Protein Science (1998), 7:1593–1602. Cambridge University Press. Printed in the USA.

An improved chemical approach toward the C-terminal sequence analysis of proteins containing all natural amino acids

KLAAS HARDEMAN,[1,2] BART SAMYN,[1] JOHAN VAN DER EYCKEN,[2] AND JOZEF VAN BEEUMEN[1]

[1]University of Gent, Department of Biochemistry, Physiology and Microbiology, Laboratory of Protein Biochemistry and Protein Engineering, K.L. Ledeganckstraat 35, B-9000 Gent, Belgium
[2]University of Gent, Department Organic Chemistry, Synthesis and Bio-organic Chemistry, Krijgslaan 281, B-9000 Gent, Belgium

(RECEIVED February 10, 1998; ACCEPTED April 28, 1998)

Abstract

An improved chemical method, capable of derivatizing all natural amino acids to their corresponding thiohydantoins, is described. This involves activation by acetyl chloride in TFA followed by derivatization with ammonium thiocyanate. Possible interference of reactive side chains was investigated by reacting *N*-acetylamino acids as well as several peptides with propionyl chloride instead of acetyl chloride. The products were characterized by PDMS mass spectrometry and ^{1}H-NMR. This chemical method allows, for the first time, complete derivatization of *N*-acetylproline to proline thiohydantoin. Applying this chemistry to peptides with a C-terminal proline, the yields for formation of proline thiohydantoin were found to be up to 60%, depending on the peptide sequence. The previous inability to derivatize C-terminal proline to thiohydantoin was thought to stem from the fact that proline cannot form the oxazolonium ion required for efficient reaction with the thiocyanate ion. However, we have found mass spectrometric evidence for the existence of a proline oxazolonium ion, under basic as well as under acidic conditions. This improvement in derivatization of C-terminal amino acids including proline is a major step forward in the development of a general chemical C-terminal sequencing method that permits the C-terminal sequence analysis of proteins of any amino acid composition.

Keywords: acetyl chloride; amino acid thiohydantoin; C-terminal sequence analysis; oxazolonium ion; proteins; thiocyanate ion

The search for a method capable of identifying the carboxy-terminal sequence of a protein has gained renewed interest in the last few years. Although such a method might provide information on naturally N-terminal–blocked proteins or prove post-translational proteolytic processing events, it is also invaluable in the design of oligonucleotide probes for gene cloning. Several approaches combining carboxypeptidase digestion with a mass spectrometric read-out have been proposed (Patterson et al., 1995; Thiede et al., 1995; Bonetto et al., 1997). Recently, the advances in mass spectrometric techniques have enabled one to obtain peptide sequences of proteins by MS/MS techniques such as collision-induced dissociation or by post-source decay (Siuzdak, 1994; Mann & Talbo, 1996; Morris et al., 1997). Although these methods can be applied at a picomol or subpicomol sensitivity, they still require the isolation of the C-terminal peptide because high molecular weight biomolecules do not fragment in a predictable way. Furthermore, peptides with amino acids having the same molecular weight cannot be analyzed with confidence (Shevchenko et al., 1997).

Although different chemical C-terminal sequencing methods have been under investigation for several years (Ward, 1986; Inglis, 1991) only one approach, known as the thiocyanate method, and first described in 1926 (Schlack & Kumpf, 1926), has led to reasonable success in the specific derivatization of the carboxy-terminal amino acid into an easily detectable amino acid derivative. This derivatization, into a thiohydantoin, involves the reaction of a protein with a thiocyanate reagent in the presence of a carboxylic acid activating reagent. Although several attempts to apply this method have been successful, the derivatization of proline has always remained difficult, and has been a major impediment to the development of a routine chemical method for the C-terminal sequence analysis of proteins.

Here we report a new approach for chemical C-terminal sequence analysis based on the derivatization procedure first described by Kubo et al. (1971). Furthermore, we present evidence for the existence of a proline oxazolonium intermediate, under

Reprint requests to: Jozef Van Beeumen, University of Gent, Department of Biochemistry, Physiology and Microbiology, Laboratory of Protein Biochemistry and Protein Engineering, K.L. Ledeganckstraat 35, B-9000 Gent, Belgium; e-mail: Jozef.vanbeeumen@rug.ac.be.

Abbreviations: PDMS, plasma desorption mass spectrometry; TFA, trifluoroacetic acid; TH, thiohydantoin; Z, benzyloxycarbonyl.

basic as well as under acidic conditions. All the 20 naturally occurring amino acids can thus form the oxazolinone or the oxazolonium ion, which reacts with the thiocyanate ion to form the peptidylisothiocyanate, followed by cyclization to the peptidylthiohydantoin (Fig. 1, route I).

Results

An improved chemical approach for the formation of thiohydantoins

Kubo et al. (1971) described the formation of proline thiohydantoin for Z-Gly-Pro and Z-Gly-Pro-Leu-Pro, and obtained a yield of 81% for the derivatization of *N*-acetylproline using the following chemistry. At 30 °C, the substrates were reacted with acetyl chloride and trifluoroacetic acid for 15 min (activation), then a 3% solution of thiocyanic acid in dioxane was added for the derivatization (60 min). These conditions were also sufficient to convert many other *N*-acetylamino acids quantitatively. In our approach, the activation is also performed with acetyl chloride in the presence of trifluoroacetic acid but followed by a derivatization with ammonium thiocyanate. The use of acetyl chloride is preferred to acetic anhydride because the latter is known to cause problems with the sequence analysis of certain amino acids (Bailey & Shively, 1990). As it is known that thiocyanic acid tends to polymerize (Bailey, 1995), we preferred to use a 1% solution of ammonium thiocyanate in acetonitrile. Ammonium thiocyanate is commercially available as a highly purified, inexpensive, stable crystalline product. It requires, however, the addition of a small amount of strong acid (0.1% trifluoroacetic acid by volume) to facilitate the formation of peptidylthiohydantoins (Inglis et al., 1992).

To discriminate unambiguously, during mass analysis, between the formation of a thiohydantoin (+41 Da) and an acetylation on any other functional group of the peptide (+42 Da), acetyl chloride was replaced for this study by propionyl chloride because such an acylation increases the mass of the peptide with 56 Da. For an efficient formation of thiohydantoin at 30 °C, the reaction time of the activation had to be prolonged from 15 to 30 min. The results for three different peptides are shown in Figure 2. For LWMRFA (822.41 Da, Fig. 2A) the peaks at 922.4 and 978.6 Da correspond to the formation of C-terminal thiohydantoin with one and two propionylations on the peptide. The peak at 881.5 Da corresponds to the nonactivated but singly propionylated peptide, and the peak at 851.4 Da is the result from precleavage: once the alanine thio-

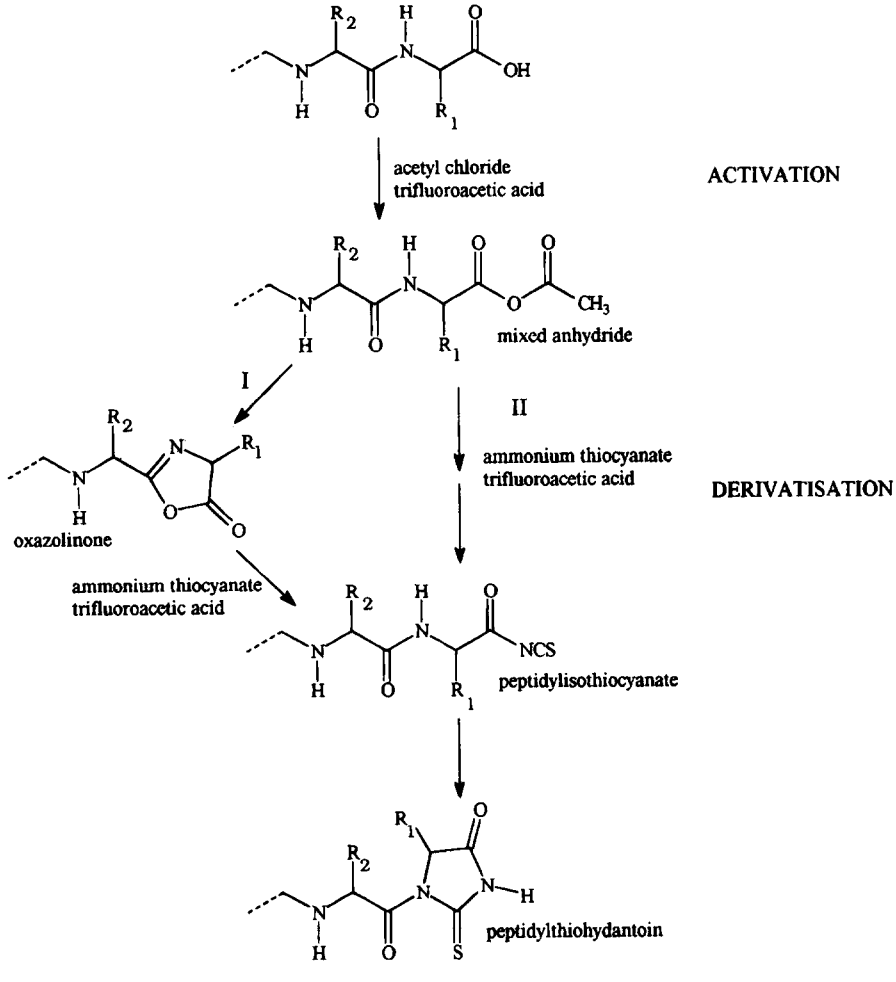

Fig. 1. The two possible reaction pathways for the formation of a peptidylthiohydantoin using the improved activation-derivatization chemistry described here.

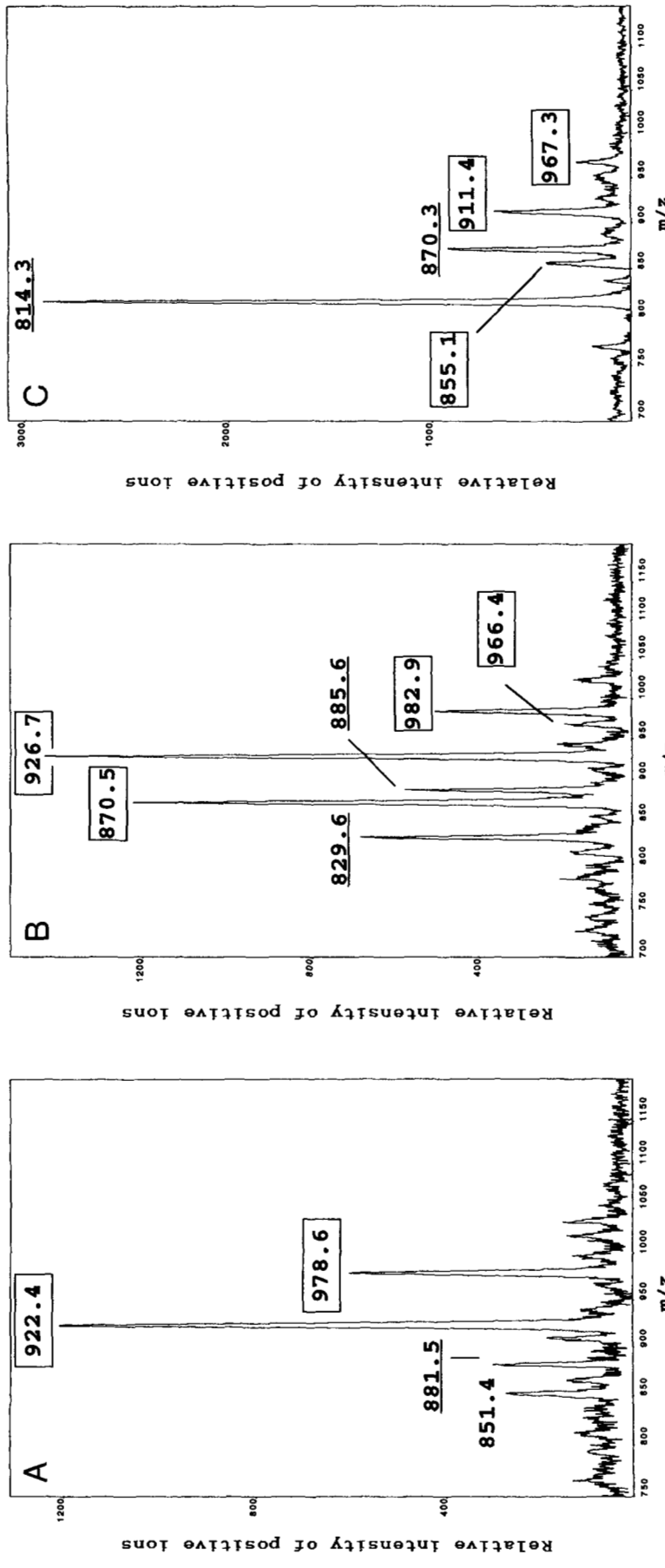

Fig. 2. PDMS mass spectra of 5 nmol of peptide LWMRFA (**A**), kemptide (**B**), and bradykinine fragment 1–7 (**C**) activated with propionyl chloride for 30 min at 30 °C and derivatized with a 1% solution of ammonium thiocyanate in acetonitrile (0.1% TFA) for 60 min at 30 °C. The mass values of the peaks that correspond to the thiohydantoin derivatives are boxed. The multiple peaks for the same thiohydantoin are due to different degrees of acylation. The mass values of the propionylated but nonactivated peptides are underlined. This way of indicating peptide derivatives is also used in the Figures 3, 5, and 6.

hydantoin is formed, it is cleaved, and thiohydantoin is formed for the penultimate phenylalanine residue (plus one propionylation on the peptide). For kemptide (LRRASLG, 771.9 Da, Fig. 2B) the peaks at 870.5, 926.7, and 982.9 Da correspond to the formation of C-terminal thiohydantoin and one to three propionylations on the peptide. The peak at 966.4 Da is caused by oxidation of the latter thiohydantoin, the sulfur atom being replaced by an oxygen atom (mass decrease of 16 Da). The peaks at 829.6 and 885.6 Da correspond to the mono- and di-propionylated peptide, respectively. The major drawback of a reaction temperature of 30 °C is the poor yield observed for thiohydantoin formation of a C-terminal proline (bradykinine fragment 1–7, RPPGFSP, 756.9 Da, Fig. 2C): the peaks at 814.3 and 870.3 Da represent one or two propionylations on the peptide. The peaks at 855.1, 911.4, and 967.3 Da correspond to the intended proline thiohydantoin and one, two, or three propionylations on the peptide. The improvement in the yield of C-terminal proline thiohydantoin formation is extensively discussed below.

Optimizing the formation of C-terminal proline thiohydantoin

For efficient formation of proline thiohydantoin, the reaction with bradykinine fragment 1–7 was performed at 55 °C. The result clearly shows that proline thiohydantoin is now readily formed (Fig. 3A). As in Figure 2C, the now major peaks at 910.5 and 966.6 Da correspond to the formation of C-terminal proline thiohydantoin together with two or three propionylations of the peptide, whereas the peaks at 814 and 870.9 Da are minimized. To exclude the possibility that the mass increase between the peaks of 870.9 and 910.5 Da was due to acetylation of the hydroxyl function of the penultimate C-terminal serine residue by acetonitrile under acidic conditions, followed by hydrolysis during sample preparation for mass spectrometric analysis, the experiment was repeated in propionitrile. Results were identical: again the desired mass increase of 42 Da was obtained. Assuming that all peptide derivatives ionize and are desorbed from the nitrocellulose target to the same extent, we estimate the yield of the thiohydantoin formation to be over 60%. Similar results, but with a lower yield of 30%, were obtained for the peptides pEKWAP and dynorphin A fragment 1–10 (YGGFLRRIRP).

The formation of thiohydantoin for the other amino acids is not compromised by this elevated reaction temperature, as demonstrated for Sar^1-angiotensine II (Sar-RVYIHPF, 1,002.2 Da, Fig. 3B). The peak at 1,060.1 Da represents a simple propionylation. The peaks at 1,101.2, 1,157.2, and 1,213.4 Da correspond to the C-terminal thiohydantoin, and one to three propionylations. The peaks at 1,197 and 1,253 Da are oxidized thiohydantoins (reactions were not performed under an inert atmosphere) with three or four propionylations.

Generality of application

To determine the yields of the thiohydantoins formed and to investigate the likely interference of the reactive side chains during the activation step, our approach was applied to *N*-acetylamino acids on a preparative scale (see Materials and methods). Reaction

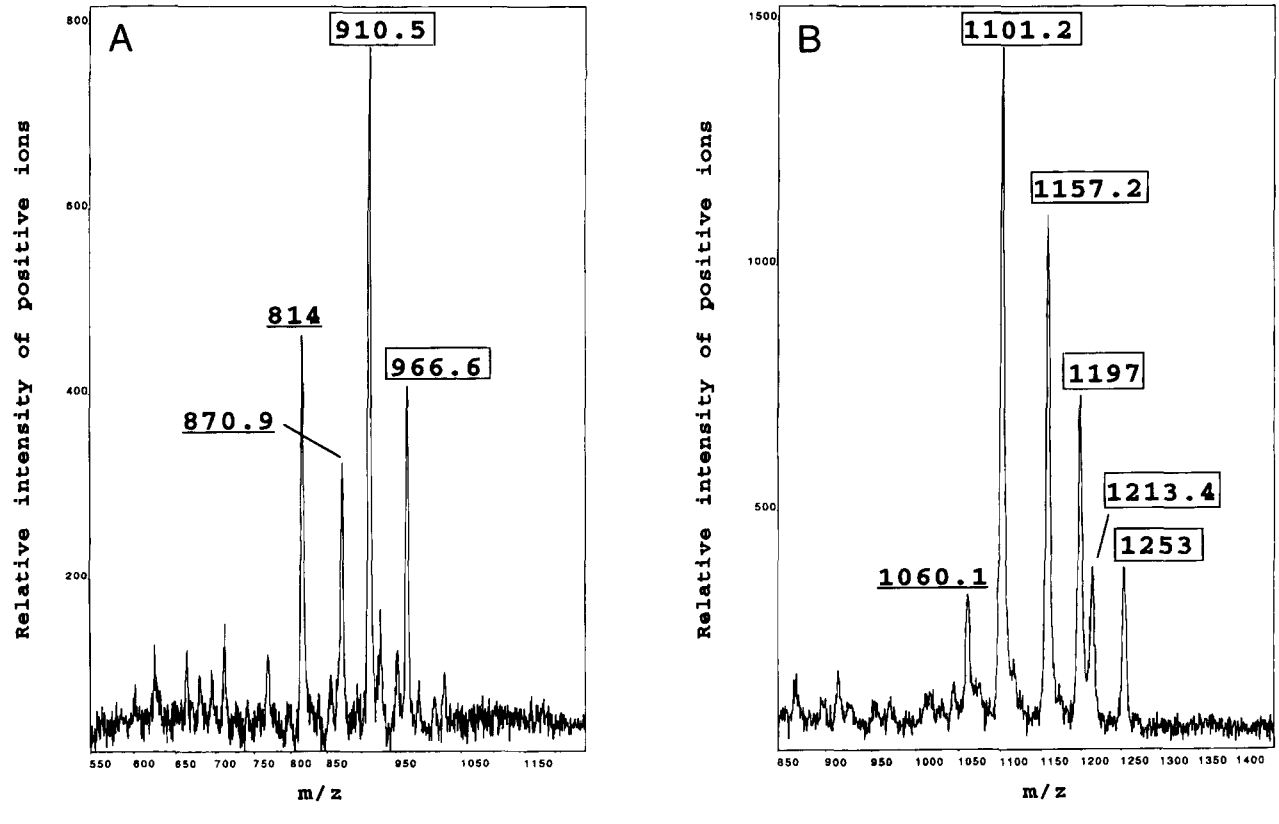

Fig. 3. PDMS mass spectra of 5 nmol of bradykinine fragment 1–7 (**A**) and Sar^1-angiotensine II (**B**) activated with propionyl chloride for 5 min at 55 °C and derivatized with a 1% solution of ammonium thiocyanate in acetonitrile (0.1% TFA) for 15 min at 55 °C.

products were purified, characterized by mass spectrometry and ^{1}H-NMR (results not shown), and the yield determined by weight. *N*-acetyl-L-alanine represented the "inert" amino acids. It was easily converted with 100% yield. Other "inert" amino acids such as Gly, Val, Leu, Ile, Phe, Asn, and Gln, and the "noninert" residues Met, Tyr, Cys, Lys, and Arg were also derivatized with 100% yield. The phenolic hydroxyl function of Tyr and the thiol function of Cys were both fully acetylated, whereas 20% of the Lys ϵ-amino function was trifluoroacetylated. The latter is due to the fact that trifluoroacetic acid and acetyl chloride form a mixed anhydride, which is attacked by the Lys ϵ-amino function. Eighteen percent of the Arg guanidine side chain was acetylated. For His, the *N*-acetylthiohydantoin was formed in 85% yield, the nitrogen atom of the imidazole side chain being acetylated. Most of the *N*-acetyl-Trp degraded during the activation. The *N*-acetylthiohydantoin was generated with only 30% yield, acetylated on the indol nitrogen.

For the first time, 100% yield for the derivatization of *N*-acetylproline to proline thiohydantoin was achieved. Proline thiohydantoin was characterized as follows: melting point (151 °C), UV-VIS spectrophotometry (λ_{max} = 233 nm, λ_{min} = 247 nm, and λ_{max} = 271 nm), mass spectrometry (m/z = 156, 128, 101, 85, 69, 59, and 41), ^{1}H-NMR (200 MHz, $CDCl_3$: δ 1.8 (m, 1H), 2.2 (m, 3H), 3.5 (ddd, J = 3.1, 8.6, and 13.2, 1H), 3.9 (ddd, app. dt, J = 8.2, 8.2, and 13.2, 1H), 4.3 (dd, J = 7.1 and 10.6, 1H), and 8.2 (bs, 1H)) and ^{13}C-NMR (200 MHz, $CDCl_3$: δ 26.7, 27.2, 47.9, 66.9, and 173.8). These spectroscopic data are identical to those of Bailey et al. (1995), except that the ^{13}C-NMR data are given for the first time (Fig. 4).

