ELSEVIER

Journal of Chromatography A, 705 (1995) 47–65

JOURNAL OF CHROMATOGRAPHY A

Review

Chemical methods of protein sequence analysis

Jerome M. Bailey

Beckman Research Institute of the City of Hope, Division of Immunology, 1450 E. Duarte Road, Duarte, CA 91010, USA

Abstract

Chemical methods of protein sequence determination are reviewed with particular emphasis on methodology for increasing the sensitivity of amino-terminal sequence analysis and on progress toward the development of an automated procedure for sequential degradation from the carboxy-terminus.

Contents

1. Introduction

Since it is the amino acid sequence which determines the manner in which the polypeptide chain can fold to form the secondary and tertiary structures necessary for biological function, determination of the protein primary structure or sequence is a necessary first step in understanding how a particular function of a protein relates to its structure.

Currently, protein sequence determination is accomplished either by sequencing the protein on an automated sequencer using chemical methods for successive degradation from the amino-terminus or by sequencing the gene for that protein using established DNA sequencing meth-

0021-9673/95/$29.00
SSDI 0021-9673(94)01250-4

odology. Although protein sequencing can be considered to be more difficult and slower than DNA sequencing, it often provides information not obtainable by the latter method. Such information includes: (1) identification of post-translational modifications which are not predictable from the gene sequence. This has become particularly important in view of the ever increasing numbers of recognized post-translational modifications [1], (2) determination of partial protein sequence information which can be used for the design of oligonucleotide probes complementary to predicted gene sequences. In many cases, these oligonucleotide probes, often obtained from low abundance proteins, have been the only route to the cloning of a particular gene, (3) characterization of recombinant proteins in order to confirm that the predicted structure conforms to the expressed product.

These applications have created an increased need for improved methods of protein sequencing from the amino-terminus which can be performed faster and with smaller quantities of sample and for a procedure capable of a sequential degradation from the carboxy-terminus.

This manuscript reviews progress in chemical methodology relating to the development of more sensitive and faster methods of amino-terminal sequence analysis and recent methodology which shows promise toward the development of a routine method for a sequential degradation from the carboxy-terminus.

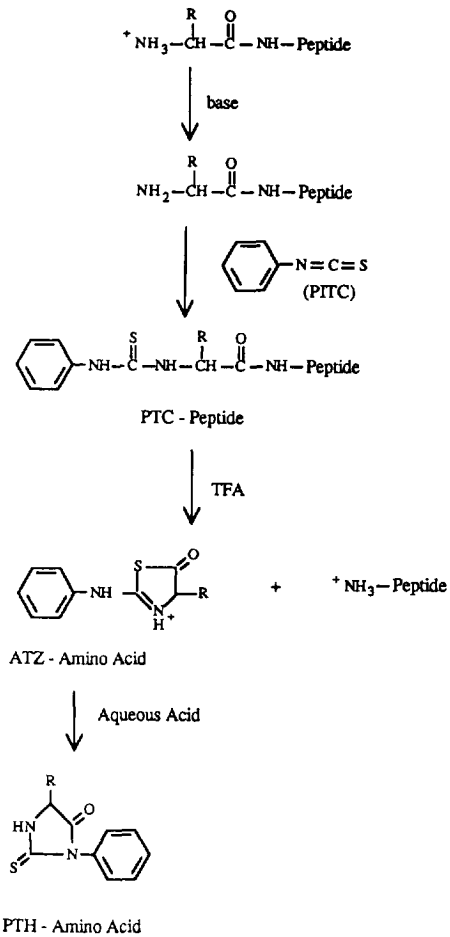

Fig. 1. Edman chemistry for amino-terminal sequence analysis.

2. Amino-terminal sequence analysis

The most widely used method of protein sequence analysis is based on the automated Edman chemistry [2]. In this method (See Fig. 1), the free amino terminus of a polypeptide is reacted with phenylisothiocyanate (PITC) in the presence of base to form a phenylthiocarbamyl peptide (PTC peptide) (this step is commonly referred to as the "coupling step"). Subsequent treatment of the modified polypeptide with acid cleaves the modified N-terminal amino acid as an anilinothiazolinone (ATZ) (this step is referred to as the "cleavage step") and leaves the shortened peptide ready for continued sequence analysis. The thiazolinone is then treated with aqueous acid in order to convert it the more thermodynamically stable phenylthiohydantoin (PTH) derivative (the "conversion step") which is then identified and quantitated by reversed-phase HPLC.

The amount of protein required for sequence analysis with current methodologies (10 to 50 pmol of sample for 10 to 20 cycles of sequence information), while steadily decreasing over the years, has not kept pace with the recent advances in protein purification techniques which

now permit isolation of 50 to 100 femtomoles of sample [3]. The major obstacles to increasing the sensitivity of protein sequencing down to the femtomolar level involve: difficulty in handling such small amounts of protein without losses; the intrinsic detectability of the released PTH amino acids which have relatively low extinction coefficients; the intrinsic background noise associated with absorbance measurements at the wavelength used for PTH amino acid detection (269 nm); the chemical background, caused by the reagents used in sequencing, visible on the HPLC chromatograms in which the PTH amino acids are detected; and the background due to PTH amino acids formed from nonspecific cleavage along the polypeptide chain. The need for a unified approach to solve these problems by redesign of the sequencing chemistry, sample supports, detection methods, and instrumentation has been discussed [4].

2.1. Instrumentation

The chemistry as originally described by Edman in 1950 is still practiced today with only minor variations in all of the commercially available protein sequenators. The performance criteria routinely expected today has been achieved primarily through advances in instrumentation. The original spinning-cup sequenator [5] was capable of sequencing 15 residues in 24 h using approximately 250 nmoles of protein. The combination of reversed-phase HPLC for detection of the PTH amino acids [6], the use of polybrene as a carrier for proteins and peptides [7,8], and modifications to the spinning cup sequenator [9–11] allowed the sensitivity to be reduced to approximately 1 nmole. The introduction of the "gas-phase" sequenator [12,13] and on-line, narrow-bore (2 mm I.D. columns) HPLC detection of the PTH amino acids, reduced the sample requirements to 10–100 picomoles [14].

Since the introduction of the gas-phase sequenator, only minor advances have been made in instrument development. The introduction of the continuous flow reactor (CFR) [15] simplified and miniaturized the reaction cartridge resulting in lower backgrounds and facilitated application of the sequencer to polyvinylidene difluoride (PVDF) membranes (commonly used for electrotransfer from protein gels). A recently described protocol for optimization of the reaction times, miniaturization of the reaction cartridge, and HPLC separation of PTH amino acids was found to permit the usual 40–50 min cycle time to be reduced to 25 min [16]. A commercial reaction cartridge (Blott cartridge, Applied Biosystems) designed specifically for samples on polymeric supports (PVDF) has been introduced and evaluated [17]. Other efforts to reduce chemical background in automated sequencers has involved the evaluation of alternative coupling bases, such as diisopropylethylamine, reduced concentrations of phenylisothiocyanate, and diode-array detection of the PTH amino acids [18]. Some more recent studies showing sequence data obtained at the 1–10 picomole level have advocated routine sequencer maintenance [19] and the use of reduced levels of polybrene and extensive precycling [20]. The successful adaptation of a microbore HPLC (Michrom BioResources) to a commercial sequencer (Applied Biosystems) showed an approximately 2–3 fold enhancement of the limit of detectability of the released PTH amino acids [21]. A compact gas-phase instrument with minimized liquid flow systems [22] was demonstrated to produce less sequencer background peaks at the 5–10 picomole level.

Recently published methods for the separation of the PTH amino acids by capillary HPLC [23] and capillary electrophoresis [24] have shown femtomolar levels of detection of the PTH amino acids, however these techniques require the use of low microliter and subnanoliter injection volumes, respectively, and are thus not of current practical value, since current automated sequencer technologies dissolve the PTH amino acids in a 50–100 μl volume for injection. Use of only a small fraction of this volume would negate any value in the increased sensitivity of detection with these methodologies until a sequencer generating the PTH amino acid in suitably low volumes is developed.

2.2. Solid supports for sequence analysis

Modern gas-phase sequenators when first described relied on the use of glass fiber filters coated with polybrene for retaining the non-covalently applied peptide or protein to be sequenced [12]. Although the proteins separated by one and two dimensional gel electrophoresis could be electroblotted onto these supports [25,26], the introduction of PVDF membranes [27] has largely superseded the use of glass fiber filters in most laboratories. The PVDF membranes were found to give superior yields for electroblotting and for sequence analysis and found to be easier to stain for the presence of protein samples [27–29]. More recent studies comparing the efficiency of sequence analysis and blotting from a series of different commercially available membranes, including glass fiber, PVDF, and polypropylene, have been described [3,30]. An alternative support for non-covalent protein or peptide immobilization consisting of a biphasic column, containing hydrophobic (reversed phase) and hydrophilic (ion-exchange) portions, has recently been introduced by Hewlett-Packard [31]. The biphasic column offers several advantages in that sample clean-up prior to sequencing is greatly facilitated and the lack of necessity for the use of polybrene to assist in sample retention greatly reduces the chemical background during detection of the PTH amino acids. Other preliminary work shows promising results with the use of Zitex (porous PTFE) supports for protein sequencing. Originally used for automated C-terminal sequencing [32], these supports have shown improved repetitive yields for N-terminal protein sequencing as compared to PVDF supports and do not require the use of polybrene for sample retention, resulting in reduced chemical backgrounds.

