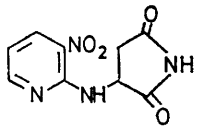
The reaction of peptides with 1-fluoro-2,6-dinitrobenzene or 2-fluoro-3-nitropyridine followed by reductive cyclisation proceeded normally to give, respectively, the



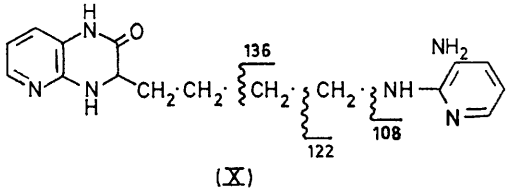
(VII)



(VIII)



(IX)



(X)

tetrahydroquinoxaline (III) and the tetrahydropyridopyrazine (VI) corresponding to the *N*-terminal amino-acid residue with release of the remaining peptide. The peptides H-Leu-Phe-OMe and H-Gly-Leu-Phe-OMe, each gave the cyclic product corresponding to the *N*-terminal amino-acid and a ninhydrin-positive residue, characterised as the 2,6-dinitrophenyl derivative of H-Phe-OMe or H-Leu-Phe-OMe, respectively. Similarly, the tetrapeptides H-Val-Gly-Leu-Phe-OMe and H-Ala-Tyr-Leu-Phe-OMe gave the cyclic products corresponding to valine and alanine, respectively.

A larger peptide was used to find to what extent the reductive cyclisation step could be repeated along a chain. With 20 μmol of the octapeptide H-Glu(OMe)-Glu(OMe)-Ala-Glu(OMe)-Glu(OMe)-Ala-Tyr-Gly-OMe and by using the nitropyridyl derivative, the amino-acid residues 1—5 were identified by comparison (t.l.c. and mass spectra) of their cyclic products (VI) with authentic samples. The final glutamic residue was only just detectable and the next residue (alanine) was not found. During the sequence other unidentified products began to accumulate, making detection of amino-acid residues increasingly difficult. Some of these products may be cyclic dipeptides: peptides are known [10] to give these substances on refluxing in acetic acid although at higher temperatures and for longer times than those used in these experiments. The use of 1-fluoro-2,6-dinitrobenzene or 2-fluoro-3-nitropyridine is not yet as satisfactory on a small scale as the Edman method. Further work aimed at decreasing the formation of by-products and extracting impurities at each step is desirable.

Mass spectrometric data are shown in the Tables. Molecular ions were produced in low abundance with 2,6-dinitrophenyl (I) and 3-nitropyridyl (IV) compounds and in some cases were accompanied by $M + 1$ ions of similar abundance. In all the compounds studied, the *C*-terminus of the amino-acid or peptide was protected by formation of the methyl ester, and extremely abundant $M - 59$ ions (loss of CO_2Me) were always present for compounds (I) and (IV). Other abundant ions corresponded either to loss of an amino-acid side-chain or to the side-chain itself. An extremely abundant ion at *m/e* 264 from the nitropyridyl derivative of lysine is probably due to a fragment of type (VIII). From the nitropyridyl derivative of asparagine, a second spectrum

TABLE 1

Mass spectra of 2,6-dinitrophenyl derivatives (I) [a]

Gly [b] 196,* 209, *255*; Val 238, 254, *297*; Phe 254,* 286, 291, 299, *345*; Trp 77, 103, 130,* 196, 254, 325, *384*; Ser 196, 226,* 254, *285*; (pyr)Orn 107, 108, 124, 136,* 150, 164, 176, 236, 252, *418*, 419; Met 61,* 75, 196, 222, 237, 238, 254, 255, 282, 283, 330; met(O_2) 222,* 254, 302, *361*; Tyr 107,* 196, 254, 255, 302, *361*, 362; Tyr(Dnp) 167, 254, 273,* 344, 360, 468, *527*; His 81,* 167, 196, 254, 276, 290, 335, 349; His(Dnp) 247, 276, 335, 442,* *501*; Leu-Phe 252,* 399, 412, 427, 443, *458*, 459; Gly-Leu-Phe 196, 291, 309,* 337, 349, 353, 459, 469, 484, *515*, 516

[a] All of type 2,6-$(NO_2)_2C_6H_3$·X·OMe with X = amino-acid or peptide residue. [b] In this and other Tables only the more prominent ions are recorded; the value corresponding to the base peak is marked by an asterisk and that corresponding to the molecular ion is italicised.

TABLE 2

Mass spectra of 3-nitropyridyl derivatives (IV) [a]

Phe 91, 139, 162, 210,* 242, 270, *301*; Gly(OMe) 115, 123, 178, 206, 210, 238,* 266, *297*; Gln 123, 140, 178,* 206, 210, 223, 233, 238, 250, 251, 265, 266, 282, 283; Asn 192,* 209, 237, 251, 252, *268*, 269; Lys 152, 178, 192, 194, 206,* 210, 264, 282, 373, *404*, 405; Cys 164, 169, 181, 192, 210, 224,* 240, 320, *379*, 380; Cys(CH_2·CO_2·CH_3) 139, 158, 190, 196, 210,* 224, 256, 270, 298, *329*; Leu-Phe 91, 152, 208,* 236, 358, 383, *414*, 415

[a] All of type 3-nitro-2-pyridyl-X(OMe) with X = amino-acid or peptide residue.