The *N*-acetyl derivatives of Ser, Thr, Asp, and Glu could not be investigated in this manner because of the instability of the reaction products on the silica gel used for thin-layer chromatography and column chromatography. Peptides containing these amino acids as C-terminal residues were, therefore, reacted under the same conditions, and the results were again verified by plasma desorption mass spectrometry. Figure 5A shows the mass spectrum for bradykinine fragment 1–6 (RPPGFS, 659.7 Da). The peaks at 813.6 and 869.4 Da correspond to the formation of serine thiohydantoin and two or three propionylations on the peptide. The peaks at 716.6 and 773.5 Da correspond to one and two propionylations. The peaks at 624.6 and 680.5 Da are not due to a thiohydantoin compound either. They represent a serine oxazolinone with a dehydrated side chain and zero to one propionylation on the peptide. The formation of an oxazolinone and the dehydratation of the serine hydroxyl function both decrease the mass of the peptide with 18 Da. A similar pattern was found for dynorphin B, containing threonine as the C-terminal amino acid (result not shown). The lower yield for serine and threonine can partially be explained by the oxazolinone with the dehydrated side chain. The α,β-unsaturated ketone formed possibly interferes with the attack of the carbonyl carbon atom by a thiocyanate ion.

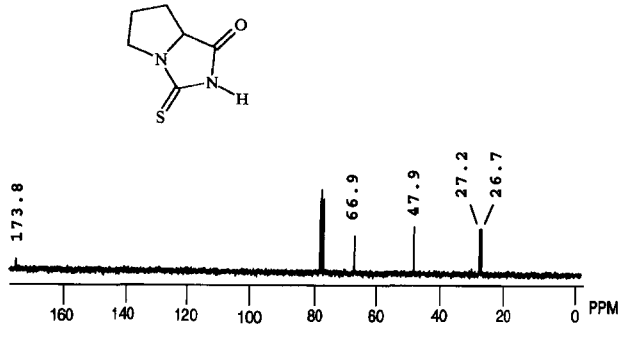

Fig. 4. ^{13}C-NMR data of proline thiohydantoin. The small peak at 173.8 PPM belongs to the carbonyl carbon atom, the peak at 66.9 PPM is caused by the α-carbon atom and the peak at 47.9 PPM by the side-chain carbon atom adjacent to the proline nitrogen. The two remaining side-chain carbon atoms are at 27.2 and 26.7 PPM. The thiocarbonyl carbon atom is not seen, and the solvent, deuterated chloroform, causes the peaks at approximately 77 PPM.

Figure 5B shows the result of an assay of thiohydantoin formation on a peptide with a C-terminal glutamic acid. The α-casein fragment 90–96 (RYLGYLE, 913.0 Da) yielded peaks at 1,024, 1,081.6, and 1,137.4 Da, which correspond to two to four propionylations without any thiohydantoin formation. The peaks at 1,008.3, 1,065.1, 1,119.9, and 1,177.5 Da represent the formation of a cyclic anhydride at the C-terminus, and two to five propionylations on the peptide. Because cyclic anhydrides are well documented not to be reactive toward ammonium thiocyanate (Stark, 1968), the mass decrease of 18 Da is thus due to cyclic anhydride and not to oxazolinone formation. Some additional peaks are due to the fact that prior to mass spectrometric analysis, the dried peptide sample was redissolved in 50% methanol. The cyclic anhydride can be opened, both by methanol and water. In the former case, a methyl ester is formed. The peak at 1,040.1 Da can be explained by such an esterification (a mass increase of 14 Da) and two propionylations on the peptide. The exact location of the esterification is unknown. Finally, the peaks at 1,104.3 and 1,160.2 Da correspond to an oxidized thiohydantoin and three to four propionylations. Similar results were obtained for tumor necrosis factor α-fragment 31–45 containing aspartic acid as C-terminal residue (result not shown).

Mechanism of thiohydantoin formation for C-terminal proline

The mechanism of peptidylthiohydantoin formation using the thiocyanate chemistry is still not very well understood. Figure 6A proves the formation of proline oxazolonium ion at the C-terminal end of bradykinine fragment 1–7 (RPPGFSP, 756.9 Da) under acidic conditions and with propionylchloride as activating agent. The reaction time was 15 min, at 55 °C. The peaks at 907.1 and 962.9 Da represent the peptide with a C-terminal proline oxazolonium ion (−17 Da) and three to four propionylations (resulting from reaction at the amino and the guanidino group of the N-terminal arginine residue, the hydroxyl group of serine, and the proline oxazolonium ion; see Fig. 7A). Considering the accuracy of 0.1% for molecular mass information obtained by plasma desorption mass spectrometry (Roepstorff et al., 1988), the decrease in mass of 17 Da could also be explained by dehydratation of serine rather than by the formation of the oxazolonium ion. However, such dehydratation makes a fourfold propionylation of the peptide impossible.

Figure 6B shows the formation of proline oxazolonium ion at the C-terminal end of bradykinine fragment under basic conditions and with diphenylchlorophosphate (Boyd et al., 1995) as the activating reagent. The reaction time is 30 min at room temperature. The major peak at 1,204.2 can be explained by the formation of an oxazolonium ion (−17 Da) and by two phosphorylations (+232 Da), one at the hydroxyl group of serine, and a second at proline oxazolonium ion. A dehydratation of serine would not allow a twofold phosphorylation of the peptide.

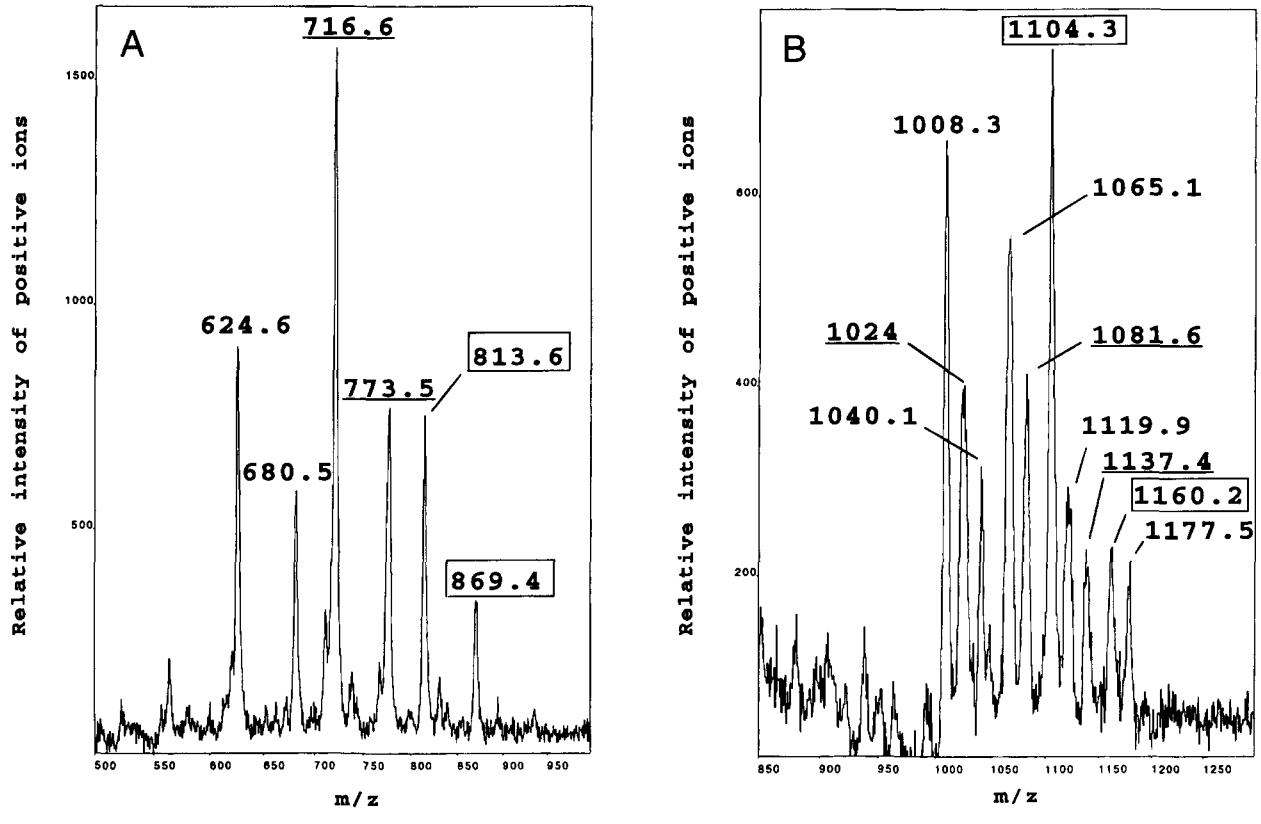

Fig. 5. PDMS mass spectra of 5 nmol of bradykinine fragment 1–6 (**A**) and α-casein fragment 90–96 (**B**) activated with propionyl chloride for 5 min at 55 °C and derivatized with a 1% solution of ammonium thiocyanate in acetonitrile (0.1% TFA) for 15 min at 55 °C.

As shown in Figure 7A, an excess of acylchloride at prolonged reaction time will result in the formation of an O-acylated oxazolonium ion. Because it takes hours at room temperature for an O-derivatized oxazolinone to react with a thiocyanate ion and to finally form thiohydantoin (Boyd et al., 1995), it is clear that for a quick and efficient formation of proline thiohydantoin the reaction time of the activating step has to be limited to avoid further reaction of the proline oxazolonium ion with the activating reagent. For instance, under acidic conditions, at 55 °C and with a reaction time of 5 min for the activating step, no fourth propionylation of bradykinine fragment 1–7 is noticed (data not shown). The peptide is then only propionylated up to three times (the N-terminus, the guanidino group of arginine, and the hydroxyl group of serine), proline oxazolonium ion is formed and, on addition of thiocyanate ions, proline thiohydantoin is readily formed as shown in Figures 3A and 7A.

Discussion

Benefits of the improved chemical approach

Progress in developing a chemical C-terminal sequencing method based on the Schlack and Kumpf (1926) thiocyanate method has been hampered by problems in forming derivatives of the hydroxy amino acids and the acidic amino acids. Perhaps the most serious shortcoming common to all C-terminal sequencing methods has been the inability to sequence through proline.

Inglis et al. (1992), explaining a yield of 70% for proline thiohydantoin formation from *N*-acetylproline under acidic conditions and in the presence of acetic anhydride, suggested that there are no mechanistic impediments to the formation of proline thiohydantoin. They proposed that the reaction may proceed via an acid-catalyzed activation of the carboxyl group, rather than through an oxazolinone or a mixed anhydride, making the carboxyl group more susceptible to nucleophilic attack by the thiocyanate ion, in a manner analogous to the acid-catalyzed formation of carboxylic acid esters. Boyd et al. also used ammonium thiocyanate in the derivatization step but activated the C-terminal carboxyl group with 1,1,3,3-tetramethylchlorouronium chloride under basic conditions (10% DIEA/ACN) (Boyd et al., 1995). The thiohydantoin then becomes S-alkylated, transforming it into a good leaving group. During the subsequent cleavage (with the thiocyanate ion), the truncated protein forms a new thiohydantoin, eliminating the need for reactivation of the C-terminal carboxyl function during successive C-terminal sequencing (Boyd et al., 1992). Application of this chemistry to several proteins showed that the repetitive yield of the sequence analysis drops notably when encountering an Asp, Glu, or His residue, and that the sequence analysis stops at the C-terminal residue preceding Ser, Thr, and Pro. An alternative approach, which combines the activation and derivatization steps, involved the use of diphenylphosphoro-isothiocyanatidate and pyridine (Bailey et al., 1992). The authors postulated that the oxazolinone was not an intermediate in the formation of peptidyl-thiohydantoin, and that this was the reason why they succeeded to

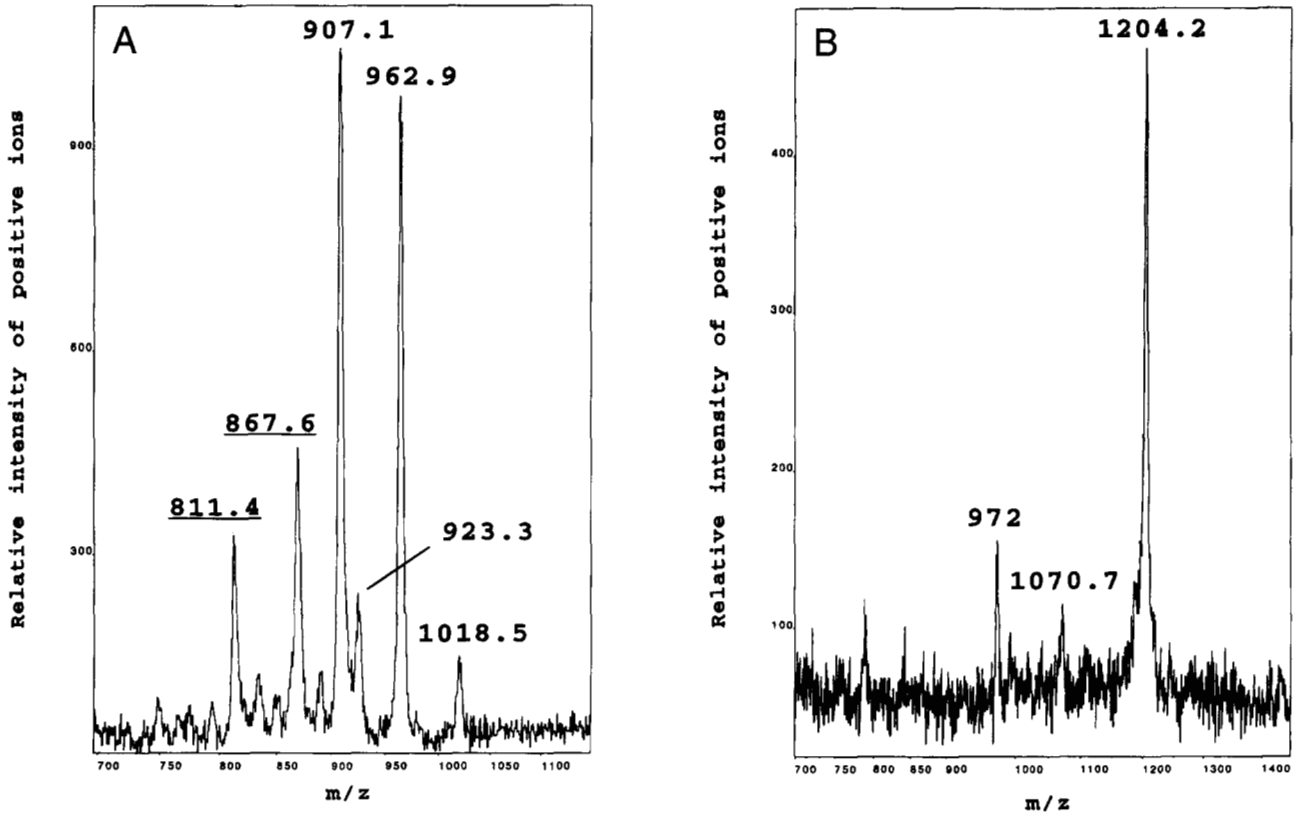

Fig. 6. A: Formation of proline oxazolonium ion at the C-terminal end of bradykinine fragment 1–7 (5 nmol) under acidic conditions and with propionylchloride as activating reagent. The reaction time was 15 min at 55 °C. **B:** Formation of proline oxazolonium ion at the C-terminal end of bradykinine fragment 1–7 (5 nmol) under basic conditions and with diphenylchlorophosphate as activating reagent. The reaction time was 30 min at room temperature.

sequence through amino acids such as Pro (Bailey et al., 1995). Sequencing results for the covalently bound tripeptides AFP and LAP as well as for ovalbumin (C-terminal sequence: Val Ser Pro) and polyproline showed, however, that the Pro thiohydantoin is formed with a yield of less than 1% (Bailey & Shively, 1994; Bailey, 1995; Bailey et al., 1995). Recently, the derivatization reagent acetyl isothiocyanate was proposed, which, in combination with triethylamine as a catalyst, can also be used for C-terminal sequence analysis (Mo et al., 1997b). These authors likewise postulated that the thiohydantoin formation does not involve an oxazolinone intermediate. They demonstrated a manual C-terminal sequence analysis of six amino acids on 7.2 nmol of a covalently attached peptide.

The improved chemical procedure for the C-terminal sequence analysis of proteins described here allows complete derivatization of the following *N*-acetylamino acids to their thiohydantoins: Gly, Ala, Val, Leu, Ile, Met, Phe, Tyr, Lys, Arg, Pro, Cys, Asn, and Gln. Acetylation of the side chain was observed to be complete for Tyr and Cys, and partially for Arg. The thiohydantoins of *N*-acetyl-His and *N*-acetyl-Trp were recovered with a yield of 85 and 30%, respectively. The side chains of both derivatives were fully acetylated. Twenty percent of the Lys ϵ-amino function was observed to be trifluoroacetylated. For the first time, a 100% yield for the derivatization of *N*-acetylproline to proline thiohydantoin was achieved. Because it is well known that acylated amino acids react more readily than the same amino acid at the C-terminal end of a protein (Inglis, 1991), the chemistry was also applied to peptides. We observed lower yields in thiohydantoin formation for peptides with one of the following C-terminal amino acids: Ser, Thr, Glu, Asp, and Pro. The lower yields for thiohydantoin formation on a C-terminal Ser or Thr can be partially explained by the formation of an unsaturated oxazolinone not capable of reaction with ammonium thiocyanate. This was also observed during the activation of a C-terminal Thr with acetic anhydride (Johnson & Scott, 1913; Bailey & Shively, 1990). The mass decrease of 18 Da observed when activating a C-terminal Asp or Glu residue is due to the formation of a cyclic anhydride, which is well documented to be unreactive toward ammonium thiocyanate (Stark, 1968). The formation of C-terminal cyclic anhydrides for Asp and Glu with acetic anhydride as activating reagent has been observed before (Nicolet, 1930; Bailey & Shively, 1990).

Thiohydantoin formation goes via the oxazolinon or the oxazolonium ion

Although the derivatization of amino acids to their corresponding thiohydantoins has been studied since the reaction was first reported (Johnson & Nicolet, 1911), the mechanism of peptidylthiohydantoin formation by an activating reagent (i.e., acetic anhydride or acetylchloride) and a thiocyanate ion is still not well understood. As the formation of oxazolinones is known to cause racemization of amino acids (Czonka & Nicolet, 1933; Carter & Stevens, 1940;

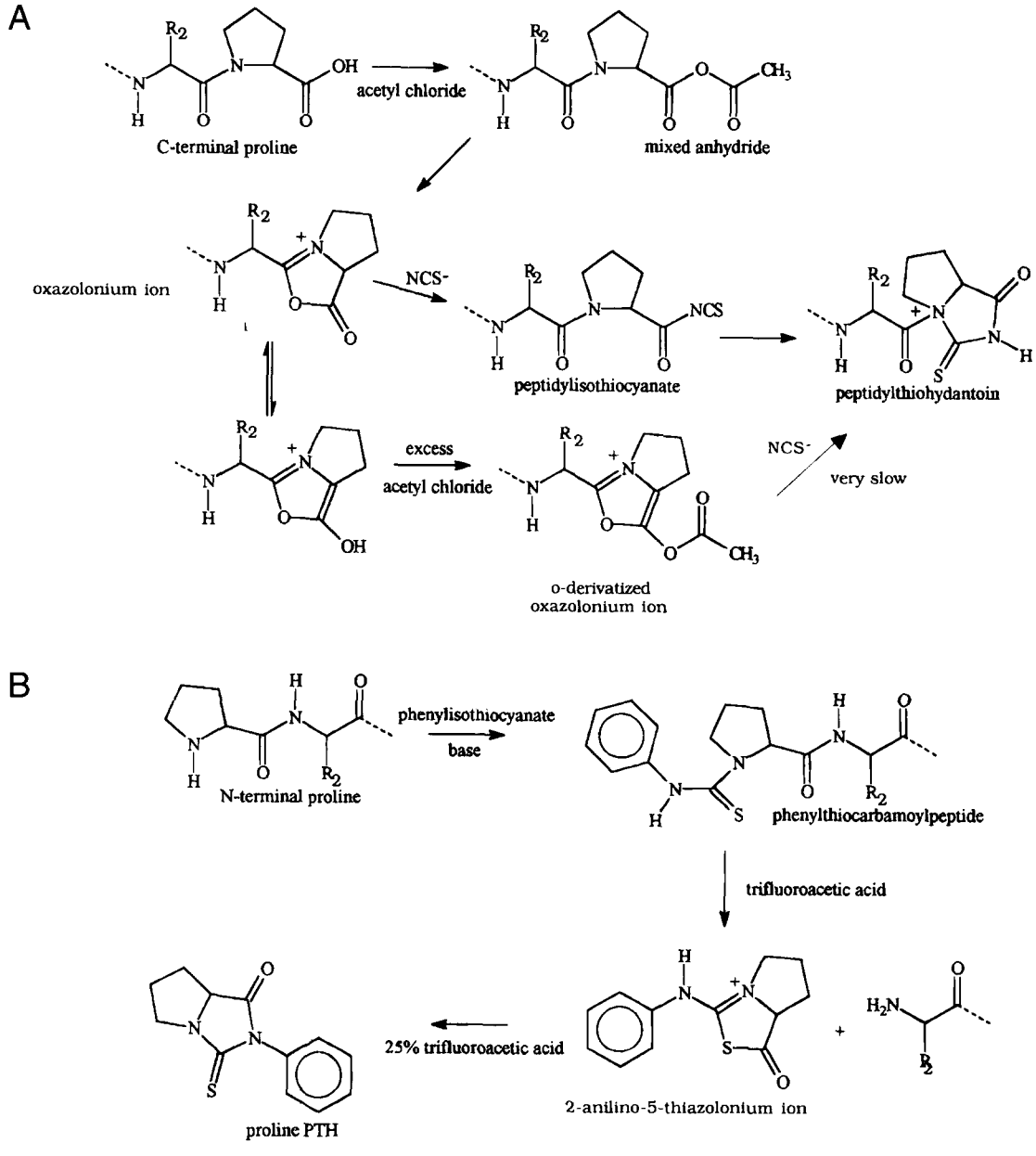

Fig. 7. A: The reaction pathway and possible side products in the formation of a C-terminal proline thiohydantoin using the activation-derivatization chemistry described here. **B:** Note the similarity between the proline oxazolonium ion as C-terminal intermediate and the proline 2-anilino-5-thiazolonium ion as intermediate in the Edman degradation.