The non-covalent methods of sample application, while convenient, involve compromises in the choice of solvents and reagents which can be utilized. The types of solvents that may be used are limited to those that are relatively non-polar and that do not cause elution of the protein or peptide. The base and acid reagents used in the coupling and cleavage step are delivered in the gas phase in order to prevent sample elution from the support. A method, termed "solid-phase sequencing", where the peptide or protein to be sequenced was covalently coupled to a solid support prior to sequence analysis was introduced by Laursen in 1966 [33]. The potential advantages to this approach stem from the fact that proteins which are covalently linked to an insoluble matrix can be easily separated from reagent and reaction by-products. The choice of solvents for sequencing can be optimized for ability to remove excess reagents without the need to worry about sample elution. This can lead to faster instrument cycle times and reduction in the amount of UV absorbing background.

Since the introduction of the solid-phase approach to N-terminal protein sequencing, several different types of functionalized supports have been described for the covalent immobilization of polypeptide samples. These include polystyrene resins, polyacrylamide resins, and glass beads substituted with aminoalkyl or aminophenyl groups [34]. Typically these amino functionalized supports are activated for protein coupling with bifunctional reagents such as phenylene diisothiocyanate (DITC). The DITC group is capable of forming a stable thiourea linkage to the support and the peptide N-terminal amino group or ε-amino group of lysine side chains. Recently glass beads derivatized with isothiocyanate, aminophenyl and aminoethylaminopropyl groups [35], glass fiber sheets functionalized with aminophenyl groups [36], and PVDF membranes derivatized with aryl amines and DITC [37] have been used for the covalent immobilization of polypeptides for N-terminal sequencing. The polypeptides are either immobilized by coupling between the epsilon-amino groups of the lysine and the isothiocyanate groups on the solid support using the established DITC chemistry or by the coupling of the carbodiimide activated C-terminal carboxyl groups of the polypeptides and the amino groups on the matrix.

Despite the potential advantages of the solid-phase sequencing approach, the method has never gained routine acceptance. This stems in

part from the difficulties in obtaining high yield, routine methods for covalent coupling, partial blockage of the N-terminal amino group and low yields of Lys when using amino groups for covalent attachment, and low yields of Asp and Glu when using carboxyl groups for covalent attachment. The commercial availability of DITC and arylamine functionalized PVDF membranes for covalent immobilization of proteins and peptides has sparked some renewed interest in the solid-phase approach [38,39]. The solid-phase methodology has been shown to be most useful for solving specific problems not accessible to gas phase sequencers, such as the determination of phosphorylation sites [40–42] and glycosylation sites [43,44].

2.3. Chemistry

Numerous attempts have been made to increase the sensitivity of Edman degradation through the use of radiolabeled, chromophoric, or fluorescent reagents. The vast majority of reagents that have been utilized for this purpose are used in place of the standard PITC reagent and have relied on the isothiocyanate group as the electrophilic group used to mediate the coupling reaction. Isothiocyanates form a stable thiourea group with the N-terminal amino acid. The sulfur atom of a thiourea is thus perfectly placed so that upon acidification a kinetically favored five-membered ring could form which specifically cleaves only the N-terminal amino acid. The use of isothiocyanates also has the advantage in that the isothiocyanate group does not cause any complications in the sequencing reaction due to reaction with amino acid side chains.

4-(N,N'-Dimethylamino)azobenzene-4'-isothiocyanate (DABITC), a highly chromophoric reagent first described by Chang et al. [45] has the potential for improving the sensitivity of detection of the amino acid derivatives, but suffers from a lack of complete derivatization of the N-terminal amino group. It thus requires a follow-up coupling with PITC to prevent serious overlap problems. This reagent has primarily been used as a manual sequencing reagent with a DABITC–PITC double coupling procedure [46], although it has been used in automated solid-phase sequencing [47]. More recently, Aebersold and co-workers [48,49] reported a DABITC–PITC solid-phase sequencing method in which proteins were immobilized on DITC-derivatized aminopropyl glass-fiber sheets. Sequence analysis was performed at the 20–50 picomole level, a substantial improvement over previous methods, but still less sensitive than current gas-phase sequence analysis. Fluorescent reagents, such as fluorescein isothiocyanate [50,51] and dansyl-containing isothiocyanates [52–56] have also been evaluated as sensitivity enhancing reagents. Although synthetic amino acid derivatives prepared using these reagents show subpicomole sensitivity by HPLC analysis, they have not surpassed the sensitivity of gas-phase Edman degradation during automated sequence analysis. In general, it has been found that the use of large bulky chromophores on the isothiocyanate reagent interferes with the efficiency of the derivatization and cleavage reactions of the Edman degradation. An examination of several isothiocyanates [57] showed that, when an isothiocyanate reagent contains an electron withdrawing group the coupling reaction is favored and the cleavage reaction is slowed, and when the isothiocyanate contains an electron donating group the coupling reaction is slowed and the cleavage reaction is favored. It was thus concluded that PITC was the optimal choice of the reagents tested for the Edman chemistry because it offered the best balance between the rates of the coupling and cleavage reactions. Most chromophoric and fluorescent compounds are relatively large when compared to a phenyl ring and when such a compound that contains a reactive isothiocyanate group is substituted for PITC, the coupling and/or cleavage reaction is kinetically disfavored by a combination of steric and electronic effects.

The use of radiolabeled reagents is not a successful approach, since radiolabeled reagents undergo autoradiodegradation which results in decreasing product yields and increasing amounts of labeled by-products. Modified phenyl isothiocyanates such as 4-(*tert.*-butoxycar-

bonylaminomethyl)-PITC, which are designed to react with post-column fluorescent reagents, have also been investigated [58] but have been found to undergo side reactions during the cleavage reaction resulting in loss of the amino group [54].

An alternative to the use of modified Edman reagents is the reaction of the anilinothiazolinone (ATZ)-amino acid intermediate with sensitivity-enhancing nucleophilic reagents (Fig. 2). The use of radiolabeled amines produced amino acid derivatives which could be detected at the femtomole level [59,60], but the handling of radioactive materials was inconvenient. Horn et al. [61] have extended earlier studies on the use

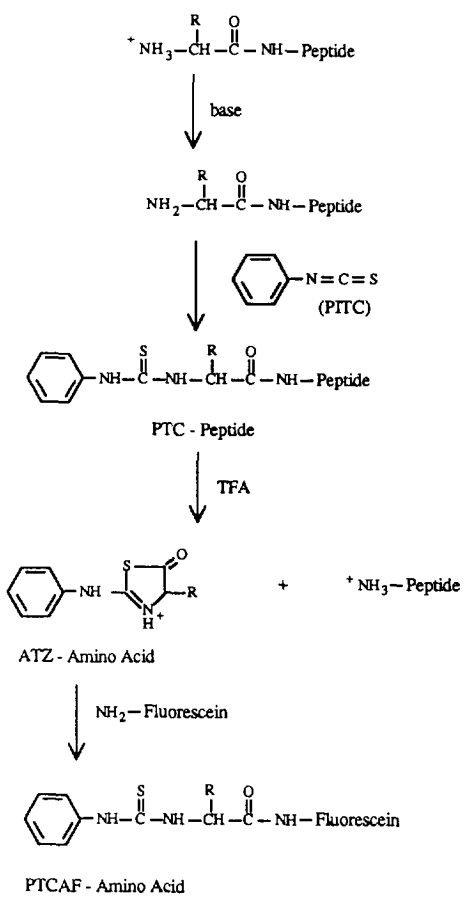

Fig. 2. Reaction of the anilinothiazolinone-amino acid intermediate with sensitivity enhancing nucleophilic reagents.

of MeOH–HCl as a conversion reagent [62] to include chromophoric or fluorophoric alcohols, resulting in the formation of phenylthiocarbamyl amino acid esters. Tsugita et al. [63] have recently reported a modification of the Edman degradation scheme, in which ATZ amino acids are reacted with 4-aminofluorescein resulting in highly fluorescent, phenylthiocarbamyl amino acid aminofluorescein amides (PTCAF-amino acids). PTCAF-amino acids were separated by reversed-phase HPLC and were detectable at the 0.1–1 femtomole level. Several known and unknown protein samples were reported to be sequenced at the 100 femtomole to 10 picomole level using an Applied Biosystems 477A sequencer. Work performed in our laboratory with this method demonstrated low yields with the hydrophilic amino acids, in particular threonine, histidine, glutamate, lysine, and glutamine, and the total lack of yield with aspartate. Recent studies concerning the aminolysis of the ATZ-amino acids by Pavlik et al. [64], showed that many of the ATZ-amino acids, in particular the hydrophilic amino acids, can rearrange so rapidly to the more thermodynamically stable PTH amino acids that by the time the ATZ-amino acid is brought over to the conversion flask of an automated instrument anywhere from 5–70% of the amino acid has already been converted. Once an ATZ-amino acid has converted to a PTH it would not be capable of reacting with aminofluorescein. A recent report describes chemistry which converts the PTH amino acids to the phenylthiocarbonyl derivatives and then back to an ATZ derivative, which in turn, is reacted with aminofluorescein [65]. This procedure was reported to solve the difficulties of derivatization with all of the amino acids, except aspartate which was found not to derivatize.

