Amino Acids (1997) 13: 13–23

Amino
Acids

Identification of post-translational modified amino acids

P. Tsiboli, G. Konstantinidis, Y. Skendros, A. Katsani, and **T. Choli-Papadopoulou**

School of Chemistry, Laboratory of Biochemistry,
Aristotle University of Thessaloniki, Greece

Accepted January 5, 1997

Summary. Methylated lysines (N^{ε}-mono-methylated, N^{ε}-di-methylated and N^{ε}-tri-methylated) have been identified after derivatization with ortho-phthaldialdehyde (OPA) by using pre-column and post-column derivatization techniques.

Also the N^{ε}-acetylated lysine and N^{ε}-formylated lysine have been identified by OPA post-column derivatization techniques but only in free form due to their instability under acidic conditions which are used for protein hydrolysis.

Additionally, all the modified amino acids mentioned above have been derivatized with DABITC/PITC, an Edman reagent, and identified as DABTH-derivatives on thin-layer polyamide sheets.

Keywords: Amino acids – Modified amino acids – HPLC – Edman reagent

Introduction

It has become evident in recent years that the majority of proteins are covalently modified during their cellular lifetime and that such modifications can influence their activity, localisation, assembly or regulation (Wold, 1981). The covalently modified amino acid residues are not necessarily recognized even during analysis of complete amino acid sequences since the majority of the naturally occuring derivatives are acid-labile and thus undetected by amino acid analysis after conventional acid hydrolysis. No generally applicable method for detection can be described, but the different modifications can be detected by using different strategies.

The acetylation of amino acids within the polypeptide chain, as distinct from acetylation of free NH_2-terminal groups, was first detected in cell nuclei as a rapid and reversible incorporation of radioactively labelled acetate into histones H3 and H4 (Allfrey et al., 1964). Acetylation of lysine α-amino groups is not limited to histones, but takes place also on other DNA-binding proteins such as the High Mobility Group (HMG) proteins (Sterner et al., 1979, 1981).

Modification of the α-amino group of a protein or of a fragment usually precludes Edman degradation. The two most common blocks of this nature are N^{α}-acetylation which can be prevented during in vitro translation (Palmiter, 1977); and aminoterminal glutamine which can sometimes be removed enzymically (Doolittle, 1977).

Protein methylation is another post-translational modification reaction which occurs in many proteins. This reaction involves side chains of various amino acids, such as lysine, arginine, histidine and glutamine for the N-methylation and glutamic and aspartic acid for O-methylesterification.

Methylated amino acids have been reported to be present in many proteins as histones or ribosomal proteins. Methylation of side-chain amino groups of lysine are common modifications.

Occasionally N^{α}-methylation has been observed (Hermodson et al., 1978), but in most cases this does not interfere with the Edman degradation. Several methylated amino acids have been identified, as 3-methyl-histidine (3-MeHis), N^{ε}-monomethyl-lysine (MeLys), N^{ε}-dimethyl-lysine (Me_2Lys) and N^{ε}-trimethyl-lysine (Me_3Lys) (Huszar and Elzinga, 1969; Kuehl and Adelstein, 1969; Trayer et al., 1968; Paik and Kim, 1971; Becam and Lederer, 1981; Delange et al., 1969; Hempel et al., 1968; Kakimoto and Akazawa, 1970).

The N^{G}, N^{G}-dimethyl-arginine residue (asymmetric dimethylarginine) has been found near the COOH-terminus of the UP1 calf thymus helix-destabilizing protein (Williams et al., 1985) as well as in several histones (Paik and Kim, 1967), in the High Mobility Group proteins (Boffa et al., 1979), in ribosomal proteins (Chang et al., 1976); and in proteins associated with heterogeneous RNA (hnRNP particles) (Boffa et al., 1977; Christensen et al., 1977; Marvil et al., 1980).

Materials and methods

Materials

The monomethylated, dimethylated and trimethylated lysine derivatives were from Calbiochem and the acetylated- and formylated-lysines were from SERVA (Heidelberg, F. R. G.).

The 4-N, N′-dimethylaminoazobenzene-4′-isothiocyanate (DABITC) was from Fluka (Buchs, Switzerland) or Pierce (Rockford, U.S.A) recrystallized from boiling acetone (Pro analysis grade, dried over molecular sieve): 1 gr dissolved in 70 ml, passed through a paper filter and allowed to cool slowly yield 0.7 g of brown needles, b.p. 169°–170°C. Trifluoroacetic acid (TFA) and phenylisothiocyanate (PITC) are necessary to be sequencing grade. Pyridine, was pro analysis grade from MERCK (Sharp and Dohme, F. R. G) distilled successively over KOH pellets, ninhydrin and KOH, b.p. 114°–116°. The water was twice glass distilled and freshly prepared. All the other chemicals were sequencing grade (n-heptane and ethylacetate) or pro analysis grade (n-butylacetate, acetic acid, toluene, n-hexane and acetone) and used without further purification.

Amino acid analysis

All glasswear used in acid hydrolysis were heated at 500°C for 6 hrs. Approximately 0.1–2.0 μg of protein was placed in a hydrolysis tube (5 mm i.d $\times$ 100 mm) and

dried under vacuum. 50 μl of 6N HCl were added and the tube was sealed under vacuum.

Hydrolysis of the L11 ribosomal protein from *E. coli* was performed for 24 hours at 110°C. The amino acids were determined, both with pre-column and post-column derivatization with o-phthaldialdehyde (OPA), by RP-HPLC and ion-exchange chromatography, respectively.

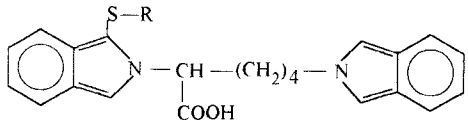
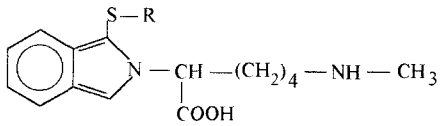
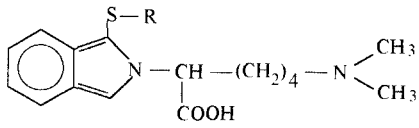
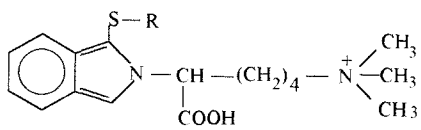
DABITC-derivatization

All N^{ε}-modified lysines are derivatized with DABITC/PITC as described in Chang et al. (1980) to ensure complete reaction of all terminal amino groups. In combination with the identification of the released derivatives by 2-dimensional thin-layer chromatography on small polyamide sheets the method allows the degradation of as little as 500 pmol of polypeptide. The detection limit of the derivatives on the sheets is 20 pmol. Standard DABTH-amino acids were kindly given from Dr. J. Y. Chang.

Results and discussion

Figure 1 shows the separation of the modified lysines, namely N^{ε}-monomethylated (K*), N^{ε}-dimethylated (K**) and N^{ε}-trimethylated (K***) by using the pre-column derivatization technique. As it is shown on the chromatogram all these derivatives are well separated. The elution of the lysine and its methylated derivatives is in agreement with their hydrophobicity after OPA-derivatization (see Table 1) namely K*** < K* < K** < K. The

Table 1. OPA-lysine derivatives (the ortho-phthaldialdehyde reacted with the lysines in the presence of a reducing agent, β-mercaptoethanol, at pH = 10)
R = CH_2—CH_2—OH

	OPA-lysine K
	OPA-N^{ε}-monomethylated lysine K*
	OPA-N^{ε}-dimethylated lysine K**
	OPA-N^{ε}-trimethylated lysine K***

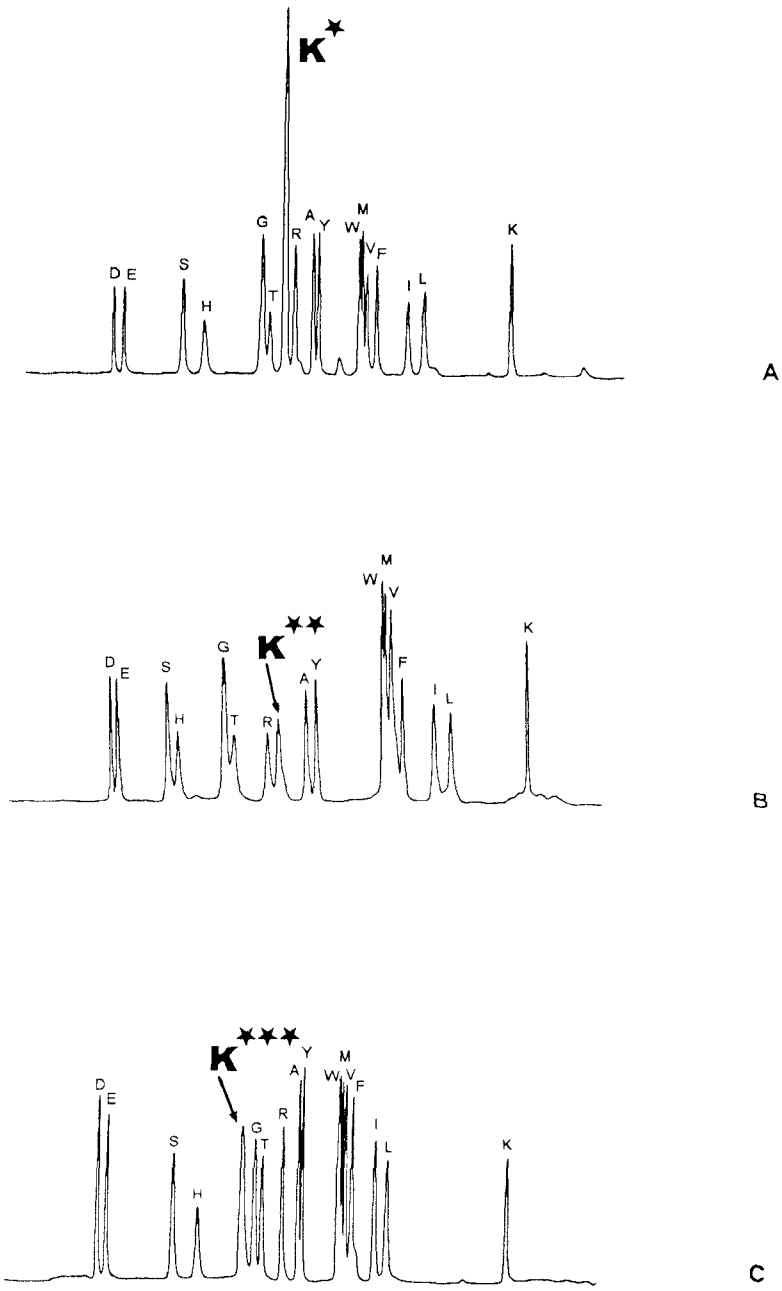


Fig. 1. Separation of 100 pmol of a reference amino acid mixture containing N^{ε}-monomethylated lysine (**A**), N^{ε}-dimethylated lysine (**B**) and N^{ε}-trimethylated-lysine (**C**), after ortho-phthaldialdehyde pre-column derivatization, by reversed-phase HPLC, using a column (250 × 4 mm) filled with Shandon Hypersil ODS 5 μm material. Buffer A was 12.5 mM Na_2HPO_4 (pH = 7.2) and Buffer B was 3% tetrahydrofuran in methanol (Ashman et al., 1985)

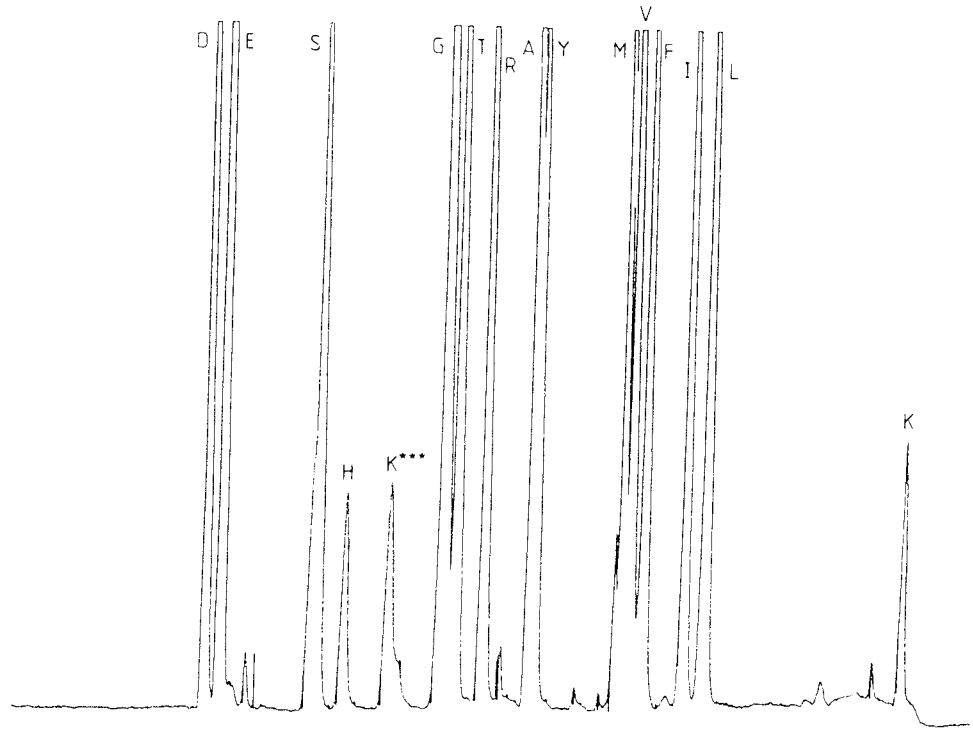


Fig. 2. Amino acid composition of protein L11 prm-1 *E. coli*, after subjection to total hydrolysis as described in Materials and methods. The separation of the amino acids was the same as described in Fig. 1. The characteristic peak for N^{ε}-trimethylated lysine (*K****) elutes at the same position as in the chromatogram of Fig. 1C

N^{ε}-trimethylated OPA-lysine due to its positive charge elutes between OPA-histidine and OPA-glycine while the N^{ε}-monomethylated and N^{ε}-dimethylated OPA derivatives elute between OPA-threonine and OPA-arginine and between OPA-arginine and OPA-alanine, respectively. All lysine derivatives mentioned above have been tested for their stability under the acidic conditions needed for the hydrolysis of proteins and peptides. All of them were stable and eluted at the same positions as the non-hydrolysed derivatives.

Figure 2 shows the amino acid analysis of the L11 ribosomal protein from prm-1 *E.coli*. The peak which appears between the OPA-histidine and OPA-glycine corresponds to the OPA-trimethylated lysine which has been found by Dognin and Wittmann-Liebold (1977) at positions 3 and 39. The amino acids, which are released after total hydrolysis (see Materials and methods), are separated as in Fig. 1, namely by the pre-column derivatization technique.

On Fig. 3 shown is the elution of the methylated OPA-lysine derivatives using the post-column derivatization technique. Only the N^{ε}-trimethylated lysine derivative (K***), (C) is well separated from the other amino acid derivatives, while the N^{ε}-monomethylated and N^{ε}-dimethylated lysine derivatives co-elute with the non-methylated lysine derivative (3A and 3B,

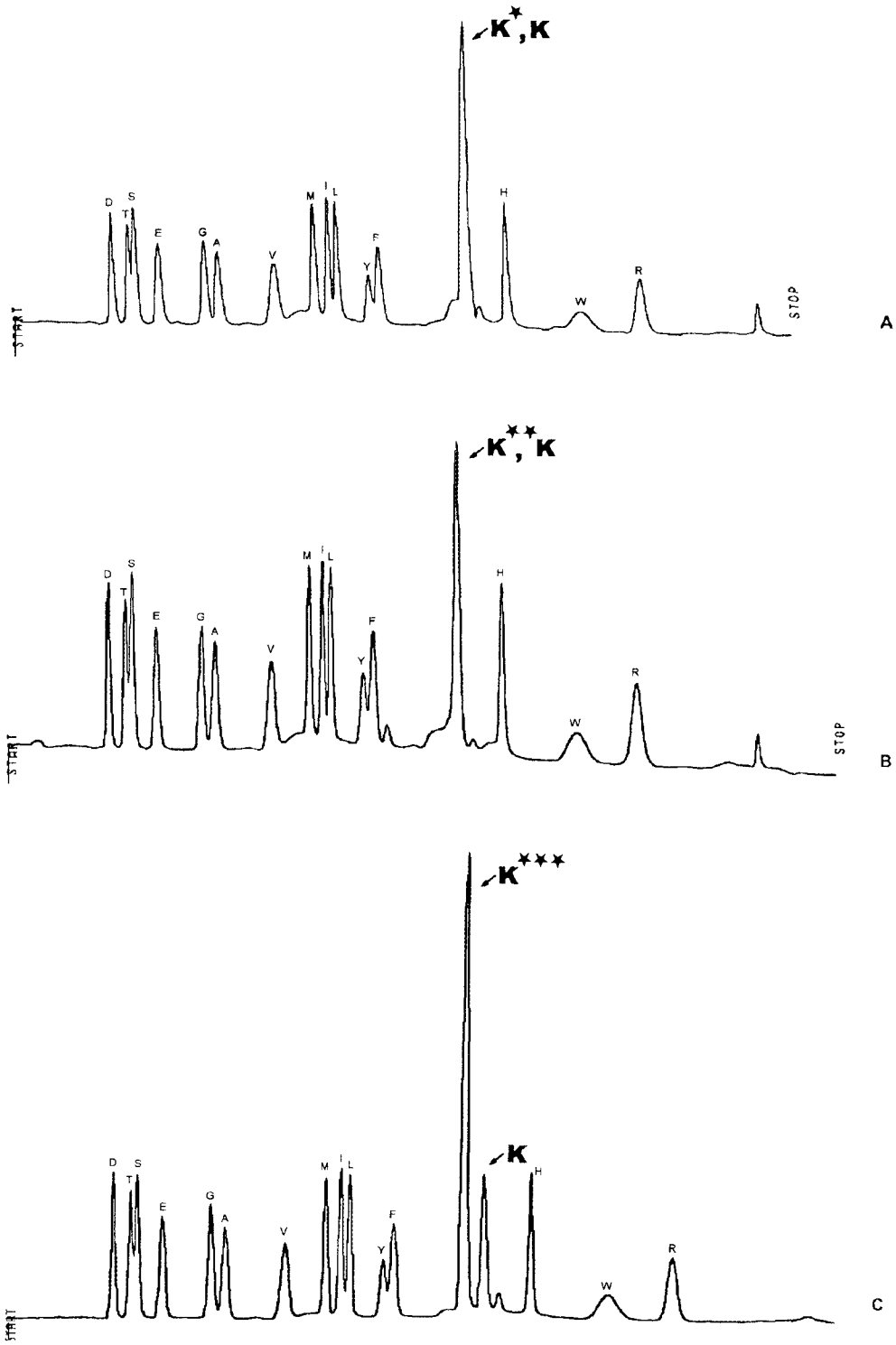


Fig. 3. Separation of 1 nmol of a reference amino acid mixture containing 2 nmoles N^{ε}-monomethylated lysine (**A**), N^{ε}-dimethylated lysine (**B**) and N^{ε}-trimethylated lysine (**C**), respectively, by ion-exchange chromatography. The column was (250 × 3 mm) 10 μM cation exchange and the conditions of the chromatography were the following: 100% 0.2N sodium dihydrogen citrate pH = 3.28, 10 min, 100% 1.0N di-sodium hydrogen citrate-1.5 sodium citrate, pH = 7.40, 22 min, 100% 1.0N sodium citrate, pH = 7.40, 28 min, 100% 0.2N sodium citrate pH = 12.00, 2 min, 100% 0.2N sodium dihydrogen citrate pH = 3.28, 0.1 min and 100% 0.2N sodium dihydrogen citrate. Reaction reagent: 50 mg OPA, 5 ml 2-mercaptoethanol, 1.5% Brij 35% aqueous solution in 500 ml H_3BO_3, pH 10.4 (pH adjusted with 5N KOH). The flow rate was 0.3 ml/min

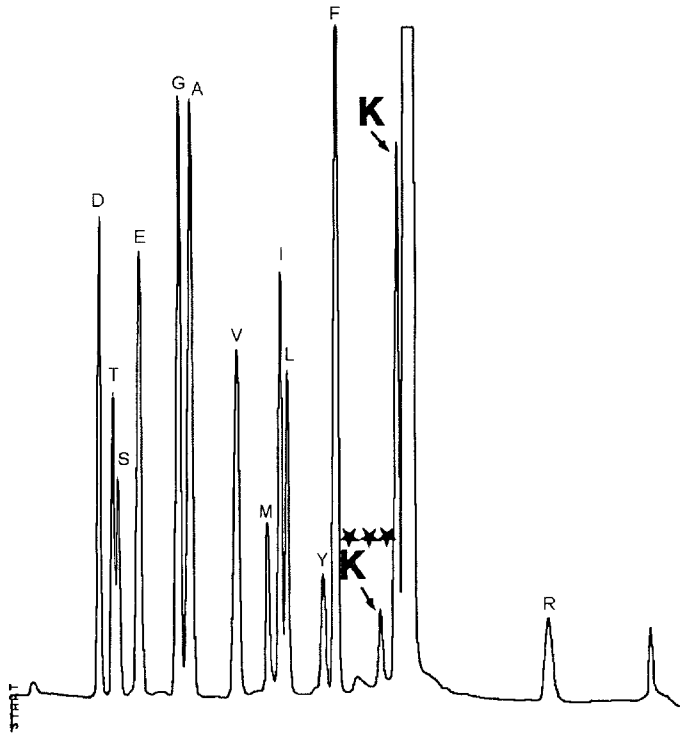


Fig. 4. Amino acid composition of protein L11 prm-1 *E. coli*. The separation of the amino acids was the same as described in Fig. 3. The characteristic peak for N^{ε}-trimethylated lysine (*K****) elutes at the same position as in chromatogram of Fig. 3C

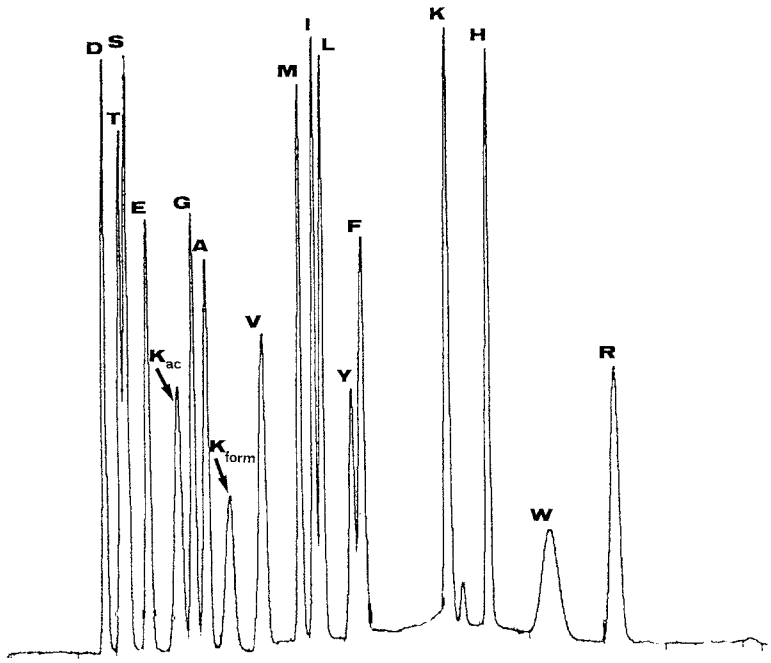


Fig. 5. Separation of 1 nmol of a reference amino acid mixture containing 2 nmoles N^{ε}-formylated lysine and N^{ε}-acetylated lysine. The conditions of the separations are the same as in Fig. 3. The arrows show the N^{ε}-formylated OPA-lysine derivative and N^{ε}-acetylated OPA-lysine derivative, respectively

respectively). Both the change of the salt concentration of the gradient used and the use of another pH gradient were unable to separate these derivatives from the non-methylated lysine derivative. On the other hand by using quite different salt and pH conditions the separation of the other amino acid derivatives is discriminated and should be avoided for a routine amino acid analysis.

Figure 4 shows the amino acid analysis of the L11- prm-1 *E. coli* ribosomal protein which possesses N$^{\varepsilon}$-trimethylated lysine (see Fig. 2), as mentioned above. The amino acids which are released after total hydrolysis (see Materials and methods), are separated, as in Fig. 3, by the post-column derivatization technique. The peak which is shown by the arrow corresponds to the

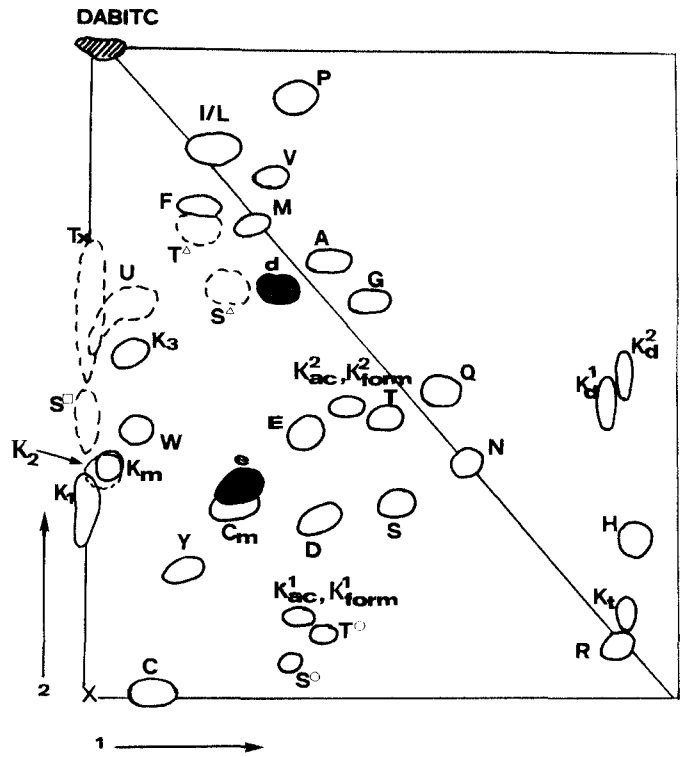


Fig. 6. Identification of DABTH-amino acid and additional spots: DABTH-amino acid derivatives of: *A* alanine; *C* cysteic acid; *Cm* carboxymethyl cysteine; *D* aspartic acid; *E* glutamic acid; *F* phenylalanine; *G* glycine; *H* histidine; *I* isoleucine; K_1 alpha-DABTH-epsilon-DABTC-lysine(red); K_2 alpha-PTH-epsilon-DABTC-lysine(blue); K_3 alpha-DABTH-epsilon-PTC-lysine(blue); K^1_{ac} alpha-DABTZ-N$^{\varepsilon}$-acetyl-L-lysine after hydrolysis of the acetyl-group(blue); K^2_{ac}, alpha-DABTH-N$^{\varepsilon}$-acetyl-L-lysine(red); $K_{form}{}^1$ alpha-DABTZ-N$^{\varepsilon}$-formyl-L-lysine, after hydrolysis of the formyl-group; $K_{form}{}^2$ alpha-DABTH-N$^{\varepsilon}$-formyl-L-lysine(red); K_m derivative of N$^{\varepsilon}$-methyl-L-lysine; $K_d{}^1$ derivative of N$^{\varepsilon}$-dimethyl-L-lysine; $K_d{}^2$ derivative of N$^{\varepsilon}$-dimethyl-L-lysine; K_t derivative of N$^{\varepsilon}$-trimethyl-L-lysine; *L* leucine; *M* methionine; *N* asparagine; *P* proline; *Q* glutamine; *R* arginine; *S* serine; S^{Δ} dehydro-serine, $S^{\square}$ polymerization product of serine; S^{o} DABTZ-serine (blue); *T* threonine; T^{Δ} dehydro-threonine; T^{o} DABTZ-threonine(blue); *Tx* blue-colored polymerization product of threonine; *U* thiourea derivative; *V* valine; *W* tryptophan; *Y* tyrosine; *d* and *e* blue-colored reference markers of DABTC-reacted diethylamine and ethanolamine

N^{ε}-trimethylated OPA-lysine derivative, which appears at the same position, just before non-methylated lysine as in chromatogram of Fig. 3C.

The N^{ε}-formylated OPA-lysine and N^{ε}-acetylated OPA-lysine as shown in Fig. 5, are well separated from the other amino acids. The N^{ε}-acetylated OPA-lysine elutes between the OPA-glutamic acid and OPA-glycine and the N^{ε}-formylated OPA-lys between OPA-alanine and OPA-valine, respectively.

It has to be mentioned that both N^{ε}-formylated lysine and the N^{ε}-acetylated lysine are unstable under the acidic conditions which are used for the hydrolysis of proteins and peptides. Namely, after a protein containing one or both of these modified lysines is hydrolysed it is impossible to identify the above refered modified lysines. However in these cases when a protein or a peptide is subjected to enzymic digestion, with protease K or carboxy-peptidases, the identification of the above refered lysine derivatives is both undoubtly possible and recommended.

Identification of DABTH-derivatives

All lysine derivatives mentioned above, namely N^{ε}-dimethylated and N^{ε}-trimethylated, as well as N^{ε}-formylated and N^{ε}-acetylated lysine are identified after derivatization with DABITC/PITC on polyamide sheets (Choli et al., 1977).

Figure 6 illustrates the resolution of the DABTH-amino acid derivatives by chromatography on stamp-sized polyamide sheets. The location of the individual derivatives is determined relative to the migration of the two markers, DABITC-reacted diethylamine and ethanolamine. The detection limit for the DABTH-amino acids on micropolyamide sheets is about 20 picomoles.

During the Edman degradation the N^{ε}-formylated and N^{ε}-acetylated lysine derivatives are also unstable under the acidic conditions which are used for the cleavage step of the degradation. However, they do not co-migrate with the normal lysine on the polyamide sheets because the coupling reaction of the NH_2-group with DABITC/PITC is performed prior to the cleavage reaction with trichloroacetic acid, while the modification with all reagents used for amino acid analysis takes place after the hydrolysis of the proteins. It means that in cases where the above method of approach for amino acid estimation is used, namely Edman degradation, the identification of N^{ε}-formylated and N^{ε}-acetylated lysine derivatives is undoubtly possible, whereas in the case of the modified amino acids released after hydrolysis the above refered derivatives of lysine remain undetected.

References

Allfrey VG, Faulkner R, Mirsky AE (1964) Acetylation and methylation of histones and their possible role in the regulation of ribonucleic acid. Proc Natl Acad Sci USA 51: 786

Ashman K, Bosserhoff A (1985) Amino acid analysis by HPLC and precolumn derivatization. In: Modern methods in protein chemistry, vol 2. Walter de Gruyter, Berlin New York, pp 155–171

Becam MA, Lederer F (1981) Amino acid sequence of the cytochrome c from the yeast *Hansenula anomala*. Identification of three methylated positions. Eur J Biochem 118: 295–302

Boffa LC, Karn J, Vivali G, Allfrey VG (1977) Distribution of NG, NG-dimethylarginine in nuclear protein fractions. Biochem Biophys Res Commun 74: 969–976

Boffa LC, Sterner R, Vivali G, Allfrey VG (1979) Post-synthetic modifications of nuclear proteins. High mobility group proteins are methylated. Biochem Biophys Res Commun 89: 1322

Chang JY, Lehmann A, Wittmann-Liebold B (1980) Analysis of dimethylamino-azobenzene-thiohydantoins of amino acids by high pressure liquid chromatography. Anal Biochem 102: 380–383

Chang FN, Navickas IJ, Chang CN, Dancis BM (1976) Methylation of ribosomal proteins in HeLa cells. Arch Biochem Biophys 172: 627–633

Choli T, Skendros Y, Katsani K (1997) A manual method for protein sequence analysis using the DABITC/PITC method. In: Kamp RM, Choli-Papadopoulou T, Wittmann-Liebold B (eds) Protein structure analysis. Springer, Berlin Heidelberg New York Tokyo, pp 137–151

Christensen ME, Beyer AL, Walker B, Le Strourgeon WM (1977) Identification of NG, NG-dimethylarginine in a nuclear protein from the lower eukaryote *Physarum polycephalum* homologous to the major proteins of mammalian 40S ribonucleoprotein particles. Biochem Biophys Res Commun 74: 621–627

Delange RJ, Glazer AN, Smith EL (1969) Presence and location of an unusual amino acid, epsilon-N-trimethyllysine, in cytochrome c of wheat germ and *Neurospora*. J Biol Chem 244: 1385

Dognin MJ, Wittmann-Liebold B (1977) The primary structure of L11, the most heavily methylated protein from *E. coli* ribosomes. FEBS Lett 84: 342

Doolittle LR, Mross GA, Fothergill LA, Doolittle RF (1977) A simple solid phase amino acid sequencer employing a thioacetylation stepwise degradation procedure. Anal Biochem 78: 491–505

Hempel VK, Lange HW, Birkofer L (1968) Epsilon-N-trimethyllysine, a new amino acid in histones. Naturwissenschaften 55: 37

Hermodson MA, Chen KCN, Buchanan TM (1978) Neisseria pili proteins: amino-terminal amino acid sequences and identification of an unusual amino acid. Biochemistry 17: 442

Huszar G, Elzinga M (1969) Epsilon-N-methyl lysine in myosin. Nature (London) 223: 834

Kakimoto Y, Akazawa S (1970) Isolation and identification of NG, NG- and NG, N′G-dimethylarginine, N-epsilon-mono-, di- and tri-methyllysine and glycosylgalactosyl- and galactosyl-delta-hydroxylysine from human urine. J Biol Chem 245: 5751

Kuehl WM, Adelstein RS (1969) Identification of epsilon-N-monomethyllysine and epsilon-N-trimethyllysine in rabbit skeletal myosin. Biochem Biophys Res Commun 37: 59

Marvil DK, Nowak L, Szer W (1980) A single-stranded nucleic acid-binding protein from *Artemia salina* I. Purification and characterization. J Biol Chem 255: 6466–6472

Paik WK, Kim S (1971) Protein methylation. Science 147: 114–119

Paik WK, Kim S (1967) