Journal of Protein Chemistry, Vol. 20, No. 7, October 2001 (© 2001)

Chemical Carboxyl-Terminal Sequence Analysis of Peptides and Proteins Using Tribenzylsilyl Isothiocyanate

Yu Liu[1] and Songping Liang[1,2]

Received April 30, 2001

A new derivatization reagent, tribenzylsilyl isothiocyanate (TBS-ITC), is applied to C-terminal peptide and protein sequencing. It has been successfully used to sequence six C-terminal residues of house apomyoglobin and a synthetic peptide at low nanomole levels. The chemistry involves activation with acetic anhydride, derivatization with TBS-ITC, and cleavage of derivatized C-terminal amino acid thiohydantoin with sodium hydroxide. The tribenzylsilyl is a bulky, electric donor group and is a good leaving group. It facilitates the nucleophilic attack of the NCS^{-1} in the coupling reaction. The efficiency for C-terminal sequencing by TBS-ITC is about the same as that of acetyl isothiocyanate (AITC), which is a derivatizing reagent for C-terminal sequencing developed by our laboratory. TBS-ITC is much more stable than AITC and trimethylsilyl isothiocyanate (TMS-ITC). TBS-ITC is a solid with relatively long shelf life, whereas AITC and TMS-ITC are liquid and not stable at room temperature.

KEY WORDS: C-terminal sequencing; tribenzylsilyl isothiocyanate (TBS-ITC); thiohydantoin.

1. INTRODUCTION

The past decade has 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 complimentary to the Edman degradation commonly used as a routine approach for N-terminal sequence analysis (Edman, 1950). As an orthogonal method to Edman degradation, C-terminal sequencing also is invaluable for the sequence analysis of proteins with naturally occurring N-terminal blocking groups, for the detection of post-transnational 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 "thiocyanate method," first described in 1926 (Schlack and Kumpf, 1926), has been the most promising due to its similarity to current methods of N-terminal sequence analysis.

The thiocyanate method, on the whole, involves an activation step using acetic anhydride to activate the carboxylic acid at the C-terminal of proteins or peptides to yield oxazolinone (Stark, G. R. 1972), followed by a derivatization step using an isothiocyanate donor reagent to convert the protein or peptide to a peptidylthiohydantoin. The derivatized amino acid is then cleaved by an acid or base to yield a shortened peptide and a thiohydantoin amino acid (THAA).[3] In Edman degradation, the amino group of the protein, owing to its nuclephilicity, attacks the isothiocyanate group of phenylisothiocyanate (PITC) and forms phenylthiohydantoin amino acid after cyclization, cleavage, and conversion. In Schlack-Kumpf degradation, which differs

[1] College of Life Sciences, Hunan Normal University, Changsha, Hunan 410081, People's Republic of China.

[2] To whom correspondence should be addressed. College of Life Sciences, Hunan Normal University, Changsha, Hunan 410081, People's Republic of China; Tel: 011-86-731-8872556; Fax: 011-86-731-8861304; e-mail: liangsp@public.cs.hn.cn

[3] Abbreviations: PITC, phenylisothiocyanate; HSCN, thiocyanic acid; NH_4SCN, ammonium thiocyanate; KSCN, potassium thiocyanate; TMS-ITC, trimethylsilyl isothiocyanate; TBS-ITC, tribenzylsilyl isothiocyanate; TBSn-ITC, tributyltinisothiocyanate; BITC, benzoyl isothiocyanate; DPP-ITC, diphenyl phosphoroisothiocyanatidate; AITC, acetyl isothiocyanate; DITC, 1,4-phenylene diisothiocyanate; DTT, dithiothreitol; TH-AA, amino acid thiohydantoin.