2.4. Use of mass spectrometry for detection

Recent advances in mass spectrometry have made the use of a mass spectrometer as a detector in place of chromatographic methods a viable option with potential advantages that include, speed, sensitivity, and the ability to analyze post-translationally modified amino acids

since detection is based on mass and not retention time. The use of mass spectrometry for protein sequencing has followed two general approaches.

The first approach involves the substitution of the phenylisothiocyanate reagent of the Edman chemistry with an alternative reagent to generate an amino acid with enhanced detectability by mass spectrometry since PTH amino acids are not detected with high sensitivity. A number of reagents for this purpose have been described in recent years, but to date none have been reduced to practice. The use of 4-nitrophenyl isothiocyanate as a modified Edman reagent to generate amino acid analogues detectable in the femtomolar range by negative ion chemical ionization mass spectrometry has been described [66]. Replacement of the phenylisothiocyanate reagent in the coupling reaction with 3-[4′(ethylene - N,N,N - trimethylamino)phenyl] - 2 - isothiocyanate was shown to produce thiohydantoin amino acid derivatives with quaternary amine groups with femtomole levels of detectability [67]. A disadvantage of this reagent is that covalent attachment of the sample to a solid support is required in order to prevent elution of the sample due to the polar nature of the coupling reagent. The use of dimethylaminopropyl isothiocyanate as a sequencing reagent was shown to proceed with kinetics equivalent to phenyl isothiocyanate and was found to generate amino acid derivatives with the easily ionized dimethylaminopropyl group [68]. A potential advantage of this reagent over reagents with permanently charged quaternary amine groups is that the dimethylaminopropyl group will not become charged until the cleavage step, thereby eliminating the need for covalent attachment of the polypeptide sample. Thiohydantoin amino acids containing the dimethylaminopropyl and trimethylaminopropyl groups were both synthesized and found to be detectable by electrospray mass spectrometry with equal sensitivity in the femtomolar range [68]. The synthesis of 5-acetoxy-2-phenylthiazole derivatives for all of the naturally occurring amino acids was recently described [69]. These amino acids are the products formed during the thiobenzoylation method of protein sequencing and were found to be detectable by gas chromatography–mass spectrometry at the femtomole level.

The second approach involving the use of mass spectrometry involves the generation of a defined series of peptide fragments (called "ladder sequencing") by performing the coupling step of the Edman chemistry in the presence of a small amount of phenylisocyanate as a chain termination reagent [70]. The ladder of peptides, each differing from the next by one amino acid, are then analyzed by matrix-assisted laser-desorption ionization, time of flight mass spectrometry (MALDI-TOF-MS). The difference in mass between each peptide corresponds to the amino acid removed, thus the sequence of the peptide is read by the mass difference between successive peaks. The chief advantages of the method are its speed and ability to analyze post-translationally modified amino acids. The disadvantages are that the method still requires low picomole amounts of sample, and thus is currently not any more sensitive than automated methods of sequence analysis based on chromatographic detection. The technique is limited to peptides with masses less than 3000–4000 daltons due to the increased mass errors of MALDI-TOF-MS above these masses, and the method can not distinguish the isobaric amino acids Leu and Ile.

Despite current limitations of the above described methodologies, continued work combined with predicted future advances in the sensitivity of detection by mass spectrometry, suggest that the combination of chemical methodology for sequencing with mass spectrometric methods of detection show much promise.

3. Carboxy-terminal sequence analysis

The last few years have seen a renewed interest in the development of a chemical method for the sequential C-terminal sequence analysis of proteins and peptides. Such a method would be analogous and complimentary to the Edman degradation commonly used for N-terminal sequence analysis [2]. It would also be invaluable for the sequence analysis of proteins

with naturally occurring N-terminal blocking groups, for the detection of post-translational processing at the carboxy-terminus of expressed gene products, and for assistance in the design of oligonucleotide probes for gene cloning. Although a number of methods have been described, the method known as the "thiocyanate method", first described in 1926 [71], has been the most widely studied and appears to offer the most promise due to its similarity to current methods of N-terminal sequence analysis. The field of C-terminal sequencing was extensively reviewed in 1991 [72], so this review will concentrate mainly on more recent developments in the field and in particular, the thiocyanate method.

3.1. Chemistry

The thiocyanate method involves the reaction of a protein or peptide with an isothiocyanate reagent, in the presence of a carboxylic acid activating reagent (such as acetic anhydride), to form a peptidylthiohydantoin (Fig. 3). The derivatized amino acid is then hydrolyzed to yield a shortened peptide or protein and a thiohydantoin amino acid. As in the Edman degradation the thiohydantoin amino acid is then identified

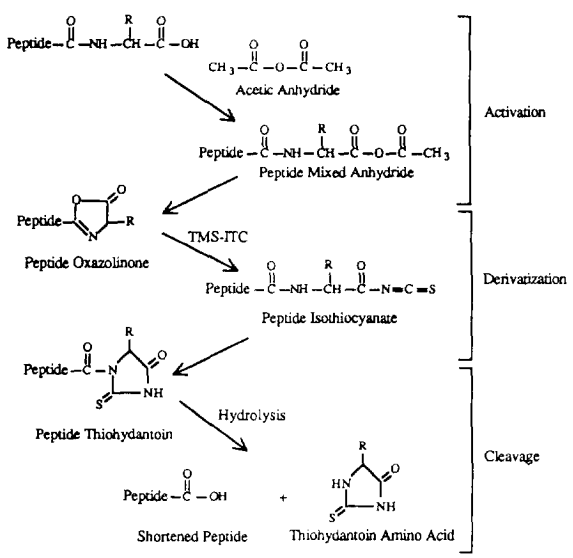

Fig. 3. Reaction scheme for the thiocyanate method of carboxy-terminal sequence analysis.

and quantitated by reversed-phase HPLC. Since the thiohydantoin amino acids produced with this methodology have UV absorption spectra and extinction coefficients similar to those of the phenylthiohydantoin amino acids formed during the Edman degradation, the sensitivity of the thiocyanate method is anticipated to be similar to that of current N-terminal methods (10–200 pmol of applied sample). Traditionally, the sequence of steps necessary for one complete cycle of the thiocyanate chemistry have been referred to as the activation, derivatization, and cleavage steps. The activation step refers to the treatment of the C-terminal carboxylic acid with a reagent that activates it for reaction with a nucleophilic reagent by converting it to an electrophilic group, such as a mixed anhydride or oxazolinone (Fig. 3). The derivatization step refers to the reaction of the activated C-terminal carboxylic acid with an isothiocyanate reagent to derivatize it to a thiohydantoin and the cleavage step refers to the specific hydrolysis of the peptidylthiohydantoin to form a shortened peptide or protein and a thiohydantoin amino acid.

A number of problems have prevented this methodology from becoming a routine method in the protein chemistry laboratory. These include the severity of the conditions used for the derivatization reaction, blockage of the shortened peptide to continued sequencing by the reagents used for the cleavage reaction, the inability to sequence through some of the amino acids commonly found in proteins (especially proline), long reaction times, large amounts of required sample, and the need to covalently couple the sample to be sequenced to a solid support.

