ELSEVIER

Journal of Chromatography B, 662 (1994) 261–278

JOURNAL OF CHROMATOGRAPHY B: BIOMEDICAL APPLICATIONS

Review

Deamidation of asparagine and glutamine residues in proteins and peptides: structural determinants and analytical methodology

Rainer Bischoff*, Hanno V.J. Kolbe

Transgène S.A., 11 Rue de Molsheim, 67082 Strasbourg Cédex, France

Abstract

Non-enzymatic deamidation of asparagine and glutamine residues in proteins and peptides are reviewed by first outlining the well-described reaction mechanism involving cyclic imide intermediates, followed by a discussion of structural features which influence the reaction rate. The second and major part describes analytical techniques that allow studying deamidation in proteins using recombinant human growth hormone and recombinant hirudin as examples. Finally, the significance of non-enzymatic deamidation with respect to the production of pharmaceutical proteins is discussed.

Contents

1. Introduction

Non-enzymatic deamidation of asparagine (Asn) and glutamine (Gln) residues in peptides and proteins is a well-documented phenomenon which may occur under physiological conditions [1,2]. While hydrolysis of the amide bond is accelerated at basic pH values, significant reaction rates have been observed at neutral pH especially when Asn is followed by a Gly residue [3–7]. As a general rule, Asn deamidates more easily than Gln in corresponding positions [8].

The deamidation reaction has been studied for various reasons and we review some of these aspects in the following. Major emphasis is given to the description of analytical methods that can

* Corresponding author.

0378-4347/94/$07.00
SSDI 0378-4347(94)00203-H

be employed to detect and characterize products as well as reactive intermediates of the deamidation reaction after first outlining the reaction mechanism involving cyclic imide intermediates and discussing structural features which facilitate deamidation. The examples of recombinant human growth hormone (rhGH) and recombinant hirudin, a specific thrombin inhibitor, will serve to outline analytical strategies which led to the localization of deamidation sites as well as to the characterization of succinimide intermediates. The review concludes with a discussion of the significance of deamidation with respect to production and stability of pharmaceutical proteins.

2. Reaction mechanism

A number of studies on synthetic peptides has established the pronounced influence of the amino acid C-terminal to a given Asn or Gln residue on their respective rates of deamidation [1–8]. An early review of this subject showed that half-lives of the internal Asn under physiological conditions (pH 7.4) ranged from 6 to 507 days and those of Gln in similar positions from 96 to 3409 days based on the analysis of 70 pentapeptides [1]. Subsequent work provided strong evidence that Gly when following Asn in synthetic penta- or hexapeptides facilitated deamidation as compared to amino acids with branched hydrophobic side chains such as Val, with half-lives differing by a factor of 70 between Val-Ser-*Asn-Gly*-Val ($t_{1/2} = 5.8$ h) and Val-Ser-*Asn-Val*-Val ($t_{1/2} = 420$ h) at pH 7.3 and 37°C [7]. Detailed quantitative analyses of the reaction products permitted further insight into the reaction mechanism which was shown to involve formation of a five-membered cyclic imide intermediate undergoing subsequent hydrolytic ring opening to give L-aspartic acid (Asp) and L-iso-aspartic acid (iso-Asp) in a ratio of about 1:3 [4,9] (Fig. 1). In addition, small amounts of D-aspartic acid and D-iso-aspartic acid were observed, indicating slow racemization of the cyclic imide. Further confirmation of this reaction mechanism was obtained when subjecting a synthetic peptide containing a cyclic imide to slightly basic pH and analyzing the reaction products. Formation of the corresponding iso-Asp and Asp containing isomers in a ratio of 3:1 proved to be in accordance with the proposed reaction mechanism [4].

More recently, isolation and characterization of cyclic imide intermediates from proteins undergoing deamidation at Asn–Gly [10] or isomerization at Asp–Gly sites [11–13] have confirmed that this reaction mechanism also takes place in larger molecules with more complex three-dimensional structures. The fact that incubation of recombinant human or bovine growth hormone (rhGH or rbGH) under physiological conditions leads to dehydration to the corresponding imide at a specific Asp–Gly site confirmed the reversible character of the reaction as proposed by Geiger and Clarke [4]) (Fig. 1). These authors also showed that at pH 7.4 and 37°C the rate of formation of the respective iso-Asp or Asp-containing peptides was 38-fold faster for Val-Tyr-Pro-*Asn-Gly*-Ala than for Val-Tyr-Pro-*Asp-Gly*-Ala indicating that deamidation at Asn–Gly sites proceeds more rapidly than isomerization at Asp–Gly sites. This is in agreement with the fact that an amide has better leaving group properties than a carboxylic acid.

It is to be expected that cyclic imides occur in numerous other proteins undergoing deamidation or isomerization according to the mechanism shown in Fig. 1 both in vitro during production, purification and storage as well as in vivo as part of their normal metabolism, since Asn–Gly sites are quite frequent in natural proteins [14]. Furthermore, cyclic imide formation at an Asp–Gly site has recently been demonstrated to occur in the joining peptide portion of pro-opiomelanocortin extracted from mouse or porcine pituitary glands [13]. The reaction mechanism depicted in Fig. 1 has thus been supported by a number of detailed in vitro studies but it can not be ruled out that other mechanisms may also occur in vivo.

Cyclic imide formation according to the mechanism of deamidation described above necessitates a significant change in local protein structure and it is not surprising that Gly and to a lesser

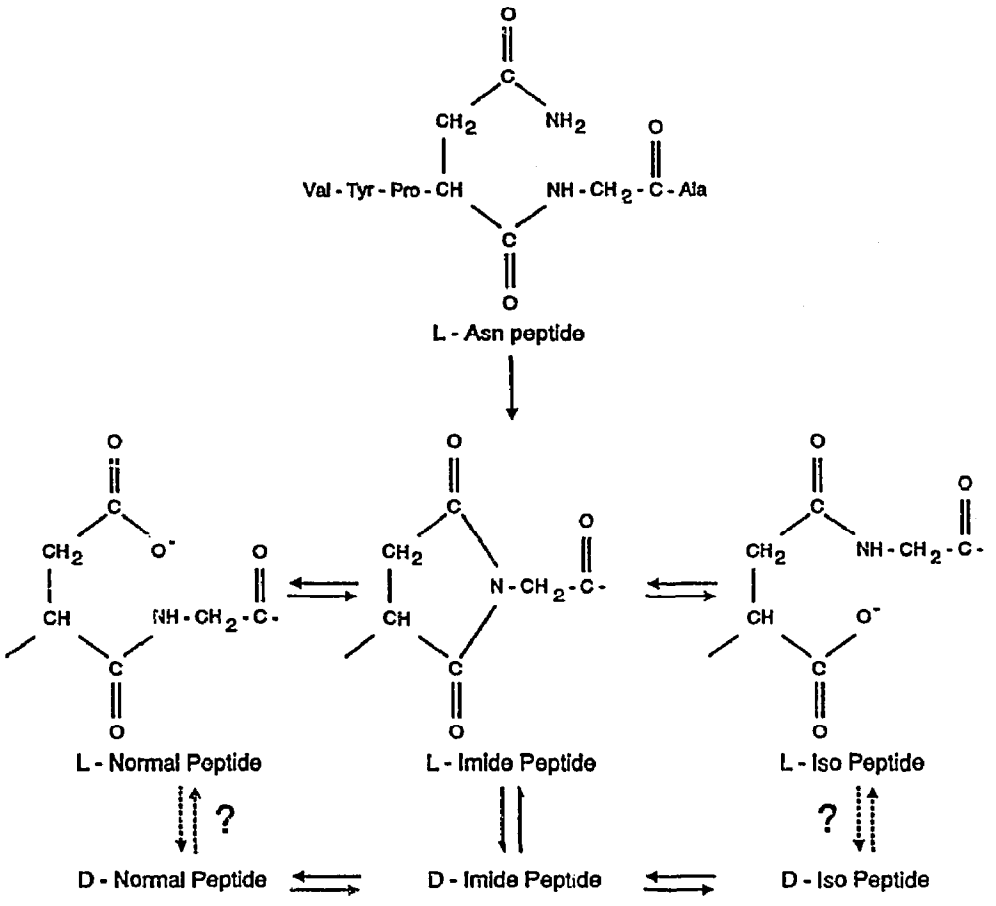

Fig. 1. Pathways for spontaneous deamidation, isomerization and racemization for aspartyl- and asparaginyl hexapeptides related to adrenocorticotropin (adapted from Ref. [4]).

extent Ser with their small side chains are more tolerant to this structural change. Also consistent with the above mechanism is that the nature of the amino acid N-terminal to Asn has only a minor influence on the rate of deamidation [7] although some controversy concerning this point remains in the literature [8]. The reaction mechanism is also consistent with the fact that Gln residues deamidate more slowly than Asn, since formation of a six-membered cyclic imide is entropically less favorable. In a few cases some cleavage of the peptide bond following an Asn residue has been observed giving rise to additional minor reaction products [4,7,12].

While the reaction mechanism of deamidation involving cyclic imides has been well established, there may exist other pathways for deamidation in cases where the surrounding amino acids do not favor cyclic imide formation. At present none of these other mechanisms has been elucidated in molecular detail although a number of hypotheses have been proposed based on known three-dimensional structures of proteins undergoing deamidation [15,16]. It may well be that deamidation in these cases is facilitated by favorable steric arrangements of functional groups providing acid–base catalysis for the hydrolysis of the amide bond and that no stable intermediates can be isolated.

3. Structural determinants

The influence of amino acid sequence on deamidation in proteins has been investigated by analyzing 1000 protein sequences from a data-

base with respect to amino acids adjacent to known deamidation sites [15]. The results obtained support most of the conclusions drawn from the peptide studies outlined above but also indicate that the influence of neighbouring groups may be more complex in proteins than in short peptides due to their three-dimensional structures. While Gly and Ser were found to be the most destabilizing C-terminal amino acids for deamidation of Asn, Ser and to a lesser extent Thr and Lys also appeared to favor deamidation of Asn when located at its N-terminal side (Fig. 2). Although effects on deamidation half-lives of the residue N-terminal to Asn and Gln were reported in an early study [8], their influence was later questioned [7] and further experimental studies will be necessary in order to elucidate this issue with respect to proteins.

A statistical evaluation of pairs and triplets of amino acids in over 1400 protein sequences [14] indicated that most of the amino acid combinations which favored deamidation occurred at lower frequencies than the expected statistical average, supporting the notion that they have been eliminated by evolutionary processes. Surprisingly, however, combinations of Asn–Gly and Glu–Asn–Gly appeared at rather high frequencies and the authors suggested that these labile sites may function as biological clocks controlling protein ageing. Studies on the A and B subunits of chicken and bovine eye lens α-crystallins have recently provided supportive evidence for such a biological role, in that deamidation of Asn-146, which occurs in the inner eye lens in the absence of protein turnover, is related to age, since αB-crystallin from a five year old bovine was partially deamidated while no deamidation was observed in calf αB-crys-

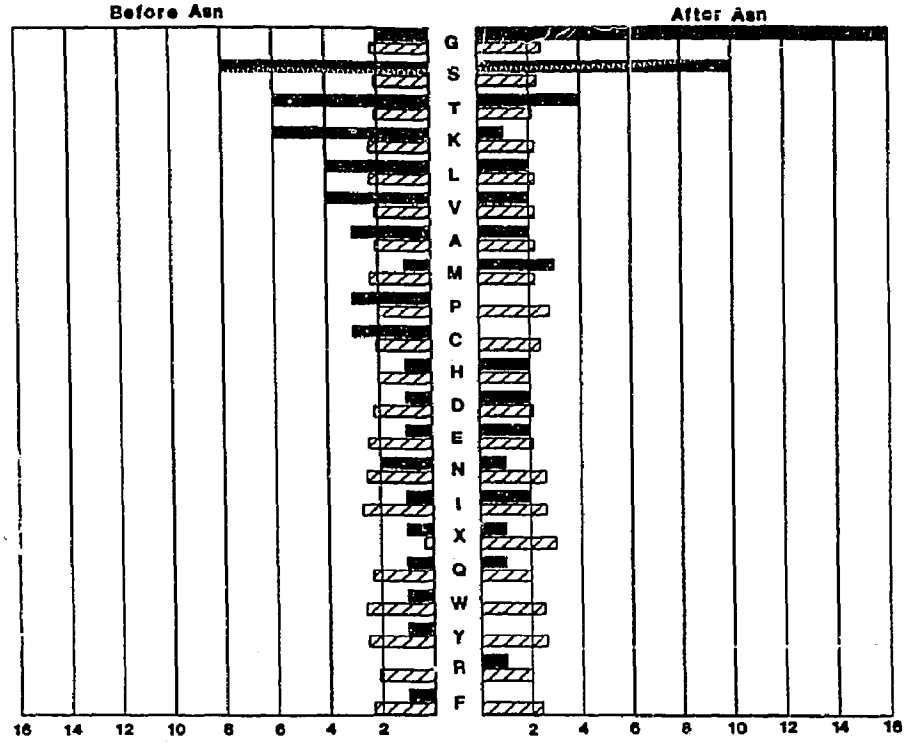

Fig. 2. Frequency with which each amino acid occurs on the amino (solid, left bars) and carboxyl (solid, right bars) sides of labile Asn residues of proteins. Hatched bars indicate expected frequency for that amino acid before or after Asn (reprinted from Ref. [15] by permission of Oxford University Press).

tallin [17]. Fig. 3A shows the amino acid sequence of the B_2 chain of bovine α-crystallin with the Asn–Gly site underlined [18].

While studies on small synthetic peptides have contributed significantly to our understanding of the deamidation reaction mechanism, it is ultimately the three-dimensional structure of a protein which determines whether deamidation will be favored or not. Structure determination by X-ray crystallography of succinimide-containing synthetic peptides has provided detailed insights into the conformational constraints that are imposed on the peptide backbone and the adjacent side chains as a result of cyclization [19]. So far, no such structural data of a succinimide intermediate in a protein have been obtained, most likely due to the instability of such intermediates. However, a number of studies on deamidation sites in proteins of known three-dimensional structures has been performed in order to derive structural models of the intermediates involved. One of the first proteins in which deamidation has been studied in detail is horse heart cytochrome c (Fig. 3B) [20–22] and Wright [15] recently proposed a reaction mechanism for deamidation of Asn-103 based on a structural model derived from the three-dimensional structure of the homologous albacore cytochrome c [23]. At least three deamidation sites were observed with Asn-103 reacting most rapidly. The C-terminal Glu-104 residue was proposed to facilitate deamidation through intramolecular assistance of its side chain carboxylic acid moiety. In addition, flexibility of the C-terminal region in the protein may contribute to the ease with which deamidation occurs, since

A

Met	Asp	Ile	Ala	Ile	His	His	Pro	Trp	Ile	Arg	Arg	Pro	Phe	Phe
Pro	Phe	His	Ser	Pro	Ser	Arg	Leu	Phe	Asp	Gln	Phe	Phe	Gly	Glu
His	Leu	Leu	Glu	Ser	Asp	Leu	Phe	Pro	Ala	Ser	Thr	Ser	Leu	Ser
Pro	Phe	Tyr	Leu	Arg	Pro	Pro	Ser	Phe	Leu	Arg	Ala	Pro	Ser	Trp
Ile	Asp	Thr	Gly	Leu	Ser	Glu	Met	Arg	Leu	Glu	Lys	Asp	Arg	Phe
Ser	Val	Asn	Leu	Asn	Val	Lys	His	Phe	Ser	Pro	Glu	Glu	Leu	Lys
Val	Lys	Val	Leu	Gly	Asp	Val	Ile	Glu	Val	His	Gly	Lys	His	Glu
Glu	Arg	Gln	Asp	Glu	His	Gly	Phe	Ile	Ser	Arg	Glu	Phe	His	Arg
Lys	Tyr	Arg	Ile	Pro	Ala	Asp	Val	Asp	Pro	Leu	Ala	Ile	Thr	Ser
Ser	Leu	Ser	Ser	Asp	Gly	Val	Leu	Thr	Val	Asn	Gly	Pro	Arg	Lys
Gln	Ala	Ser	Gly	Pro	Glu	Arg	Thr	Ile	Pro	Ile	Thr	Arg	Glu	Glu
Lys	Pro	Ala	Val	Thr	Ala	Ala	Pro	Lys	Lys					

B

Gly	Asp	Val	Glu	Lys	Gly	Lys	Lys	Ile	Phe	Val	Gln	Lys	Cys	Ala
Gln	Cys	His	Thr	Val	Glu	Lys	Gly	Gly	Lys	His	Lys	Thr	Gly	Pro
Asn	Leu	His	Gly	Leu	Phe	Gly	Arg	Lys	Thr	Gly	Gln	Ala	Pro	Gly
Phe	Thr	Tyr	Thr	Asp	Ala	Asn	Lys	Asn	Lys	Gly	Ile	Thr	Trp	Lys
Glu	Glu	Thr	Leu	Met	Glu	Tyr	Leu	Glu	Asn	Pro	Lys	Lys	Tyr	Ile
Pro	Gly	Thr	Lys	Met	Ile	Phe	Ala	Gly	Ile	Lys	Lys	Lys	Thr	Glu
Arg	Glu	Asp	Leu	Ile	Ala	Tyr	Leu	Lys	Lys	Ala	Thr	Asn	Glu	

Fig. 3. (A) Amino acid sequence of α-crystallin B_2 chain from bovine [18]. The Asn-146,Gly-147 site involved in age-related deamidation is underlined. (B) Amino acid sequence of cytochrome c from horse heart [20]. The Asn-103,Glu-104^{COOH} site which is involved in deamidation and the second deamidation site at Asn-54 are underlined. Sequences are represented with 15 amino acid residues per line.

reaction rates in the related synthetic tetrapeptide Gly–Thr–Asn–Glu (half-life at pH 7.4 and 37°C: 16 days) were found to be similar to those in the protein [15,24]. This early work on cytochrome c showed also that deamidation is an ordered process where transformation of Asn-103 into Asp favors the subsequent deamidation of Asn-54 adding another level of complexity to the situation in proteins where structural changes induced by deamidation of one residue may influence further deamidation reactions.

Recombinant hirudin is a more recent example of a protein where deamidation has been studied in detail and the three-dimensional structure of the hirudin–human α-thrombin complex has been determined by X-ray crystallography. (Fig. 4) [25–27]. Treatment of the recombinant hirudin variant rHV2-Lys47 (referred to as hirudin throughout this review) at pH 9 resulted in deamidation of Asn-53 and Asn-33 which are both followed by Gly residues (Fig. 5) [28]. Prolonged treatment under these conditions led to further deamidation at Asn-52 presumably influenced by the newly formed Asp-53. In view of the mechanistic studies described above, it was most likely that deamidation at Asn-33 and Asn-53 occurred via cyclic imide intermediates which was later confirmed experimentally [10]. Fig. 4 shows that both Asn-53 and Asn-33 as well as Asn-52 are located on the surface of the molecule in solvent-exposed loop or extended structures thus facilitating formation of the more rigid imide ring structures. Partial solution structures of free hirudin by NMR-spectroscopy [29,30] indicated also that the regions around Asn-33 and Asn-53 are highly flexible, as their mobility prevented acquisition of high resolution NMR data. These examples show that deamidation rate constants determined for short synthetic peptides may serve as guidelines to look for "hot-spots" in a protein of known sequence (e.g. hirudin) but that other mechanisms may play a

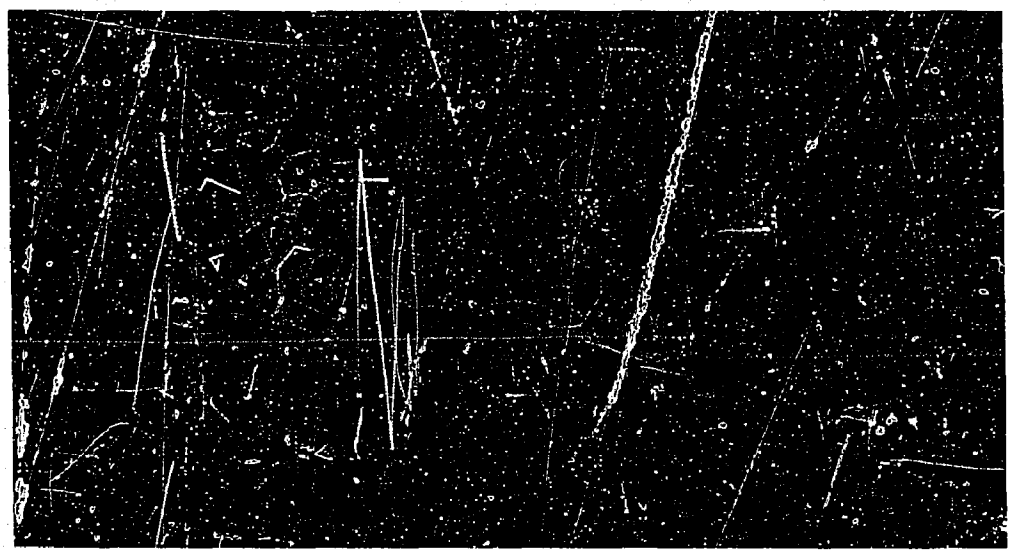

Fig. 4. Representation of the three-dimensional structure of rHV2-Lys-47 as determined by X-ray crystallography of its complex with human α-thrombin [25,26]. The human α-thrombin structure has been removed for clarity. The seven Asn residues (Asn-12,20,26,33,47,52 and 53) are highlighted in green and Asn-33, Asn-52, and Asn-53 are numbered. The highly flexible C-terminal tail (containing Asn-52 and 53) and loop around Asn-33 are shown to the right and the bottom left of the figure, respectively. The region around Asn-33 (residues 32–35) was found to be disordered upon crystallographic [25,26] and NMR [29,30] analysis. The C-terminal tail (residues 49–65) was well defined in the crystal structure of the complex (except for residues 52–54) but disordered in the solution structure of free hirudin [29,30]. Disulfide bonds are represented in white.

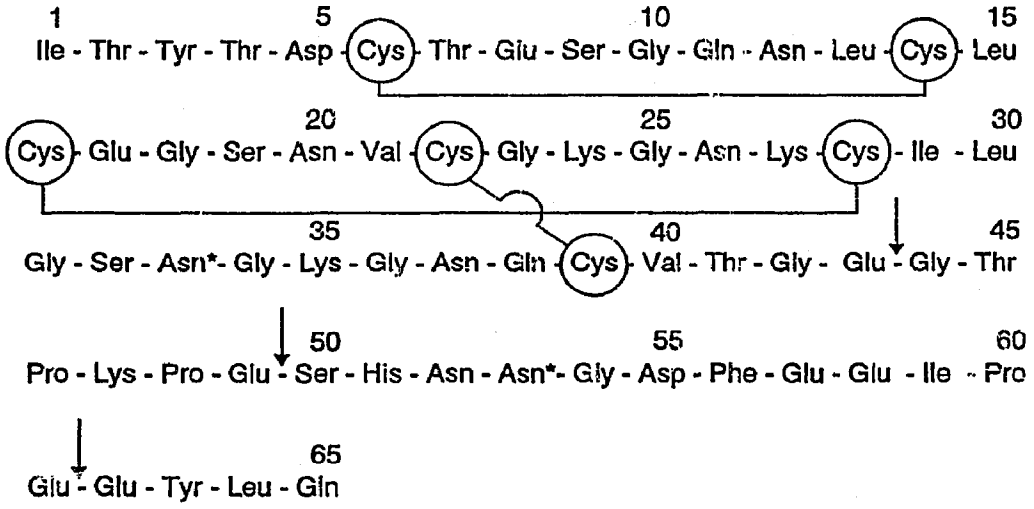

Fig. 5. Primary structure of hirudin variant rHV2-Lys-47 including disulfide bonds. The two Asn residues deamidating most rapidly (Asn-33 and Asn-53) are marked (*), and the cleavage sites for V8 protease are indicated by arrows. Cleavage was found after Glu-43, Glu-49 and Glu-61 but not following Glu-8, Glu-17, Glu-57, Glu-58 and Glu-62 (reprinted with permission from Ref. [10]. Copyright 1993, American Chemical Society).

role depending on the structural environment of a particular Asn residue (e.g. cytochrome c).

4. Analytical methods

4.1. Detection and isolation of deamidated forms and cyclic imide intermediates

Deamidation of Asn and Gln residues introduces additional negative charges into a protein. Separations of deamidated forms from the parent molecule are therefore mainly based on techniques which are sensitive to changes in surface charge such as anion-exchange high-performance liquid chromatography (anion-exchange HPLC) [11,28], cation-exchange chromatography [31], isoelectric focusing (IEF) [32], alkaline urea gel electrophoresis [17,33], capillary electrophoresis coupled on-line to electrospray mass spectrometry [34] and thin layer chromatography on silica plates [7]. Reversed-phase HPLC which is based on hydrophobic interactions between solutes and the stationary phase has also been employed as a separation technique [4,35] with selectivity adjustments being made by modifying the mobile phase from the customary acidic conditions (0.1% trifluoroacetic acid, pH ≈ 2) to, for example, a pH 6.5 mobile phase (30 m*M* sodium phosphate). All of these separation methods are complementary and are often used in combination to resolve deamidated forms of peptides or proteins.

A first indication as to whether a protein is partially deamidated or undergoing deamidation during cell culture, purification or storage can be obtained by IEF in polyacrylamide gels [36]. This technique has a high resolution with respect to a protein's isoelectric point (p*I*) which is related to its overall surface charge. Performing IEF in immobilized pH gradients has resulted in excellent resolution of protein variants differing by as little as 0.01 pH units in their p*I*s [37]. Such high resolution is generally not required when studying deamidation processes, since the resulting forms differ from each other and the original protein in their p*I*s by ca. 0.1 pH units [32] (Fig. 6; Table 1). However, interpretation of results obtained by IEF can be ambiguous, since most natural and recombinant proteins appear to be heterogenous in IEF for various reasons such as different amounts of sialic acid in glycoproteins, the presence of phosphorylated and/or sulfated amino acids, N- or C-terminal modifications, lipidation etc. It is thus not easy to discriminate whether an additional band on an

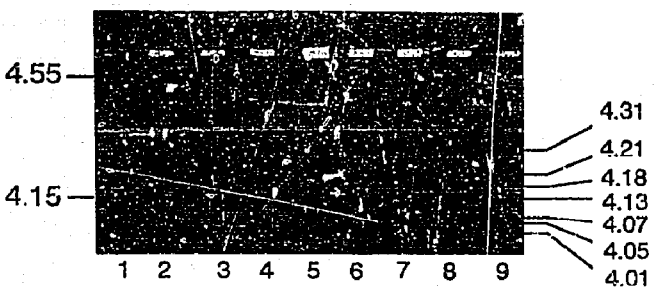

Fig. 6. Isoelectric focusing of recombinant hirudin (rHV2-Lys-47) in an immobilized pH gradient between pH 3.8–4.8 after incubation in 0.1 *M* Tris-HCl, pH 9, at 4°C for 4, 5, 6, 7, 10 and 12 days, respectively (lanes 8–3). Lane (1) p*I* standards and lanes (2) and (9) purified hirudin (reprinted with permission from Ref. [32]. VCH Verlagsgesellschaft mbH).

IEF gel is due to deamidation or other modifications. Incubation of the protein at different pH values may allow discrimination between these possibilities, since deamidation progresses faster under basic conditions but proof as to which modification was at the origin of the observed shift in p*I* is only possible once the different forms have been isolated in pure form and in sufficient amounts to perform detailed structural analyses.

Combining IEF with immunoblotting techniques has the advantage that protein heterogeneity can be analyzed early in a production process, for example by taking samples during culture of a recombinant microorganism or cell line. Controlling a purification process with this methodology is also of interest, since it may identify steps where deamidation or other protein modifications occur and thus allow the manipulation of different process parameters in order to minimize these reactions. The limitation of IEF with subsequent immunoblotting or staining of protein bands is that the method will only provide semi-quantitative results, since the linear range of detection is rather limited. More recent developments of IEF in narrow diameter capillaries promise to overcome this problem, since proteins are detected by ultraviolet light absorbance and may be further characterized by mass spectrometry [34,38–42]. Capillary IEF may thus prove to be a useful tool for quantitative deamidation studies in the future.

Table 1
Calculated isoelectric points of hirudin and its deamidated forms

Number of deamidation sites	Calculated p*I*
None (Asn-33,52,53)	4.30
One (Asp-33 or Asp-53, Asn-52)	4.17
Two (Asp-33 and Asp-53, Asn-52)	4.06
Three (Asp-33,52,53)	3.97

IEF has only found limited application as a preparative technique although the use of immobilized pH gradients allows the separation of milligram-quantities of protein. Anion-exchange HPLC is more commonly used for this purpose although it generally shows lower resolution than IEF. In work related to hirudin, Tuong et al. [28] employed anion-exchange HPLC to partially separate 11 different deamidated forms which were generated upon incubation at pH 9 (see Fig. 9), and Johnson et al. used both anion-exchange and reversed-phase HPLC to study deamidation of recombinant human growth hormone (rhGH) [35]. These examples will be discussed in more detail in section 4.2.

Isolation and characterization of cyclic imide intermediates of the deamidation reaction have been more difficult primarily because these compounds are sensitive to basic pH, and their physical characteristics are very similar to the parent molecule (e.g. no difference in charge). In the case of hirudin, formation of a cyclic imide at position Asn-53,Gly-54 changed the p*I* by as little as 0.01 pH unit, a difference that is barely detectable by IEF even in a very shallow immobilized pH gradient. In the few detailed studies that have been performed on cyclic imides in proteins, isolation was based on their slightly increased hydrophobicity upon reversed-phase HPLC [10,11,13] (Fig. 7). Identification of this unusual modification in purified proteins has only recently become possible due to the advent of mass spectrometric ionization techniques which allow precise mass measurements of high-molecular-mass compounds [43,44].

Peptide mapping by reversed-phase HPLC is a powerful tool to detect and analyze cyclic imide intermediates. The appearance of additional

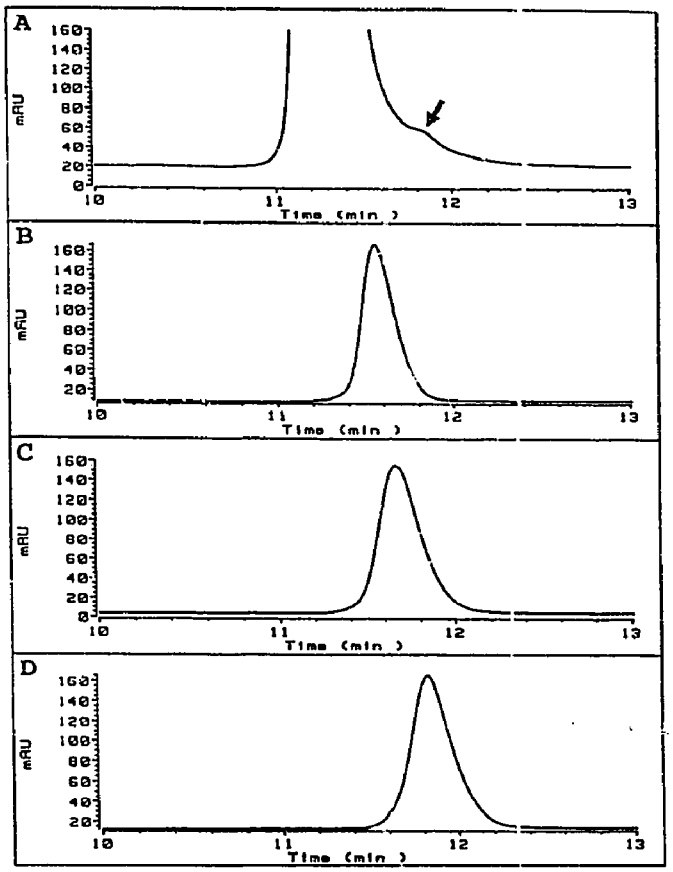

Fig. 7. Reversed-phase HPLC analysis of identical amounts of hirudin (rHV2-Lys-47) and two succinimide containing forms. (A) Starting material from which the succinimide containing forms (depicted by an arrow) were prepared. (B) Highly purified hirudin after elimination of succinimide intermediates (t_R = 11.55 min). (C) Hirudin form containing a succinimide between Asn-33 and Gly-34 (t_R = 11.66 min). (D) Hirudin form containing a succinimide between Asn-53 and Gly-54 (t_R = 11.83 min) (reprinted with permission from Ref. [10]. Copyright 1993, American Chemical Society).

peaks in a tryptic peptide map of rhGH pointed to a modification which was subsequently identified as a cyclic imide between Asp-130 Gly-131 using mass spectrometry [11]. Digestion of hirudin with endoprotease Glu-C (V8 protease) and analysis of the generated peptides by reversed-phase HPLC indicated a modification in one of the peptide fragments (residues 50–61) which was later identified as a cyclic imide between Asn-53 and Gly54 by mass spectrometric analysis (Fig. 8) [10]. It is to be expected that the more widespread use of reversed-phase HPLC coupled on-line with electrospray mass spectrometry [45,46] for peptide mapping will further contribute to the detection of modifications such as succinimides in a larger number of both natural and recombinant proteins. Novel approaches to separating proteins and peptides such as capillary zone electrophoresis will also find more widespread applications as a complement to reversed-phase HPLC to analyze protein modifications. Although capillary electrophoresis is a powerful separation technique with multiple applications in biochemical research [47], it is still quite tedious to obtain structural information from the separated compounds because of the difficulty to recover sufficient material for subsequent analyses. Recent advances in on-line coupling of capillary electrophoresis to mass spectrometry may, however, overcome this limitation [34,48–50].

4.2. Identification of deamidation sites and iso-aspartic acid linkages

The localization of deamidation sites in a protein or peptide necessitates analytical methods with resolution at the amino acid level. Two methods most often employed are amino acid sequencing by automated Edman degradation and more recently mass spectrometry with collision-activated dissociation (CAD). Furthermore, enzymatic methylation of iso-aspartic acid linkages by protein isoaspartyl methyltransferase (PIMT, EC 2.1.1.77) has provided a sensitive and specific way of localizing such unusual structures in deamidated proteins and peptides [31,35,51–53].

Hirudin and rhGH may serve as examples to outline the strategies for identifying multiple deamidation sites and cyclic imide intermediates in proteins [10,28,35]. Hirudin, when incubated at pH 9 and 37°C to accelerate deamidation, produces a number of deamidated forms that have been partially resolved by anion-exchange HPLC, resulting in eleven fractions (Fig. 9). Each of these fractions was subjected to analysis by liquid secondary-ion mass spectrometry (LSIMS) after chemical esterification of carboxylic acid groups in methanolic hydrochloric

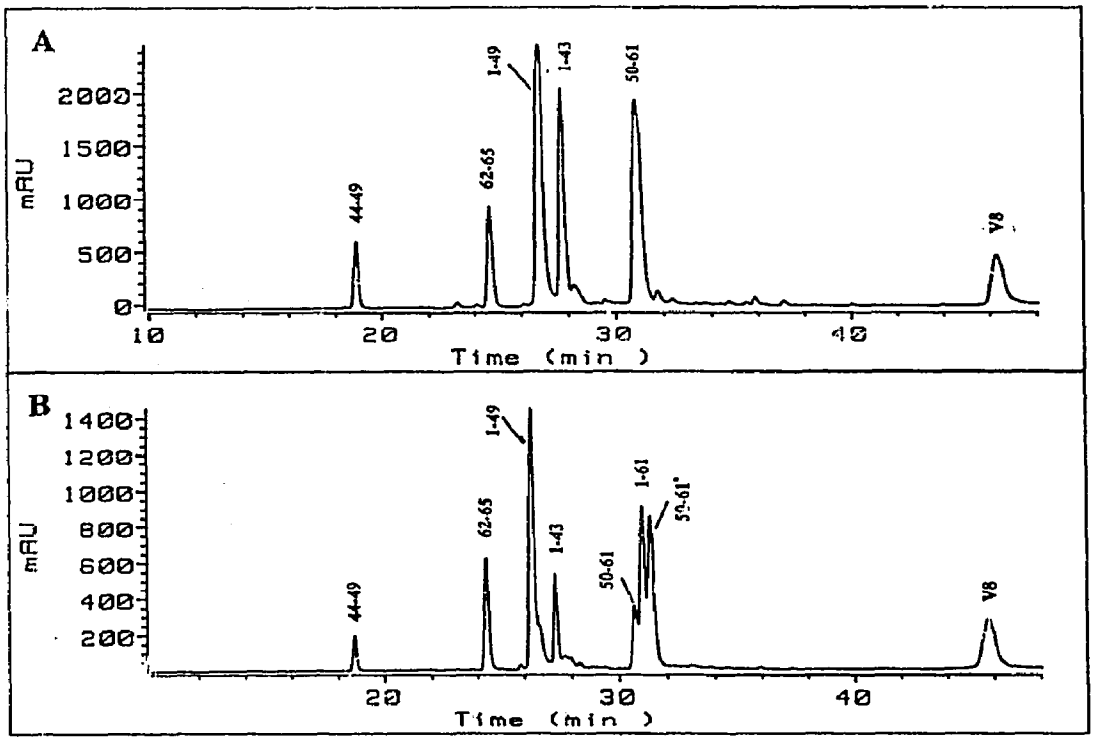

Fig. 8. Peptide mapping of (A) hirudin (rHV2-Lys-47) and (B) the form containing a succinimide between Asn-53 and Gly-54 by reversed-phase HPLC after digestion with V8 protease. Peptide fragments are indicated showing that the position of the fragment containing the cyclic imide (50–61') has been shifted to a slightly longer retention time. Cleavage at Glu-43 and Glu-49 was found to be somewhat less efficient for the succinimide-containing form giving rise to fragment 1–61.

acid. Esterification of hirudin gave a measured mass of 7061.5 Da which corresponds closely to its molecular mass plus 11 additional methylesters (calculated mass: 7061.8 Da, see Table 2) indicating that the reaction conditions did not induce artefactual deamidations or cyclic imide formations at a measurable level. This allowed determination of the number of carboxylic acid groups in each form based on the 15 additional mass units for each methylester due to the conversion of a COOH into a $COCH_3$-group showing that fractions F1–F4 contained a single additional acidic group, F5–F7 two and F8–F11 three such groups (Table 2, Ref. [28]). In addition, fragmentation of the deamidated and methylated forms by CAD provided sufficient sequence information to localize the deamidation site in fractions F3 and F4 to Asn-53 and to show that Asn-52 was the third deamidation site appearing in fractions F8–F11. The second deamidation site (Asn-33) was identified by amino acid sequence analysis of fraction F2. The above studies provided evidence that deamidation was most rapid at the two Asn–Gly sites (Asn-33 and Asn-53) while Asn-52 was deamidated much more slowly following deamidation of Asn-53. Interestingly, none of the other four Asn (Asn-12,20,26 and 37) or three Gln residues [Gln-11,38 and 65 (C-terminus)] in hirudin were found to be deamidated even after 168 h at pH 9 and 37°C (see Fig. 5). This was in agreement with the stabilizing effects of Leu-13 and Val-21 with their branched hydrophobic side chains on Asn-12 and Asn-20 and the fact that all of the four Asn residues are located in the compact N-terminal part of hirudin close to intramolecular disulfide bonds (see Fig. 4) which limits the conformational flexibility necessary to undergo succinimide formation. The greater stability of the Gln residues towards deamidation was in accordance with results of the peptide studies mentioned earlier.

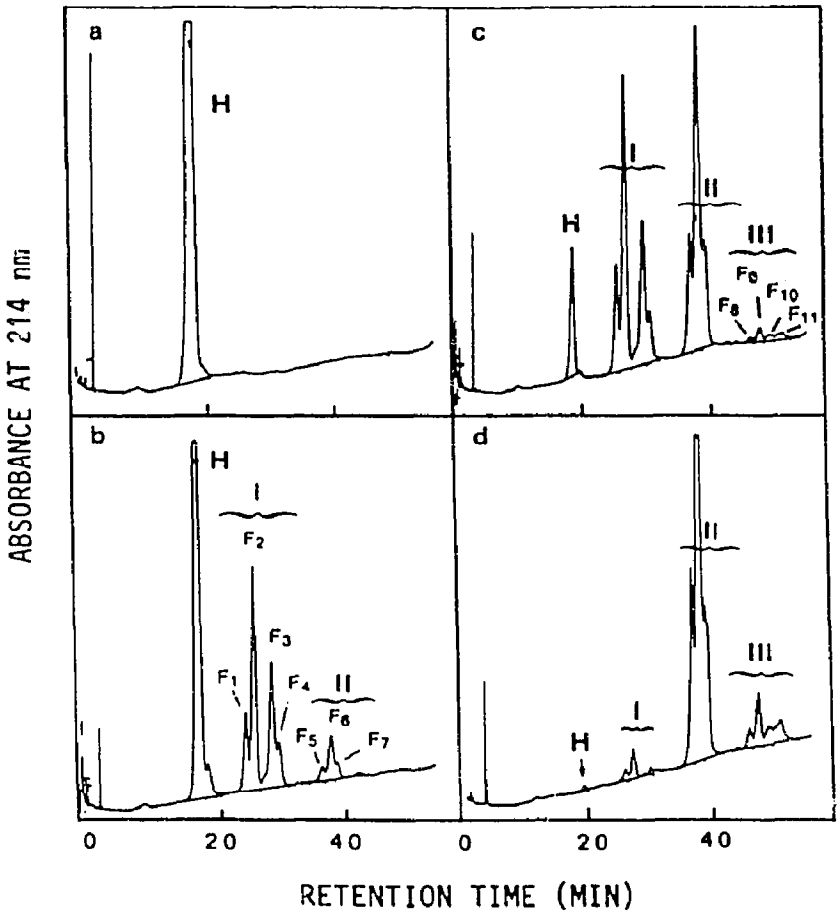

Fig. 9. Anion-exchange HPLC of hirudin (rHV2-Lys-47) after incubation in 0.17 *M* Na_2HPO_4, pH 9, at 37°C for 0 h (a), 6 h (b), 24 h (c), and 168 h (d). Three groups of peaks with one (I), two (II) and three (III) deamidation sites are labelled and individual fractions (F1–F11) are marked (see Table 2) (reprinted with permission from Ref. [28]. Copyright 1992, American Chemical Society).

The formation of iso-Asp linkages via cyclic imides in peptides and proteins can be selectively detected by methylation with PIMT. The use of radiolabeled *S*-adenosyl-L-[*methyl*-^{3}H]methionine provides a means of detecting methylated peptides or proteins and to determine the number of incorporated methyl-groups. Enzymatic methylation has been used to follow formation of iso-Asp upon aging of natural and recombinant human growth hormone at pH 7.4 and 37°C [35]. The results of these studies showed that both the natural and the recombinant form of the protein generate 1.8 methyl-accepting sites per day and per 100 molecules resulting in approximately 27% of the molecules containing one iso-Asp linkage after 14 days of aging. Tryptic mapping of methylated, aged rhGH by RP-HPLC showed the generation of two principal methyl-accepting sites (Fig. 10). The corresponding tryptic peptide fragments were identified by amino acid composition analysis and their identity confirmed either by coelution with synthetic counterparts containing the iso-Asp linkage (site 1) or by fast atom bombardment mass spectrometry at high resolution (monoisotopic mass measurements) and amino acid sequencing (site 2). The characteristic 3:1 ratio between the iso-Asp- and Asp-containing peptides was in accordance with previous observations made on synthetic peptides [4] and indi-

Table 2
Molecular mass determinations of hirudin and its eleven deamidated forms before and after carboxymethylation (see Fig. 9) [28]

Before carboxymethylation			After carboxymethylation		
Fraction	Measured MH^+ (amu)	ΔM^a (amu)	Measured MH^+ (amu)	ΔM^a (amu)	Number of deamidated residues
hirudin	6907.5		7061.5[b]		0
F_1	6907.2	−0.3	7076.5	15.0	1
F_2	6907.5	0.0	7076.4	14.9	1
F_3	6907.9	0.4	7076.4	14.9	1
F_4	6907.8	0.3	7076.7	15.2	1
F_5	6908.7	1.2	7090.9	29.4	2
F_6	6908.7	1.2	7091.1	29.6	2
F_7	6908.7	1.2	7091.7	30.2	2
F_8	6909.9	2.4	7106.8	45.3	3
F_9	6909.8	2.3	7106.8	45.3	3
F_{10}	6909.8	2.3	7107.0	45.5	3
F_{11}	6909.5	2.0	7106.4	44.9	3

[a]Mass difference with respect to hirudin.
[b]Expected value for hirudin (11 carboxylic acid groups) 7061.8.

cated that a succinimide intermediate was involved (see Fig. 1). Site 1 was generated by isomerization of Asp-130 which is followed by Gly and site 2 by deamidation at Asn-149,Ser-150 in rhGH corroborating the fact that Ser facilitates deamidation via the cyclic imide mechanism. Asn-152, on the contrary, which is followed by an Asp residue was not found to be deamidated although it resided within the same tryptic fragment of rhGH thus underlining the pronounced influence of local sequence on the rate of deamidation. Iso-Asp-149 was the major methyl-accepting site in the protein accounting for 58% of the incorporated radioactivity as compared to 22% for site 1. The remaining 20% of incorporated radioactivity was assigned to a number of smaller peaks in the chromatogram of which some were present in the starting material before aging (see Fig. 10, upper trace). Results on Asn–Ser-containing peptides have shown that peptide cleavage after an Asn residue may occur under slightly basic conditions (pH 7.3) with the reaction rate being ca. 10-fold slower than the deamidation rate [7].

4.3. Characterization of reactive succinimide intermediates

While reactive cyclic imide intermediates have been shown to be implicated in deamidation reactions at Asn residues in small synthetic peptides, there was until recently no direct evidence for such intermediates in preparations of purified recombinant or natural proteins. The advent of new ionization methods for mass spectrometry of high-molecular-mass compounds has had a significant impact on the biochemical analysis of proteins and the presence of cyclic imides in purified proteins was demonstrated in a few cases using this technique [10,11,13]. Recombinant human and bovine growth hormones, recombinant hirudin and natural pro-opiomelanocortin have been studied in particular detail with respect to deamidation and cyclic imide formation. Recombinant hirudin and rhGH will serve as examples to describe the analytical methodology which led to the identification of this protein modification. Teshima et al. [11] focused on the above mentioned Asp-

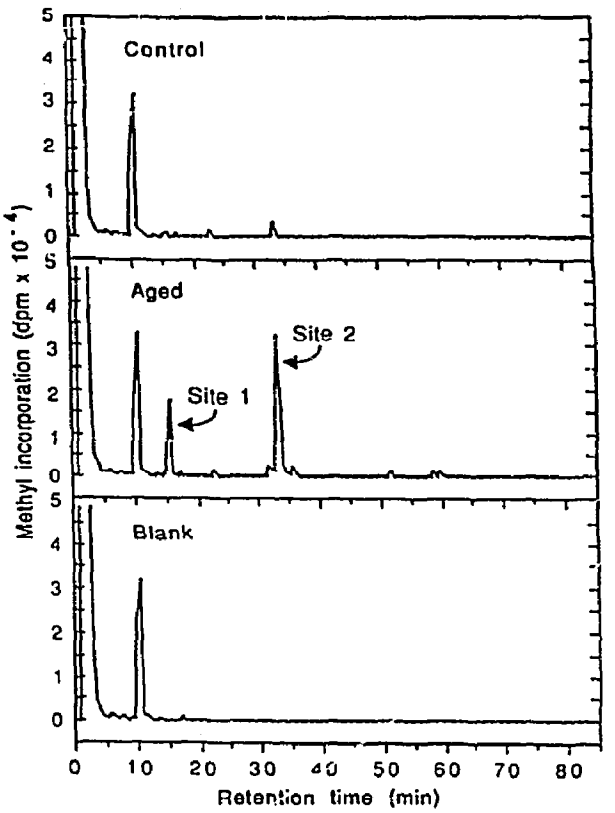

Fig. 10. Reversed-phase HPLC of tryptic fragments of aged and control recombinant human growth hormone (Met-rhGH) with or without prior methylation by protein carboxyl methyltransferase. Met-rhGH had been aged for 14 days at pH 7.4, 37°C. Methylation was performed at a protein concentration of 10 μM using 5 μM protein carboxyl methyltransferase and 200 μM [^{3}H]*S*-adenosyl-L-methionine (200 dpm/pmol). The lower panel shows methylation of a blank trypsin digest. Site 1 corresponds to iso-Asp-130 and site 2 to iso-Asp-149 (reprinted with permission from Ref. [35]. Copyright The American Society for Biochemistry and Molecular Biology).

130,Gly-131 site in rhGH where isomerization to iso-Asp had been observed previously [35]. Aging of rhGH was this time performed in the lyophilized state (the protein was lyophilized from pH 7.2 phosphate buffer) at 45°C and the cyclic imide isolated by a combination of anion-exchange and reversed-phase HPLC. Tryptic digestion at pH 8.2 and analysis of the resulting peptides by reversed-phase HPLC revealed two additional peaks in the aged protein. The more prominent of the two peaks was subsequently identified as fragment Leu-128–Glu–Asp–Gly–Ser–Pro–Arg-134 (measured mass: 755.3 Da; expected mass: 773.3 Da) by fast atom bombardment mass spectrometry with CAD. The mass reduction of 18 mass units was in agreement with the loss of a water molecule. Dehydration was subsequently localized to the Asp–Gly site by tandem mass spectrometry of fragments originating from the molecular ion (m/z 755.3). These results together with the earlier studies by Johnson et al. [35] established that iso-Asp formation at this site involved a cyclic imide intermediate which was present as a by-product in the lyophilized protein during accelerated stability studies.

Recombinant hirudin is another example where detailed analyses applying mass spectrometry led to identification of two cyclic imide intermediates that were present at levels well below 1% in the purified protein [10]. Intermediates were isolated by reversed-phase HPLC employing extremely shallow acetonitrile gradient slopes (see Fig. 7) and site 1 was identified by LSIMS with CAD in the ion source of the mass spectrometer (Fig. 11). These analyses detected a mass difference of ca. 17 mass units between the expected mass and the measured mass of hirudin and localized it to the Asn-53,Gly-54 site in agreement with the loss of an NH_3 moiety from the Asn residue upon cyclic imide formation (see Figs. 1 and 5). Furthermore direct analysis of an endoprotease Glu-C digest of the two hirudin by-products in comparison with the parent molecule by electrospray mass spectrometry confirmed that one cyclic imide was localized in the C-terminal (residues 50–61) and another in the N-terminal part (residues 1–43) of the molecule (Fig. 12B and C). Amino acid sequence analyses confirmed the position of one cyclic imide at Asn-53,Gly-54 (major site accounting for approximately 66% of the by-products) and identified the second site to reside at Asn-33,Gly-34, since Edman degradation did not proceed beyond these sites with appreciable efficiency presumably due to iso-Asp formation as a result of the strongly acidic and basic reaction conditions during sequencing (iso-Asp is refractory to Edman degradation). As these sites had previously been shown to deamidate most

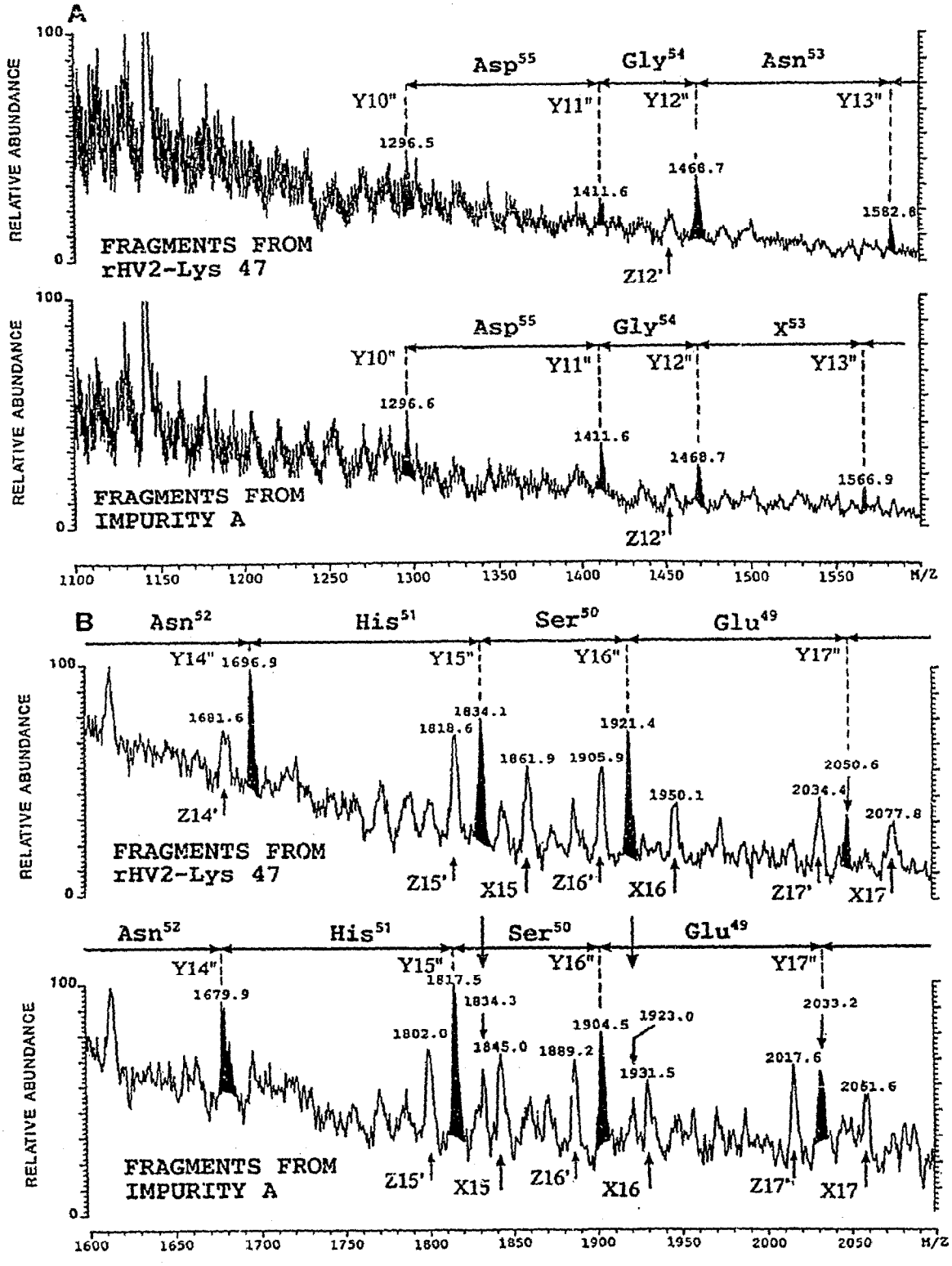

Fig. 11. Liquid secondary-ion mass spectrometry (LSIMS) of (A) hirudin (rHV2-Lys-47) and (B) a mixture of the two cyclic imide containing forms (impurity A; cyclic imides at Asn-33,Gly-34 and Asn-53,Gly-54) showing C-terminal fragments due to CAD in the ion source (nomenclature of fragments according to 54). Fragments of type Y" (filled-in peaks), Z' and X allowed reading of the protein sequence between Asp-55 and Glu-49. The difference of 15.9 mass units for fragment Y13" derived from hirudin (1582.8) and impurity A (1566.9), respectively, corresponds to Asn-53 (reprinted with permission from Ref. [10]. Copyright 1993, American Chemical Society).

