Research Article

Journal of PeptideScience

Received: 15 March 2011 Revised: 11 October 2011 Accepted: 14 October 2011 Published online in Wiley Online Library: 7 November 2011

(wileyonlinelibrary.com) DOI 10.1002/psc.1429

Synthesis and analytical investigation of C-terminally modified peptide aldehydes and ketone: application to oxime ligation

Corinne Buré,[a,b]* Philippe Marceau,[a] Hervé Meudal[a] and Agnès F. Delmas[a]

C-terminally modified peptides aldehyde (glycinal and alpha-oxo aldehyde peptides) and ketone (pyruvic acid-containing peptide) were synthesised to get new insights into the mechanism of acido-catalysed oxime ligation. Their tetrahedral hydrated forms were investigated in solution and in the gas phase, using NMR and in-source collision-induced dissociation mass spectrometry, respectively, and the kinetics of the oximation reactions followed using analytical HPLC. The results obtained confirmed that the first step of the oximation reaction was the limiting step for the pyruvic acid-containing peptides because of the steric effect and of the carbon angular strain of the ketone. The second step is the determining step for the aldehyde peptides because the basicity of the oxygen of the hydroxyl function of the tetrahedral form is greater for glycinal than for alpha-oxo aldehyde. These data strongly suggest that the hydrated form of the aldehyde partner has to be considered when oxime reactions are performed in aqueous buffer. Copyright © 2011 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: peptide aldehyde; peptide ketone; ESI/mass spectrometry; NMR; oxime ligation; HPLC kinetic

Introduction

The chemoselective ligation, i.e. condensation of unprotected peptides in aqueous medium, is a useful method allowing the assembly of biomolecules. In this technique, aldehyde group [1] is the electrophilic partner for the formation of oxime [2,3], hydrazone [2,4–6] and thiazolidine [2,7–9] bond. Among these ligations, the oxime ether is the most largely used chemistry because of the *in vitro* stability of the oxime bond in a wide range around the physiological pH [2,10,11] and the high efficiency and selectivity of oximation reactions [12–14]. The oximation reactions have been applied to the ligation of peptides with peptides or proteins [15], with carbohydrates [16,17] and with oligonucleotides [18]. Oxime chemoselective ligation that requires minimal chemical handling is carried out under optimal conditions at moderate acidic pH without any degradation or unwanted side reactions [10,11,19]. Moreover, to accelerate oxime ligation, aniline was recently used as a nucleophilic catalyst [5]. Indeed, aniline allowed the use of much lower concentrations of reagents and/or reaction at a pH closer to neutral.

In previous reports, we observed that under slightly acidic conditions [10,11], the methyloxime-forming ligation proceeded at a faster rate than the α-oxo oxime, using model peptides [20], immunologically relevant peptides [21] or glycopeptides [22]. These acido-catalysed oxime reactions were assumed to involve two equilibria, with a carbinolamine as the tetrahedral intermediate (Scheme 1) [19]. Although a great body of literature deals with the kinetics of carbonyl compound reactions with nucleophiles such as aminooxy or hydrazine, no data are available concerning differences in reactivity of peptide aldehydes whose electrophilicity is directed by the substituent of the carbonyl. Moreover, although there are evidences that the unhydrated rather than the hydrated aldehyde is the reactive species in the aminooxy addition [23], the hydration degree of the peptide aldehyde partner is scarcely documented in reports related to oxime ligation reaction.

In this study, we revisited the kinetics of the oximation reaction using two aldehyde-containing peptide sequences, glycinal and alpha-oxo aldehyde peptides, which only differ by the substituents of the electrophilic carbonyl group and a unique aminooxy-containing peptide as the nucleophilic partner (Table 1) in buffered aqueous conditions at pH 4.6. A pyruvic acid-containing peptide, i.e. a ketone-containing peptide, known to usually induce incomplete oximation reaction was included in this study. To get further insight into the oxime ligation reaction, the hydration of each aldehyde-containing and pyruvic acid-containing peptide was analysed using NMR and in-source collision-induced dissociation (CID) using electrospray mass spectrometry (ESI-MS). The aim was to extend our knowledge in the oximation reactions to better understand why the methyloxime-forming ligation proceeded faster than the ketooxime-forming ligation [21,22] and, thus, to improve strategies to synthesise artificial proteins.

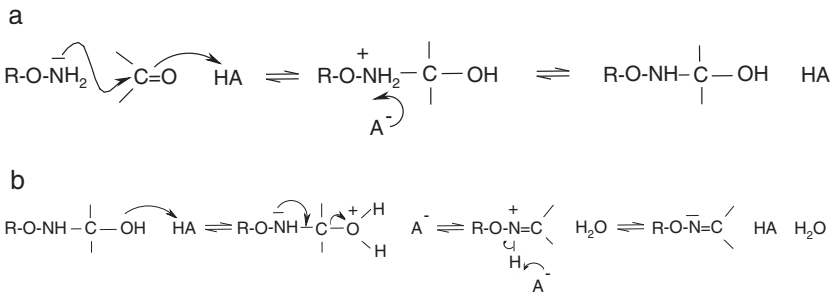
Results and Discussion

In this paper, we reinvestigated the reactivity of peptide aldehydes (glycinal and alpha-oxo aldehyde peptides) and ketone (pyruvic acid-containing peptide) towards an aminooxy-containing peptide

* *Correspondence to: Corinne Buré, Université de Bordeaux, Chimie et Biologie des Membranes et des Nanoobjets (CBMN) – UMR 5248 Centre de Génomique Fonctionnelle BP 68, Université Bordeaux 2 Victor Segalen, 146, rue Léo Saignat 33076, Bordeaux Cedex, France. E-mail: c.bure@cbmn.u-bordeaux.fr*

a *Centre de Biophysique Moléculaire CNRS UPR4301, affiliated to the University of Orléans and INSERM, rue Charles Sadron 45071 Orléans cedex 2, France*

b *Present address: Chimie et Biologie des Membranes et des Nanoobjets (CBMN) – UMR 5248 Centre de Génomique Fonctionnelle BP 68, Université Bordeaux 2 Victor Segalen, 146, rue Léo Saignat 33076 Bordeaux Cedex, France*



Scheme 1. Reaction mechanism of oximation reaction.