3.1.1. The derivatization reaction

Although the derivatization of amino acids to their corresponding thiohydantoins has been studied since this reaction was first reported in 1911 [73], the mechanism of peptidylthiohydantoin formation by acetic anhydride and thiocyanate ions is still not well understood. Early experiments confirmed the assumption that thiocyanic acid is first formed when acetic anhydride and acetic acid interact with ammonium thiocyanate and it is the thiocyanic acid which

actually reacts to form the thiohydantoin [74]. The differing ability of the various salts of thiocyanic acid to form a thiohydantoin was reasoned to result from their ability to form thiocyanic acid on interaction with acetic anhydride and acetic acid [75]. When more convenient methods for the preparation of thiocyanic acid became available, thiocyanic acid, in the presence of acetic anhydride, was found to be more reactive for the formation of 2-thiohydantoins than were the thiocyanate salts. As a result, thiocyanic acid has been used by Kubo et al. [76] and more recently by Inglis et al. [77] for the sequential degradation of peptides from the C-terminus. However, one of the principal drawbacks of thiocyanic acid is that it tends to be self reactive, even at ambient temperature, and quickly loses its ability to derivatize the peptide. Furthermore, these polymeric thiocyanic acid products are intensely UV absorbing at the wavelengths used for thiohydantoin detection and subsequently interfere with the HPLC identification of the released thiohydantoin amino acid. The instability of the free thiocyanic acid presents difficulties when the chemistry is automated, since the reagents needs to be stable to storage in a reagent bottle at room temperature. As discussed by Inglis et al. [78], one way to help stabilize the thiocyanic acid is to refrigerate it while in the automated instrument. A less costly and more convenient solution to the problems caused by the use of thiocyanic acid was presented with the introduction of trimethylsilylisothiocyanate (TMS-ITC) (Fig. 3) for derivatization of the C-terminal amino acid to a thiohydantoin [79]. The trimethylsilyl group offered two advantages, (1) it stabilized the thiocyanate sufficiently so that self reaction was no longer a problem, and (2) it did not compromise the ability of the thiocyanate to form thiohydantoins. This is consistent with the observation that silylated amines have often been found to be better nucleophiles than the corresponding unsubstituted amines [80].

The use of guanidine thiocyanate [81], benzoyl isothiocyanate [82] and tributyltin isothiocyanate [83] have also been described for the formation of peptidylthiohydantoins. Both guanidine thiocyanate and tributyltin isothiocyanate require the use of a chloride containing activation reagent, such as acetyl chloride or FMOC-chloride. The chloride ion is required for release of the thiocyanate ion from the guanidine or tributyltin molecule. Once released the thiocyanate ion is then free to react with the peptide oxazolinone to form a peptidylthiohydantoin. Benzoyl isothiocyanate (BITC) offers the potential advantage in that it does not require a separate activating reagent. Reaction of BITC with a C-terminal carboxylate was found to form a benzoic acid mixed anhydride which then required the addition of pyridine (or other catalyst such as triazine, imidazole or tetrazole) for formation of an oxazolinone. Once formed the oxazolinone can then react with the isothiocyanate ion, liberated when BITC reacted with the C-terminal carboxylate, to form a peptidylthiohydantoin. Although, all three of these reagents can effectively form peptidylthiohydantoins, they have not been found to offer any significant advantages over the acetic anhydride/TMS-ITC chemistry in terms of reaction times, yields, and ability to derivatize all twenty common occurring amino acids.

The intermediate involved in thiohydantoin formation has been a subject of study for many years. An oxazolinone was postulated to be a necessary intermediate during the synthesis of amino acid thiohydantoins with acetic anhydride and ammonium thiocyanate when this reaction was first studied [74]. The racemization of the C-terminal amino acids observed on reaction with acetic anhydride and TMS-ITC [84] suggests that reaction of peptides with acetic anhydride forms a peptide oxazolinone (Fig. 3). This is consistent with the above postulated mechanism. The formation of oxazolinones is known to cause racemization of amino acids [85–87]. Further evidence of an oxazolinone intermediate during the formation of amino acid thiohydantoins was obtained by Csonka and Nicolet [85]. Additional studies described by Cornforth [88] actually demonstrated the formation of an oxazolinone intermediate in the formation of thiohydantoins by the combined use of absorption spectra and polarimetry to follow the rate of oxazolinone

formation. In fact, once the oxazolinone was formed, the reaction with isothiocyanic acid was found to be facile enough to occur readily at 0°C in the case of 2-phenyl-4-benzyl-5-oxazolinone [88].

A number of reagents for activation of the carboxylic acid other than acetic anhydride have been described. The use of acetic anhydride was shown to cause problems with the sequence analysis of certain amino acids [84] and found to contribute to difficulties in the sequence analysis of protein samples [89]. The use of 9-fluorenylmethyl chloroformate [89], 2-ethyl-5′-phenylisoxazoliumsulfonate [90], 1,1,3,3-tetramethyl-chlorouronium chloride [91], and 2-halo-1-methylpyridinium salts [92] have all been described.

An alternative approach for the formation of peptidylthiohydantoins involved the use of diphenyl phosphoroisothiocyanatidate (DPP-ITC) and pyridine [93]. This chemistry combined the activation and derivatization steps, eliminating the need for a separate activation step. Although the use of DPP-ITC for C-terminal sequencing was first described by Kenner et al. in 1953 [94], it was never actively pursued, most likely since it took several days for the reaction to go to completion. These authors proposed a reaction mechanism whereby isothiocyanate ion was liberated by exchange with the C-terminal carboxylate forming an acyl phosphate. The acyl phosphate was in turn postulated to be attacked by the isothiocyanate ion to form the desired acyl isothiocyanate and diphenyl phosphate ion. Work in our laboratory suggested an alternative mechanism involving first attack by the C-terminal carboxylate (formed by base treatment of the polypeptide) on the phosphate of DPP-ITC to form a pentacovalent acylphosphorylisothiocyanate. Introduction of pyridine, imidazole, triazine, or tetrazole was then found to promote the rapid rearrangement of the pentacovalent acylphosphorylisothiocyanate to the acylisothiocyanate, with concomitant release of diphenylphosphate, which then rapidly cyclized to a thiohydantoin. The entire reaction was found to be complete in less than 10 min at 50°C [93]. Pyridine has recently been shown to promote a similar type of reaction by removing carbon dioxide from the mixed anhydride formed on reaction of an amino acid carboxylate with ethyl chloroformate to form an amino acid ester [95]. The success of this reaction was dependent on pyridine being present in a molar excess with respect to the ethyl chloroformate. A similar mechanism, as that proposed by us for the reaction of DPP-ITC with carboxylates to form an acylisothiocyanate, was proposed for the reaction of diphenylphosphoryl azide with peptide carboxylates to form an acyl azide [96]. These authors postulated that the diphenylphosphoryl azide reacts with the C-terminal carboxylate to form a pentacovalent phosphorus compound (which was noted to be resistant to racemization) and rearranged via an intramolecular mechanism to form an acyl azide. The use of the DPP-ITC/pyridine reaction, by elimination of the oxazolinone intermediate, was found to result in faster cycle times and improved the ability to sequence previously difficult amino acids, in particular, aspartate, glutamate, serine, threonine, and proline (see below).

3.1.2. The cleavage reaction

The cleavage reaction has been extensively studied since the thiocyanate chemistry for C-terminal degradation was first proposed by Schlack and Kumpf in 1926 [71]. In their procedure, strong base (1 *M* sodium hydroxide) was used to liberate the amino acid thiohydantoin and generate a new carboxyl-terminal amino acid. Since this work was published, numerous groups have tried to reduce the severity of the conditions required in order to apply this chemistry to the sequential degradation of proteins from the carboxyl terminal end. Lesser concentrations of sodium hydroxide [97,98] than originally used by Schlack and Kumpf [71] and of barium hydroxide [99] were found to effectively cleave peptidylthiohydantoins. Other groups [100,101] used acidic conditions based on the original procedure used by Johnson and Nicolet [73] for the de-acetylation of amino acid thiohydantoins. These authors added concentrated hydrochloric acid to the coupling solution to cause cleavage of the peptidylthiohydantoin

bond. Unlike hydroxide which was shown to cause breakdown of the thiohydantoin amino acids [102], hydrochloric acid was shown not to destroy the amino acid thiohydantoins [103,104]. Cromwell and Stark [105] showed that concentrated hydrochloric acid could be used to cleave the thiohydantoin amino acid at room temperature. The major drawback with this procedure was that when applied to proteins no more than two or three cycles could be performed. Yamashita [106] found that cleavage of petidylthiohydantoins could be done in a repetitive manner with a protonated cation-exchange resin. Stark [104] reported that certain organic bases such as morpholine or piperidine could be substituted for sodium hydroxide, and along the same lines, Kubo et al. [76] reported that aqueous triethylamine (0.5 *M*) could be used to effectively cleave peptidylthiohydantoins. Stark [104] appeared to have solved the cleavage problem by introducing acetohydroxamic acid in aqueous pyridine at pH 8.2 as a cleavage reagent. This reagent was shown to rapidly and specifically cleave peptidylthiohydantoins at room temperature and at mild pH. However, a more recent study found that, although acetohydroxamate is an excellent cleavage reagent for the first amino acid, it forms a stable peptidyl hydroxamate ester which is difficult to hydrolyze, and which can partially or completely block (depending on the conditions employed) the shortened peptide from further sequencing [84]. Primary amines, such as N-butylamine in trifluoroethanol, were also found to be excellent cleavage reagents by Inglis et al. [77], but were recently shown by Hawke and Boyd [107] and Inglis et al. [78] to permit only one cycle of sequencing since they form a stable amide at the C-terminus of the shortened peptide, effectively blocking the shortened peptide to further sequencing.

