Gel phase ^{13}C n.m.r. spectroscopy as an analytical method in solid (gel) phase peptide synthesis

R. Epton and P. Goddard
Department of Physical Sciences, The Polytechnic, Wulfruna Street, Wolverhampton WV1 1LY, UK

and K. J. Ivin
Department of Chemistry, The Queen's University of Belfast, Belfast BT9 5AG, UK
(*Received* 2 *October* 1980)

In solid (gel) phase peptide synthesis, the lack of practical analytical methods to effect non-destructive, stage-by-stage characterization of the growing, matrix-bound, peptide chains is a serious problem. Each peptide-chain-elongation cycle involves removal of an *N*-terminal protecting group from the matrix-bound peptide chain (deprotection) followed by coupling, of the *N*-terminal amino group so exposed, with a further activated, *N*-protected, amino acid. It is essential to ensure that both the deprotection and coupling steps are quantitative. The coupling step can be monitored to completion by withdrawal of a few beads from the solid (gel) phase reactor and testing these destructively for primary amino groups using the highly sensitive fluorescamine reaction[1]. Unfortunately, there is no corresponding test to check for quantitative removal of the matrix-bound peptide *N*-terminal protecting groups. It is inevitable that any residual protected amino groups will become exposed in a subsequent chain-elongation cycle. Erroneous peptide sequences will then result.

Recently, Manatt *et al.*[2] have reported a method of monitoring both *N*-terminal amino-acid coupling and matrix-bound peptide *N*-terminal deprotection steps by ^{19}F n.m.r. spectroscopy. For this work, a strategy of solid (gel) phase peptide synthesis, using a crosslinked polystyrene matrix and purpose-synthesized fluorobenzyloxycarbonyl- and 2-(4-fluorophenyl)-2-propyloxycarbonyl amino acids, was devised. Manatt *et al.* undertook their ^{19}F n.m.r. study in the light of some earlier, encouraging, ^{13}C n.m.r. results for the solvent-swollen crosslinked polystyrene resins which had been derivatized with chloromethyl or hydroxymethyl groups[3].

Of the non-destructive spectroscopic techniques applicable to peptides in free solution, ^{13}C n.m.r. is potentially the most informative because it is possible to assign unequivocally most of the lines in the ^{13}C n.m.r. spectra of quite complex peptides. Leibfritz *et al.*[4] have demonstrated that high resolution ^{13}C n.m.r. spectroscopy may be applied to constitutional and conformational studies on peptides covalently bound to specially-developed polyoxyethylene supports. The peptide component of these conjugates was synthesized stepwise on the polymer support using solution phase methods. Purification at intermediate stages was assisted by crystallization of the polymer.

Here we report our preliminary observations on the direct application of high resolution ^{13}C n.m.r. spectroscopy to characterize intermediate stages in polymer-supported peptide synthesis on a phenolic crosslinked poly(acryloylmorpholine)-based matrix[5]. This support has the advantage over crosslinked polystyrenes that it undergoes gelation (constrained dissolution) best in highly polar organic solvents which are also good solvents for protected peptides. This is a very desirable situation both from the point of view of synthesis and of characterization by ^{13}C n.m.r.

The copolymer support matrix was synthesized by radical copolymerization of *N*-acryloylmorpholine (1 mol), *N*,*N*′-diacryloyl piperazine (0.05 mol) (as crosslinking agent), and *N*-[2-(4-acetoxyphenyl)ethyl]-acrylamide (0.2 mol) (to provide the starting point for building the peptide chain). The ^{13}C n.m.r. spectra were obtained using a Bruker WH90 *FT* n.m.r. spectrometer operating at 22.63 MHz, with broad band irradiation to decouple the protons. Satisfactory signal-to-noise ratio was achieved using 30 000–150 000 pulses at 0.7 s intervals with a sample of 0.15 g swollen in 2 cm^3 DMSO-d_6. Line positions were referred to the centre DMSO-d_6 line, taken as 39.67 ppm ($\equiv$TMS = 0). The spectral width was 6024 Hz (1 channel $\equiv$ 0.07 ppm).