Table 1. Sequences used in this study

Model peptide	
1	H-Tyr-Lys-Ala-Gly-Leu-Gly-Ala-NH-CH_2-CO-H
2	H-Tyr-Lys-Ala-Gly-Leu-Ala-NH-$(CH_2)_2$-NH-CO-CO-H
3	H-Tyr-Lys-Ala-Gly-Leu-Gly-NH-$(CH_2)_2$-NH-CO-CO-CH_3
MUC1-derived peptide	
4	H-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-NH-CH(CH_3)-CO-NH-CH_2-CO-H
5	H-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-NH-CH_2-CH_2-NH-CO-CO-H
6	H-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-NH-CH_2-CH_2-NH-CO-CO-CH_3
Aoa-peptide	
7	H_2N-O-CH_2-CO-Ala-Leu-Lys-Trp-Ser-Leu-Ala-OH

in buffered aqueous conditions at pH 4.6. According to Sayer *et al.*, studies of the detailed mechanism of the addition of nitrogen nucleophiles to carbonyl compounds have demonstrated the generality of the mechanism outlined in Scheme 1 [19]. In acidic medium, the first step corresponds to a nucleophilic addition of the amine on the carbonyl compound affording a tetrahedral carbon as a carbinolamine. It followed a concerted mechanism because, in our case, the amine function is weakly basic. The second step corresponds to the protonation of the hydroxyl leading to the dehydration of the carbinolamine into an oxime. Despite the importance of the oxime ligation in bioconjugation of macromolecules [24] and a voluminous literature dealing with the kinetics of reaction of carbonyl compound with aminooxy or hydrazine-containing compounds, no data are available concerning a comparison of the reactivity of peptides whose electrophilic moiety is directed by the substituent of the carbonyl. Taking advantage of our previous experience in C-terminally modified peptides [25–27], we efficiently synthesised peptides bearing a glycinal, an alpha-oxo aldehyde or a pyruvic acid moiety, to study the addition to water giving the hydrated form of the peptide and the addition to aminooxy-containing peptide giving the oxime ether compound.

Peptide Design and Synthesis

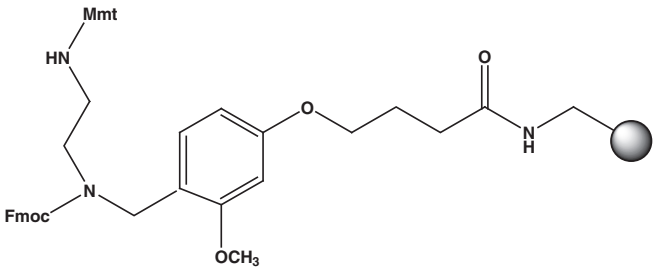
Two series of C-terminally modified peptide aldehydes and ketone were synthesised (Table 1). The first series correspond to model peptides **1–3**. The peptide sequence of the second series (peptides **4–6**) was derived from the repeat unit of the glycoprotein MUC1 [28] whose tumour-associated form is a good target for immune intervention [29] and was part of our ongoing MUC1-based tumour vaccine programme [21,22].

In each series of peptides, besides the pyruvic acid-containing peptide (peptides **3** and **6**), two aldehyde peptides, one glycinal (peptides **1** and **4)** and one alpha-oxo aldehyde peptide (peptides **2** and **5**) were designed to mainly differ in the α-substituent of the carbonyl function to modulate their electronic properties (Table 1). A great care was taken to design electrophilic peptides with similar molecular masses for in-source fragmentation of the hydrated aldehyde being not impacted by the peptide length [30]. Moreover, to avoid fragmentation that could interfere with the studied phenomenon such as losses of H_2O or NH_3 from the C-terminus as -COOH or -$CONH_2$, respectively [30], the aldehyde and ketone group were installed at the C-terminus. One tyrosinyl and one lysyl residues were included in the sequence of the model peptides **1–3** to help following the kinetics of the oximation reaction using UV-HPLC and to facilitate the peptide ionisation using ESI/MS, respectively. The biologically relevant peptides **4–6** of the second series (Table 1), derived from the sequence of the repeat unit of the MUC1 protein, are slightly different in molecular masses to keep unmodified the primary sequence.

The synthesis of glycinal peptides **1** and **4** were carried out as previously described. The glycinal group was introduced at the C-terminus by nucleophilic displacement of the ester bond between the peptide and the phenylacetamidomethyl (PAM) linker with aminoacetaldehyde-dimethylacetal [25]. The PEGA (polyethylene glycol dimethylacrylamide) resin was used as a polymeric matrix to facilitate the aminolysis of a peptide longer than ten amino acids [26].

To introduce the alpha-oxo aldehyde and the pyruvic acid at the C-terminus of a peptide sequence, we took advantage of the commercially available NovaTag™ resin (Novabiochem-Merck, Nottingham, UK) (Scheme 2) that was designed for bis-labelled peptides at the N and C-terminal side for FRET (Fluorescence Resonance Energy Transfer) experiments [31]. The elimination of the Mmt ((4-methoxyphenyl)diphenylmethyl) group under very mild acidic condition enables the installation of the alpha-oxo aldehyde and the pyruvic acid group at the C-terminus by coupling a serine and pyruvic acid, respectively. The solid phase

Journal of PeptideScience



Scheme 2. NovaTag™ resin.

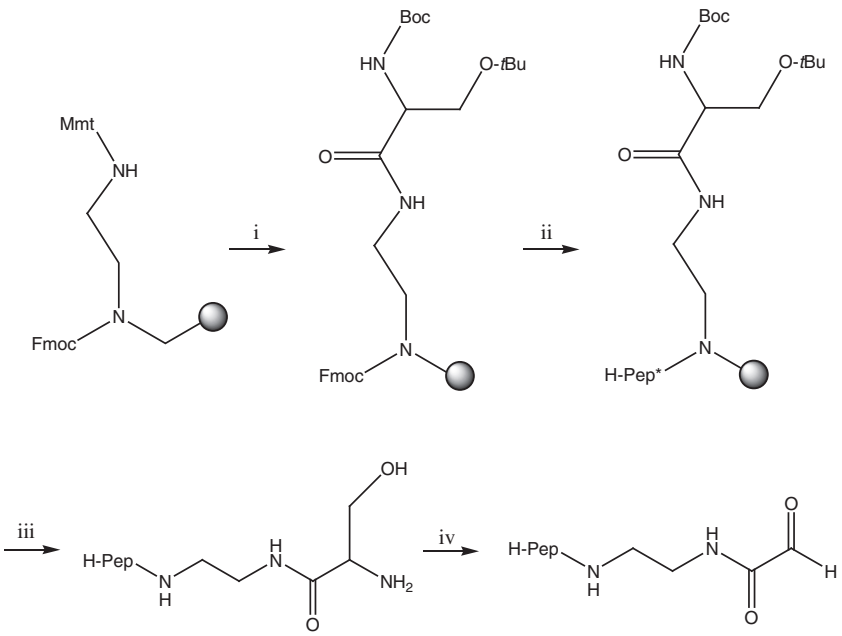
synthesis of alpha-oxo aldehyde peptides **2** and **5** is depicted in Scheme 3. After removal of the Mmt group, the seryl residue was introduced as Boc-Ser(*t*Bu)-OH with HBTU (2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate)/DIEA (diisopropylethylamine) as coupling reagent [32]. Completion of the reaction was checked using Kaiser's test [33]. After removal of the Fmoc group, the first amino acid was coupled using HATU ((O-(benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate))/DIEA as an efficient coupling reagent on hindered amine [34], and the completion of the reaction was checked using the choranil test for secondary amine [35]. The elongation was then carried out using the Fmoc/*t*Bu strategy, and HBTU was used as coupling reagent. After cleavage of the peptide from the resin with TFA treatment, the crude peptide was oxidised using $NaIO_4$ to afford an alpha-oxo aldehyde at the C-terminus. Peptides **2** and **5** were purified using HPLC and obtained in 50 and 60% overall yield, respectively. This straightforward method is an interesting alternative to that developed by Melnyk *et al.* [36] with the advantage of affording a masked -CO-CHO.