rapidly at pH 9 [28], detection of succinimide intermediates in these two positions was in agreement with the reaction mechanism depicted in Fig. 1. The above examples show that cyclic imides may constitute by-products in purified pharmaceutical proteins known to be susceptible

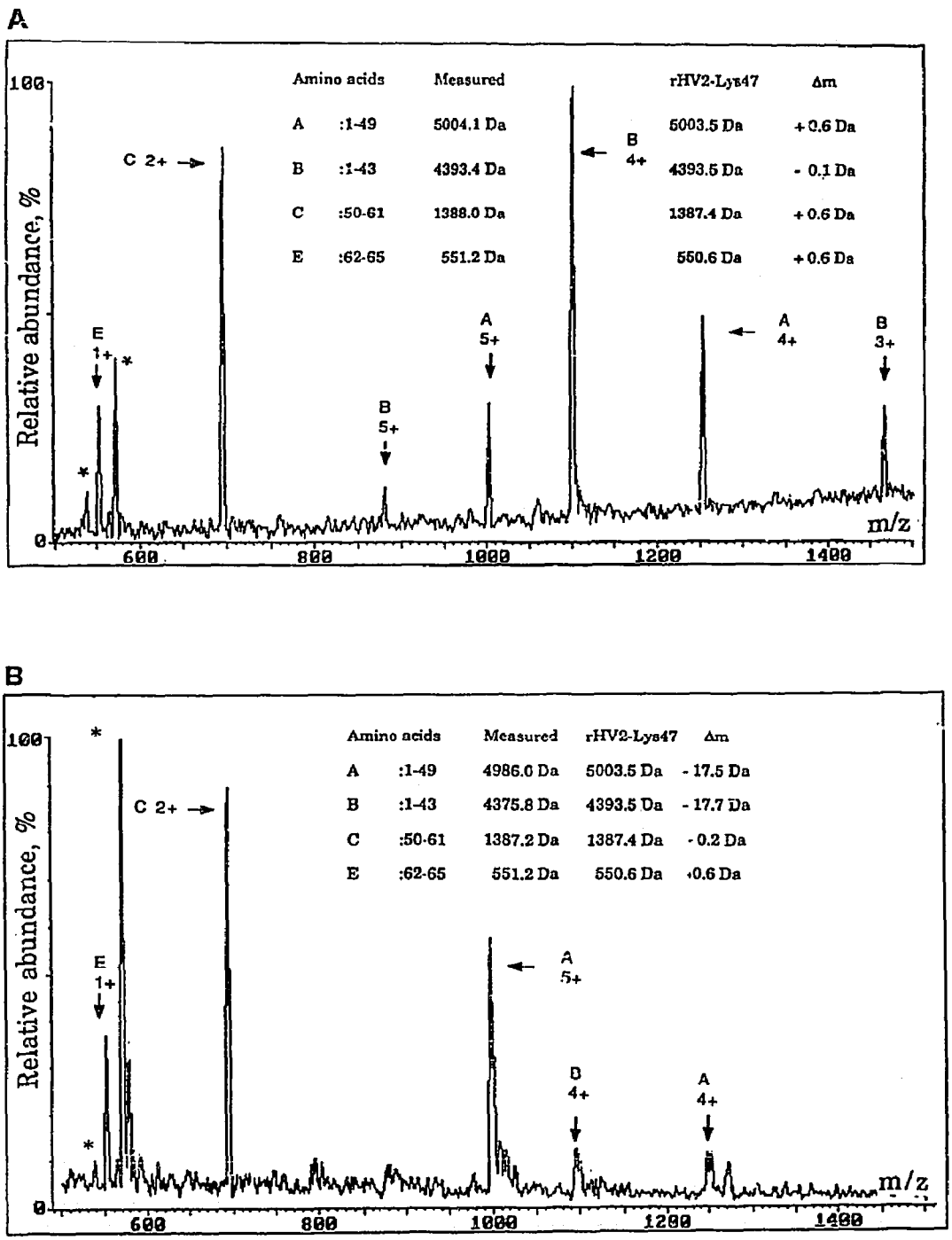

Fig. 12. Electrospray mass spectrometry (ESMS) of V8 proteolytic digests of (A) hirudin (rHV2-Lys-47), (B) hirudin form with a succinimide between Asn-33 and Gly-34, and (C) hirudin form with a succinimide between Asn-53 and Gly-54. The measured molecular masses of fragments corresponding to labeled peaks are listed in comparison to expected values derived from hirudin. Peaks labeled with asterisks were derived from the internal calibration standard gramicidin S. Fragment D (residues 44–49) was not observed (reprinted with permission from Ref. [10]. Copyright 1993, American Chemical Society).

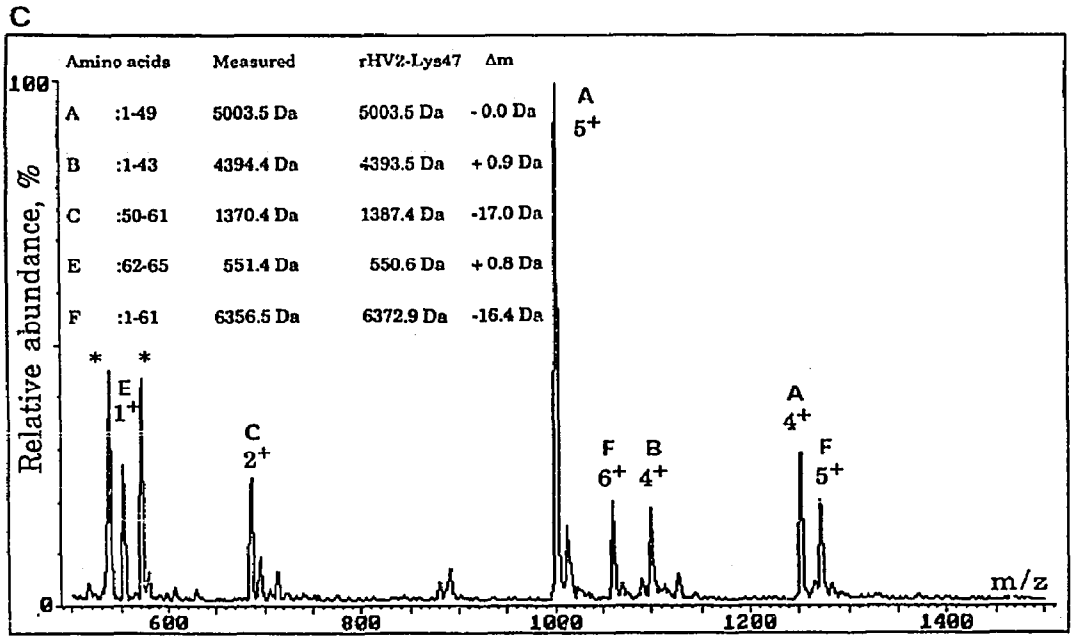

Fig. 12 (*continued*)

to deamidation at Asn or isomerization at Asp residues.

5. Conclusions

Clinical use of recombinant proteins as therapeutic agents has become widespread and control of consistent quality is of utmost importance. This represents a considerable challenge to analytical biochemists since biological macromolecules can be modified in numerous ways as a result of the different reactivities of their functional groups [55]. Analytical methods which allow detection of protein heterogeneity, isolation of modified forms for further analyses and characterization of the modification that leads to the observed change in physical or biological properties of a given protein are therefore crucial.

The present review is mainly concerned with deamidation of Asn residues and with isomerization at Asp–Gly sites as one of the more common modifications found in both natural and recombinant proteins [2]. Deamidation of Asn and isomerization of Asp residues can be considered to be natural modifications of proteins, since they have been observed in numerous proteins isolated from natural sources, notwithstanding the fact that it is often difficult to appreciate whether the modification was present in the starting material or whether it was due to the purification conditions or subsequent handling of the protein. An interesting insight into this question has been obtained through early work on dipeptides secreted into the urine of normal individuals [56,57] where by far the major dipeptide found was iso-Asp–Gly. This unusual dipeptide, which survived due to its resistance to proteolytic attack, can only be the result of ongoing deamidation and/or isomerization of proteins in the organism and thus has to be considered part of the normal turnover of proteins. The recent isolation of a succinimide intermediate of a natural protein supports this view [13]. Proteins which are used as parenteral drugs may thus undergo deamidation reactions as part of their in vivo metabolism and it has to be assumed that cyclic imide intermediates are formed and hydrolyzed during their residence in the organism.