The ^{13}C n.m.r. spectrum of the parent (*O*-acetylated phenolic poly(acryloylmorpholine)-based bead matrix (acetoxyphenyl content 1.0 mmol g^{-1}) is shown in *Figure* 1. Peaks due to almost all the carbons in groups pendant on the polymer backbone are readily discerned and have been assigned by comparison with the spectra of the acrylic monomers. The carbons nearest to the main chain give broad peaks ($\mathbf{b}_M$, $\mathbf{c}_M$, $\mathbf{d}_M$, $\mathbf{j}_M$ and $\mathbf{j}'_M$) owing to the more restricted mobility of these carbons. Those carbons which do not bear hydrogens give comparatively low peak intensities ($\mathbf{g}_M$, $\mathbf{h}_M$ and $\mathbf{i}_M$) owing to the low nuclear Overhauser enhancement and longer relaxation times of such carbons. The backbone carbons cannot be seen because (a) they fall in the 35–42 ppm region which is obscured by DMSO-d_6, (b) the intensity will be spread over several lines resulting from tacticity effects, (c) the lines will be broadened by the relatively low mobility of the backbone.

In solid (gel) phase peptide synthesis, the poly(acryloylmorpholine)-based bead matrix is first de-*O*-acetylated (activated) by reaction with morpholine. This exposes the pendant phenolic residues to which a suitable Boc amino acid may be attached, *via* a phenyl ester linkage, by diisopropylcarbodiimide-mediated coupling in the presence of 4-(*N*,*N*-dimethylamino)-pyridine as catalyst. Residual phenolic hydroxyl groups are then re-*O*-acetylated by the action of an excess of acetic anhydride in the presence of triethylamine as catalyst. The attached Boc amino acid will thus become the *C* terminal amino acid of the target peptide sequence.

Figure 2 shows the ^{13}C n.m.r. spectrum of the DMSO-d_6 swollen phenolic poly(acryloylmorpholine)-based matrix after substitution with Boc leucine ($\sim$0.5 mmol g^{-1} by amino acid analysis) and re-*O*-acetylation as outlined

0032-3861/80/1367-05$02.00

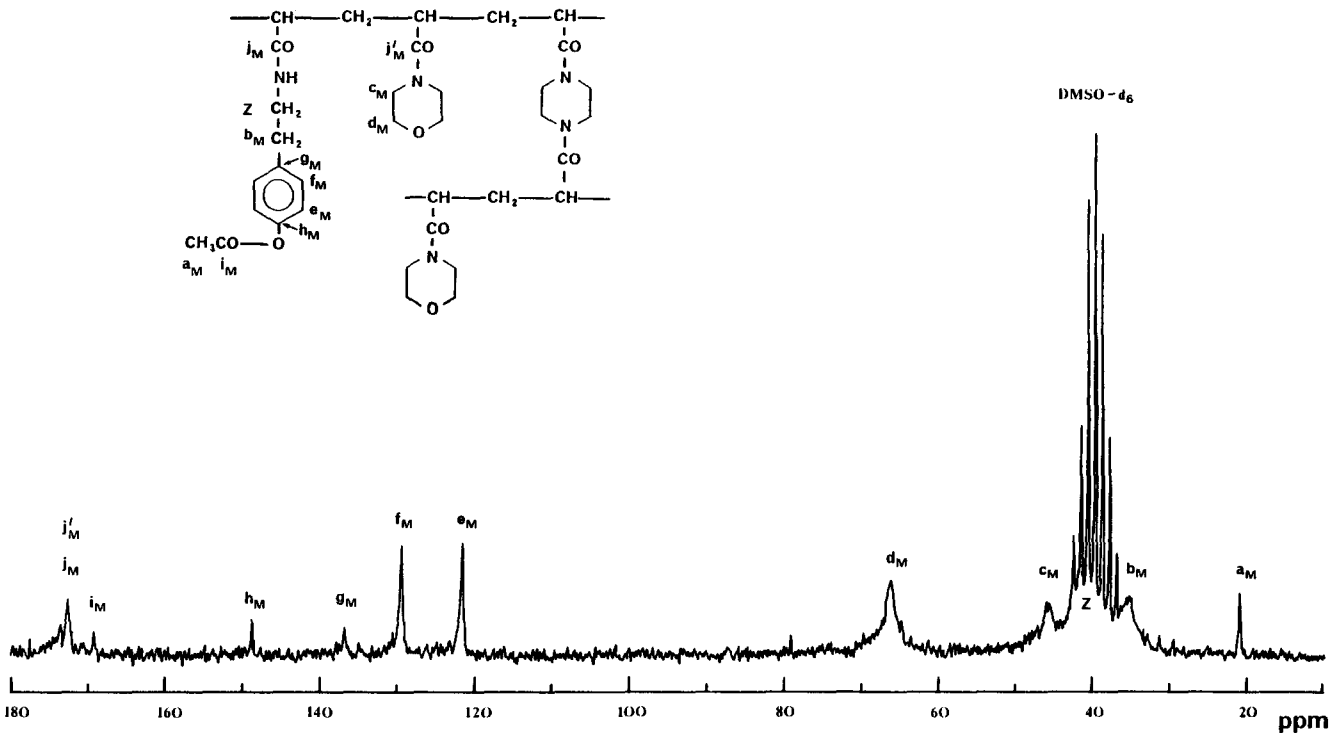


Figure 1 ^{13}C n.m.r. spectrum of parent phenolic matrix swollen in DMSO-d_6; 86 772 pulses; line positions in ppm (those for the corresponding carbons in the acrylic monomers in $CDCl_3$ are given in brackets): a_M 20.89 (21.06); b_M 35.12 (34.90); c_M 45.78 (46.08, 42.31: split into two by restricted rotation about C–N); d_M 66.32 (66.61); e_M 121.75 (121.60); f_M 129.55 (129.66); g_M 137.03 (136.67); h_M 149.05 (149.28); i_M 169.33 (169.69); j_M, j'_M 172.64 (165.92, 165.07); Z concealed by DMSO (40.75). [The unsaturated carbons in the acetate monomer give peaks at 125.95, 131.09 and those in the morpholine monomer give peaks at 127.38, 127.71]

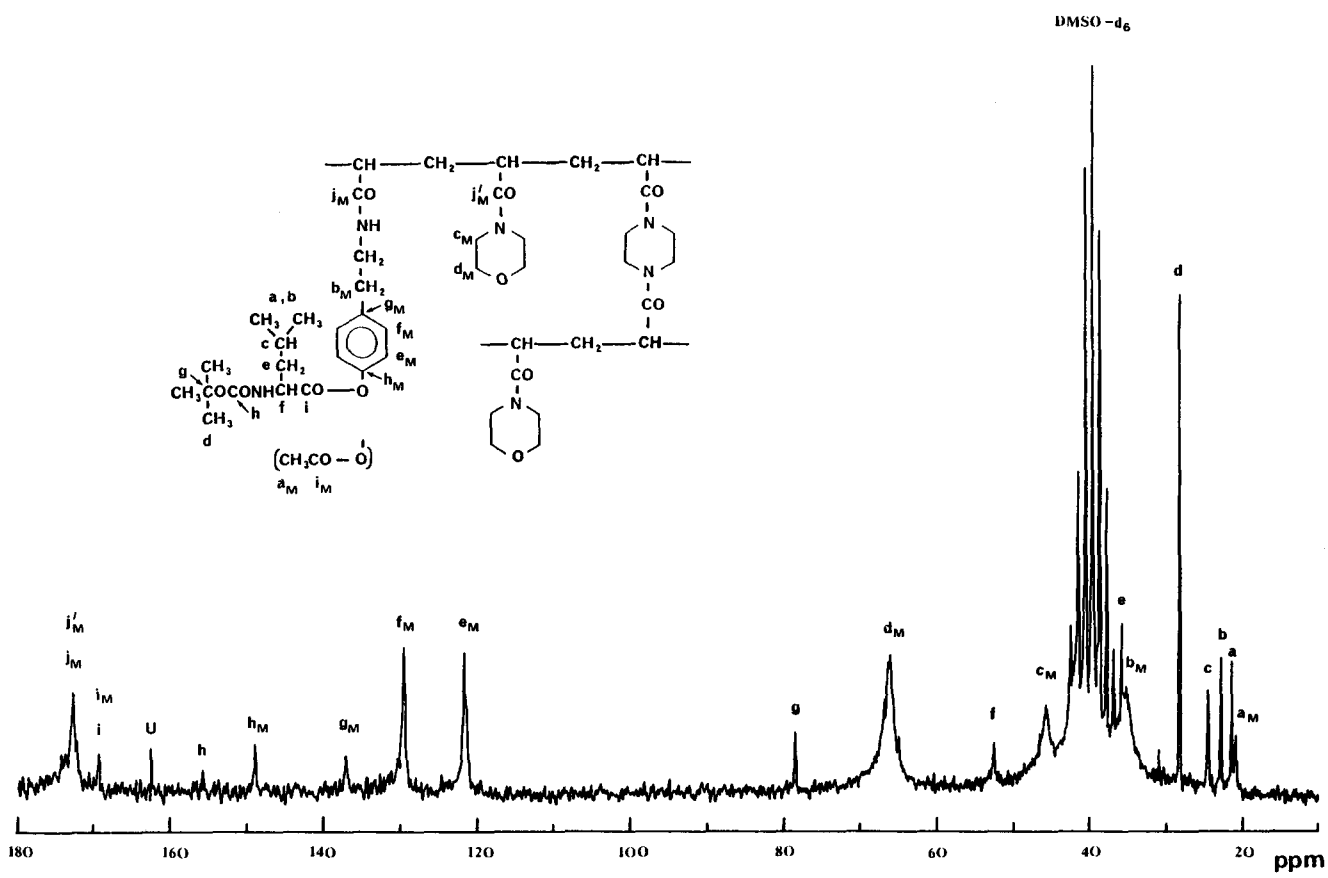


Figure 2 ^{13}C n.m.r. spectrum of re-*O*-acetylated Boc leucine substituted matrix swollen in DMSO-d_6; 130 407 pulses; line positions in ppm (those for Boc leucine in DMSO given in brackets): a 21.41 (21.41); b 22.77 (23.03); c 24.46 (24.53); d 28.23 (28.36); e 35.92 (39.99); f 52.46 (52.02); g 78.53 (78.14); h 155.84 (155.80); i 169.17 (174.78); a_M, ... j_M, j'_M matrix carbons, see *Figure 1*; U unidentified