The synthesis of C-terminal modified pyruvic acid-containing peptides **3** and **6** has already been reported [27]. Pyruvic acid was cleanly coupled at the C-terminus provided that no base or reducing agent was used for the synthetic pathway [27]. It was thus introduced after peptide elongation, the N-terminal amino acid being N^{α}-Boc protected.

The synthesis of the aminooxy-containing peptide **7** was carried out using Fmoc SPPS (Solid-Phase Peptide Synthesis) elongation. Boc-(aminooxy)acetic acid derivative (Boc-Aoa) was coupled using DCC/HOBt [37] to minimise *N*-overacylation of the NH-O group [37–39]. Because of the high sensitivity of aminooxy to trace amount of aldehydes and ketone, the peptide was precipitated into high-grade diethyl ether [40].

Kinetics of the Oxime Formation between Aldehyde and Ketone-Containing Peptides 1–6 and Aminooxy-Containing Peptide 7

The reactivity of the different electrophilic peptides **1–6** in oxime ligation reaction was evaluated at pH 4.6 by following the formation of the oxime bond using analytical HPLC. Figure 1 depicted the percentage of the ligation product as a function of time. The resulting data were fit to the second-order rate equation, leading to calculation of the ligation rate k_{obs}. The half-time ($t_{1/2}$) of the reaction corresponds to the time where 50% of ligation product was obtained compared with the maximum ligation product. For model peptides (Figure 1(A)), $t_{1/2}$ is 63 ($k_{obs} = 0.35 \pm 0.02\ ^{M^{-1}}\ s^{-1}$), 125 ($k_{obs} = 0.176 \pm 0.009\ ^{M^{-1}}\ s^{-1}$) and 421 min ($k_{obs} = 0.052 \pm 0.005\ ^{M^{-1}}\ s^{-1}$), corresponding to glycinal, alpha-oxo aldehyde and pyruvic acid-containing peptide (Table 2). Likewise, for MUC1-derived peptides (Figure 1(B)), $t_{1/2}$ is 4 ($k_{obs} = 6.8 \pm 0.7\ ^{M^{-1}}\ s^{-1}$), 150 ($k_{obs} = 0.146 \pm 0.009\ ^{M^{-1}}\ s^{-1}$) and 185 min ($k_{obs} = 0.12 \pm 0.03\ ^{M^{-1}}\ s^{-1}$), corresponding to glycinal, alpha-oxo aldehyde and pyruvic acid-containing peptide (Table 2). Globally, in the two series of peptides tested here, the pyruvic acid-containing peptide has the highest half-time and the smallest ligation rate that led to the slowest oximation reaction. The smallest half-time and the highest ligation rate obtained for glycinal



Pep: - YKAGLA **2**
- PPAHGVTSAPDTRPAPGST **5**

Scheme 3. Synthesis of alpha-oxo aldehyde peptides **2** and **5** using NovaTag™ resin. (i) 1 M HOBt in trifluoroethanol/CH_2Cl_2 and Boc-Ser(*t*Bu)-OH, HBTU, DIEA. (ii) 20% piperidine in DMF; Fmoc-Thr(*t*Bu)-OH, HATU, DIEA; and Fmoc/*t*Bu elongation. (iii) TFA treatment. (iv) $NaIO_4$.

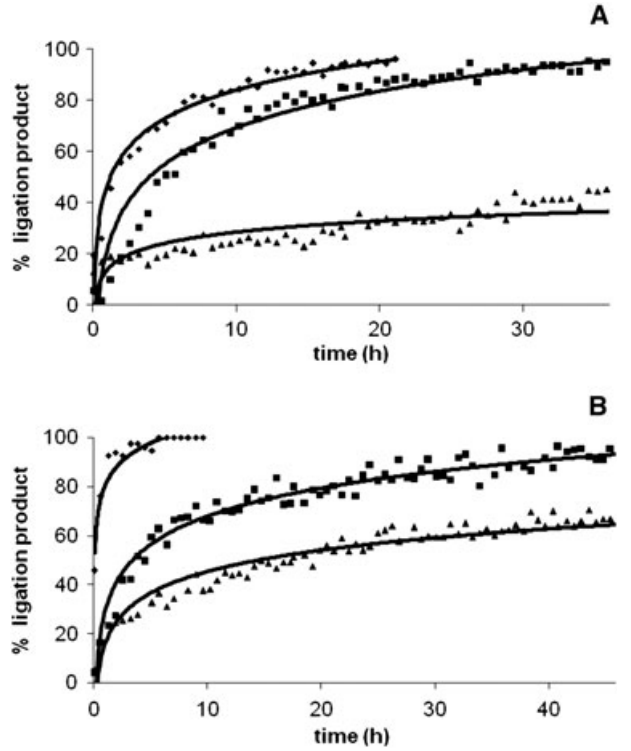


Figure 1. Kinetic of the oxime ligation. The percentage of the ligation product is in the function of time for (A) the model peptides (♦: glycinal peptide **1**; ■: alpha-oxo aldehyde peptide **2**; ▲: pyruvic acid-containing peptide **3**) and (B) the MUC1-derived peptides (♦: glycinal peptide **4**; ■: alpha-oxo aldehyde peptide **5**; ▲: pyruvic acid-containing peptide **6**).

Table 2. Half-time of the ligation reaction and ligation rate measured by HPLC

	$t_{1/2}$ (min)	k_{obs} ($^{M-1}$ s^{-1})
Model peptide		
1	63	$0.35 \pm 0.0.2$
2	125	0.176 ± 0.009
3	421	0.052 ± 0.005
MUC1-derived peptide		
4	4	6.8 ± 0.7
5	150	0.146 ± 0.009
6	185	0.12 ± 0.03

peptides confirm our previous observations that the glycinal peptides led more quickly to the oxime ether compound than the alpha-oxo aldehyde peptides. When comparing the two series of peptides, model peptides appear to have a different ligation rate than MUC1-derived peptides. For glycinal peptides, the oxime ligation is much more rapid with MUC1-derived peptide ($t_{1/2} = 4$ min) than with model peptide ($t_{1/2} = 63$ min). A similar phenomenon is observed for pyruvic acid-containing peptide because the oxime ligation is faster with MUC1-derived peptide ($t_{1/2} = 185$ min) than with model peptide ($t_{1/2} = 421$ min). The difference between the two series that could be noticed is the difference of length. The reaction seems to be faster when the peptide is longer. Surprisingly, the oxime ligation is quite similar for alpha-oxo aldehyde peptides ($t_{1/2} = 150$ min for MUC1-derived peptide and $t_{1/2} = 125$ min for model peptide). This difference should be explained by another phenomenon, which could be a slight variation in pH (because of the accuracy of the pH measure) after the mixture of peptides to form the oxime, implicating a difference in reaction rate, because the oximation reaction is very sensitive to pH. Indeed, it proceeds faster at a low pH and very slowly with a pH close to neutral [2,10]. We have previously shown that the oximation reaction implicated that the MUC1-pyruvic acid-containing peptide occurred faster at pH 2 than at pH 4.6 [27].

Hydrated Form of Model Peptides 1–3 and MUC1-Derived Peptides 4–6 Studied using NMR and In-Source CID ESI-MS

According to literature, the most reactive carbonyl compounds to addition of nitrogen nucleophiles are generally the most hydrated in aqueous solution [23,41]. The percentage of the hydrated form of glycinal, alpha-oxo aldehyde and pyruvic acid-containing peptide was determined using NMR (Table 2). For the two series of peptides, the hydration of the modified peptides under oximation reaction conditions at pH 4.6 is ranked as follows: alpha-oxo aldehyde peptide **2** and **5** > glycinal peptide **1** and **4** >> pyruvic acid-containing peptide **3** and **6**. This hydration difference can be explained because in function of the pH, the hydration of the aldehyde-containing or ketone-containing peptides could vary as for keto acids [42]. Indeed, the hydrated form of keto acids is favoured at low pH (below the p*K*a of the carboxylic acid group), whereas the keto form is favoured at neutral pH. In this way, pH 4.6 is the more favoured for the hydration of aldehyde-containing peptide than for ketone-containing peptide. Moreover, the percentage of the hydrated forms of glycinal, alpha-oxo aldehyde and pyruvic acid-containing peptide are the same for model peptides than for MUC1-derived peptide. Indeed, as model peptides have the same chemical environment near the C-terminal function than the MUC1-derived peptides, the NMR results are close.

We have previously shown that the equilibrium occurring in solution between a peptide aldehyde and its tetrahedral hydrated form, the *gem*-diol, is retained in an electrospray source during the ionisation process [43]. Increasing the voltage applied to the cone increases the energy of ions as they go through the ESI atmospheric pressure/vacuum interface and progressively induced the fragmentation of the hydrated form into an isobaric ion of the aldehyde [30,43]. This energy-dependent ESI-MS methodology enables to compare the stability of the ionised tetrahedral hydrated form of different peptide aldehydes provided that their masses and charges were similar [43]. For each model peptide, the molecular peptide MH^+ ion and its hydrated form, the $[M + H_2O + H]^+$ ion, were observed on the mass spectra (data not shown). In-source CID mass spectra were recorded by increasing the sample cone voltage from 10 to 35 V in 5-V increments. The percentages of the hydrated forms of the peptides were calculated as mentioned in the experimental part and expressed as a function of the cone voltage (Figure 2). The hydrated forms of pyruvic acid-containing peptides **3** and **6** were hardly detected clearly indicating different hydration behaviour compared with the aldehyde peptides. The gas phase stability of the hydrated forms of the electrophilic peptides can be ranked as follows: alpha-oxo aldehyde peptide > glycinal peptide > pyruvic acid-containing peptide. Because of the steric effect and of the carbon angular strain, the functional carbon of a pyruvic acid is less accessible to nucleophilic addition than that of aldehydes. The first step of the oxime ligation is then unfavoured

Journal of PeptideScience

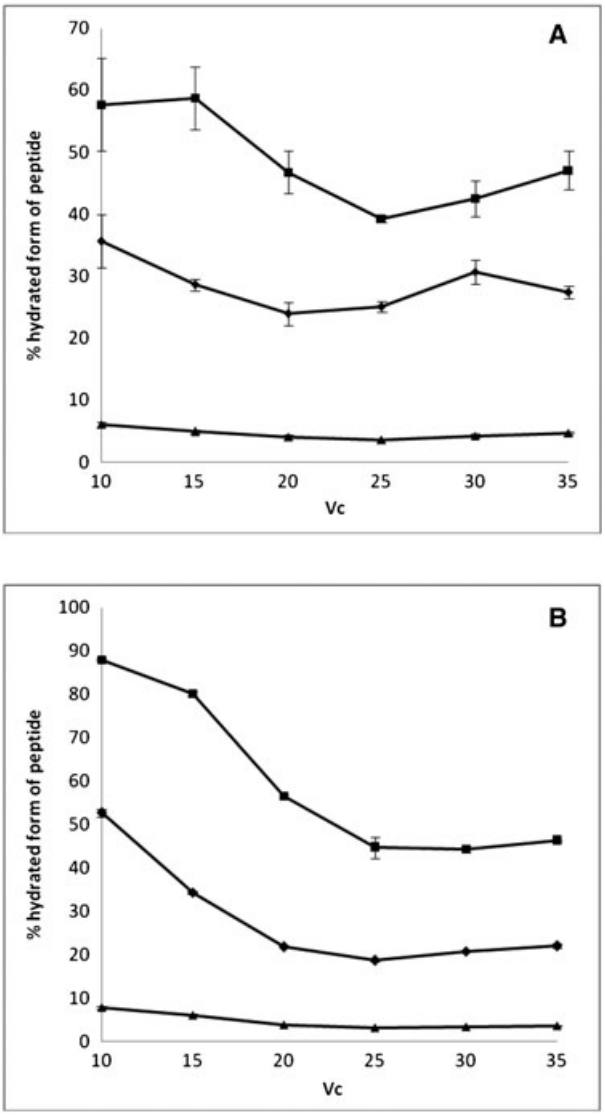


Figure 2. In-source CID ESI-MS of aldehyde-containing and pyruvic acid-containing peptides. The percentage of the hydrated form of the peptide is in the function of cone voltage Vc for (A) the model peptides (♦: glycinal peptide **1**; ■: alpha-oxo aldehyde peptide **2**; ▲: pyruvic acid-containing peptide **3**) and (B) the MUC1-derived peptides (♦: glycinal peptide **4**; ■: alpha-oxo aldehyde peptide **5**; ▲: pyruvic acid-containing peptide **6**).

and is the limiting step for pyruvic acid-containing peptides in the oximation reactions. Contrary to these ketone peptides, the glycinal and the alpha-oxo aldehyde peptides are nearly equally and totally hydrated in solution, as judged using NMR, indicating that the nucleophilic addition is not a limiting step. This observation is in accordance with the study of Sayer *et al.* [19] on the oxime bond formation.