0277-8033/01/1000-0535/0

from Edman degradation, the nitrogen atom of derivatizing reagent, owing to its nucleophilicity, attacks the carboxyl group of the protein or peptide and forms thiohydantoin amino acid after cyclization and cleavage. Obviously, it is very critical to generate peptidyl thiohydantoin effectively so that the nitrogen atom of the isothiocyanate group appears highly nucleophilic as a derivatizing reagent in C-terminal sequencing. Despite the fact that it is an isothiocyanate reagent, PITC cannot be applied for C-terminal sequencing due to its poor nucleophilicity. Originally, Schlack and Kumpf used HSCN as a derivatizing reagent for thiohydantoin formation. Inglis *et al.* (1989, 1991) succeeded in employing HSCN and NH_4SCN for C-terminal sequencing. Bailey *et al.* (1990, 1992a) reported their application of TMS-ITC. Shenoy *et al.* (1993) developed a new derivatizing reagent, TBSn-ITC, for thiohydantooin synthesis. Hawke and Boyd (1995) successfully applied BITC to produced thiohydantoins. DPP-ITC was also proved to be an effective derivatizing reagent for C-terminal (Bailey et al., 1992b, Miller *et al.*, 1995).

In previous studies, we have developed a method using acetyl isothiocyanate (AITC) (Mo *et al.*, 1997) to derivatize the C-terminal amino acid to a thiohydantoin and sodium hydroxide for specific hydrolysis of the derivatized C-terminal amino acid. A major limitation of this approach is that AITC is unstable at room temperature and one needs to make a fresh sample of AITC before performing the C-terminal sequencing. Here, we report a novel derivatizing reagent for C-terminal sequencing of proteins and peptides: tribenzylsilyl isothiocyanate (TBS-ITC). This reagent is in a solid state and is highly stable at room temperature. TBS-ITC has been successfully used for C-terminal sequencing of a synthetic model peptide and horse apomyogloboin at low nanomole levels. Here, we describe the method of the preparation of TBS-ITC and the result of the C-terminal sequencing using this reagent.

2. MATERIALS AND METHODS

2.1. Materials

KSCN, PITC, DTT, horse apomioglobin, and AP glass beads were obtained from Sigma (St. Louis, MO). TMS-ITC, sodium trimethylsilanolate, tribenzylsililyl chloride were purchased from Aldrich (Milwaukee, WI), trifluoroacetic acid (TFA). Acetic anhydride, acetic acid, hydrochloric acid triethylanmine, pyridine, sodium hydroxide, acetontrile, methanol, and dimethylformamide (DMF) were analytical grade. Polyethylene columns for manual C-terminal sequencing were from Bio-Rad (San Francisco, CA). The symmetry RP-HPLC columns used in this research were from Waters (Milford, MA). The model peptide with the sequence of NH_2-KNYEDAL-GFL-COOH was synthesized by solid phase Fmoc chemistry through a manual approach as described in our previous paper (MO *et al.*, 1997).

2.2. Preparation of Tribenzylsilyl Isothiocyanate

TBS-ITC was prepared from tribenzylsilyl chloride and potassium thiocyanate. Potassium thiocyanate (0.24 mol) was dissolved in acetonitrile (500 mL) with continuously stirring. The addition of tribenzylsilyl chloride (0.2 mol) in 500 mL of dichloromethane caused immediate deposition of a white solid (KCl). After intermittent shaking for 1 h at 50°C, the precipitate was removed and the solution was lyophilized to a white powder. Then, the white powder was dissolved in dichloromethane, the precipitate was removed, and the filtrate was evaporated to a white powder under vacuum.

2.3. Covalent Attachment of Peptide and Proteins to DITC Glass Beads

Peptide and protein ware covalently attached to DITC glass beads and quantitated as described by Liang and Laursen (1990), with minor modification. The purified peptide or protein sample (100 nmol) was dissolved in 0.6 mL of coupling buffer (0.2 M Na_2HPO_4, pH 4.6) and added to a test tube containing DITC glass beads (100 mg). After incubating at 55°C in a water bath for 1 h, the glass beads were washed with a coupling buffer (2×3 mL), water (2×3 mL), methanol, (2×3 mL), and acetontrile (2×3 mL), and blown dry with nitrogen. The wash was collected and analyzed by HPLC for determination of coupling yield.