Goodman & Levine, 1964), the racemization of the C-terminal amino acids observed on reaction of peptides with an activating reagent and a thiocyanate ion (Bailey & Shively, 1990) suggests that the peptides do form oxazolinones as intermediates. Boyd et al. (1995) monitored the reaction of acetyl-alanyl-alanyl-alanine with different activating reagents by nuclear magnetic resonance spectroscopy. An oxazolinone was observed to form immediately at the C-terminus and to react via a keto-enol tautomerization with the excess carboxy group activating reagent. Although an oxazolinone was already postulated in 1913 as a necessary intermediate during the synthesis of amino acid thiohydantoins with acetic anhydride and ammonium thiocyanate (Johnson & Scott, 1913), the existence of an oxazolonium ion was, until now, only described for *N*-acetylsarcosine (Cornforth & Elliott, 1950) and *N*-benzoylsarcosine (O'Brien & Niemann, 1957) (sarcosine = *N*-methylglycine).

Although several reports have described the derivatization of proline, as the free amino acid or as the C-terminal amino acid of a peptide, to a thiohydantoin (Kubo et al., 1971; Yamashita & Ishikawa, 1971; Inglis et al., 1992; Mo et al., 1997a), others have been unable to obtain evidence for the formation of proline thiohydantoin. Different explanations for the unreactivity of proline have been proposed. Some report that the secondary amino group does not allow the formation of the required oxazolinone (Matsuo

et al., 1966; Holcomb et al., 1968). Stark pointed out that cyclization of proline thiohydantoin would require quaternization of the nitrogen atom involved in the peptide bond (Stark, 1968; Boyd et al., 1992). Bailey and Shively (1990) proposed that the lack of reactivity might be due to the insufficient nucleophilic character of the amide nitrogen of proline to attack the linear isothiocyanate or that the side chain of proline causes steric hindrance preventing cyclization of the five-membered thiohydantoin ring. As mentioned above, Inglis et al. (1992) postulated that there is no formation of an oxazolon whatsoever.

We have now proven the existence of a proline oxazolonium ion under the acidic activation conditions described here, as well as under basic conditions. The possible reactions involved in the formation of a C-terminal proline thiohydantoin are shown in Figure 7A. Note the similarity between the proline oxazolonium ion as C-terminal intermediate and the proline 2-anilino-5-thiazolonium as intermediate in the Edman degradation (Fig. 7B). All amino acids can form the oxazolinone or the oxazolonium ion, which reacts with the thiocyanate ion to the peptidylisothiocyanate, followed by cyclization to the peptidylthiohydantoin (Fig. 1, pathway I).

We are currently optimizing this chemical approach on a model 477A sequencer in an attempt to finally develop a routine automated method for the C-terminal sequence analysis of proteins capable of identifying the 20 naturally occurring amino acids.

Materials and methods

Reagents and materials

All peptides and *N*-acetylamino acids were purchased from Sigma (St. Louis, Missouri). Acetyl chloride, propionyl chloride, acetonitrile, ethyl acetate, hexane, deuterated chloroform, and diphenyl chlorophosphate were from Aldrich (Milwaukee, Wisconsin). Acetyl chloride and propionyl chloride were redistilled from quinoline prior to use; acetonitrile and propionitrile were redistilled from phosphorus pentoxide. Ammonium thiocyanate was from Acros (Geel, Belgium). Trifluoroacetic acid and *N,N*-diisopropylethylamine was from Perkin–Elmer (Applied Biosystems Division, Foster City, California), methanol from Biosolve Ltd (Valkenswaard, The Netherlands). We used Milli-Q water (Millipore, Bedford, Massachusetts) and nitrocellulose sheets from Biorad (Trans-Blot transfer medium 0.45 micron; Hercules, California). Column chromatography was performed using silica gel (Merck 230-400 mesh; Darmstadt, Germany). Thin-layer chromatography was run on glass plates precoated with silica gel (Merck pre-coated TLC plates SIL G-25 UV_{245} 0.25 mm; Darmstadt, Germany).

Instruments

Mass spectra of the *N*-acetylthiohydantoins were recorded at 70 eV on a Finnigan 4000. Nuclear magnetic resonance spectra were recorded in deuterated chloroform at 200 MHz on a Varian Gemini 200 with tetramethyl silane (TMS) as internal standard. We used the Varian Cary 3E spectrophotometer to record UV-VIS spectra. The melting point for proline thiohydantoin given is uncorrected. Molecular masses of the derivatized peptides were obtained by plasma desorption mass spectrometry on a Bio-Ion 20 K time of flight instrument (Bio-Ion, Uppsala, Sweden). Samples were prepared by adsorption on a nitrocellulose layer in methanol:water (1:1, by vol) for 15 min. The applied acceleration voltage was 15 kV for 1×10^6 counts (stop detector).

Derivatization of N-acetylamino acids to N-acetylthiohydantoins

By way of example, *N*-acetyl-L-alanine (1.5 mmol, 200 mg) was weighed in a flask. Trifluoroacetic acid (600 μL) and acetyl chloride (550 μL) were added and the mixture was stirred at 55 °C for 5 min under an inert atmosphere of nitrogen. Ammonium thiocyanate (233 mg) was added as a 1% solution in acetonitrile (acidified with trifluoroacetic acid to 0.1% by volume) followed by stirring of the mixture at 55 °C for 15 min under an inert atmosphere of nitrogen. The solvent was removed by rotatory evaporation yielding a crude yellowish powder. Ethyl acetate was added, and the white (ammonium chloride) precipitate was filtered off. Thin-layer chromatography, using ethyl acetate:hexane (3:7) as eluent, of the ethyl acetate fraction revealed only one product, which was purified by column chromatography with the same eluent ($Rf = 0.11$). The yield was determined by weight. All *N*-acetylamino acids were treated in the same way.

Formation of peptidylthiohydantoin

Trifluoroacetic acid (2 μL) and propionyl chloride (25 μL) were added to 5 nmol peptide in an Eppendorf. The mixture was incubated at 55 °C for 5 min or at 30 °C for 30 min. Twenty-five microliters of a 1% solution of ammonium thiocyanate in acetonitrile (acidified with trifluoroacetic acid to 0.1% by vol) was added, and the mixture was incubated at 55 °C for 15 min or at 30 °C for 60 min. The reaction mixture was taken to dryness in a vacuum centrifuge. The sample was redissolved in 15 μL water:methanol (1:1) for mass spectrometric analysis.

Formation of proline oxazolonium ion under basic conditions

Ten microliters of a 10% solution of *N,N*-diisopropylethylamine in acetonitrile and 10 μL of a 5% solution of diphenyl chlorophosphate in acetonitrile were added to 5 nmol peptide in an Eppendorf. The mixture was stirred at room temperature for 30 min. Twelve microliters of water were added, followed by stirring of the mixture at room temperature for 5 min. The reaction was taken to dryness in a vacuum centrifuge and the sample was redissolved in 15 μL water:methanol (1:1) for mass spectrometric analysis. After adsorption on a nitrocellulose layer the sample was washed with a little water to remove contaminating salts.