Even at the lowest cone voltage, that is Vc = 10 V (Table 3), the hydration degree of peptides is underestimated when compared with the hydration degree measured in solution using NMR (Table 2). This indicates that even if the fragmentation is minimised at such a low cone voltage, it still occurs. In contrast to the hydration degree determined using NMR (Table 2), which is the same for a given type of aldehyde whatever the peptide series, the hydration degree of the series of model peptides is more underestimated than that of the series of the peptide derived from MUC1 (Tables 2, 3). This is in accordance with the

Table 3. Percentages of hydrated forms of the peptides measured using NMR and ESI-MS at Vc = 10 V

	NMR (%)	ESI-MS (%)
Model peptide		
1	93	36
2	100	58
3	40	6
MUC1-derived peptide		
4	92	52
5	100	87
6	38	8

fragmentation in the gas phase that is facilitated for ions of lower *m/z* value [44]. More interesting, in gas phase, the underestimation of hydration degree of the glycinal peptide is more dramatic than that of the alpha-oxo aldehyde peptide suggesting that its ionised tetrahedral hydrated form is less stable than that of the alpha-oxo aldehyde peptide. The dehydration of the tetrahedral form of the aldehyde depends mainly on the protonation of the oxygen of the hydroxyl function whose glycinal is more basic than that of the alpha-oxo aldehyde owing to electronic effect. This hydrated form can impair the nucleophilic attack of the aminooxy peptide in the oxime reaction. Moreover, the stability of the hydrated form of the aldehyde can be indicative of the stability of the carbinolamine intermediate in the oxime reaction that cannot be isolated. In the cases studied here, the stability of the tetrahedral forms related to the alpha-oxo aldehyde correlated well with their lesser reactivity in the oximation reaction showed by their highest half-time and their smallest ligation rate.

Conclusion

In this paper, we reinvestigated the reactivity of peptide aldehydes and ketone towards an aminooxy-containing peptide in buffered aqueous conditions at pH 4.6. According to Sayer *et al.* [19], in acidic medium, the first step corresponds to a nucleophilic addition of the amine on the carbonyl compound affording a tetrahedral carbon as a carbinolamine. It followed a concerted mechanism because in our case, the amine function is weakly basic. The second step corresponds to the protonation of the hydroxyl leading to the dehydration of the carbinolamine into an oxime.

Surprisingly, the oximation ligation has a different rate with MUC1-derived peptide than with model peptide, except for alpha-oxo aldehyde peptide. As the oximation ligation is sensitive to pH, this difference could be attributed to a slight variation of pH (because of the accuracy of the pH measure) when the peptides were mixed during the oximation ligation. For the two series of peptides, NMR results were close because the structure environment of the C-terminal function is the same, whereas ESI-MS results showed a difference in stability of the hydrated form ion. That could be as a result of the difference of length because the fragmentation in ESI-MS is related to the length of the peptide.

For the two series of peptides, in NMR and in ESI-MS, the hydration of the modified peptides under oximation reaction conditions at pH 4.6 is ranked as follows: alpha-oxo aldehyde peptide **2** and **5** > glycinal peptide **1** and **4** >> pyruvic acid-containing peptide **3** and **6**. As expected, we showed that for pyruvic

acid-containing peptides, the oxime ligation is less rapid compared with peptide aldehydes. The peptide ketones were also less hydrated than the peptide aldehydes studied here because of the steric effect and of the carbon angular strain. The first step of the oxime ligation is then unfavoured and is the limiting step for pyruvic acid-containing peptides in the oximation reactions. The second step of the reaction is determining for aldehyde-containing peptides. The dehydration of the tetrahedral form of the aldehyde depends mainly on the protonation of the oxygen of the hydroxyl function whose glycinal is more basic than the one of the alpha-oxo aldehyde owing to electronic effect. This hydrated form can impair the nucleophilic attack of the aminooxy peptide in the oxime reaction.

This study gives an insight in the mechanism of oximation reaction and strongly suggests that the hydrated form of the aldehyde partner has to be considered when oxime reactions are performed in aqueous buffer. These results would open new perspectives for the design of reactive peptides for bioconjugation and the synthesis of artificial proteins.

Experimental Part

General

Organic solvents were from SDS (Peypin, France) or Carlo Erba (Val de Reuil, France), with DCM, NMP (N-methyl-2-pyrrolidone) and piperidine of synthesis grade and MeCN and MeOH of HPLC grade. Diethyl ether was from SDS (Peypin, France) and the higher purity (>99.5) from Acros (Noisy-le-Grand, France). DMF was from Applied Biosystems (Courtaboeuf, France). TFA was from SDS (Peypin, France). Water was purified on a Milli-Q reagent system (Millipore, St Quentin Fallavier, France). Boc-Ala-Φ-CH_2-COOH was from Polypeptide (Strasbourg, France). NovaTag™ resin and PEGA resin were purchased from Novabiochem-Merck (Nottingham, UK). Fmoc-protected amino acids were obtained from Novabiochem-Merck. Boc-Ser(*t*Bu)-OH was from Bachem (Bubendorf, Switzerland). Aminoacetaldehyde-dimethylacetal and aminooxy acetic acid (Aoa) were from Sigma (St Quentin Fallavier, France). N^{α}-Boc–protected Aoa was obtained according to the process presented by Offord *et al.* [45]. Coupling reagents were purchased from commercial sources and were of the highest purity available.

Analytical and semi-preparative RP-HPLC were performed using a LaChrom 7000 system equipped with a Merck-Hitachi L-7100 pump and a Merck-Hitachi L-7455 diode array detector with an Merck-Hitachi D-7000 interface. The column used were a C18 column, nucleosil 300 Å (5 μm, 250 × 4.6 mm) for the analytical separations or a C18 column, nucleosil 300 Å (5 μm, 250 × 10.5 mm) for purification. Peptides were eluted with a linear gradient of MeCN/H_2O/0.1% TFA. Buffer A was water containing 0.1% TFA, and buffer B was MeCN containing 0.1% TFA. The elution followed at 275 nm.

General procedure for automated solid phase synthesis is as follows.

Solid phase peptide synthesis was run on an automated synthesiser 433A from Applied Biosystems using Fmoc/*t*Bu chemistry at 0.1-mmole scale with HBTU/HOBt as coupling reagents. Tenfold excess was used for protected amino acids and coupling reagents. The side-chain protecting groups used were Lys(Boc), Ser(*t*Bu), Trp(Boc) and Tyr(*t*Bu). The 0.1-mmole scale programme purchased from the manufacturer was used, with a single coupling followed by capping with acetic anhydride solution.

General procedure for manual coupling is as follows.

Protected amino acid (10 equiv) and HATU (10 equiv) were dissolved in DMF. The solution was then transferred on resin (1 equiv) placed in a fritted syringe, and iPr$_2$NEt (12 equiv) was added after 5 min of stirring. After 2 h, the reactants were removed by filtration and the resin was washed with DMF (three times).

Electrospray mass spectrometry

The ESI-MS analyses were performed on a triple quadrupole mass spectrometer (Quattro II, Micromass, Manchester, UK) equipped with a nebuliser-assisted electrospray source. The high-flow nebuliser was operated in standard mode with N_2 as both nebulising (20 l/h) and drying (250 l/h) gas. A voltage difference of 3 kV was applied between the capillary and the counter electrode. In-source CID mass spectra were recorded by increasing the sample cone voltage from 10 to 35 V in 5-V increments. The ion source was kept at 80 °C. Instrument control and data analysis were accomplished using MassLynx application software, version 3.4, from Micromass (Manchester, UK). Calibration of the mass spectrometer was performed using myoglobin.

Each peptide was dissolved in 80/20 acetonitrile/5 mM ammonium acetate adjusted to pH 4.6 with acetic acid. Three microlitres of each peptide was injected into a Rheodyne model 7125 injection valve with a 10 μl sample loop. The peptide was eluted into the mass spectrometer using 80/20 acetonitrile/H_2O as mobile phase delivered using an isocratic LC-10 AD pump (Shimadzu, Les Ulis, France) at a flow rate of 20 μl min^{-1}. ESI mass spectra were obtained in the positive ion mode. Monoisotopic and average masses were obtained for model peptides and MUC1-derived peptides, respectively.

The percentages of the hydrated and non-hydrated forms of the peptides were calculated by mean of the peak's intensity at Vc = 10 V as

$$\%\ \text{hydrated form} = \{I[M + H_2O + zH]^{z+}/[I(M + zH)^{z+} + I(M + H_2O + zH)^{z+}]\} \times 100,$$

where z is the ion charge, M is the molecular mass of the ion and I is the peak intensity on the mass spectrum.

Nuclear magnetic resonance

The NMR analyses were performed on an Avance-500 MHz (Bruker, Bremen, Germany) spectrometer equipped with a 5-mm Broadband Inverse probe. The peptides were dissolved in D_2O at pH 4.6 and analysed at a temperature of 25 °C. The percentages of the hydrated and non-hydrated forms of the peptides were calculated by integration of peaks corresponding to the chemical function on NMR spectra. According to Andersson *et al.* [46], the chemical shift of the proton of the aldehyde function (-C<u>H</u>O) is around 9.5 ppm and the chemical shift of the proton of the hydrated aldehyde function (-C<u>H</u>(OH)$_2$) is around 5.2 ppm. In the case of the ketone peptide, according to the chemical shift tables, the chemical shift of the protons of the methyl-ketone function (-CO-C<u>H</u>$_3$) is around 2.4 ppm and the chemical shift of the protons of the hydrated methyl-ketone function (-C(OH)$_2$-C<u>H</u>$_3$) is around 1.5 ppm.

Peptide synthesis

*Peptide glycinal (**1**) and (**4**)*

The syntheses were carried out as previously described [25,26].

- H-Tyr-Lys-Ala-Gly-Leu-Gly-Ala-NH-CH_2-CO-H (**1**): R_t 12.2 min (gradient 0–30% B over 25 min with A: H_2O : MeCN

Journal of PeptideScience

95/5 + 0.1% TFA; B: H_2O : MeCN 5/95 + 0.1% TFA); ESI-MS: [MH^+] 720.4 (calculated for $C_{33}H_{53}N_9O_9$ 719.8).
- H-Muc-NH-CH(CH_3)-CO-NH-CH_2-CO-H (**4**): R_t 12.2 min (gradient 10–35% over 40 min with A: H_2O + 0.1% TFA; B: MeCN + 0.1% TFA); ESI-MS : [MH^+] 1928.3 (calculated for $C_{82}H_{130}N_{26}O_{28}$ 1928.1).

*Alpha-oxo aldehyde peptides (**2**) and (**5**)*

NovaTag™ resin (0.1 mmol, 256 mg) was introduced in a fritted syringe and treated with 1 M HOBt in DCM/TFE 50:50 (2 ml) two times for 1 h. After DMF rinse (three times), Boc-Ser(*t*Bu)-OH (10 equiv, 1 mmol, 261 mg) was manually coupled following the general procedure. The resin was rinsed with DMF (three times) and treated with piperidine/DMF 20:80 (2 ml) for 30 min (two times). Fmoc-Ala-OH was manually coupled. The resin was then introduced into a reactor for solid phase synthesis. The elongation was conducted as described in the general procedure. Then, the peptidyl-resin was introduced into a fritted syringe and treated with TFA/Phenol/H_2O/iPr_3SiH, 88:5:5:2 for 2 h. After filtration and precipitation of the peptide in cold diethyl ether, the peptide was recovered using centrifugation.

- H-Tyr-Lys-Ala-Gly-Leu-Gly-NH-$(CH_2)_2$-NH-CO-CO-H (**2**): R_t 17.6 min (gradient 3–28% B over 25 min with A: H_2O : MeCN 95/5 + 0.1% TFA; B: H_2O : MeCN 5/95 + 0.1% TFA); ESI-MS: [MH^+] 720.5 (calculated for $C_{33}H_{53}N_9O_9$ 719.8).
- H-Muc-NH-CO-CO-H (**5**): R_t 11.9 min (gradient 10–35% B over 25 min with A: H_2O + 0.1% TFA; B: MeCN + 0.1% TFA); ESI-MS: [MH^+] 1914.3 (calculated for $C_{81}H_{128}N_{26}O_{28}$ 1914.1).

*Pyruvic acid-containing peptides (**3**) and (**6**)*

The synthesis was carried out as previously described [27].

- H-Tyr-Lys-Ala-Gly-Leu-Gly-NH-$(CH_2)_2$-NH-CO-CO-CH_3 (**3**): R_t 17.8 min (gradient 5–40% B over 40 min, with A: H_2O : MeCN 95/5 + 0.1% TFA, B: H_2O : MeCN 5/95 + 0.1% TFA); ESI-MS: [MH^+] 720.4 (calculated for $C_{33}H_{53}N_9O_9$ 719.8).
- H-Muc-NH-CH_2-CH_2-NH-CO-CO-CH_3 (**6**): R_t 15.0 min (gradient 10–35% over 40 min with A: H_2O + 0.1% TFA; B: MeCN + 0.1% TFA); ESI-MS: [MH^+] 1928.6 (calculated for $C_{82}H_{130}N_{26}O_{28}$ 1928.1).

*H_2N-O-CH_2-CO-Ala-Leu-Lys-Trp-Ser-Leu-Ala-OH (**7**)*

Starting from Fmoc-Ala-Wang resin (119 mg, 0.1 mmol), the elongation was performed according to the general procedure. Peptidyl-resin (0.04 mmol) were introduced in a fritted syringe, and Boc-Aoa-OH (10 equiv, 95 mg, 0.5 mmol) was manually coupled using the general procedure. Half of the peptidyl-resin was washed with DCM (three times) and treated with TFA/Phenol/H_2O/iPr_3SiH, 88:5:5:2. Peptide was then precipitated and washed with ice-cold diethyl ether without carbonyl-containing compounds. The crude peptide was recovered using centrifugation. R_t 12.9 min (gradient: 20–50% B over 27 min, with A: H_2O : MeCN 95/5 + 0.1% TFA, B: H_2O : MeCN 5/95 + 0.1% TFA); ESI-MS: [MH^+] 861.6 (calculated for $C_{40}H_{64}N_{10}O_{11}$ 861.0).

Kinetics of the oxime formation

The oximation reaction was carried out in 0.1 M NaAcO, pH 4.6 at 25 °C [3]. Each aldehyde or ketone peptide (1.13 mM) was in a 1.5-fold excess over the aminooxy-containing peptide (0.76 mM). Every 40 min, aliquots were analysed using micro-HPLC (CapLC) for 37 h. The percentage of ligation product calculated using integration of the HPLC peaks at 280 nm was expressed as a function of time. Micro-HPLC analyses were performed using a CapLC (Waters, Manchester, UK) equipped with a C18 column, Symmetry 100 Å (3.5 μm, 0.32 × 50 mm). Detection was achieved on a diode array detector from 200 to 350 nm. Peptides were eluted at 5 μl min^{-1}. The mobile phase was composed of buffer A, H_2O + 0.1% TFA, and buffer B, MeCN + 0.1% TFA. The elution gradient was 10–30% B for 20 min and 30–100% B for 3 min.

Acknowledgements

We thank the Association pour la Recherche sur le Cancer (ARC) for financing the CapLC instrument. A. D. thanks Dr Vincent Aucagne for helpful discussions.

References

1 Melnyk O, Fehrentz J A, Martinez J, Gras-Masse H Functionalization of peptides and proteins by aldehyde or keto groups. *Biopolymers* 2000; **55**: 165–186.
2 Shao J, Tam JP. Unprotected peptides as building blocks for the synthesis of peptide dendrimers with oxime, hydrazone, and thiazolidine linkages. *J. Am. Chem. Soc.* 1995; **117**: 3893–3899.
3 Rose K Facile synthesis of homogeneous artificial proteins. *J. Am. Chem. Soc.* 1994; **116**: 30–33.
4 Brunel FM, Lewis JD, Destito G, Steinmetz NF., Manchester M, Stuhlmann H, Dawson PE. Hydrazone ligation strategy to assemble multifunctional viral nanoparticles for cell imaging and tumor targeting. *Nano Lett.* 2010; **10**: 1093–1097.
5 Dirksen A, Dawson PE. Rapid oxime and hydrazone ligations with aromatic aldehydes for biomolecular labeling. *Bioconjug. Chem.* 2008; **19**: 2543–2548.
6 Bourel-Bonnet L, Pécheur EI, Grandjean C, Blanpain A, Baust T, Melnyk O, Hoflack Y, Gras-Masse H. Anchorage of synthetic peptides onto liposomes via hydrazone and alpha-oxo hydrazone bonds. Preliminary functional investigations. *Bioconjug. Chem.* 2005; **16**: 450–457.
7 Dheur J, Ollivier N, Melnyk O Synthesis of thiazolidine thioester peptides and acceleration of native chemical ligation. *Org. Lett.* 2011; **13**: 1560–1563.
8 Forget D, Boturyn D, Defrancq E, Lhomme J, Dumy P. Highly efficient synthesis of peptide-oligonucleotide conjugates: chemoselective oxime and thiazolidine formation. *Chemistry* 2001; **7**: 3976–3984.
9 Liu CF, Tam JP. Peptide segment ligation strategy without use of protecting groups. *Proc. Natl. Acad. Sci. U. S. A.* 1994; **91**: 6584–6588.
10 Rose K, Zeng W, Regamey PO, Chernushevich IV, Standing KG, Gaertner HF. Natural peptides as building blocks for the synthesis of large protein-like molecules with hydrazone and oxime linkages. *Bioconjugate Chem.* 1996; **7**: 552–556.
11 Kalia J, Raines RT. Hydrolytic stability of hydrazones and oximes. *Angew. Chem. Int. Ed.* 2008; **47**: 7523–7526.
12 Canne LE, Ferré d'Amaré AR, Burley SK, Kent SBH. Total chemical synthesis of a unique transcription factor-related protein: cMyc-. Max. *J. Am. Chem. Soc.* 1995; **117**: 2998–3007.
13 Kochendoerfer GG, Tack JM, Cressman S. Total chemical synthesis of a 27 kDa TASP protein derived from the MscL ion channel of *M. tuberculosis* by ketoxime-forming ligation. *Bioconjugate Chem.* 2002; **13**: 474–480.
14 Boturyn D, Coll JL, Garanger E, Favrot MC, Dumy P. Template assembled cyclopeptides as multimeric system for integrin targeting and endocytosis. *J. Am. Chem. Soc.* 2004; **126**: 5730–5739.
15 Hackenberger CP, Schwarzer D. Chemoselective ligation and modification strategies for peptides and proteins. *Angew. Chem. Int. Ed Engl.* 2008; **47**: 10030–10074.
16 Hang HC, Bertozzi CR. Chemoselective approaches to glycoprotein assembly. *Acc. Chem. Res.* 2001; **34**: 727–736.
17 Renaudet O, Dumy P. Chemoselectively template-assembled glycoconjugates as mimics for multivalent presentation of carbohydrates. *Org. Lett.* 2003; **5**: 243–246.
18 For a review see: Singh Y, Murat P, Defrancq E. Recent developments in oligonucleotide conjugation. *Chem. Soc. Rev.* 2010; **39**: 2054–2070.
19 Sayer JM, Pinsky B, Schonbrunn A, Washtien W. Mechanism of carbinolamine formation. *J. Am. Chem. Soc.* 1974; **96**: 7998–8009.

20 Lelièvre D, Buré C, Laot F, Delmas A. Synthesis of peptide di-aldehyde precursor for successive orthogonal chemoselective ligations *via* oxime bond. *Tetrahedron Lett.* 2001; **42**: 235–238.
21 Cremer GA, Bureaud N, Lelièvre D, Piller V, Piller F, Delmas A. Synthesis of branched oxime-linked peptide mimetics of the MUC1 containing a universal T-helper epitope. *Chem. Eur. J.* 2004; **10**: 6353–6360.
22 Cremer GA, Bureaud N, Piller V, Kunz H, Piller F, Delmas AF. Synthesis and biological evaluation of a multiantigenic Tn/TF-containing glycopeptide mimic of the tumor-related MUC1 glycoprotein. *Chem. Med. Chem.* 2006; **9**: 965–968.
23 Sander EG, Jencks WP. Equilibria for additions to the carbonyl group. *J. Am. Chem. Soc.* 1968; **90**: 6154–6162.
24 Lim RK, Lin Q. Bioorthogonal chemistry: recent progress and future directions. *Chem. Commun.* 2010; **46**: 1589–1600.
25 Lelièvre D, Chabane H, Delmas A. Simple and efficient solid-phase synthesis of unprotected peptide aldehyde for peptide segment ligation. *Tetrahedron Lett.* 1998; **39**: 9675–9678.
26 Lelièvre D, Turpin O, El Kazzouli S, Delmas A. Influence of polar support for the synthesis of large C-terminal peptide aldehyde: application to chemoselective ligation. *Tetrahedron* 2002; **58**: 5525–5533.
27 Marceau P, Buré C, Delmas AF. Efficient synthesis of C-terminal modified peptide ketones for chemical ligations. *Bioorg. Med. Chem. Lett.* 2005; **15**: 5442–5445.
28 Gendler SJ, Lancaster CA, Taylor-Papadimitriou J, Duhig T, Peat N, Burchell J, Pemberton L, Lalani EN, Wilson D. Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin. *J. Biol. Chem.* 1990; **265**: 15286–15293.
29 Baldus SE, Engelmann K, Hanisch FG. MUC1 and the MUCs: a family of human mucins with impact in cancer biology. *Crit. Rev. Clin. Lab. Sci.* 2004; **41**: 189–231.
30 Buré C, Le Falher G, Lange C, Delmas A. Fragmentation study of peptide acetals and aldehydes using in-source collision-induced dissociation. *J. Mass Spectrom.* 2004; **39**: 817–823.
31 Beythien J, White PD. A solid phase linker strategy for the direct synthesis of EDANS-labelled peptide substrates. *Tetrahedron Lett.* 2005; **46**: 101–104.
32 Dourtoglou V, Gross B, Lambropoulou V, Zioudrou C. O-benzotriazolyl-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate as coupling reagent for the synthesis of peptides of biological interest. *Synthesis* 1984; **7**: 572–574.
33 Kaiser E, Colescott RL, Bossinger CD, Cook PI. Color test for detection of free terminal amino groups in solid-phase synthesis of peptides. *Anal. Biochem.* 1970; **34**: 595–598.
34 Carpino LA. 1-Hydroxy-7-azabenzotriazole. An efficient peptide coupling additive. *J. Am. Chem. Soc.* 1993; **115**: 4397–4398.
35 Vojkovsky T. Detection of secondary amines on solid phase. *Pept. Res.* 1995; **8**: 236–237.
36 Melnyk O, Fruchart JS, Grandjean C, Gras-Masse H. Tartric acid-based linker for the solid-phase synthesis of C-terminal peptide alpha-oxo aldehydes. *J. Org. Chem.* 2001; **66**: 4153–4160.
37 Decostaire IP, Lelièvre D, Zhang HH, Delmas AF. Controlling the outcome of overacylation of N-protected aminooxyacetic acid during the synthesis of an aminooxy-peptide for chemical ligation. *Tetrahedron Lett.* 2006; **47**: 7057–7060.
38 Jimenez-Castells C, de la Torre BG, Gallego RG, Andreu D. Optimized synthesis of aminooxy-peptides as glycoprobe precursors for surface-based sugar-protein interaction studies. *Bioorg. Med. Chem. Lett.* 2007; **17**: 5155–5158.
39 Mezö G, Szabó I, Kertész I, Hegedüs R, Orbán E, Leurs U, Bösze S, Halmos G, Manea M. Efficient synthesis of an (aminooxy) acetylated-somatostatin derivative using (aminooxy)acetic acid as a 'carbonyl capture' reagent. *J. Pept. Sci.* 2010; **17**: 39–46.
40 Buré C, Lelièvre D, Delmas A. Identification by mass spectrometry and tandem mass spectrometry of by-products from an orthogonal peptide ligation by oxime bonds. *Rapid Commun. Mass Spectrom.* 2000; **14**: 2158–2164.
41 Malpica A, Calzadilla M. Correlation analysis between equilibria for addition of semicarbazide and hydroxylamine to several carbonyl compounds and their equilibria for addition of water. *J. Phys. Org. Chem.* 2003; **16**: 202–204.
42 Cooper AJL, Redfield AG. Proton magnetic resonance studies of α-keto acids. *J. Biol. Chem.* 1975; **250**: 527–532.
43 Buré C, Gobert W, Lelièvre D, Delmas A. In-source fragmentation of peptide aldehydes and acetals: influence of peptide length and charge state. *J. Mass Spectrom.* 2001; **36**: 1149–1155.
44 Buré C, Boujard O, Bertrand M, Lange C, Delmas A. Collision-induced dissociation of peptide thioesters: the influence of peptide length on fragmentation. *Eur. J. Mass Spectrom.* 2005; **11**: 31–34.
45 Offord RE, Gaertner HF, Wells TN, Proudfoot AE. Synthesis and evaluation of fluorescent chemokines labeled at the amino terminal. *Methods Enzymol.* 1997; **287**: 348–369.
46 Andersson L, Isley TC, Wolfenden R. α-Aminoaldehydes: transition state analogue inhibitors of leucine aminopeptidase. *Biochemistry* 1982; **21**: 4177–4180.