Pharmaceutical production has to address the

question of partial deamidation and isomerization during the production process or during long-term storage, since function, stability or immunogenicity of a given protein may be altered [58]. In order to control or to avoid deamidations by altering conditions of production or storage, it is necessary to apply a range of analytical techniques that allow to follow the generation of such forms rapidly and specifically. The examples of recombinant proteins described in this review summarize analytical strategies on how to approach this task but further work will be needed to render these analyses less time-consuming and costly. On-line techniques such as reversed-phase HPLC or capillary electrophoresis coupled to mass spectrometry are promising approaches which may contribute to reach this goal and improvements in these areas will allow to perform detailed structural analyses of proteins in a time frame that is compatible with routine quality control. In cases where deamidation and/or isomerization cannot be completely avoided it may be necessary to assure that the modified forms do not differ significantly from the parent molecule in critical pharmacological properties.

Acknowledgements

The authors thank Drs. M. Courtney and C. Roitsch for their encouragement and Dr. D. Black for careful reading of the manuscript. The help of Dr. E. Mathieu and B. Heller with computer graphics and the art work is gratefully acknowledged.

References

[1] A.B. Robinson and C. Rudd, *Curr. Top. Cell Regul.*, 8 (1974) 248.
[2] H.T. Wright, *CRC Crit. Rev. Biochem.*, 26 (1991).
[3] P. Bornstein and G. Balian, in C.H.W. Hirs and S.N. Timasheff (Editors), *Methods in Enzymology*, 47 (1977) 132, Academic Press, *San Diego*, CA..
[4] T. Geiger and S. Clarke, *J. Biol. Chem.*, 262 (1987) 785.
[5] R. Lura and V. Schirch, *Biochemistry*, 27 (1988) 7671.
[6] R.C. Stephenson and S. Clarke, *J. Biol. Chem.*, 264 (1989) 6164.
[7] R. Tyler-Cross and V. Schirch, *J. Biol. Chem.*, 266 (1991) 22549.
[8] A.B. Robinson, J.W. Scotchler and J.H. McKerrow, *J. Am. Chem. Soc.*, 95 (1973) 8156.
[9] Y.C. Meinwald, E.R. Stimson and H.A. Scheraga, *Int. J. Peptide Protein Res.*, 28 (1986) 79.
[10] R. Bischoff, P. Lepage, M. Jaquinod, G. Cauet, M. Acker-Klein, D. Clesse, M. Laporte, A. Bayol, A. Van Dorsselaer and C. Roitsch, *Biochemistry*, 32 (1993) 725.
[11] G. Teshima, J.T. Stults, V. Ling and E. Canova-Davis, *J. Biol. Chem.*, 266 (1991) 13544.
[12] B.M. Violand, M.R. Schlittler, P.C. Toren and N.R. Siegel, *J. Prot. Chem.*, 9 (1990) 109.
[13] K. Toney, A. Bateman, C. Gagnon and H.P.J. Bennett, *J. Biol. Chem.*, 268 (1993) 1024.
[14] A.B. Robinson and L.R. Robinson, *Proc. Natl. Acad. Sci. U.S.A.*, 88 (1991) 8880.
[15] H.T. Wright, *Protein Eng.*, 4 (1991) 283.
[16] A.A. Kossiakoff, *Science*, 240 (1988) 191.
[17] P.T.J.A. Groenen, M.J.P. van Dongen, C.E.M. Voorter, H. Bloemendahl and W.W. de Jong, *FEBS Lett.*, 322 (1993) 69.
[18] F.J. van der Ouderaa, W.W. de Jong, A. Hilderink and H. Bloemendal, *Eur. J. Biochem.*, 49 (1974) 157.
[19] S. Capasso, L. Mazzarella, F. Sica and A. Zagari, *Biopolymers*, 28 (1989) 139.
[20] E. Margoliash, E.L. Smith, G. Kreil and H. Tuppy, *Nature*, 192 (1961) 1121.
[21] T. Flatmark, *Acta Chem. Scand.*, 18 (1964) 1656.
[22] T. Flatmark and K. Sletten, *J. Biol. Chem.*, 243 (1968) 1623.
[23] T. Takano, B.L. Trus, N. Mandel, G. Mandel, O.B. Kallai, R. Swanson and R.E. Dickerson, *J. Biol. Chem.*, 252 (1977) 776.
[24] A.B. Robinson, J.H. McKerrow and M. Legaz, *Int. J. Peptide Protein Res.*, 6 (1974) 31.
[25] T.J. Rydel, K.G. Ravichandran, A. Tulinsky, W. Bode, R. Huber, C. Roitsch and J.W. Fenton II, *Science*, 249 (1991) 277.
[26] T.J. Rydel, A. Tulinsky, W. Bode and R. Huber, *J. Mol. Biol.*, 221 (1991) 583.
[27] M.G. Grütter, J.P. Priestle, J. Rahuel, H. Grossenbacher, W. Bode, J. Hofsteenge and S.R. Stone, *EMBO J.*, 9 (1990) 2361.
[28] A. Tuong, M. Maftouh, C. Ponthus, O. Whitechurch, C. Roitsch and C. Picard, *Biochemistry*, 31 (1992) 8291.
[29] P.J.M. Folkers, G.M. Clore, P.C. Driscoll, J. Dodt, S. Köhler and A.M. Gronenborn, *Biochemistry*, 28 (1989) 2601.
[30] H. Haruyama and K. Wüthrich, *Biochemistry*, 28 (1989) 4301.
[31] D.W. Aswad, *J. Biol. Chem.*, 259 (1984) 10714.
[32] R. Bischoff, D. Roecklin and C. Roitsch, *Electrophoresis*, 13 (1992) 214.
[33] W.W. de Jong, A. Zweers, M. Versteeg and E.C. Nuy-Terwindt, *Eur. J. Biochem.*, 141 (1984) 131.

[34] K. Tsuji, L. Baczynski and G.E. Bronson, *Anal. Chem.*, 64 (1992) 1864.
[35] B.A. Johnson, J.M. Shirokawa, W.S. Hancock, M.W. Spellmann, L.J. Basa and D.W. Aswad, *J. Biol. Chem.*, 264 (1989) 14262.
[36] B. An der Lan and A. Chrambach, in B.D. Hames and D. Rickwood (Editors), *Gel Electrophoresis of Proteins–A Practical Approach*, IRL Press, Oxford, 1981, p. 157,
[37] P.G. Righetti, in R.H. Burdon and P.H. van Knippenberg (Editors), *Laboratory Techniques in Biochemistry and Molecular Biology*, 1990, Elsevier, Amsterdam.
[38] S. Hjerten and M.D. Zhu, *J. Chromatogr.*, 346 (1985) 265.
[39] J.R. Mazzeo and I.S. Krull, *Anal. Chem.*, 63 (1991) 2852.
[40] J.R. Mazzeo and I.S. Krull, *J. Chromatogr.*, 606 (1992) 291.
[41] X.-W. Yao, D. Wu and F. Regnier, *J. Chromatogr.*, 636 (1993) 21.
[42] G. Mandrup, *J. Chromatogr.*, 604 (1992) 267.
[43] J.A. McCloskey (Editor), *Methods in Enzymology*, Vol. 193, 1990, Academic Press, San Diego, CA.
[44] S.A. Carr, M.E. Hemling, M.F. Bean and G.D. Roberts, *Anal. Chem.*, 63 (1991) 2802.
[45] M.E. Hemling, G.D. Roberts, W. Johnson, S.A. Carr and T.R. Covey, *Biomed. Environ. Mass Spectrom.*, 19 (1990) 677.
[46] V. Ling, A.W. Guzetta, E. Canova-Davis, J.T. Stults, W.S. Hancock, T.R. Covey and B.I. Shushan, *Anal. Chem.*, 63 (1991) 2909.
[47] S.F.Y. Li, *Capillary Electrophoresis. Principles, Practice and Applications (Journal of Chromatography Library Ser.*, Vol. 52), Elsevier, Amsterdam, 1992.
[48] J.A. Olivares, N.T. Nguyen, C.R. Yonker and R.D. Smith, *Anal. Chem.*, 59 (1987) 1230.
[49] R.D. Smith, J.A. Olivares, N.T. Nguyen and H.R. Udseth, *Anal. Chem.*, 60 (1988) 436.
[50] W.M.A. Niessen, U.R. Tjaden and J. van der Greef, *J. Chromatogr.*, 636 (1993) 3.
[51] S. Clarke, *Annu. Rev. Biochem.*, 54 (1985) 479.
[52] E.D. Murray Jr. and S. Clarke, *J. Biol. Chem.*, 259 (1984) 10722.
[53] D.W. Aswad and B.A. Johnson, *Trends Biochem. Sci.*, 12 (1987) 155.
[54] P. Roepstorff and J. Fohlmann, *Biomed. Mass Spectrom.*, 11 (1984) 601.
[55] J.J. Harding, *Adv. Protein Chem.*, 37 (1985) 247.
[56] J.J. Pisano, E. Prado and J. Freedman, *Arch. Biochem. Biophys.*, 117 (1966) 394.
[57] F.E. Dorer, E.E. Haley and D.L. Buchanan, *Biochemistry*, 5 (1966) 3236.
[58] D.T.-Y. Liu, *Trends Biotechnol.*, 10 (1992) 364.