2.4. Preparation of Amino Acid Thiohydantoins

The thiohydantoins derivatives from the 20 free amino acids were prepared by using acetyl chloride as activating reagent and TMS-ITC as derivatizing reagent according to the procedure described by Mo *et al.* (1997). Typically, the amino acid (0.05 mmol) was heated with acetyl chloride (0.2 mL) and acetyl acid (0.03 mL) in the presence of TMS-ITC (0.02 mL) for 30 min at 80°C. The resulting 1-acetyl, 2-thiohydantoin amino acid was evaporated to a yellow gelatinoid under vacuum. Deacetylation

was performed by 12 N HC1(0.1 mL) and stirring at shorted intervals for 3 h at room temperature. The solution was lyophilized to a small amount of yellow powder and then purified by RP-HPLC. The proline thiohydantoin was synthesized by using TFA, rather than acetic acid, to stabilize the intermediate of quaternary ammonium, and the deacetylation procedure was omitted. The preparation of tryptophan thiohydantoin was carried out at 40°C for 1 h with 6N HCl as the cleavage reagent to prevent from possible destruction at high temperature or in the highly concentrated acid. To avoid the unwanted formation of triacetylanhydroarginine, the Arg-acetyl thiohydantoin was produced at room temperature.

2.5. Mass Spectrometry

Electron-impact mass spectrometry was carried out on the Shimadzu QP5000 mass spectrometer operating at 70eV(probe) and at accelerating potential of 8kV. Dried TBS-ITC product or amino acid thiohydantoin were analyzed in direct sampling mode. The instrument was mass measured against perfluorotributylamino (PFTBA) and results were the average of scans across with the mass range m/z 370–40 at interval of 0.5 s.

2.6. Infrared Spectra

Infrared spectra was carried out on the Nicolet Instrument 510P infrared spectrometer system. The sample tablet with dried potassium bromide and the results were the average of scans across with the range 4000 cm^{-1}–400 cm^{-1}.

2.7. Separation of Amino Acid Thiohydantoins by HPLC

Reversed-Phase HPLC was performed with a Waters Symmetry C18 (3.9 × 150 mm) column on a Waters 2010 system, which consists of two 510 pumps, a U6k injector, and a 996-photodiode array detector. The column was eluted with buffer A (0.01% aqueous ammonium acetic, pH 4.6) followed by a discontinuous gradient to buffer B (100% acetronitrile) at 40°C at a flow rate of 0.7 mL/min. The gradient used was as follows: 0% B for 4 min, 0%–16% B over 10 min, 16%–32% B over 3 min, 32% B for 4 min, 32%–45% B over 3 min, and 45%–0% B over 1 min. Both the amino acid thiohydantoins as reference standards and the sequencing products were separated by Waters analysis HPLC system and monitored at 266 nm.

2.8. Manual Procedure for C-Terminal Sequencing with TBS-ITC

Model peptide or protein covalently attached to DITC glass beads (5 mg) was placed into a Bio-Rad polyethylene column with a sieve plate and a bottom cap. Step 1: Acetic anhydride (200 μL) was added, and the column was capped, vortexed, and incubated at 60°C for 10 min. Excess reagents were removed and glass beads were washed repeatedly with methanol and acetonitrile. Step 2: TBS-ITC (200 μL, 0.5 M in 80% acetonitrile and 20% dichloromethane) were added, and the column was capped, vortexed, and incubated at 60°C for 20 min. Excess reagents were removed and glass beads were washed with methanol and acetonitrile. Steps 1 and 2 were repeated. After drawing off all liquid with a vacuum pump, 0.12 N NaOH in 30% methanol containing 0.01% DTT (200 μL) was added and agitated occasionally for 5 min at room temperature. The filtrate was collected and glass beads were washed with 20% acetonitrile (100 μL) and merged into the cleavage solution. The solution was neutralized with 6% in acetic acid 20% acetonitrile (100 μL) and stirred. Finally, the sequencing product (20 μL) was injected for HPLC analysis.

