

Synthesis of glyoxylyl peptides using an Fmoc-protected α,α' -diaminoacetic acid derivative

SAMIA FAR and OLEG MELNYK*

UMR 8525 CNRS-Université de Lille 2, Institut de Biologie de Lille, 1 rue du Professeur Calmette, 59021 Lille, France

Received 20 August 2004; Revised 9 September 2004; Accepted 14 September 2004

Abstract: The synthesis of glyoxylyl peptides by coupling the masked glyoxylic acid derivative (FmocNH)₂CHCO₂H, **1**, to a peptidyl resin assembled using Fmoc/*tert*-butyl chemistry has been described recently. Deprotection and cleavage of the peptide from the solid support using TFA was followed by unmasking of the glyoxylyl group in solution in the presence of DBU. [] The glyoxylyl peptide was thus generated using non-oxidizing conditions by comparison with the method based on the periodic oxidation of a seryl-precursor. However, base treatment of the (FmocNH)₂CHCO₂-peptide led to the formation of a byproduct besides the desired glyoxylyl peptide. This paper describes an optimized procedure for unmasking the Fmoc-protected α,α' -diaminoacetic acid moiety in solution which suppressed byproduct formation. Also presented is a series of experiments that permitted a structure and a mechanism of formation for the byproduct to be suggested. Copyright © 2005 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: glyoxylyl peptide; α,α' -diaminoacetic acid; Fmoc; Edman degradation; MALDI-TOF

INTRODUCTION

Aldehyde chemistry allows the convergent synthesis of large peptidic constructs starting from unprotected peptide fragments, due to the ability of the aldehyde function to react chemoselectively with a variety of naturally occurring amino acids [1]. Aldehyde-functionalized peptides can be ligated with the β -aminothiol moiety of cysteine or the β -aminoalcohol moiety of serine or threonine to give the thiazolidine [2] or pseudoproline linkages, respectively [3]. Hydrazones [4] or oximes [5] can also be formed by reaction with hydrazine or hydroxylamine derivatives. In this context, the chemistry of the glyoxylyl group has been studied extensively due to its stability and reactivity.

The glyoxylyl group is usually generated in solution by periodate oxidation of *N*-terminal serine or threonine residues [6]. The control of pH and the use of scavengers such as methionine or dimethylsulfide was often utilized to minimize the oxidation of sensitive amino acids but oxidation of methionine or cysteine residues cannot usually be suppressed. To overcome this problem, a novel Fmoc-protected α,α' -diaminoacetic acid derivative **1** (FmocNH)₂CHCO₂H was designed that allowed the introduction of an α -oxo aldehyde functionality into a peptide using non-oxidizing conditions [7]. This derivative could be synthesized easily and introduced on a peptide chain after Fmoc/*tert*-butyl solid phase peptide synthesis [8]. Deprotection and cleavage of the peptide from the solid support using TFA was followed by unmasking of the glyoxylyl group in

the presence of DBU. In a model experiment using peptide (FmocNH)₂CHCO-ILKEPVHGV-NH₂ **2**, a minor byproduct could be isolated in addition to the major glyoxylyl peptide **3** CHO-CO-ILKEPVHGV-NH₂. MALDI-TOF analysis of the byproduct showed a monoisotopic peak [M + H]⁺ of approximately twice the value of that of the target glyoxylyl peptide, thus suggesting a dimeric structure.

This article reports the effect of the base structure and concentration on the amount of byproduct formed in the deprotection step and an optimized procedure for unmasking the glyoxylyl group without byproduct formation. In addition, a series of experiments is reported that led to a structure and mechanism of formation for the byproduct.

MATERIALS AND METHODS

General

MALDI-TOF spectra were recorded on a PerSeptive Biosystems Voyager-DE STR spectrometer. Analytical RP-HPLC separations were performed on a Waters Alliance 2695 separation module. Anhydrous DMF (99.8%) was purchased from Aldrich and *O*-benzylhydroxylamine hydrochloride (99%) from Acros.