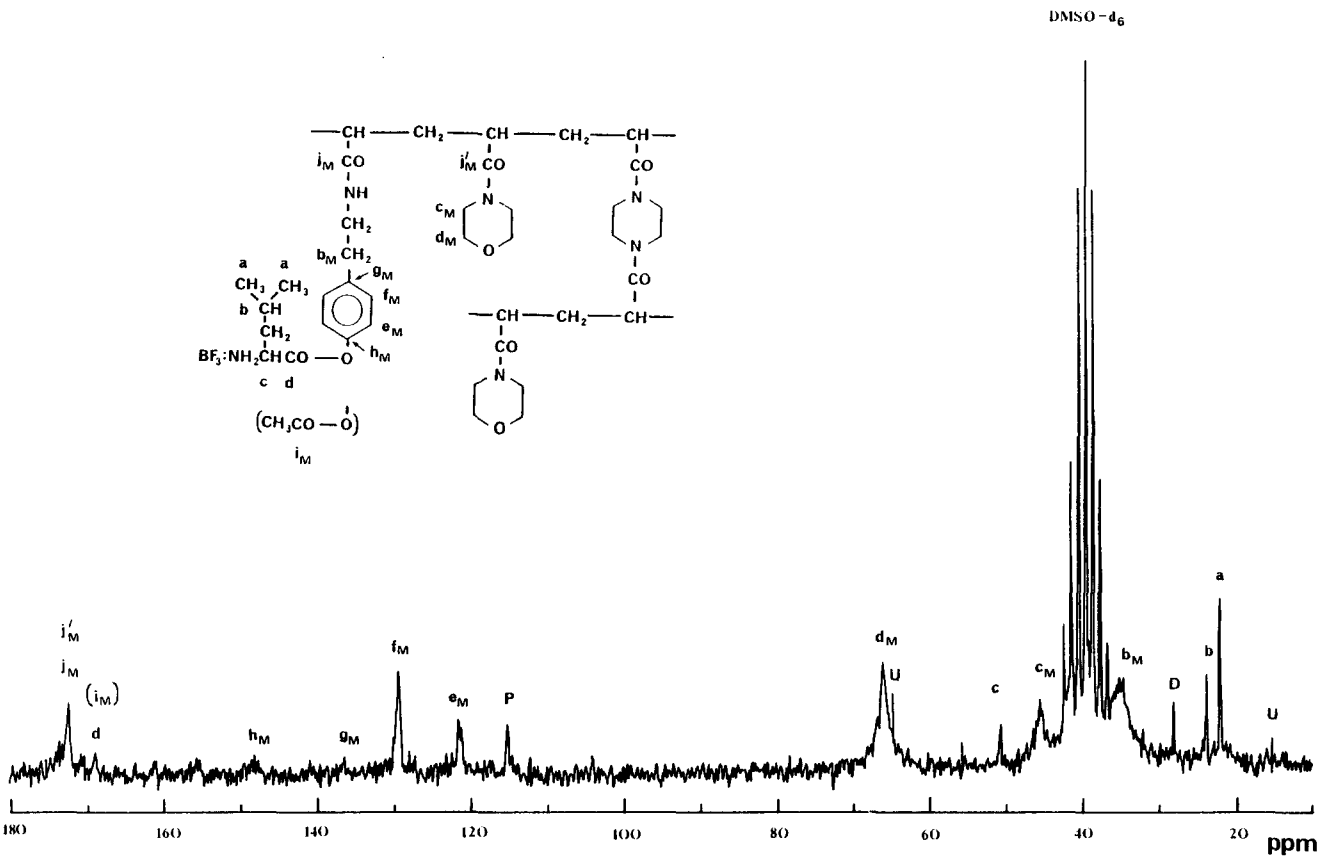


Figure 3 ^{13}C n.m.r. spectrum of $BF_3.Et_2O$/benzyl alcohol-treated Boc leucine substituted matrix, swollen in DMSO-d_6; about 30 000 pulses; line positions in ppm (those for leucine in $BF_3.Et_2O$/DMSO-d_6 given in brackets): a 22.19 (22.06, 22.38); b 23.94 (23.94); c 50.78 (50.78); d 169.17; b_M . . . . . j_M, j'_M matrix carbons see *Figure 1*; D 28.23 Boc; P 115.25 aromatic carbons *ortho* to OH; the peaks due to the other aromatic carbons are obscured by peak f_M or lost in noise; U 15.22, 64.95 ether impurity from work-up

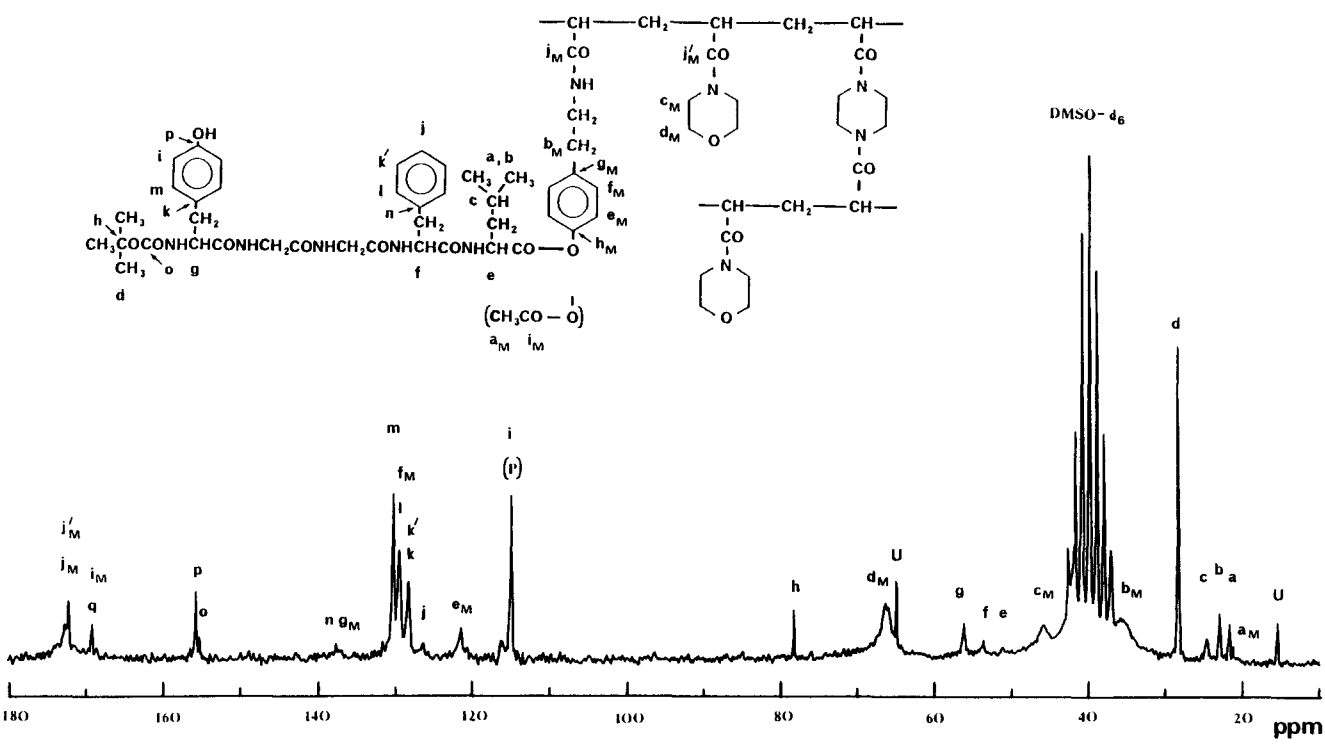


Figure 4 ^{13}C n.m.r. spectrum of DMSO-d_6-swollen phenolic matrix on which the peptide sequence Boc–Tyr–Gly–Gly–Phe–Leu has been built up; 299 784 pulses; line positions in ppm (those for corresponding Boc amino acids in DMSO are given in brackets): a 21.50 (21.41); b 22.76 (23.03); c 24.47 (24.53); d 28.17 (28.36); e 51.17 (52.02); f 53.81 (55.27); g 56.23 (55.59); h 78.21 (78.27); i 114.93 (115.06); j 126.56 (126.50); k, k' 128.25 (128.12, 128.32); l 129.36 (129.29); m 130.14 (130.14); n 137.78 (138.20); o 155.37 (156.55); p 155.81 (156.00); q 169.13 (one of the carbonyls); a_M . . . . . j_M, j'_M matrix carbons see *Figure 1*; U ether impurity from work-up: (P) (see *Figure 3*) obscured by peak i

above. In addition to the peaks due to the parent matrix, the expected leucine peaks (**a, b, c, e, f** and **i**) and Boc peaks (**d, g** and **h**) are all well resolved and readily assigned. The geminal terminal methyl groups of the leucine side-chain are magnetically non-equivalent owing to their proximity to the asymmetric carbon atom and so give separate peaks (**a** and **b**). From the peptide chemist's point of view, the fact that the major sample peak in the spectrum is that due to the three magnetically equivalent methyl carbons, $-C(\underline{C}H_3)_3$, of the Boc group is most important because Boc is by far the most widely used *N*-terminal protecting group in solid (gel) phase peptide synthesis.

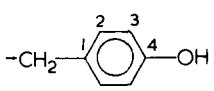
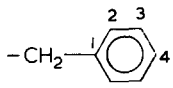
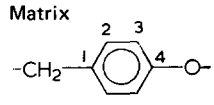
Removal of the *N*-terminal Boc group from the amino acid residues pendant on the phenolic poly(acryloylmorpholine)-based matrix may be effected with 10% v/v $BF_3.Et_2O$ benzyl in alcohol solution. The amino terminals are 'liberated' in the form of a boron trifluoride complex. In this form, the deprotected matrix still swells very well in highly polar organic solvents. *Figure* 3 shows the ^{13}C n.m.r. spectrum, in DMSO-d_6, of the Boc leucine substituted poly(acryloylmorpholine) resin after treatment with a 10% v/v solution of $BF_3.Et_2O$ at 25°C for 0.5 h.

The spectrum shows several interesting features. The geminal terminal methyl groups of the leucine side-chain now give only a single peak (**a**) at 22.19 ppm. Data in support of this assignment was obtained by running the ^{13}C n.m.r. spectrum of leucine, solubilized in DMSO-d_6 by addition of a few drops of $BF_3.Et_2O$. In this spectrum the geminal methyl groups are resolved into closely spaced peaks at 22.06 and 22.38 ppm. The peak (**c**) (50.78), due to the leucine asymmetric carbon, is upfield from the corresponding peak (**f**) (52.46) in the previous spectrum. Deshielding of the asymmetric carbon is clearly less effective in the case of the adjacent amino group associated with BF_3 than it is in the case of the Boc-protected amino group. Two peaks (**D** and **P**) were unexpected. Peak **D** is due to the three magnetically equivalent methyl carbons, $-C(\underline{C}H_3)_3$ of residual Boc groups. By comparison of the intensities of the $-C(\underline{C}H_3)_3$ peaks (**d** and **D**) in *Figures* 2 and 3 with the intensity of the $-\underline{C}H(CH_3)_2$ peaks of the leucine side-chains in the respective spectra, it is estimated that about 11% of the Boc groups remain after 'deprotection'. Leucine peaks in the original positions (**a, b** and **c** in *Figure* 2), of this order of magnitude are just evident in *Figure* 3. Peak **P** is due to aromatic carbons *ortho* to (phenolic)–OH (see *Figure* 4). This is consistent with loss of acetyl groups and/or leucine residues from the support by BF_3 catalysed transesterification to the benzyl alcohol used as solvent. Cleavage of the relatively unhindered acetyl group is the more likely and it is significant that the peak (a_M), due to CH_3CO, in *Figure* 2 is barely discernible above the noise in *Figure* 3. Loss of Boc amino acids from the matrix by transesterification has been shown previously to be slow by ^{14}C labelling studies*.

Following *N*-terminal Boc removal from the matrix bound amino acid, selected target peptides may be built up, in stepwise manner, by a series of chain elongation cycles. Each cycle consists of two steps: firstly, reaction of the amino acid or peptide substituted matrix, as its boron trifluoride complex, with an appropriate Boc amino acid in the presence of diisopropylcarbodiimide, 1-hydroxybenzotriazole and *N*-methylmorpholine, thereby adding the new Boc amino acid units; secondly, *N*-terminal deprotection by the boron trifluoride method. 1-Hydroxybenzotriazole acts as both a catalyst and as an *O*-acylation suppressor. In the latter role it should minimize coupling of Boc amino acids to exposed phenolic –OH groups.

* In a model experiment[6], matrix bound *N*-acetyl-[1-^{14}C]-glycine has been treated with 10% v/v $BF_3.Et_2O$ in benzyl alcohol at 25°C. Over 16 h, only 4.1% loss of matrix-bound amino acid took place.

Table 1 ^{13}C n.m.r. line positions[a] for Boc protected, matrix-bound, peptide subunits of [Leu] enkephalin

Structural feature		Boc–Leu–OC_6H_4–	Boc–Phe–Leu–OC_6H_4–	Boc–Gly_1–Gly_2––Phe–Leu––OC_6H_4~	Boc–Tyr–Gly_1––Gly_2–Phe––Leu–OC_6H_4~
Boc	$(\underline{C}H_3)_3CO$	28.23	28.17	28.36	28.17
	$(CH_3)_3\underline{C}O$	78.53	78.05	78.53	78.21
Tyr	NH$\underline{C}$HCO				56.23
	C^3				114.93
	C^1				128.25
	C^2				130.14
	C^4				155.81
Gly_1	NH$\underline{C}H_2$CO			43.70	*b*
Gly_2	NH$\underline{C}H_2$CO			*b*	*b*
Phe	NH$\underline{C}$HCO		55.51	53.80	53.81
	C^4		126.22	126.44	126.56
	C^3		128.19	128.19	128.25
	C^2		129.35	129.36	129.36
	C^1		137.88	137.71	137.78
Leu	$CH(\underline{C}H_3)_2$	21.41	21.47	21.60	21.50
		22.77	22.90	22.84	22.76
	$CH\underline{C}H(CH_3)_2$	24.46	24.33	24.46	24.47
	$CH\underline{C}H_2CH$	35.92	*b*	*b*	*b*
	NH$\underline{C}$HCO	52.46	50.93	51.03	51.17
Matrix	C^3	121.69	121.49	121.49	121.43
	C^2	129.49	129.35	129.36	129.36
	C^1	137.04	137.00	137.14	136.94
	C^4	148.87			

[a] Relative to DMSO-d_6 centre line as 39.67 ppm; [b] concealed by DMSO peaks

The ^{13}C n.m.r. line positions for the intermediate matrix-bound peptide units, *via* which the leucine residues have been converted to a protected form of the opioid peptide [Leu] enkephalin, are given in *Table* 1. The synthesis involved coupling and deprotection cycles incorporating, in turn, Boc phenylalanine, Boc glycyl glycine and Boc tyrosine. The spectrum of the final protected matrix-bound [Leu] enkephalin is shown in *Figure* 4. As the peptide chain is lengthened, the signals from the carbons nearer the support tend to diminish in height and to broaden. The peptide unit at the end of the chain, next to Boc, always gives the sharpest lines, which is understandable in terms of decreased mobility of the nuclei nearer the support. When the next amino acid unit is added, the NHCHO signal from the unit previously adjacent to Boc tends to move upfield. This may be seen in *Table* 1 for the shifts 52.46→50.93 (Leu) and 55.51→53.80 (Phe); also by the disappearance of the 43.70 (Gly_1) peak into the DMSO peaks when the Tyr unit is added. This is a normal end-group effect. Leu, Tyr, Phe, as well as Boc and $-C_6H_4OCO-$ peaks are readily identified when present. However, the DMSO-d_6 peaks obscure the peak from Gly_1 unless this is the end amino acid unit. The carbonyl peaks are weak and it is difficult to make assignments from the spectra at present available.

Finally, we may say that the application of gel phase ^{13}C n.m.r. spectroscopy to solid (gel) phase peptide synthesis yields directly hitherto inaccessible information on the removal of Boc groups and on the build-up of the peptide chain. Almost certainly, the removal of other protecting groups such as Bpoc, Ddz and Fmoc could also readily be checked. The use of a higher frequency *FT* n.m.r. instrument should permit greater sensitivity and/or reduction in required instrument time. ^{13}C n.m.r. has the advantage that there is no need to use purpose-synthesized (e.g. fluorinated) protecting groups to facilitate spectroscopic monitoring. It seems probable that the technique of gel phase ^{13}C n.m.r. spectroscopy will be applicable to many analytical problems encountered in solid (gel) phase peptide synthesis on solvent-swollen supports, possibly including the classical polystyrene-based Merrifield resins[3].

Abbreviations

Boc, *N*-tert-butyloxycarbonyl
Bpoc, 2-(*p*-biphenylyl)-2-propyloxycarbonyl
Ddz, 2-(3,5-dimethoxyphenyl)-2-propyloxycarbonyl
Fmoc, 9-fluorenylmethyloxycarbonyl

References

1 Felix, A. M. and Jimenez, M. H. *Anal. Biochem.* 1973, **52**, 377
2 Manatt, S. L., Amsden, C. F., Bettison, C. A., Frazer, W. T., Gudman, J. T., Lenk, B. E., Lubetich, J. F., McNelly, E. A., Smith, S. C., Templeton, D. J. and Pinnell, R. P. *Tetrahedron Letters*, 1980, **21**, 1397
3 Manatt, S. L., Horowitz, D., Horowitz, R. and Pinnell, R. P. *Anal. Chem.* 1980, **52**, 1529
4 Leibfritz, D., Mayr, W., Oekonomopulos, R. and Jung, G. *Tetrahedron*, 1978, **34**, 2045
5 Epton, R., Goddard, P., Marr, G., McLaren, J. V. and Morgan, G. J. *Polymer* 1979, **20**, 1444
6 Epton, R., Hocart, S. J. and Marr, G. unpublished work