3. RESULTS AND DISCUSSION

3.1. Preparation and Identification of TBS-ITC

Preparation of TBS-ITC proceeded with tribenzylsilyl chloride and potassium thiocyanate (cf. Materials and Methods). In nonpolar solvent, as a double-site-reactive group, SCN^- possesses a soft nucleophilic center, N, and a hard nucleophilic center, S. Tribenzylsillyl chloride is a reactive compound because the covalent bond between Si and Cl is very weak. A soft electrophilic center in tribenzylsilyl group is formed after Cl leaving. According to the soft and hard theory of acids and alkalis, the tribenzylsillyl group prefers reacting with N rather than S. The only product from tribenzylsilyl group and thiocyanate group will be TBS-ITC. The mass spectrum of the product is shown in Fig. 1. The molecular ion peak of TBS-ITC was observed at m/z 359, and fragment peaks of m/z 268, m/z 91, and m/z 58 were intense. The peaks of m/z 268, m/z 209, and m/z 91 were assigned to $[M\text{-}C_7H_7]^+$, $[M\text{-}C_7H_7NCS]^+$ and $[C_7H_7]^+$. The infrared spectrum of product is shown in Fig. 2. The peak of 2067 cm^{-1} was assigned to $-N = C = S$, because only isothiocyanate has weighty absorption peak from 2100 cm^{-1} to 2000 cm^{-1} in infrared spectrum. The mass spectrum was the key for identifying the structure of TBS-ITC together with infrared spectrum.

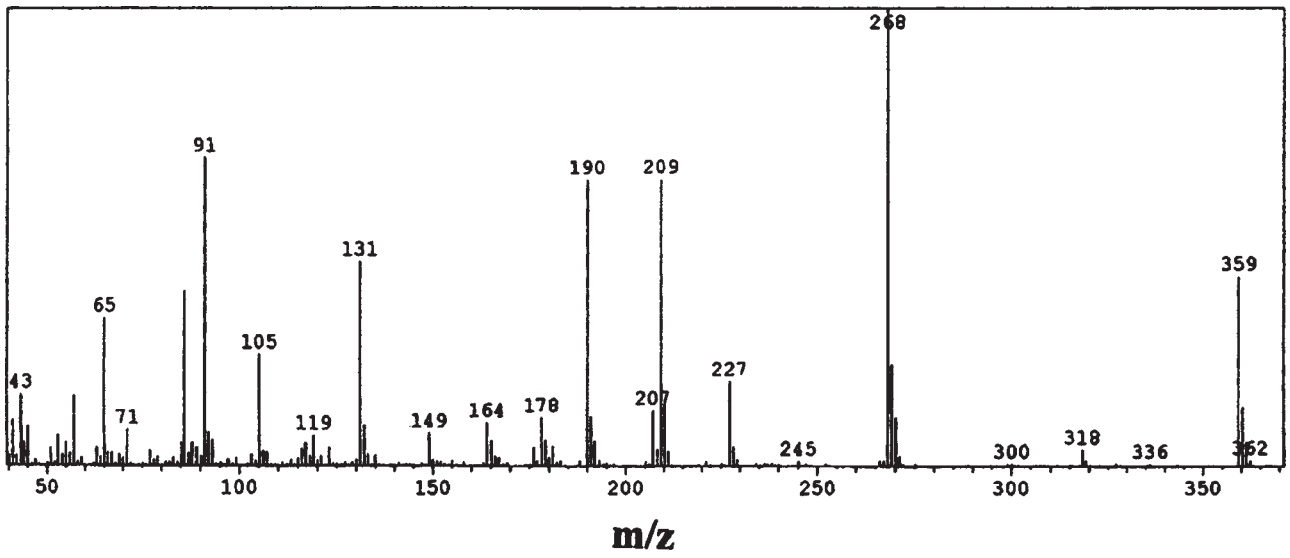

Fig. 1. Mass spectrum of TBS-ITC. Mass spectrometry was carried out on the Shimadzu QP5000 mass spectrometer operating at 70eV (probe) and at accelerating potential of 8kV. The molecular ion peak of TBS-ITC was observed at m/z 359. The peaks of m/z 268, m/z 209, and m/z 91 were assigned to $[M\text{-}C_7H_7]^+$, $[M\text{-}C_7H_7NCS]^+$, and $[C_7H_7]^+$.