Synthesis of (FmocNH)₂CHCO₂H **1**

9*H*-fluorenylmethylcarbamate (1.00 g, 4.18 mmol), solid glyoxylic acid monohydrate (192 mg, 2.09 mmol) and TosOH (3.8 mg, 20 μ mol) were dissolved in toluene (200 ml) under vigorous stirring in a two-neck flask equipped with a Dean-Stark trap and a condenser on top of it. The mixture was refluxed for 2 h under vigorous stirring and the formation of a fluffy white solid was observed during water elimination.

*Correspondence to: Oleg Melnyk, UMR 8525 CNRS, Institut de Biologie de Lille, 1 rue du Professeur Calmette, 59021 Lille, France; e-mail: oleg.melnyk@ibl.fr

The reaction mixture was then cooled to room temperature. The white solid was filtered, washed with toluene and dried *in vacuo* to give compound **1** (657 mg, 1.23 mmol) in 59% yield.

Mp 216°–218°C; IR (KBr) ν (cm⁻¹) 3301, 1726, 1690; ¹H NMR (300 MHz, DMF-*d*₇, TMS as internal reference) δ in ppm 4.17 (m, 6H, CH₂ + CH Fmoc), 5.59 (t, 1H, *J* = 7.8 Hz, CH-COOH), 7.19 (t, 4H, *J* = 7.2 Hz, CH_{Ar}), 7.29 (t, 4H, *J* = 7.2 Hz, CH_{Ar}), 7.65 (d, 4H, *J* = 7.2 Hz, CH_{Ar}), 7.79 (d, 4H, *J* = 7.2 Hz, CH_{Ar}), 8.13 (d, 2H, *J* = 8.1 Hz, NH); ¹³C NMR (75 MHz, DMF-*d*₇) δ 47.6, 60.8, 67.3, 120.7, 126.1, 127.8, 128.4, 141.8, 144.8, 156.5, 170.6; FAB-MS *m/z* 535.3 [M + H]⁺, 557.3 [M + Na]⁺; HRMS (FAB) *m/z* calcd for [M + H]⁺C₃₂H₂₇N₂O₆: 535.1869 found 535.1865 (–0.7 ppm).

Isolation of compound 4. Peptide **2** (100 mg, 57.0 μ mol) was dissolved in DMF (20 ml). DBU (80.0 μ l, 0.57 mmol) was added dropwise to the solution. The reaction mixture was stirred at room temperature for 30 min. Then 3 ml of acetic acid and 1 ml of water were added to the reaction mixture which was stirred for a further 5 min and then concentrated *in vacuo*. The crude residue was purified by RP-HPLC on a C18 Nucleosil 4.6 \times 250 mm column (eluent A: water containing 0.05% TFA by vol, eluent B: acetonitrile/water 4/1 by vol. containing 0.05% TFA by vol., linear gradient 0–100% B in 30 min, 1 ml/min, detection at 215 nm) to give 40.0 mg (54%) of peptide **3** and 7.0 mg (12%) of compound **4**.

Peptide **3**. Selected ¹H NMR data (300 MHz, DMF-*d*₇, TMS as internal reference) δ in ppm 5.21 (s, 0.7 H, CO-CH(OH)₂), 9.41 (s, 0.3 H, CO-CHO); selected ¹³C NMR data (75 MHz, DMF-*d*₇) δ 88.8 (CH(OH)₂), 189.3 (COCHO); MALDI-TOF *m/z* [M + H]⁺ calcd 1046.59, found 1046.52 (matrix: α -cyano-4-hydroxycinnamic acid in CH₃CN/H₂O 50/50 by vol. containing 0.1% TFA by vol).

Compound **4**: MALDI-TOF *m/z* [M + H]⁺ found 2090.42 (same matrix as for **3**).

