Journal of Chromatography, 314 (1984) 467–470
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 17,191

Note

Identification of phenylthiohydantoin derivatives of substituted amino acids encountered during amino acid sequence analysis by high-performance liquid chromatography*

AJIT S. BHOWN* and J. CLAUDE BENNETT
Division of Clinical Immunology and Rheumatology, Department of Medicine, University of Alabama in Birmingham, Birmingham, AL 35294 (U.S.A.)
(First received June 17th, 1984; revised manuscript received September 5th, 1984)

Mounting evidence indicates that methylated amino acids are the constituents of naturally occurring proteins and that they provide a new molecule for structure–function hypotheses[1]. These derivatives of amino acids, although capable of undergoing the Edman degradation to form their anilinothiazolinone (ATZ) and phenylthiohydantoin (PTH) derivatives, are difficult to identify by usual methods of PTH amino acid identification. Chen *et al.*[2] could establish the presence of N-α-monomethylalanine at the amino terminus of ribosomal protein S11 by mass spectrometry and amino acid analysis only. Although numerous methods have been developed to identify the PTH derivatives of unsubstituted amino acids by reversed-phase high-performance liquid chromatography (RP-HPLC) no such method is available for the identification of PTH derivatives of monomethylated amino acids. We report a RP-HPLC method to separate and identify the PTH derivatives of monomethylated amino acids obtained during amino acid sequence analysis. SAR-angiotensin-II has been employed as a model peptide.

MATERIALS AND METHODS

N-Monomethyl derivatives of tryptophan, glutamic acid, phenylalanine, aspartic acid, alanine, leucine, lysine, valine, and glycine, α-methylated methionine, serine and histidine and O-methylated tyrosine and threonine were obtained from Pierce Chemical (Rockford, IL, U.S.A.) and/or Sigma (St. Louis, MO, U.S.A.). SAR-angiotensin-II was purchased from Vega Chemical (Tucson, AZ, U.S.A.). Sequencer chemicals were supplied by the Beckman Instruments (Palo Alto, CA, U.S.A.). Methanol and the reversed-phase column for HPLC were the products of MCB (Cincinnati, OH, U.S.A.) and IBM (Wallingford, CT, U.S.A.), respectively.

A modified[3] Beckman 890C automated sequencer was employed for preparing the standard PTH derivatives of monomethylated amino acids and for sequencing the model peptide SAR-angiotensin-II. Perkin-Elmer (Norwalk, CO, U.S.A.) and

* Presented at the *5th Internaional Conference on Methods of Protein Sequence Analysis, Cambridge, U.K., 1984.*

0021-9673/84/$03.00

Waters (Milford, MA, U.S.A.) HPLC systems were used for the identification of released amino acids.

Monomethylated amino acids were dissolved in 0.1 *M* hydrochloric acid (10 mg/ml). A 400-μl volume of this solution was added to the sequencer cup, the sample was dried by using the sample application subroutine (SAS) program (Beckman; part No. 02772) and subjected to one cycle of Edman degradation as described earlier[4], except that polybrene was omitted.

The ATZ derivative was converted to its PTH form by heating with 1.0 *M* hydrochloric acid containing 15 mg/l of dithiothreitol for 10 min at 80°C under nitrogen. The PTH derivative was extracted twice with 1.0 ml of ethyl acetate, the organic layer was separated, pooled and dried under nitrogen. The dried residue was then dissolved in a desired volume of initial buffer of HPLC for identification and quantitation. SAR-angiotensin-II was sequenced exactly as described above except that 2 mg of polybrene was included for each run.