TABLE 3

Mass spectra of the 3-R-substituted quinoxalines (III) [a]

H(Gly) 146(14), 147(16), 191(24), 192(28), *193*(100); Me(Ala) 146(22), 192(100), 193(13), 206(41), *207*(57); Me_2CH(Val) 146(31), 192(100), 193(29), *235*(23); Me_2CH·CH_2(Leu) 146(64), 192(100), 193(20), *249*(16); $PhCH_2$(Phe) 91(89), 146(34), 192(100), *283*(11); indol-3-ylmethyl(Trp) 77(14), 130(100), 146(6), 192(3), 193(4), *322*(14); $HOCH_2$(Ser) 146(57), 192(100), 193(48), 223(22); $MeSCH_2$·CH_2(Met) 146(69), 192(100), 193(18), 221(31), *267*(18)

[a] Data of the form R(amino-acid residue) *m/e* (% relative abundance).

TABLE 4

Mass spectra of the 3-R-substituted pyrido[2 3-*b*]-pyrazines (VI) [a]

Me(Ala) 93(28), 120(100), 134(15), 148(78), 162(8), *163*(75); $PhCH_2$(Phe) 91(20), 93(14), 119(5), 120(27), 148(100), *239*(10); MeO_2C·CH_2·CH_2 [Glu(OMe)] 93(34), 119(16), 120(100), 148(96), 161(36), 162(13), 175(25), 176(5), 203(14), 204(9), *235*(18); H_2N·CO·CH_2 (Asn) 120(151), 148(75), 161(100), 162(48), 189(20), *206*(30); 3-amino-2-pyridyl-NH·CH_2·CH_2 (Lys) 108(25), 109(100), 119(9), 120(19), 122(54), 135(45), 136(25), 148(22), 176(5), 202(12), 312(50); (Gln) [b] 120(100), 148(47), 161(10), 174(7). 175(7), *203*(37)

[a] Data of the form R(amino-acid residue) *m/e* (% relative abundance). [b] Material underwent further cyclisation to pyroglutamyl derivative; see (VII).

[10] J. D. Baty, R. A. W. Johnstone, and T. J. Povall, *J.C.S. Chem. Comm.*, 1973, 392.

was obtained after the compound had been in the mass spectrometer for a few minutes; this showed a molecular ion at *m/e* 236, probably due to the succinimide (IX) formed by a thermolytic reaction in the spectrometer.

The tetrahydroquinoxalines (III) gave simple mass spectra having abundant molecular ions and relatively few but prominent fragment ions. Two ions at *m/e* 192 and 146 were present in all cases due respectively to loss of the amino-acid side-chain and of both side-chain and nitro-group. Less satisfactory spectra were obtained for the disubstituted histidine and pyro-ornithine derivatives. No molecular ions were present but the *m/e* 192 fragment was abundant. In the case of the histidine derivative, an $M + 1$ ion was present and in that of the ornithine derivative, ions due to the side-chain were present (*m/e* 136, 150, 164, and 176).

The tetrahydropyridopyrazines (VI) also gave simple mass spectra with abundant molecular ions; two other abundant ions at *m/e* 148 and 120 were always found and correspond to loss of R and R + CO from the molecular ion. The compound (VI; R = CH_2·CO·NH_2) derived from asparagine gave an abundant ion at *m/e* 161 (loss of $CONH_2$ + H), and the glutamine derivative gave a molecular ion at *m/e* 203 rather than at the expected value of *m/e* 220 (loss of NH_3), showing that a pyroglutamyl compound (VII) had been formed. In the case of the disubstituted lysine compound (X), the base peak was at *m/e* 109 and other abundant ions were due to side-chain fragmentation.