Trypsin digestion of compound 4. Compound **4** was solubilized in water at 1 mg/ml concentration and 1 μ l of this solution was mixed with 10 μ l of 0.1 M tris(hydroxymethyl)amino-methane hydrochloride buffered at pH 8.3 and 1 μ l of a 40 μ g/ml trypsin solution in water (Promega, 15000 U/mg, enzyme/peptide ratio of 1/25 w/w). The mixture was stirred overnight at 37°C then a 1 μ l-sample of the mixture was mixed with 1 μ l of 2,5-dihydrobenzoic acid 20 mg/ml in acetonitrile/10 mM aqueous dibasic ammonium citrate 1/1 by vol for MALDI-TOF analysis. Three peaks at *m/z* (monoisotopic) 636.36, 855.58 and 1472.90 were observed (Figure 4).

Reaction with O-benzylhydroxylamine. 1.0 mg of compound **4** was dissolved in 82 μ l of a 0.047 M O-benzylhydroxylamine hydrochloride solution in 0.1 M sodium acetate buffer (pH 4.7). The mixture was stirred at 37°C and monitored by RP-HPLC on a C18 Nucleosil 4.6 \times 250 mm column (eluent A: water containing 0.05% TFA by vol., eluent B: acetonitrile/water 4/1 by vol. containing 0.05% TFA by vol., linear gradient 0–100% B in 30 min, 1 ml/min, detection at 215 nm). Three peaks with retention times of 15.1, 20.9 and 22.1 min (Figure 5) were collected and analysed by MALDI-TOF (*m/z* monoisotopic 1047.74, 2196.60 and 1166.85, respectively).

Microsequencing by Edman degradation. Automated Edman degradation was performed on an Applied Biosystems 492 Procise sequencer with an on-line Perkin Elmer

Applied Biosystems Model 140C PTH Amino Acid Analyzer. Phenylthiohydantoin-amino acids were identified by chromatography on a C18 column (Spheri-5 PTH, 2.1 \times 220 mm).

Study of Base Nature/Concentration Effects on Byproduct 4 Formation

Peptide and base stock solutions preparation. Peptide solution: A stock solution of peptide **2** (3.2 mM) was prepared by dissolving 5.0 mg (2.0 μ mol) of peptide in 900 μ l of anhydrous DMF (99.8%, Aldrich).

Base solutions: 1 ml stock solutions of the bases DBU, piperidine and morpholine (0.87 M) were prepared by adding anhydrous DMF to 129 μ l of DBU, 85.5 μ l of piperidine and 75.5 μ l of morpholine. Dilutions of 0.144 M (1/6), 0.288 M (1/3), 0.432 M (1/2) and 0.576 M (2/3) in DMF were prepared from each stock solution.

Typical unmasking procedure. For each base, 10 μ l of each dilution as well as 10 μ l of stock solution were added to 5 aliquots of 90 μ l (0.29 μ mol, 500 μ g) of peptide solution in 600 μ l microcentrifuge tubes containing microstirring bars. After stirring for 30 min at room temperature, 400 μ l of H₂O/CH₃CO₂H (7/1 by vol) was added to each solution and the mixtures were stirred for a further 5 min.

RP-HPLC monitoring. 20 μ l of each solution was injected on a C18 Delta Pak 3.9 \times 300 mm column using the above eluents and the following experimental conditions: 0% B for 15 min then linear gradient 0–100% B in 30 min, 1 ml/min, detection at 215 nm.

RESULTS AND DISCUSSION

A recent study [7] reported the synthesis of a glyoxylyl peptide **3** CHO-CO-ILKEPVHGV-NH₂ obtained by coupling reagent **1** to a peptidyl resin assembled using Fmoc/*tert*-butyl chemistry. After deprotection and cleavage from the solid support using TFA, removal of the Fmoc groups was performed in the presence of DBU in DMF (Figure 1). Base treatment generated the desired glyoxylyl peptide **3** but also a byproduct **4** eluting slightly later during RP-HPLC analysis (Figure 2). In a control experiment, treatment of glyoxylyl peptide **3** with DBU using the same experimental conditions failed to give any compound **4**.

Peptide **4** was isolated by preparative RP-HPLC and analysed by MALDI-TOF. The observation of a peak at *m/z* 2090.42 (monoisotopic [M + H]⁺) suggested the covalent association of two peptide chains. NMR analysis failed to give any interesting data concerning the structure of the link separating the two peptide units. However, the characteristic protons of the α -oxo aldehyde function were not observed in the ¹H spectrum (for **3** in DMF-*d*₇, CH(OH)₂ was found at 5.2 ppm (hydrated form) and CHO appeared at 9.4 ppm [7], thus suggesting the implication of the α -oxo aldehyde functions in the dimerization process. Moreover, all the amino acids displayed typical NMR

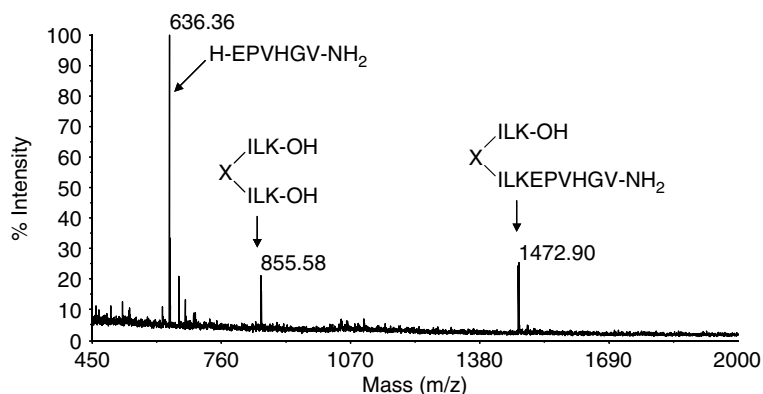


Figure 4 MALDI-TOF spectrum of the trypsin digest of compound **4**.

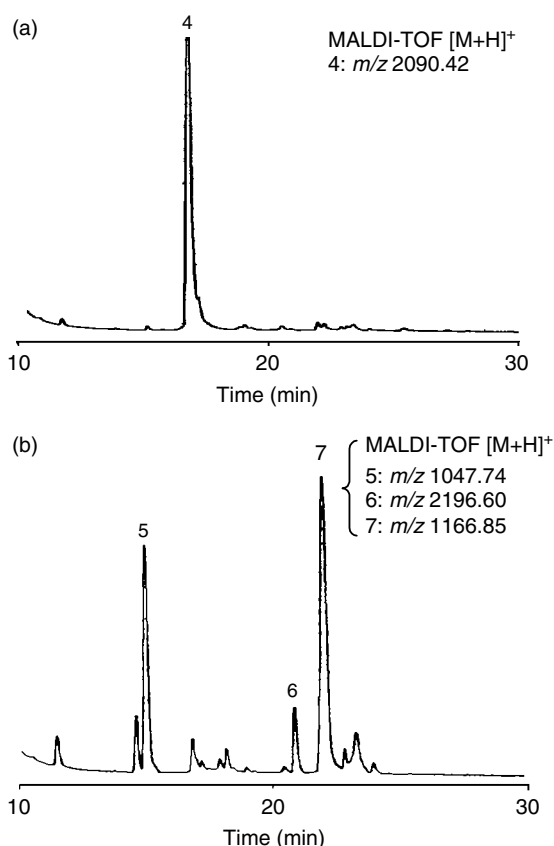


Figure 5 Reaction of byproduct **4** with *O*-benzylhydroxylamine, leading to peptides **5**, **6** and **7**. RP-HPLC profiles of (a) by-product **4** before addition of *O*-benzylhydroxylamine and (b) the crude reaction mixture after 24 h. C18 Nucleosil column (4.6 × 250 mm). Eluent A: water containing 0.05% TFA by vol., eluent B: acetonitrile/water 4/1 by vol. containing 0.05% TFA by vol., linear gradient 0–100% B in 30 min, 1 ml/min, detection at 215 nm.

Hypothetical Structure of Peptide **4**

All these data taken together are consistent with the hypothetical structure for peptide **4** described in Figure 6. This structure, based on the general model depicted in Figure 3, is composed of two peptide

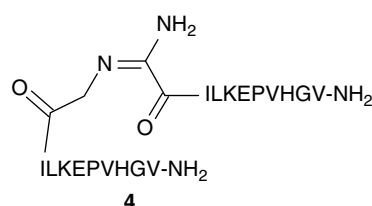


Figure 6 Probable structure of byproduct **4**.

units linked through an amidine function [9] with an undefined stereochemistry. The amidine group can be generated as shown in Figure 7. Fmoc removal gives amination intermediate **8** which can lead to imine **9** after elimination of ammonia. Intermediate **11** can be the product of self-condensation of the imine **9**, but other pathways may operate in parallel. Then, base-catalysed isomerization of **11** could produce amidine **4**. The greatest stability of amidine **4** compared with imine intermediate could displace the equilibrium toward the formation of **4**.

The deprotection was performed in anhydrous DMF. However, the partial *in situ* hydrolysis of amination **8** into glyoxylyl peptide **3** by traces of water cannot be excluded. To test whether compound **4** could be formed by reaction of glyoxylyl peptide **3** with amination **8**, an equimolar mixture of Fmoc-protected peptide **2** and glyoxylyl peptide **3** was submitted to the deprotection conditions (5 equivalents DBU, 30 min in DMF then addition of aqueous acetic acid). If no reaction occurred between **3** and amination **8** during the deprotection step, the ratio of the HPLC peak areas **4/3** should be 0.14, i.e. half the value obtained for the control experiment (without added glyoxylyl peptide **3**, ratio **4/3**: 0.28). Interestingly, the ratio of the HPLC peak areas compound **4**/peptide **3** was found to be 0.11. The minor difference between this experimental value and 0.14, the calculated ratio in the case of no effect of added **3**, showed that compound **4** was probably not formed by reaction of peptide **3** with amination **8**.

The products formed by reaction of compound **4** with *O*-benzylhydroxylamine can be easily rationalized

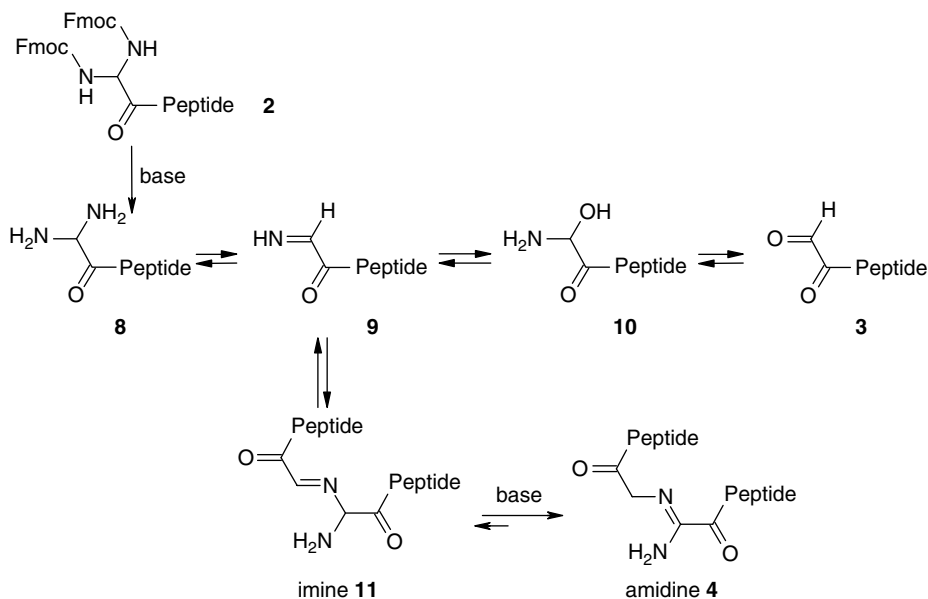


Figure 7 Possible mechanism for the formation of byproduct **4**.

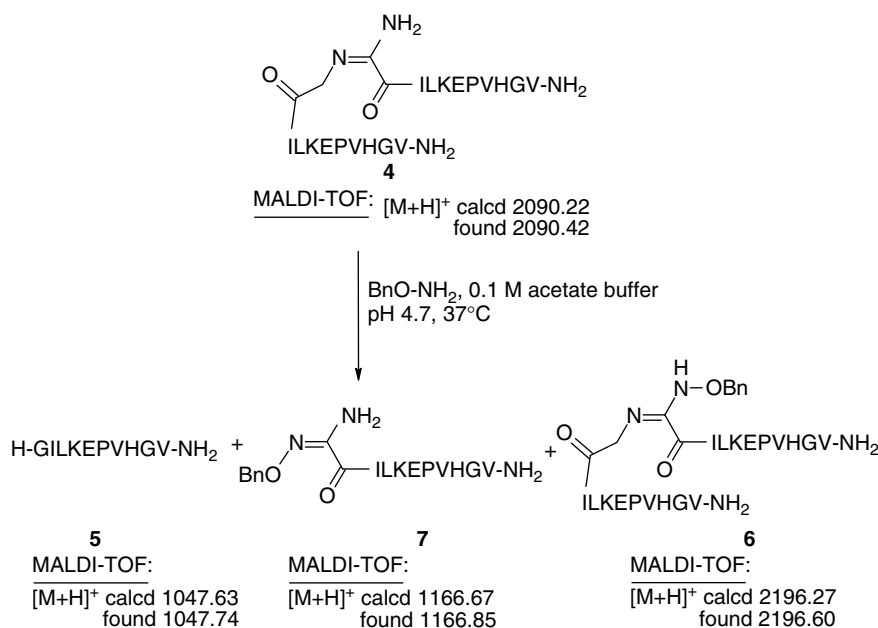


Figure 8 Reaction of byproduct **4** with *O*-benzylhydroxylamine.

based on this structure (Figure 8). In particular, amidine **4** contains a masked glycol residue, which can be liberated by reaction with the hydroxylamine.

Optimization of the Deprotection Step

The hypothetical mechanism described in Figure 7 suggested that both the concentration of the Fmoc-protected peptide **2** and the nature of the base used in the deprotection could have an impact on the yield of the byproduct **4**.

In a first series of experiments, different bases (DBU, piperidine or morpholine) were examined at

the peptide concentration used in the preliminary experiments (5 mg/ml). The Fmoc-protected peptide **2** was incubated with increasing amounts of each base (5, 10, 15, 20 and 30 equivalents 30 min in DMF). The reactions were quenched with aqueous acetic acid and then analysed by RP-HPLC. Figure 9 shows the RP-HPLC profiles obtained using 5 equivalents of base.

Using morpholine (Figure 9C, 5 equivalents of base), the conversion was found to be low. After 30 min, glyoxylyl peptide **3** was barely detectable in the chromatogram. Increasing the base concentrations up to 30 equivalents did not permit the complete removal

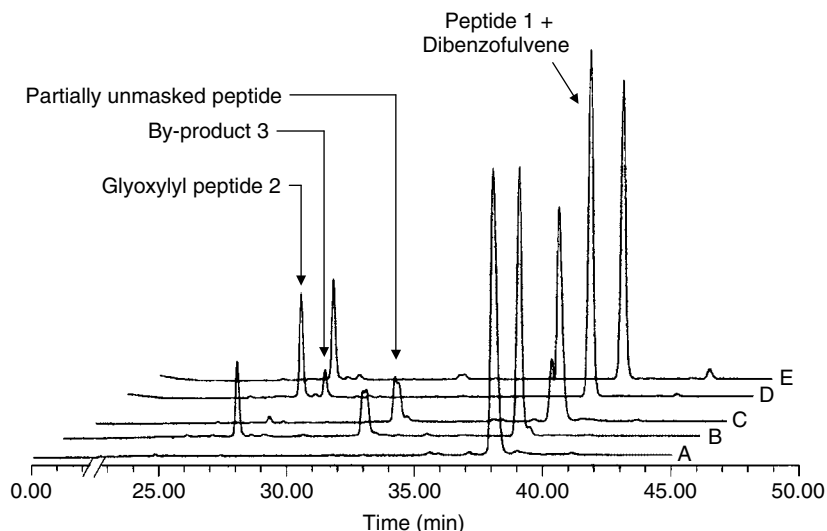


Figure 9 RP-HPLC profiles of (A) masked peptide **2**; (B) crude reaction mixture of peptide **2** (5 mg/ml) with 5 eq. of piperidine; (C) crude reaction mixture of peptide **2** (5 mg/ml) with 5 eq. of morpholine; (D) crude reaction mixture of peptide **2** (5 mg/ml) with 5 eq. of DBU; (E) as for D at a peptide concentration of 1 mg/ml. C18 Delta Pak column (3.9 × 300 mm). Eluent A: water containing 0.05% TFA by vol., eluent B: acetonitrile/water 4/1 by vol. containing 0.05% TFA by vol., 0% B for 15 min then linear gradient 0–100% B in 30 min, 1 ml/min, detection at 215 nm.

of the Fmoc groups (data not shown). The product eluting between peptide **2** and glyoxylyl peptide **3** corresponded to the removal of only one Fmoc group. Indeed, analysis of the intermediate by MALDI-TOF spectrometry gave a peak at m/z 1267.93 which corresponded to the loss of one Fmoc-NH₂ moiety. This peak can be attributed to imine Fmoc-N=CH-CO-ILKEPVHGV-NH₂ (m/z [M + H]⁺ calcd 1267.68) that could be formed by elimination of ammonia from intermediate FmocNH(H₂N)CH-CO-ILKEPVHGV-NH₂ during MALDI-TOF analysis. However, the direct formation of the imine in the reaction medium cannot be excluded.

Using 5 equivalents of piperidine (Figure 9B), the conversion of peptide **2** into glyoxylyl peptide **3** was improved compared with morpholine but still incomplete. 20 equivalents was found to be necessary for a complete unmasking after 30 min of reaction (data not shown). Interestingly, amidine **4** was not observed during these experiments.

As expected, rapid deprotection was observed with 5 equivalents of DBU (Figure 9D). Contrary to experience with piperidine, the intermediate corresponding to the removal of only one Fmoc group was not observed. However, formation of byproduct **4** was clearly visible. Interestingly, the amount of **4** relative to glyoxylyl peptide **3** was found to decrease by increasing DBU concentration, as shown in Figure 10. Byproduct **4** was not formed when using 30 equiv. of DBU.

The above first series of experiments were performed at 5 mg/ml peptide concentration. Next, the effect of peptide concentration on byproduct formation was examined using DBU. Interestingly, at 1 mg/ml and using different amounts of base (5 to 30 equivalents),

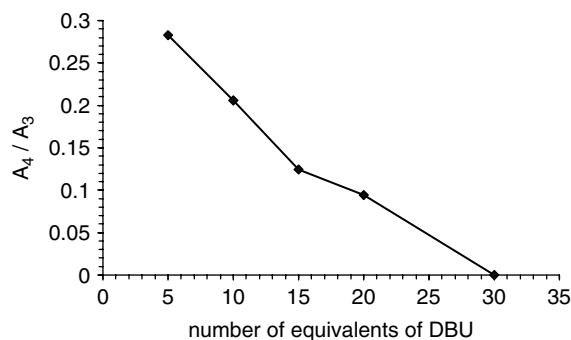


Figure 10 Unmasking of peptide **2** using DBU in DMF at a peptide concentration of 5 mg/ml. A₄ is the RP-HPLC peak area of byproduct **4** and A₃ the RP-HPLC peak area of glyoxylyl peptide **3**. The ratio of RP-HPLC peak areas A₄/A₃ is plotted as a function of the number of equivalents of DBU used in the deprotection reaction.

the deprotection proceeded cleanly without byproduct formation. Therefore, the optimal concentrations for performing glyoxylyl group unmasking are 1 mg/ml for peptide **2** and 5 equivalents of DBU (Figure 9E for the corresponding RP-HPLC profile).

The effects of peptide concentration and of the pK_a of the base used during the deprotection step are in accord with the mechanism proposed in Figure 7. Indeed, lowering the peptide concentration is expected to disfavor the reaction of two molecules of imine **9** to give intermediate **11**, which can further be converted into amidine **4** by base-catalysed isomerization. Lowering the pK_a of the base is expected to disfavor the isomerization of the intermediate **11** into amidine **4**. Finally, the effect of base concentration can

be explained by considering the condensation step **9** → **11**, which is expected to be acid-catalysed. Thus, high base concentrations should disfavor the pathway leading to intermediate **11**.

CONCLUSION

This paper describes the usefulness of (FmocNH)₂CHCO₂H for the preparation of glyoxylyl peptides which are valuable building blocks for the assembly of large molecules based on ligation chemistry.

Deprotection of the Fmoc-protected precursor in solution in the presence of base (morpholine, piperidine, DBU) was examined. The deprotection in the presence of morpholine or piperidine required a large excess of base. On the other hand, DBU permitted the removal of the Fmoc groups by using only 5 equivalents of base. However, the formation of an amidine byproduct could be observed at high peptide concentration (5 mg/ml) that could be suppressed easily by performing the deprotection at a lower concentration (1 mg/ml). The hypothetical structure of the byproduct, which was composed of two peptide chains linked through their *N*-termini by an α -oxo-amidine group, was supported by the results of enzymatic, chemical reactions and MALDI-TOF analyses.

Acknowledgements

We gratefully acknowledge le Ministère de l'Economie, des Finances et de l'Industrie (GenHomme, Coquepuce

Project), CNRS, Université de Lille 2 and Institut Pasteur de Lille for financial support. We thank Yves Lemoine for useful discussions, Hervé Drobecq for MALDI-TOF experiments, Gérard Montagne for NMR analyses and Rémi Desmet for technical assistance.

REFERENCES

1. Melnyk O, Fehrentz JA, Martinez J, Gras-Masse H. Functionalization of peptides and proteins by aldehyde or keto groups. *Peptide Sci.* 2000; **55**: 165–186.
2. Spetzler JC, Tam JP. Unprotected peptides as building blocks for branched peptides and peptide dendrimers. *Int. J. Peptide Protein Res.* 1995; **45**: 78–85.
3. Tam JP, Rao C, Liu CF, Shao J. Specificity and formation of unusual amino acids of an amide ligation strategy for unprotected peptides. *Int. J. Peptide Protein Res.* 1995; **45**: 209–216.
4. King TP, Zhao SW, Lam T. Preparation of protein conjugates via intermolecular hydrazone linkage. *Biochemistry* 1986; **25**: 5774–5779.
5. Rose K. Facile synthesis of homogeneous artificial proteins. *J. Am. Chem. Soc.* 1994; **116**: 30–33.
6. Geoghegan KF, Stroh JG. Site-directed conjugation of nonpeptide groups to peptides and proteins via periodate oxidation of a 2-amino alcohol. Application to modification at *N*-terminal serine. *Bioconjugate Chem.* 1992; **3**: 138–146.
7. Far S, Melnyk O. A novel α, α' -diaminoacetic acid derivative for the introduction of the α -oxo aldehyde functionality into peptides. *Tetrahedron Lett.* 2004; **45**: 1271–1273.
8. Fields GB, Noble RL. Solid phase peptide synthesis utilizing 9-fluorenylmethoxy carbonyl amino acids. *Int. J. Peptide Protein Res.* 1990; **35**: 161–214.
9. Häfelinger G, Kuske FKH. General and theoretical aspects of amidines and related compounds. In *The Chemistry of Amidines and Imidates, volume 2*, Patai S, Rappoport Z (eds). Wiley: Chichester, 1991; 1–100.