Analyses of the PTH derivatives of monomethylated and unsubstituted amino acids (obtained from the model peptide) were performed on Perkin-Elmer Series 4 microprocessor-controlled solvent delivery system and Waters HPLC system equipped with two 6000A solvent delivery pumps, WISP and a model 440 dual-channel absorbance detector. An IBM 5-μm ODS column (150 × 4.5 mm; part No. 8635328) and an Altex Ultrasphere ODS 5-μm column (150 × 4.5 mm; part No. 235330) obtained from Beckman Instruments were employed for separation of the PTH derivatives of the substituted and unsubstituted amino acids respectively. Solvents A annd B were constituted as follows: solvent A, 0.04 *M* sodium acetate (pH 3.72) containing 50 μl acetone per liter and filtered through a 0.2-μm filter; solvent B, 100% methanol (OmniSolv, MCB) containing 250 μl of glacial acetic acid per liter. The methanol was prefiltered by the manufacturer. For the separation and identification of the PTH derivatives of substituted amino acids the IBM column was developed for 2 min at initial condition (68% A + 32% B) after injection followed by a 10-min linear gradient to final condition (48% A + 52% B) and held at final condition for 8 min. The flow-rate was 1.5 ml/min. The Altex column was employed as described earlier[4] to identify the PTH derivatives of the unsubstituted amino acids obtained after sequential degradation of SAR-angiotensin-II.

RESULTS AND DISCUSSION

Fig. 1A shows the separation of the PTH derivatives of thirteen monomethylated amino acids using the program described above. Although a baseline separation is achieved within 20 min, some spurious peaks are observed, which are due to the impourities present in the substituted amino acids. The most noticeable is in aspartic acid which always appears as a doublet.

The first cycle from Edman degradation of SAR-angiotensin-II was analysed on the IBM and the Altex columns and the results are shown in Figs. 2A and 2B, respectively. HPLC analysis on the IBM column indicated that residue at cycle 1 is N-monomethylglycine (sarcosine). On the other hand, analysis by HPLC, on the Altex column, indicated the presence of alanine residue at cycle 1. The correct residue is N-monomethylglycine (sarcosine, hence the name SAR-angiotensin-II) and not alanine. The remaining cycles (not shown) from this model peptide when analyzed

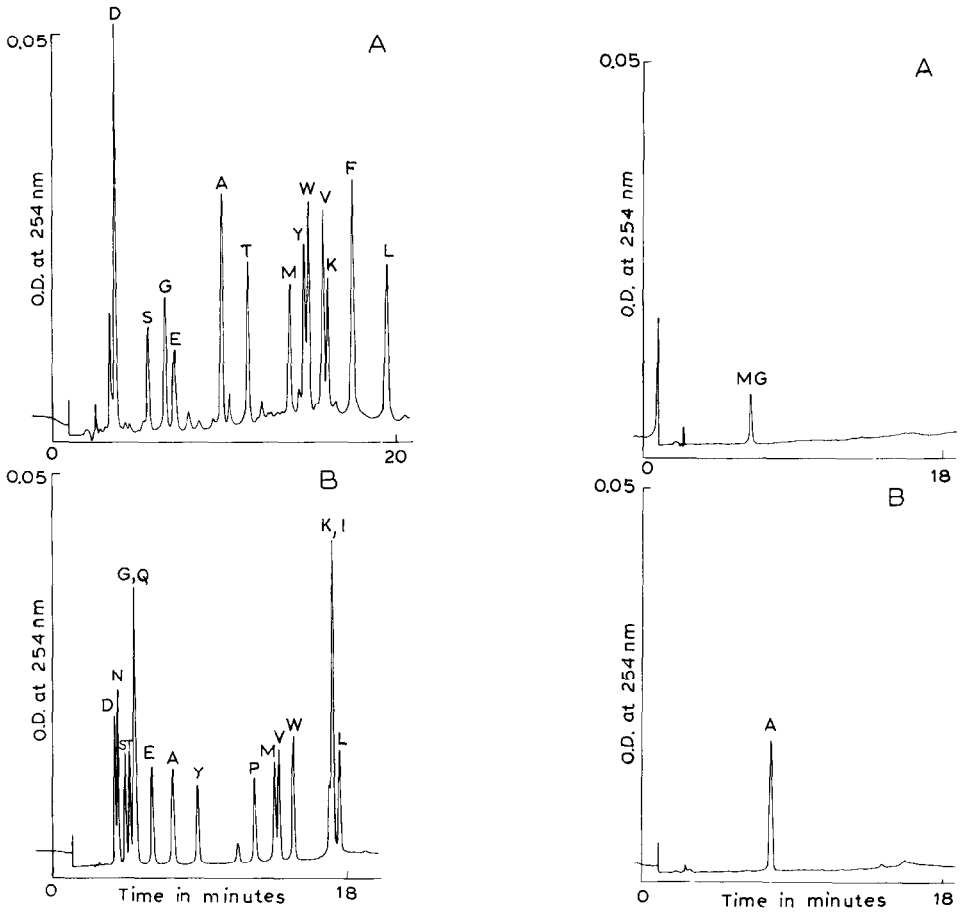

Fig. 1. Separation of standard PTH-derivatives of substituted (A) and unsubstituted (B) amino acids on the IBM column. Abbreviations: D = aspartic acid, S = serine, G = glycine, E = glutamic acid, A = alanine, T = threonine, M = methionine, Y = tyrosine, W = tryptophan, V = valine, K = lysine, F = phenylalanine, L = leucine, Q = glutamine, I = isoleucine, W = tryptophan, P = proline, N = asparagine. See text for details.

Fig. 2. HPLC tracing of the first cycle of SAR-angiotensin-II, with an IBM column (A) and an Altex column (B). For details see text. Abbreviations: MG = methylated glycine, A = alanine.

on the IBM column failed to show peaks corresponding to any of the standard peaks on this (IBM) column. The complete sequence of SAR-angiotensin-II was established when the sequencer cycles (other than the first) were analyzed on the Altex column.

Fig. 1B shows the separation of the PTH derivatives of unsubstituted amino acids on the IBM column under the conditions used for separating PTH derivatives of the substituted amino acids. A comparison of the retention times of PTH derivatives of substituted and unsubstituted amino acids clearly indicates that, as expected, the former being more hydropohobic elute later than their corresponding unsubstituted forms (Fig. 1B). Although the order in which these PTH derivatives elute is

unchanged, some of the methylated forms exhibit an altered elution pattern (Fig. 1B). These are threonine, tyrosine and valine which are eluted late and methionine which are eluted early with respect to tyrosine. This probably indicates that an increase in hydrophobicity, by introducing a methyl group on the molecule, is not an additive but a configurational effect. It is interesting to note that N-monomethylglycine (the amino terminal amino acid of SAR-angiotensin-II) elutes as PTH-alanine on the Altex column. This is not unusual because of the structural similarity between N-monomethylglycine and alanine. During the course of this investigation it has been noticed that PTH-derivatives of monomethylated amino acids degrade at a faster rate than their corresponding non-methylated forms and, furthermore, we could never obtain a stable derivative of histidine.

These results indicate that, at times, a residue may be misinterpreted, particularly those which have similar molecular structure *e.g.* N-methylglycine and alanine. Furthermore, because of the evidence indicating the presence of such substituted amino acids at the amino terminus of naturally occurring proteins it is recommended that the sequencer cycles be closely monitored for the possible presence of such substituted amino acids especially when no residue could be identified because of the absence of a peak or the presence of an unidentifiable peak. Finally, since the coupling efficiency of monomethylated amino acids used in this study is unknown, it was difficult to quantitate the results. However, it has been observed that the peak height response has a linear relationship with the amount injected on the HPLC column. This indicates the efficacy of HPLC identification but not of the Edman chemistry.

ACKNOWLEDGEMENTS

We thank Karen Buttler and Lisa Kallman for their expert technical assistance as well as to Brenda Gilmore for typing this manuscript. The work was supported by Grants AM-03555 and AM-20614.

REFERENCES

1 G. C. Barrett, *Amino acids, Peptides and Proteins*, Vol. 10, Chemical Society, London, 1979, p. 3.
2 R. Chen, J. Brosius, B. Wittmann-Leibold and W. Schafer, *J. Mol. Biol.*, 111 (1977) 173.
3 A. S. Bhown, T. W. Cornelius, J. E. Mole, J. D. Lynn, W. A. Tidwell and J. C. Bennett, *Anal. Biochem.*, 102 (1980) 35.
4 A. S. Bhown, J. E. Mole and J. C. Bennett, *Anal. Biochem.*, 110 (1981) 355.