In another study by Inglis et al. [78], an aqueous solution of potassium hydroxide containing 33% methanol and dithioerythritol (DTE) or dithiothreitol (DTT) was employed for the cleavage reaction. Although the presence of DTE or DTT has been shown to protect the released thiohydantoin amino acid from degradation under these basic conditions [77,108], the mercapto group of these molecules has also been shown to form an adduct at the C-terminus of the shortened peptide resulting in the formation of a percentage of shortened peptide blocked to further degradation at each cycle [108]. This may partially explain the decrease in the repetitive yield with each cycle observed by Inglis et al. [77].

A dilute solution of aqueous triethylamine was demonstrated to cause rapid and quantitative cleavage of a wide variety of peptidylthiohydantoins in the solution phase [84], but when applied to automated sequencing on the solid phase, several problems became apparent which resulted in poor repetitive yields and increased background [109,110].

A new cleavage reagent, sodium trimethylsilanolate, was introduced and demonstrated to cause rapid and quantitative cleavage of peptidylthiohydantoins [110]. This reagent was found not to suffer from the limitations associated with previous reagents, such as: the generation of UV absorbing background peaks which interfere with identification of the released thiohydantoin amino acids, partial or complete blockage of the shortened peptide to continued sequencing, and destruction of the released thiohydantoin amino acid. Use of sodium trimethylsilanolate in alcoholic solvents and under an inert atmosphere, such as argon, has permitted the extended sequencing of peptides and proteins, covalently coupled to polyethylene supports and non-covalently applied to Zitex strips, respectively [32,93].

Another approach described involves the treatment of the peptidylthiohydantoin with an alkylating agent in the presence of base, followed by cleavage with trimethylsilyl isothiocyanate in the presence of trifluoroacetic acid [91]. There are two potential advantages of this approach. The first is that the need to reform the C-terminal carboxylate on the shortened peptide is obviated since the shortened peptide is derivatized to a peptidylthiohydantoin simultaneous with cleavage. The second is that the alkylation step could be used to introduce a highly chromophoric or fluorescent group that

would permit a more sensitive detection of the thiohydantoin amino acids. However, in practice, this approach still suffers from several disadvantages. The initial derivatization to a peptidylthiohydantoin in cycle 1 still requires the use of a separate activation and isothiocyanate reagent limiting the number of amino acids which can be successfully derivatized to a thiohydantoin; several of the amino acid side chains have been found to interfere with the alkylation reaction thereby permitting successful sequencing of only 13 of the 20 naturally occurring amino acids with this approach [91]; the method requires covalent attachment of protein samples to a solid support in order to prevent loss of the sample; and the alkylated thiohydantoin amino acids are unstable and must be analyzed immediately after formation. Despite these limitations, the method still shows promise, and with continued work and incorporation of advances from other laboratories, it is likely that the limitations associated with this approach can be solved.

3.2. Solid supports for C-terminal sequence analysis

Most of the studies involving C-terminal sequence analysis have employed samples which are covalently coupled to a solid support. The use of glass beads for the covalent immobilization of peptide samples [77,79,89,111–113] suffers from chemical instability especially during the basic conditions used for the cleavage step in the thiocyanate chemistry. The use of carboxylic acid modified PVDF [110], DITC-activated amino PVDF [78,108], a disuccinimidoyl carbonate polyamide resin [107] and aminomethylpolystyrene beads [91] have also been considered as alternatives to silica supports. In our own work, both the glass and PVDF-based supports were found to be unstable in the presence of the cleavage reagent, sodium trimethylsilanolate [83]. A more suitable support was found to be a carboxylic acid modified polyethylene film [83].

As discussed above for N-terminal sequence analysis, there are several problems with the use of covalent methods for C-terminal sequence analysis. These include the lack of a routine method for high yield covalent coupling. This stems from the fact that each protein or peptide sample often has quite unique properties and it is difficult to develop a procedure which works well for all of the samples. Additionally, low yields or no yields of lysine often result when covalently coupling to peptide amino groups [93]. Recent work demonstrating the C-terminal sequence analysis of proteins non-covalently applied to Zitex supports has demonstrated that the use these supports can permit the routine C-terminal sequence analysis of polypeptides without the need for covalent coupling, simplifying the procedure and eliminating the loss of sample all too often associated with covalent coupling procedures [32,93].

3.3. Proline

The derivatization of C-terminal proline to a thiohydantoin has been a major impediment to the development of a routine method of C-terminal sequence analysis of proteins and peptides. Since the method was first described in 1926 [71], the derivatization of C-terminal proline has been problematic. While over the years a few investigators have reported the derivatization of proline, either with the free amino acid or on a peptide, to a thiohydantoin [76,77,114], others have been unable to obtain any experimental evidence for the formation of a thiohydantoin derivative of proline [84,99,103,104,115]. As pointed out by Stark [104], cyclization to form a peptidylthiohydantoin with proline would require the quaternization of the imino nitrogen, thereby potentially resulting in the simultaneous cyclization and cleavage of C-terminal proline. This would lead to a gap at proline, since the method would continue on to the next residue. Recently, Inglis and De Luca [116] described the successful synthesis of thiohydantoin proline from N-acetylproline. This was done by the one-pot reaction of acetic anhydride, acetic acid, trifluoroacetic acid, and ammonium thiocyanate

with N-acetyl proline. Experiments performed in our laboratory reproduced this work and developed a scaled up procedure which permitted several hundred milligrams of thiohydantoin proline to be produced [117]. However, application of this procedure to a tripeptide containing a C-terminal proline by Inglis and De Luca [116] and similar attempts in our laboratory showed that this method did not work well for peptide samples. Application of the procedure to the tripeptide, N-acetyl-Ala-Phe-Pro, in our laboratory, found that thiohydantoin proline was formed in low yield (approx. 1–2% of theoretical). Recovery of the peptide products after the reaction revealed that approximately half of the starting peptide was unchanged and the remaining half had been decarboxylated at the C-terminus, thereby blocking it to C-terminal sequence analysis. This was most likely caused by the high concentration of trifluoroacetic acid, the excess of acetic anhydride present, and the high temperature (80°C) at which the the reaction was performed.

The poor reaction with C-terminal proline most likely stems from the fact that proline cannot form the necessary oxazolinone for efficient reaction with the isothiocyanate. Work in our laboratory has solved this problem by the use of diphenyl phosphoroisothiocyanatidate and pyridine. Reaction of this reagent with C-terminal proline directly forms the acylisothiocyanate without the need for oxazolinone formation. Once the acylisothiocyanate is formed, the addition of either liquid or gas phase acid followed by water releases the proline as a thiohydantoin amino acid derivative [117]. Unlike thiohydantoin formation with the other 19 naturally occurring amino acids, C-terminal proline thiohydantoin requires the addition of acid to provide a hydrogen ion for protonation of the thiohydantoin ring nitrogen. This step is necessary for stabilization of the proline thiohydantoin ring. The resulting quaternary amine containing thiohydantoin can then be readily hydrolyzed to a shortened peptide and thiohydantoin proline by introduction of water vapor or by the addition of sodium trimethylsilanolate (the reagent normally used for cleavage of peptidylthiohydantoins). The automation of this chemistry has allowed proline to be analyzed in a sequential fashion without affecting the chemical degradation of the other amino acids [117].

3.4. Instrumentation

The instrumentation used for automated C-terminal sequencing has involved only minor modifications to existing commercial protein sequenators. Instruments which have been utilized for automated C-terminal sequence analysis include: modules of the Knauer sequencer [81], the commercially available Applied Biosystems Model 477A [91], and the commercially available Hewlett-Packard Protein Sequencing system [117]. A recent manuscript describes the construction of a compact sequencer designed specifically for automated C-terminal sequence analysis [32]. In all of these instruments the "conversion flask" is used as a holding vessel for the thiohydantoin amino acid. After the cleavage reaction the thiohydantoin amino acid is transferred to the conversion flask, where it is then dried and then re-dissolved in a solvent suitable for on-line HPLC injection.

3.5. Examples of automated C-terminal sequence analysis

The chemistry and instrumentation developed in our laboratory now permits the derivatization and identification of all of the twenty naturally occurring amino acids. Derivatization of the C-terminal amino acid to a thiohydantoin is accomplished with diphenyl phophoroisothiocyanatidate (liquid phase) and pyridine (gas phase). The peptide is then extensively washed with ethylacetate and acetonitrile to remove reaction by-products. The peptide is then treated briefly with gas-phase trifluoroacetic acid, followed by water vapor in case the C-terminal residue is a proline (this treatment has no effect on residues which are not proline). The derivatized amino acid is then specifically cleaved with sodium tri-

methylsilanolate to generate a shortened peptide or protein which is ready for continued sequencing. The thiohydantoin amino acid derivative is then quantitated and identified by reversed-phase HPLC. Peptides are covalently attached to carboxylic acid-modified polyethylene film prior to sequencing and proteins are non-covalently applied to Zitex strips. Automated sequencing was performed on a compact sequencer constructed in our laboratory [32].

Fig. 4 shows the reversed-phase separation of the thiohydantoin amino acid standards (400 pmol) [117]. The extinction coefficients of the thiohydantoins [approx. 17 500 [105]) are similar to the phenylthiohydantoins and therefore would be expected to have similar limits of detectability. The major difficulty with the HPLC separation of the thiohydantoin amino acids has been the greater hydrophilicity of the thiohydantoins as compared to the phenylthiohydantoins. Most commercially available C_{18} reversed-phase HPLC columns are not ideal for the separation of such polar molecules. This is reflected most with the thiohydantoin derivatives of the polar amino acids such as Asp, Asn, Gly, and Gln. These derivatives are poorly retained, even in the absence of organic solvents, and consequently these residues are poorly resolved and often elute as broad peaks which are difficult to identify and quantitate, especially under automated sequencing conditions. As with PTH-Ile, the TH-Ile analogue elutes as a doublet due to the formation of the allo-isomer during synthesis. This results in the formation of two diastereomeric derivatives of TH-Ile which are resolved during the separation. Although Thr and Ser sequence in good yield, the nature of the derivatives formed is not well understood. The threonine derivative obtained does not co-elute with a dehydrothreonine standard (labelled as T in Fig. 4), but rather elutes 2 min earlier. It has equivalent UV absorbance at 265 and 319 nm, suggesting that it may not be a simple thiohydantoin threonine. The same UV absorbance properties were observed with the serine analogue. The elution position of the Thr and Ser analogues is shown with arrows in Fig. 4. C-Terminal cysteine is also readily sequenced with this methodology, forming a derivative indistinguishable from that formed with serine. The lysine standard shown in Fig. 4 is acetylated at the ε amino group during the synthesis of the analogue. However, during actual automated sequencing, the epsilon acetyl analogue is not formed unless the protein sample is treated with acetic anhydride prior to sequencing. In the absence of any pretreatment lysine forms a thiohydantoin derivative which co-elutes with TH-Phe (see Fig. 5, cycle 3).

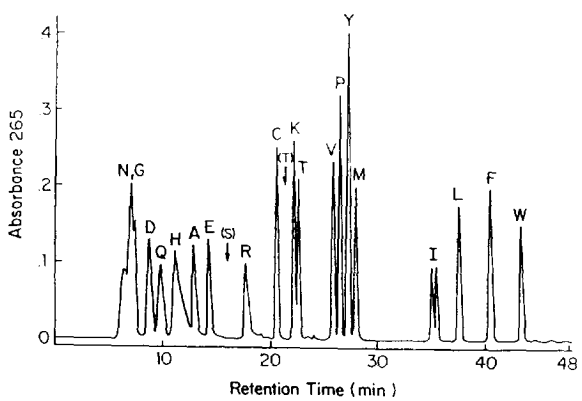

Fig. 4. High-performance liquid chromatographic separation of the amino acid thiohydantoin standards. The amino acid thiohydantoins standards (400 pmol) were separated on a C_{18} Reliasil column [117]. Absorbance was monitored at 265 nm at 0.4 AUFS. The thiohydantoin derivative of cysteine is S-methylated and the side ε-amino group of the lysine derivative is acetylated. The elution positions of the Thr and Ser analogues formed during sequencing are indicated with arrows.

Figs. 5 and 6 show three cycles of automated sequencing of hemoglobin α-chain (4.1 nmol) and bovine serum albumin (700 pmol), respectively. Both proteins were non-covalently applied to Zitex strips. Fig. 7 shows three cycles of sequence analysis of the tripeptide, LAP (15 nmol), covalently coupled to carboxylic acid-modified polyethylene. The low yields of leucine in cycle 3 are due to covalent attachment of this amino acid to the solid support [93]. This methodology currently is capable of providing 2–3 cycles of C-terminal sequence information with

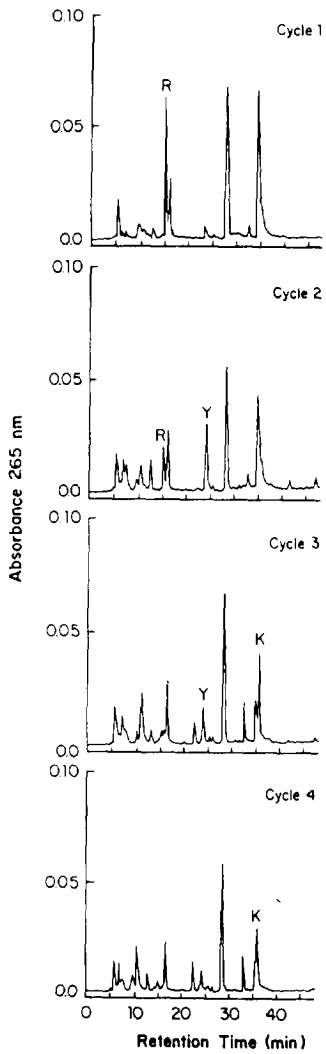

Fig. 5. Automated C-terminal sequencing of hemoglobin α-chain (4.1 nmol) non-covalently applied to Zitex. The sequence of the C-terminus is –Lys-Tyr-Arg.

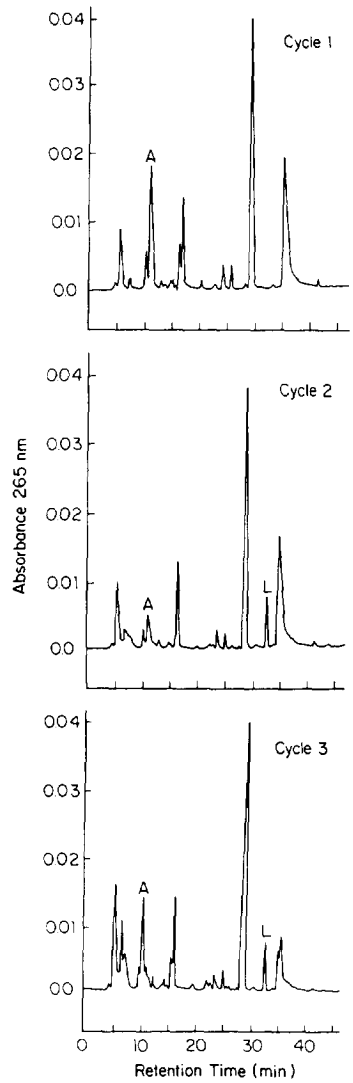

Fig. 6. Automated C-terminal sequencing of bovine serum albumin (700 pmol) non-covalently applied to Zitex. The sequence of the C-terminus is –Ala-Leu-Ala.

noncovalently applied proteins and 5–6 cycles on most covalently attached peptides.

4. Conclusions

There are a number of promising chemical approaches to improving the sensitivity of sequence analysis from the amino-terminus. The recent advances in the use of fluorescent detection and the use of mass spectrometry for detection of the released amino acid derivatives both show promise and suggest that a more sensitive and faster method for amino-terminal sequencing will be possible in the not too distant future.

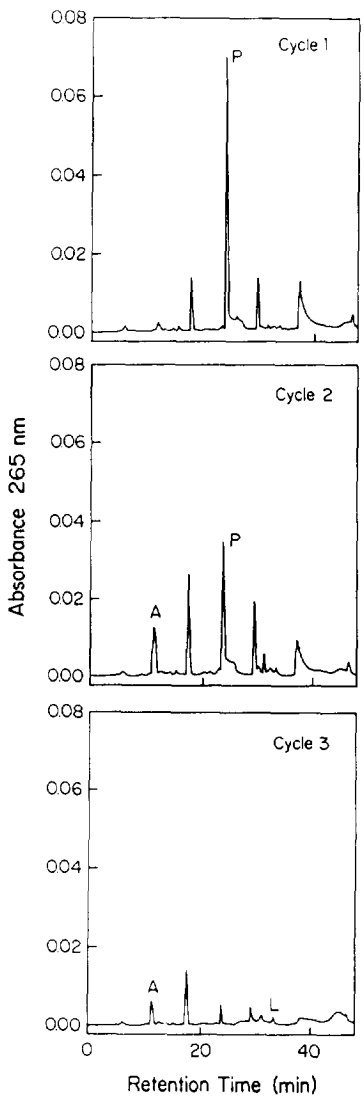

Fig. 7. Automated C-terminal sequencing of the tripeptide, LAP (15 nmol) covalently coupled to carboxylic acid modified polyethylene.

Advances in the development of chemistry for automated C-terminal sequencing now permit for the first time the ability to sequentially degrade all twenty of the common amino acids, an essential first step in the development of a routine automated procedure. The ability to sequence proteins which can be non-covalently applied to solid supports and the adaptability of the chemistry to a wide range of commercially available protein sequenators is anticipated to speed the development of this long awaited procedure.

Acknowledgments

Appreciation is expressed to Dr. John Shively for helpful discussions and to Oanh Tu and Gilbert Issai for technical assistance. This work was supported by NIH grant GM46022.

References

[1] R.G. Krishna, and F. Wold, in K. Imahori and F. Sakiyama (Editors), *Methods in Protein Sequence Analysis*, Plenum, 1993, pp. 167–172.
[2] P. Edman, *Acta Chem. Scand.*, 4 (1950) 283–293.
[3] R.J. Simpson, R.L. Moritz, G.S. Begg, M.R. Rubira and E.C. Nice, *Anal. Biochem.*, 177 (1989) 221–236.
[4] S. Kent, R. Hood, R. Abersold, D. Teplow, L. Smith, V. Farnsworth, P. Cartier, W. Hines, P. Hughes and C. Dodd, *Biotechniques*, 5 (1987) 314–321.
[5] P. Edman and G. Begg, *Eur. J. Biochem.*, 1 (1967) 80–91.
[6] C.L. Zimmermann, E. Appella and J.J. Pisano, *Anal. Biochem.* 77 (1977) 569–573.
[7] G.E. Tarr, J.F. Beecher, M. Bell and D.J. McKean, *Anal. Biochem.*, 84 (1978) 622–627.
[8] D.G. Klapper, C.E. Wilde and J.D. Capra, *Anal. Biochem.*, 85 (1978) 126–131.
[9] B. Wittmann-Liebold, *Hoppe-Seyler's J. Physiol. Chem.*, 354 (1973) 1415–1431.
[10] M.W. Hunkapiller and L.E. Hood, *Biochemistry*, 17 (1978) 2124–2133.
[11] J.E. Shively, *Methods Enzymol.*, 79 (1981) 31–48.
[12] R.M. Hewick, M.W. Hunkapiller, L.E. Hood and W.J. Dreyer, *J. Biol. Chem.*, 256 (1981) 7990–7997.
[13] D.H. Hawke, D.C. Harris and J.E. Shively, *Anal. Biochem.*, 147 (1985) 315–330.
[14] R.J. Simpson, J.A. Smith, R.L. Moritz, M.J. O'Hare, P.S. Rudland, J.R. Morrison, C.J. Lloyd, B. Grego, A.W. Burgess and E.C. Nice, *Eur. J. Biochem.*, 153 (1985) 629–637.
[15] J.E. Shively, P. Miller and M. Ronk, *Anal. Biochem.*, 163 (1987) 517–529.
[16] N.F. Totty, M.D. Waterfield and J.J. Hsuan, *Protein Science*, 1 (1992) 1215–1224.
[17] D.F. Reim, P. Hembach and D.W. Speicher, in R.H. Angeletti (Editor), *Techniques in Protein Chemistry: III*, Academic Press, 1992, pp. 53–60.
[18] V. Farnsworth, W. Carson and H. Krutzsch, *Peptide Research*, 4 (1991) 245–251.

[19] D. Atherton, J. Fernandez, M. DeMott, L. Andrews and S.M. Mische, in R.H. Angeletti (Editor), *Techniques in Protein Chemistry: IV*, Academic Press, 1993, pp. 409–418.
[20] H. Erdjument-Bromage, S. Geromanos, A. Chodera and P. Tempst, in R.H. Angeletti (Editor), *Techniques in Protein Chemistry: IV*, Academic Press, 1993, pp. 419–426.
[21] R.W. Blacher and J.H., Wieser, in R.H. Angeletti (Editor), *Techniques in Protein Chemistry: IV*, Academic Press, 1993, pp. 427–433.
[22] J. Calaycay, M. Rusnak and J.E. Shively, *Anal. Biochem.*, 192 (1991) 23–31.
[23] R.L. Moritz and R.J. Simpson, in K. Imahori and F. Sakiyama (Editors), *Methods in Protein Sequence Analysis*, Plenum, 1993, pp. 3–10.
[24] K.C. Waldron and N.J. Dovichi, *Anal. Chem.*, 64 (1992) 1396–1399.
[25] R.H. Aebersold, D.B. Teplow, L.E. Hood and S.B.H. Kent, *J. Biol. Chem.*, 261 (1986) 4229–4238.
[26] J. Vandekerckhove, G. Bauw, M. Puype, J. Van Damme and M. Van Montagu, *Eur. J. Biochem.*, 152 (1985) 9–19.
[27] P. Matsudaira, *J. Biol. Chem.*, 261 (1987) 10035–10038.
[28] Q. Xu and J.E. Shively, *Anal. Biochem.*, 170 (1988) 19–30.
[29] M. Moos, N.Y. Nguyen and T. Liu, *J. Biol. Chem.*, 263 (1988) 6005–6008.
[30] C. Eckerskorn and F. Lottspeich, *Electrophoresis*, 14 (1993) 831–838.
[31] K. Granlund-Moyer, *Am. Biotech. Lab.*, (1992) 34.
[32] J.M. Bailey, M. Rusnak and J.E. Shively, *Analytical Biochemistry*, 212 (1993) 366–374.
[33] R.A. Laursen, *J. Am. Chem. Soc.*, 88 (1966) 5344–5346.
[34] R.A. Laursen and W. Machleidt, *Methods Biochem. Anal.*, 26 (1980) 201–284.
[35] Liang Song-Ping and R.A. Laursen, *Anal. Biochem.*, 188 (1990) 366–373.
[36] R.H. Aebersold, D.G. Pipes, R.E.H. Wettenhall, H. Nika and L.E. Hood, *Anal. Biochem.*, 187 (1990) 56–65.
[37] D.J.C. Pappin, J.M. Coull and H. Koester, in J.J. Villafranca (Editor), *Current Research in Protein Chemistry*, Academic Press, 1990, pp. 191–202.
[38] J.M. Coull, D.J.C. Pappin, J. Mark, R. Aebersold and H. Koster, *Anal. Biochem.*, 194 (1991) 110–120.
[39] A. Admon and D.S. King, in R.H. Angeletti (Editor), *Techniques in Protein Chemistry: III*, Academic Press, 1992, pp. 45–51.
[40] R. Aebersold, J.D. Watts, H.D. Morrison and E.J. Bures, *Anal. Biochem.*, 199 (1991) 51–60.
[41] R.E.H. Wettenhall, R.H. Aebersold and L.E. Hood, *Method Enzymol.*, 201 (1991) 186–199.
[42] S. Sullivan and T.W. Wong, *Anal. Biochem.*, 197 (1991) 65–68.
[43] J.L. Abernethy, Y. Wang, A.E. Eckhardt and R.L. Hill, in R.H. Angeletti (Editor), *Techniques in Protein Chemistry: III*, Academic Press, 1992, pp. 277–286.
[44] P.J. Neane and F.P. Barry, in R.H. Angeletti (Editor), *Techniques in Protein Chemistry: IV*, Academic Press, 1993, pp. 153–159.
[45] J.Y. Chang, E.H. Creaser and K.W. Bentley, *Biochem. J.*, 153 (1976) 607–611.
[46] J.Y. Chang, D. Brauer and B. Wittmann-Liebold, *FEBS Lett.*, 93 (1978) 205–214.
[47] J. Salnikow, A. Lehmann and B. Wittmann-Liebold, in M. Elzinga (Editor), *Methods in Protein Sequence Analysis*, Humana Press, Clifton, NJ, 1982, pp. 181–188.
[48] R.H. Aebersold, G.D. Pipes, H. Nika, L.E. Hood and S.B. Kent, *Biochemistry*, 27 (1988) 6860–6867.
[49] R.H. Aebersold, H. Nika, G.D. Pipes, R.E.H. Wettenhall, S.M. Clark, L.E. Hood and S.B.H. Kent, in B. Wittmann-Liebold (Editor), *Methods in Protein Sequence Analysis*, Springer-Verlag, Berlin, 1989, pp. 79–97.
[50] H. Maeda and H. Kawauchi, *Biochem. Biophys. Res. Commun.*, 31 (1968) 188–192.
[51] K. Muramoto, H. Kamiya and H. Kawauchi, *Anal. Biochem.*, 141 (1984) 446–450.
[52] H. Hirano and B. Wittmann-Liebold, *Biol. Chem. Hoppe-Seyler*, 367 (1986) 1259–1265.
[53] H. Hirano and B. Wittmann-Liebold, in B. Wittmann-Liebold (Editor), *Methods in Protein Sequence Analysis*, Springer-Verlag, Berlin, 1989, pp. 42–51.
[54] S.W. Jin, G.X. Chen, Z. Palacz and B. Wittmann-Liebold, *FEBS Lett.*, 198 (1986) 150–154.
[55] S.W. Jin, S.Z. Xu, S.L. Zhang and T.B. Tang, in: B. Wittmann-Liebold (Editor), *Methods in Protein Sequence Analysis*, Springer-Verlag, Berling, 1989, pp. 34–41.
[56] J. Salnikow, Z. Palacz and B. Wittmann-Liebold, in K.A. Walsh (Editor), *Methods in Protein Sequence Analysis – 1986*, Humana Press, Clifton, NJ, 1987, pp. 247–260.
[57] G.E. Tarr, in: J.E. Shively (Editor), *Methods of Protein Microcharacterization*, Humana Press, Clifton, NJ, 1986, pp. 155–194.
[58] J.J. L'Italien and S.B.H. Kent, *J. Chromatogr.*, 283 (1984) 149–156.
[59] J.K. Inman and E. Appella, *Methods Enzymol.*, 47 (1977) 374–385.
[60] A. Tsugita, I. Arai, M. Kamo and C.S. Jone, *J. Biochem.*, 103 (1988) 399–401.
[61] M.J. Horn, S.L. Early and S.G. Magil, in T.E. Hugli (Editor), *Techniques in Protein Chemistry*, Academic Press, San Diego, CA, 1989, 51–58.
[62] M.N. Margolies, A. Brauer, C. Oman, D.G. Klapper and M.J. Horn, in M. Elzinga (Editor), *Methods in Protein Sequence Analysis*, Humana Press, Clifton, NJ, 1989, pp. 189–203.
[63] A. Tsugita, M. Kamo, C.S. Jone and N. Shikama, *J. Biochem.*, 106 (1989) 60–65.