Formation of proline oxazolonium ion under acidic conditions

Trifluoroacetic acid (2 μL) and propionyl chloride (25 μL) were added to 5 nmol peptide in an Eppendorf. The mixture was incubated at 55 °C for 15 min and was then taken to dryness in a vacuum centrifuge. The sample was redissolved in 15 μL water:methanol (1:1) for mass spectrometric analysis.

Acknowledgments

J.V.B. is indebted to the "Bijzonder Onderzoeksfonds" of the University of Gent for grant 01104593. J.V.D.E. thanks the Fund for Scientific Research–Flanders (F.W.O.) and the "Ministerie voor Wetenschapsbeleid" for finan-

cial assistance. K.H. thanks the "Vlaams Instituut ter bevordering van het Wetenschappelijk Technologisch onderzoek in de Industrie" (I.W.T.) for a research grant. Terrance E. Meyer (Department of Biochemistry, University of Arizona, Tucson, Arizona) is thanked for critical comments on the manuscript.

References

Bailey JM. 1995. Chemical methods of protein sequence analysis. *J Chromatogr A 705*:47–65.

Bailey JM, Nikfarjam F, Shenoy NR, Shively JE. 1992. Automated carboxy-terminal sequence analysis of peptides and proteins using diphenyl phosporoisothiocyanatidate. *Protein Sci 1*:1622–1633.

Bailey JM, Shively JE. 1990. Carboxy-terminal sequencing: Formation and hydrolysis of C-terminal peptidylthiohydantoins. *Biochemistry 29*:3145–3156.

Bailey JM, Shively JE. 1994. A chemical method for the C-terminal sequence analysis of proteins. In: *Methods: A companion to methods in enzymology*, vol 6. Academic Press Inc. pp 334–350.

Bailey JM, Tu O, Issai G, Ha A, Shively JE. 1995. Automated carboxy-terminal sequence analysis of polypeptides containing C-terminal proline. *Anal Biochem 224*:588–596.

Bonetto V, Bergman A-C, Jörnvall H, Sillard R. 1997. C-terminal sequence analysis of peptides and proteins using carboxypeptidases and mass spectrometry after derivatization of Lys and Cys residues. *Anal Chem 69*:1315–1319.

Boyd VL, Bozzini M, Guga PJ, DeFranco RJ, Yuan P-M. 1995. Activation of the carboxy terminus of a peptide for carboxy-terminal sequencing. *J Org Chem 60*:2581–2587.

Boyd VL, Bozzini M, Zon G, Noble RL, Mattaliano RJ. 1992. Sequencing of peptides and proteins from the carboxy terminus. *Anal Biochem 206*:344–352.

Carter HE, Stevens CM. 1940. Azlactones II. Azlactone formation in glacial and in aqueous acetic acid and preparation of benzoyl-α-aminocrotonic acid azlactone II. *J Biol Chem 133*:117–128.

Cornforth JW, Elliott DF. 1950. Mechanism of the Dakin and West reaction. *Science 112*:534–535.

Czonka FA, Nicolet BH. 1933. The preparation of optically active thiohydantoins and the racemization of amino acids as their azlactones. *J Biol Chem 99*:213–216.

Goodman M, Levine L. 1964. Peptide synthesis via active esters. IV. Racemization and ring-opening reactions of optically active oxazolones. *J Am Chem Soc 86*:2918–2922.

Holcomb GN, James SA, Ward DN. 1968. A critical evaluation of the selective tritiation method of determining C-terminal amino acids and its application to luteinizing hormone. *Biochemistry 4*:1291–1296.

Inglis AS. 1991. Chemical procedures for C-terminal sequencing of peptides and proteins. *Anal Biochem 195*:183–196.

Inglis AS, Duncan MW, Adams P, Tseng A. 1992. Formation of proline thiohydantoin with ammonium thiocyanate: Progress towards a viable C-terminal amino-acid sequencing procedure. *J Biochem Biophys Methods 25*:163–171.

Johnson TB, Nicolet BH. 1911. Hydantoins: The synthesis of 2-thiohydantoin. *J Am Chem Soc 33*:1973–1978.

Johnson TB, Scott WM. 1913. Hydantoins: The action of ammonium thiocyanate on lactone-anhydrides of acyl-amino acids. *J Am Chem Soc 35*:1136–1143.

Kubo H, Nakajima T, Tamura Z. 1971. Formation of thiohydantoin derivative of proline from C-terminal of peptides. *Chem Pharm Bull 19*:210–211.

Mann M, Talbo G. 1996. Developments in matrix-assisted laser desorption/ionisation peptide mass spectrometry. *Curr Opin Biotechnol 7*:11–19.

Matsuo H, Fujimoto Y, Tatsuno T. 1966. A novel method for the determination of C-terminal amino acids in polypeptides by selective tritium labelling. *Biochem Biophys Res Commun 22*:69–74.

Mo B, Li J, Liang S. 1997a. A method for preparation of amino acid thiohydantoins from free amino acids activated by acetyl chloride for development of protein C-terminal sequencing. *Anal Biochem 249*:207–211.

Mo B, Li J, Liang S. 1997b. Chemical carboxy-terminal sequence analysis of peptides using acetyl isothiocyanate. *Anal Biochem 252*:169–176.

Morris HR, Paxton T, Panico M, McDowell R, Dell A. 1997. A novel geometry mass spectrometer, the Q-TOF, for low-femtomole/attomole-range biopolymer sequencing. *J Protein Chem 16*:469–479.

Nicolet BH. 1930. Interpretation of the dehydration of acetylglutamic acid by means of glutamylthiohydantoin derivatives. *J Am Chem Soc 52*:1192–1195.

O'Brien JL, Niemann C. 1957. The formation and reaction of certain oxazolonium ions. *J Am Chem Soc 79*:80–85.

Patterson DH, Tarr GE, Regnier FE, Martin SA. 1995. C-terminal ladder sequencing via matrix-assisted laser desorption mass spectrometry coupled with carboxypeptidase Y time-dependent and concentration dependent digestions. *Anal Chem 67*:3971–3978.

Roepstorff P, Nielsen PF, Klarskov K, Højrup P. 1988. Applications of plasma desorption mass spectrometry in peptide and protein chemistry. *Biomed Environ Mass Spectrom 16*:9–18.

Schlack P, Kumpf W. 1926. Uber eine neue Methode zur Ermittlung der Konstitution von Peptiden. *Physiol Chem 154*:125–170.

Shevchenko A, Chernushevich I, Ens W, Standing KG, Thomson B, Wilm M, Mann M. 1997. Rapid "de novo" peptide sequencing by a combination of nanoelectrospray, isotopic labelling and a quadrupole/time-of flight mass spectrometer. *Rapid Commun Mass Spectrom 11*:1015–1024.

Siuzdak G. 1994. The emergence of mass spectrometry in biochemical science. *Proc Natl Acad Sci USA 91*:11290–11297.

Stark GR. 1968. Sequential degradation of peptides from their carboxyl termini with ammoniumthiocyanate and acetic anhydride. *Biochemistry 7*:1796–1807.

Thiede B, Wittmann-Liebold B, Bienert M, Krause E. 1995. MALDI-MS for C-terminal sequence determination of peptides and proteins degraded by carboxypeptidase Y and P. *FEBS Lett 357*:65–69.

Ward CW. 1986. Carboxyl terminal sequence analysis. In: Darbre A, ed. *Practical protein chemistry, a handbook*. New York: John Wiley & Sons Ltd. pp 491–525.

Yamashita S, Ishikawa N. 1971. Sequential degradation of polypeptides from the carboxyl ands II. Application to polypeptides. *Proc Hoshi Coll Pharm 13*:136–138.