EXPERIMENTAL

Mass spectra were recorded at 70 eV on an A.E.I. MS 12 spectrometer by direct insertion. All t.l.c. was carried out on silica gel. R_F Values for 2,6-dinitrophenyl (Dnp) derivatives of amino-acid methyl esters (Dnp-NH·CHR·CO_2Me) in benzene–methanol (either 15 : 2 or 15 : 1) were respectively: Gly 0.77, 0.70; Val 0.78, 0.72; Phe 0.83, 0.75; Trp 0.68, 0.60; Met 0.73, 0.72; Ser 0.50, 0.38; (pyr)Orn 0.66, 0.55; Tyr 0.53, 0.39; tyr(Dnp) 0.70, 0.64; His 0.39, 0.17; his(Dnp) 0.72, 0.71. Similarly R_F values for 3-nitropyridyl (Np) derivatives (Np–NH·CHR·CO_2Me) were: Leu 0.80, 0.74; Phe 0.80, 0.72; Glu(OMe) 0.73, 0.64; Asn 0.21, 0.11; Gln 0.24, 0.13; cys(Np) 0.77, 0.73; Lys(Np) 0.75, 0.69; cys(CH_2·CO_2Me) 0.72, 0.67; and for tetrahydroquinoxolines (III) were: (derived from) Gly 0.43, 0.29; Ala 0.44, 0.39; Val 0.50, 0.41; Leu 0.52, 0.47; Phe 0.53, 0.44; Trp 0.35, 0.24; Tyr 0.24, 0.12; Met 0.49, 0.36; Ser 0.21, 0.08.

2 6-*Dinitrophenyl Derivatives* (I).—The hydrochloride of serine methyl ester (3 mg) was shaken for 2 h at 20 °C in a solution of 1-fluoro-2,6-dinitrobenzene (7 mg) in aqueous 1% sodium hydrogen carbonate (1 ml) and ethanol (1 ml). The solution was acidified (2M-HCl) and the yellow product extracted with ethyl acetate and purified by t.l.c. [benzene–methanol (6 : 1)]. Other amino-acid and peptide esters were treated similarly.

3-*Nitro-2-pyridyl Derivatives* (IV).—The hydrochloride of lysine methyl ester (5 mg) was shaken for 2 h at 20 °C in ethanol–water (80 : 20; 3 ml; adjusted to pH 9 with Et_3N) containing 2-fluoro-3-nitropyridine (15 mg). The product was poured into ethyl acetate–water (1 : 1; 20 ml) and the organic phase was dried and evaporated to give the yellow bis-3-nitropyridyl derivative, purified by t.l.c. [benzene–methanol (15 : 2)]. Other amino-acid and peptide esters were treated similarly.

Reduction and Cyclisation Steps.—The 2,6-dinitrophenyl derivative of tryptophan methyl ester (4 mg) was heated under reflux in ethanol (4 ml) with cyclohexene (0.05 ml) and 10% palladium–charcoal (5 mg) for 60 min. The yellow solution became orange and was filtered from the catalyst to give 3,4-dihydro-3-(indol-3-ylmethyl)-5-nitroquinoxalin-2(1*H*)-one (III; R = indol-3-ylmethyl), purified by t.l.c. [benzene–methanol (6 : 1)]. As a further example, the 3-nitropyridyl derivative of asparagine methyl ester (3 mg) was heated under reflux in ethanol–ethyl acetate (1 : 1; 3 ml) in the presence of cyclohexene (0.05 ml), 10% palladium–charcoal (5 mg), and acetic acid (1 ml) for 45 min. The yellow solution became colourless and was filtered from catalyst to give 3-carbamoylmethyl-3,4-dihydropyrido[2,3-*b*]pyrazin-2(1*H*)-one (VI; R = CH_2·CO·NH_2) purified by t.l.c. [benzene–methanol (6 : 1)]. All other reductions and cyclisations were carried out similarly; in the absence of acetic acid the 3-nitropyridyl derivatives were quickly reduced but cyclised slowly so that the intermediate 3-aminopyridyl compounds (V) could be isolated.

Sequencing of an Octapeptide.—The protected octapeptide Boc-Glu(OBu^t)-Glu(OBu^t)-Ala-Glu(OBu^t)-Glu(OBu^t)-Ala-Tyr-Gly-OMe (25 mg) * was treated with 90% trifluoroacetic acid followed by methanolic hydrogen chloride to give the hydrochloride of the peptide H-Glu(OMe)-Glu(OMe)-Ala-Glu(OMe)-Glu(OMe)-Ala-Tyr-Gly-OMe; the latter was heated under reflux in 90% ethanol (5 ml) with triethylamine (0.05 ml) and 2-fluoro-3-nitropyridine (16 mg) for 35 min. The solvents were evaporated off to give a yellow solid which was heated under reflux in ethanol (4.3 ml) and acetic acid (0.7 ml) with cyclohexene (0.05 ml) and 10% palladium–charcoal (10 mg) for 2 h. The catalyst was filtered off and the solution evaporated to leave a residue (A) which was extracted with dichloromethane. The extract was filtered and evaporated *in vacuo* to give the pyridopyrazine (VI; R = CH_2·CH_2·CO_2Me), identified by t.l.c. and mass spectral comparison with an authentic sample. The derivatisation and reduction–cyclisation steps were repeated with residue A to give successively the pyridopyrazines from Glu(OMe) twice, Ala, Glu(OMe) twice, and Glu(OMe); at the sixth cycle, no pyridopyrazine corresponding to alanine was detected.

[5/148 *Received*, 23*rd January*, 1975]

* Kindly donated by Professor G. W. Kenner.