[64] M. Pavlik, Z. Voburka, I. Kluh and V. Kostka, *Anal. Biochem.*, 201 (1992) 9–16.
[65] V. Farnsworth and K. Steinberg, *Anal. Biochem.*, 215 (1993) 190–199.
[66] S. Waidyanatha and R.J. Anderegg, *39th ASMS Conference Proceedings*, (1991) 1400–1401.
[67] R. Aebersold, E.J. Bures, M. Namchuk, M.H. Goghari, B. Shushan and T.C. Covey, *Protein Science*, 1 (1992) 494–503.
[68] J.M. Bailey, O. Tu, C. Basic, G. Issai and J.E. Shively, in J.W. Crabb (Editor), *Techniques in Protein Chemistry V*, Academic Press, 1994, pp. 169–178.
[69] M.L. Stolowitz, C.-S. Kim, S.R. Marsh and L. Hood, in K. Imahori and F. Sakiyama (Editors), *Methods in Protein Sequence Analysis*, Plenum, 1993, pp. 37–44.
[70] B. Chait, R. Wang, R.C. Beavis and S.B.H. Kent, *Science*, 262 (1993) 89–92.
[71] P. Schlack and W. Kumpf, *Z. Physiol. Chem.*, 154 (1926) 125–170.
[72] A.S. Inglis, *Anal. Biochem.*, 195 (1991) 183–196.
[73] T.B. Johnson and B.H. Nicolet, *J. Am. Chem. Soc.*, 33 (1911) 1973–1978.
[74] T.B. Johnson and W.M. Scott, *J. Am. Chem. Soc.*, 35 (1913) 1136–1143.
[75] T.B. Johnson, A.J. Hill and B.H. Bailey, *J. Am. Chem. Soc.*, 37 (1915) 2406–2416.
[76] H. Kubo, T. Nakajima and Z. Tamura, *Chem. Pharm. Bull.*, 19 (1971) 210–211.
[77] A.S. Inglis, J.F.K. Wilshire, F. Casagranda and R.L. Laslett, in B. Wittmann-Liebold (Editor), *Methods in Protein Sequence Analysis*, Springer, 1989, pp. 137–144.
[78] A.S. Inglis, R.L. Moritz, G.S. Begg, G.E. Reid, R.J. Simpson, H. Graffunder, L. Matschull and B. Wittmann-Liebold, in H. Jornvall, J.O. Hoog and A.-M. Gustavsson (Editors), *Met. Protein Sequence Analysis*, Birkhauser-Verlag, Basel, 1991, pp. 23–34.
[79] D.H. Hawke, H.-W. Lahm, J.E. Shively and C.W. Todd, *Anal. Biochem.*, 166 (1987) 298–307.
[80] I. Fleming, in D.N. Jones (Editor), *Comprehensive Organic Chemistry*, Vol. 3, Pergamon, 1979, pp. 541–686.
[81] B. Wittmann-Liebold, L. Matschull, U. Pilling, H.-A. Bradaczek and H. Graffunder, in H. Jornvall, J.O. Hoog and A.-M. Gustavsson (Editors), *Met. Protein Sequence Analysis*, Birkhauser-Verlag, Basel, 1991, pp. 9–22.
[82] D.H. Hawke and V.L. Boyd, in J.J. Villafranca (Editor), *Techniques in Protein Chemistry: II*, Academic Press, 1991, pp. 107–114.
[83] N.S. Shenoy, J.M. Bailey and J.E. Shively, *Protein Science*, 1 (1992) 58–67.
[84] J.M. Bailey and J.E. Shively, *Biochemistry*, 29 (1990) 3145–3156.
[85] F.A. Csonka and B.H. Nicolet, *J. Biol. Chem.*, 99 (1933) 213–216.
[86] H.E. Carter and C.M. Stevens, *J. Biol. Chem.*, 133 (1940) 117–128.
[87] M. Goodman and L. Levine, *J. Am. Chem. Soc.*, 86 (1964) 2918–2922.
[88] J.W. Cornforth, *The Chemistry of Penicillin*, Princeton University Press, 1949, pp. 688–848.
[89] H.-A. Bradaczek and B. Wittmann-Liebold, *J. Prot. Chem.*, 11 (1992) 365.
[90] V.L. Boyd, D.H. Hawke and T.G. Gieser, *Tetrahedron Lett.*, 31 (1990) 3849–3852.
[91] V.L. Boyd, M. Bozzini, G. Zon, R.L. Noble and R.J. Mattaliano, *Anal. Biochem.*, 206 (1992) 344–352.
[92] N.R. Shenoy, J.E. Shively and J.M. Bailey, *J. Protein Chemistry*, 12 (1993) 195–205.
[93] J.M. Bailey, F. Nikfarjam, N.S. Shenoy and J.E. Shively, *Protein Science*, 1 (1992) 1622–1633.
[94] G.W. Kenner, H.G. Khorana and R.J. Stedman, *J. Chem. Soc. (London)*, (1953) 673–678.
[95] P. Husek, *J. Chromatogr.*, 552 (1991) 289–299.
[96] T. Shioiri and S.-I. Yamada, *Chem. Pharm. Bull.*, 22 (1974) 855–858.
[97] S.G. Waley and J.J. Watson, *J. Chem. Soc.*, (1951) 2394–2397.
[98] A. Kjaer and P. Eriksen, *Acta Chem. Scand.*, 6 (1952) 448–450.
[99] R.A. Turner and G. Schmerzler, *Biochim. Biophys. Acta.*, 13 (1954) 553–559.
[100] J. Tibbs, *Nature*, 168 (1951) 910.
[101] V.H. Baptist and H.B. Bull, *J. Am. Chem. Soc.*, 75 (1953) 1727–1729.
[102] E. Scoffone and A. Turco, *Ric. Sci.*, 26 (1956) 865–871.
[103] S.W. Fox, T.L. Hurst, J.F. Griffith and O. Underwood, *J. Am. Chem. Soc.*, 77 (1955) 3119–3122.
[104] G.R. Stark, *Biochem.*, 7 (1968) 1796–1807.
[105] L.D. Cromwell and G.R. Stark, *Biochem.*, 8 (1969) 4735–4740.
[106] S. Yamashita, *Biochem. Biophys. Acta.*, 229 (1971) 301–309.
[107] D. Hawke and V.L. Boyd, in H. Jornvall, J.O. Hoog and A.-M. Gustavsson (Editors), *Met. Protein Sequence Analysis*, Birkhauser-Verlag, Basel, 1991, pp. 35–45.
[108] C.G. Miller, C.-T. Kong and J.E. Shively, in T.E. Hugli (Editor), *Techniques in Protein Chemistry*, Academic Press, 1989, pp. 67–78.
[109] J.M. Bailey, N.S. Shenoy, M. Ronk and J.E. Shively, *Protein Science*, 1 (1992) 68–80.
[110] J.M. Bailey and J.E. Shively, in J.J. Villafranca (Editor), *Techniques in Protein Chemistry: II*, Academic Press, 1991, 115–129.
[111] M.J. Williams and B. Kassell, *FEBS Lett.*, 54 (1975) 353–357.
[112] M. Rangarajan and A. Darbre, *Biochem. J.*, 157 (1976) 307–316.
[113] J.L. Meuth, D.E. Harris, F.E. Dwulet, M.L. Crowl-Powers and F.R.N. Gurd, *Biochem.*, 21 (1982) 3750–3757.
[114] S. Yamashita and N. Ishikawa, *Proc. Hoshi. Pharm.*, 13 (1971) 136–138.

[115] T. Suzuki, K.-D. Song, Y. Itagaki and K. Tuzimura, *Org. Mass Spectrometry*, 11 (1976) 557–568.
[116] A.S. Inglis and C. De Luca, in K. Imahori and F. Sakiyama (Editors), *Methods in Protein Sequence Analysis*, Plenum, 1993, pp. 71–78.
[117] J.M. Bailey and J.E. Shively, in J.E. Shively (Editor), *Micromethods for Protein Structural Analysis: A Companion to Methods in Enzymology*, Academic Press, 1994, in press.