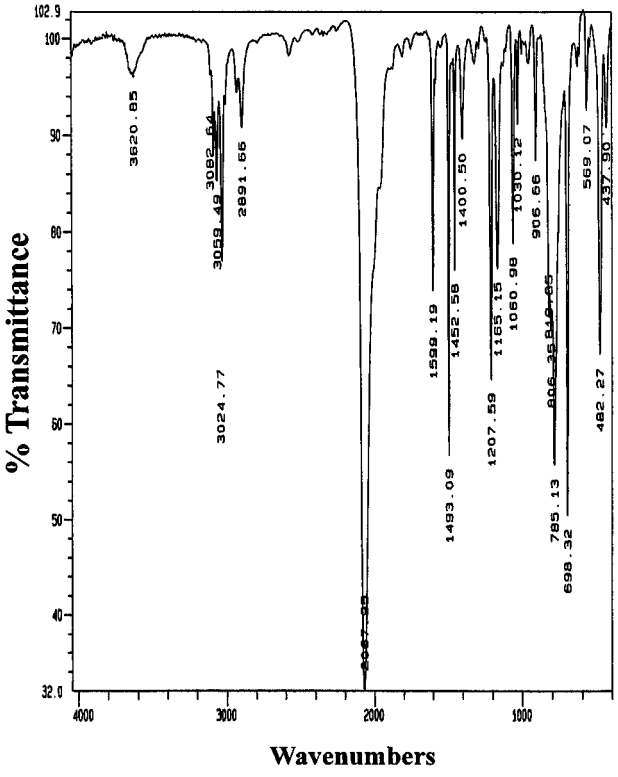

Fig. 2. Infrared spectra of TBS-ITC. Infrared spectra was carried out on the Nicolet Instrument 510P infrared spectrometer system. The peak of 2067cm^{-1} was assigned to $-N = C = S$.

3.2. Preparation and Identification of Amino Acid Thiohydantoins

Amino acid thiohydantoins are required as reference standards for the development of C-terminal protein sequencing based on thiohydantoin chemistry. The carboxyl group of a free amino acid must be activated by a suitable reagent before derivatizing to isothiocyanate ester, which is necessary to form amino acid thiohydantoin. To avoid a side reaction, the amino group of the amino acid must be protected because that amino group is more active than the carboxyl group. Therefore, it is essential to perform activation of the carboxyl group and protection of the amino group before reaction with the derivatizing reagent. An ideal activating reagent should be able to achieve the two goals simultaneously. Acetic anhydride has been used as an activating reagent in the classical method. To activate the carboxyl group and protect the amino group more effectively, we selected acetyl chloride as the activating reagent because acetyl chloride has much higher reactivity than acetic anhydride. In our procedure, the free amino acids, rather than acetyl amino acids, were used because acetyl chloride can acetylate the amino group quickly and can greatly reduce the possibility of a side reaction.

The extinction coefficients of thiohydantoins are approximately 18,000 in ethanol at 266 nm; hence they may be detected readily at the low picomole levels in HPLC

with microbore HPLC columns. The thiohydantoins of the amino acids Asp, Asn, Glu, Gln, Gly, and Ser are not strongly retained on the normal C18 columns in the presence of organic solvents. To achieve the elution of the polar thiohydantoins in a reasonable time period, it was necessary a to add a small amount of ammonium acetic to buffer A. The column was eluted for 4 min with buffer A and then flowed by a discontinuous gradient to buffer B. For each amino acid thiohydantoin, a single peak with a characteristic retention time was observed by reversed-phase HPLC. The separation chromatogram of all 20 amino acid thiohydantoins together is shown in Fig. 3.

3.4. Manual Procedure for C-terminal Sequencing with TBS-ITC

Fig. 4 shows the first six cycles of model peptide (5 nmol) from the C-terminus sequenced by using TBS-ITC as the derivatizing reagent, acetic anhydride as the activation reagent, and 0.12N NaOH in 30% methanol containing 0.01% DTT as the cleavage reagent. The initial yield and the repetitive yields were 62.5% and 72.1%, respectively, which indicated the potential of TBS-ITC chemistry. The peptide was covalently attached to DITC glass beads before the C-terminal sequencing with TBS-ITC. The peptidyl oxazolinone was formed with acetic anhydride at 60°C for 20 min. The peptidyl thiohydantoin was formed by the treatment of the covalently attached peptide with TBS-ITC at 60°C for 20 min. Cleavage of the derivatized C-terminal amino acid was accomplished with 0.12N NaOH in 30% methanol, containing 0.01% DTT, for 5 min. Comparison of results at different temperatures indicated that 60°C is preferable. The times for activation and derivatization steps are usually about 10 min and 20 min, respectively. The selection of cleavage reagents was optimized by employing 0.12N NaOH in 30% methanol, 0.16N NaOH in 30% methanol, 1% aqueous piperidine, 6N HCl, and 50% TFA in water. The most efficient cleavage was achieved with 0.12N NaOH in 30% methanol at room temperature for 5 min. The limitation of cleavage reaction with alkali was the low stability of amino acid thiohydantoins. Thiol reagents were reported as antioxidant for cleavage. It was effective to stabilize amino acid thiohydantoins by addition of DTT. Fig. 5 shows the first six cycles of C-terminal sequencing of horse apomyoglobin (5 nmol) by TBS-ITC chemistry. The amino acid thiohydantoin signals of the first five cycles are clean and unambiguous. There is a large decrease in the repetitive yield at the sixth cycle, in which the thiohydantoin of glutamic acid recovered was very low.

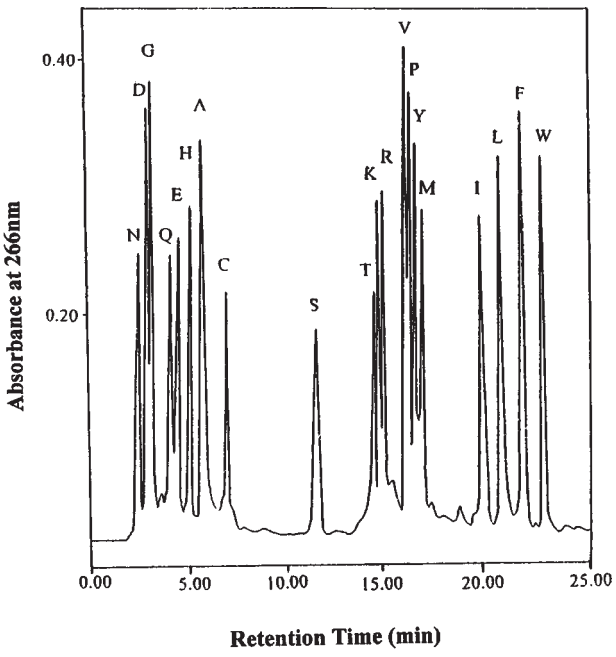

Fig. 3. HPLC separation of aminoacid thiohydantoins. Reversed-Phase HPLC was performed with a Waters Symmetry C18 (3.9 × 150 mm) column. The column was eluted with buffer A (0.01% aqueous ammonium acetic, pH 4.6) followed by a discontinuous gradient to buffer B (100% acetronitrile) at 40°C at a flow rate of 0.7 mL/min. The gradient used was as follows: 0% B for 4 min, 0%–16% B over 10 min, 16%–32% B over 3 min, 32% B for 4 min, 32%–45% B over 3 min, and 45%–0% B over 1 min.

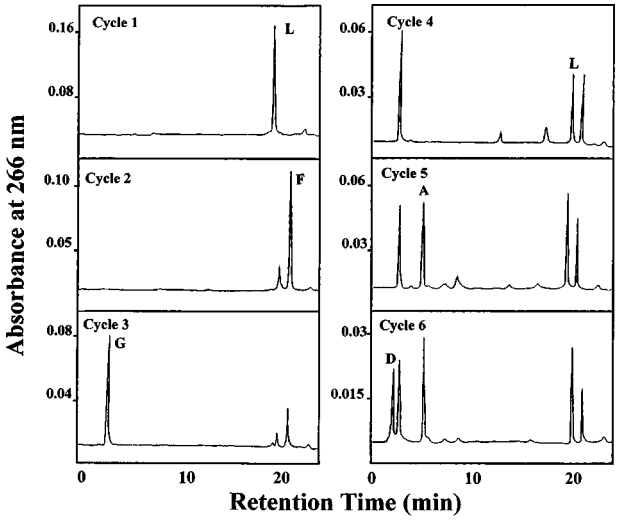

Fig. 4. Manual C-terminal sequencing of the model peptide covalently attached to DITC glass beads using TBS-ITC.

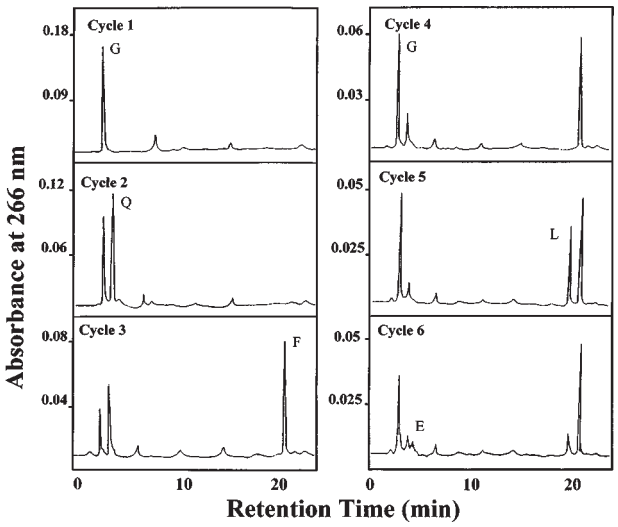

Fig. 5. Manual C-terminal sequencing of horse apomyogloboin (5 nmol) covalently attached to DITC glass beads using TBS-ITC.

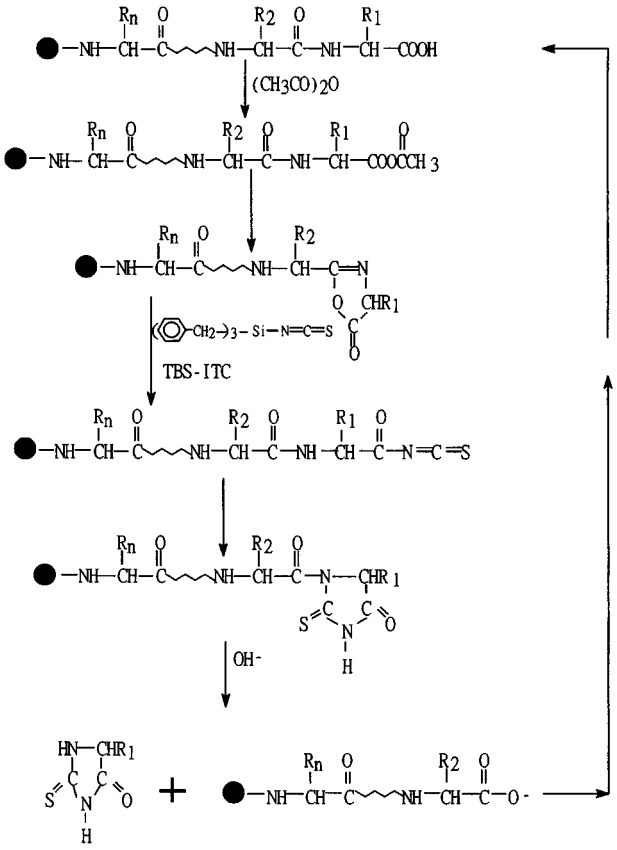

Fig. 6. Postulated reaction scheme of derivatization for C-terminal sequencing using TBS-ITC chemistry.

The chemical mechanism of thiohydantoin with TBS-ITC is outlined in Fig. 6. Each cycle of degradation consists of three steps: activation of the C-terminal carboxylic group, derivatization to a thiohydantoin amino acid, and the hydrolysis of the derivatized C-terminal amino acid to an amino acid thiohydantoin with a shorted polypeptide or protein. In classical thiohydantoin chemistry, acetic anhydride is necessary to form the requisite oxazolinone for further reaction. In a more recent paper, Boyd *et al.,* (1995) studied the activation reaction with nuclear magnetic resonance spectroscopy and found that oxazolinone was the only valid product to form thiohydantoin. In TBS-ITC chemistry, the tribenzylsilyl is a bulky, electric donor group and is a good leaving group. It facilitates the nucleophilic attack of the NCS^{-1} in the coupling reaction. The efficiency for C-terminal sequencing by TBS-ITC nearly similar to that of acetyl isothiocyanate (AITC), which is a derivatiziing reagent for C-terminal sequencing developed by our lab. TBS-ITC is much more stable than AITC and TMS-ITC. TBS-ITC is a solid at room temperature, with a relatively long shelf life, whereas AITC and TMS-ITC are liquid and are not stable at room temperature.

The current limitations of this method include a limited number of cycles due to the relatively low repetitive yield, and proline residue was not found to be capable of derivatization. Application of this chemistry to the solid phase and to automation has the potential to permit extended sequential degradation from the C-terminus of peptides and proteins with minimal sample loss. Advantages of the solid phase approach include the ability to use reagents and solvents optimal for sequencing without causing sample washout as well as the capability to wash the sample efficiently to remove reaction by-products that might otherwise interfere with identification of the released thiohydantoin amino acids.

Our method offers a possibility of a stable and routine approach for C-terminal sequencing of peptides and proteins. Work is in progress on the optimization of reaction condition and instrumental automation of this chemistry.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation of China, under contract No. 39990600, and a grant from the United Research Centre of Life Science of Hunan Province.

REFERENCES

Bailey, J. M. and Shively, J. E. (1990). *Biochemistry* **29**, 3145–3156.

Bailey, J. M., Shenoy, N. R., Ronk, M., and Shively, J. E. (1992a). *Protein Sci.* **1**, 68–80.

Bailey, J. M., Nikfarjam, F., Shenoy, N. R., and Shively, J. E. (1992b). *Protein Sci.* **1**, 1622–1633.

Bilian, M. O., Li, J., and Liang, S. P. (1997). *Anal. Biochem.* **252**, 169–176.

Boyd, V. L., Bozzini, M. L., Guga, P. J., Defranco, R. J., Yuan, P. M., Loudon, G. M., and Nguyen, D. (1995). *J. Org. Chem.* **60**, 2581–2587.

Edman, P. (1950). *Acta Chem. Scand.* **4**, 283–293.

Hawke, D. H. and Boyd, V. L. (1991). In *Methods in Protein Sequence Analysis* (Jörnval, H., Hoog, J.-O., and Gustavsson, A.-M., Eds.), Birkhausen Verlag, Basel, pp. 35–45.

Inglis, A. S., Wilshire, J. F. K., Casagranda, F., and Laslett, R. L. (1989). In *Methods in Protein Sequence Analysis* (Wittmann-Liebold, B., Ed.), Springer Verlag, Berlin, pp. 137–144.

Inglis, A. S., Moritz, R. L., Begg, G. S., Reid, G. E., Simpson, R. J., Graffunder, H., Matschun, L., and Wittmann-Liebold, B. (1991). In *Methods in Protein Sequence Analysis* (Jörnval, H., Hoog, J.-O., and Gustavsson, A.-M., Eds.), Brikhausen Verlag, Basel, pp. 23–24.

Liang, S. P. and Laursen, R. A. (1990). *Anal. Biochem*, **188**, 366–373.

Miller, C. G., Bailey, J. M., Hawke, D. H., Early, S., and Tso, J. (1995). In *Methods in Protein Structure Analysis* (Atassi, M. Z., and Appella, E., Eds.), Plenum Press, New York, pp. 119–129.

Mo, B. L., Li, J., and Liang, S. P. (1997). *Anal. Biochem.* **249**, 207–211.

Schlack, P. and Kumpf, W. (1926). *Z. Physiol. Chem.* **154**, 125–170.

Shenoy, N. R., Shively, J. E., and Bailey, J. M. (1993). *J. Protein Chem.* **12**, 195–205.

Stark, G. R. (1972). In *Methods in Enzymology* Vol. 25 (Hirs, C. H. W., and Timasheff, S. N., Eds.), Academic Press, San Diego, CA, pp. 369–384.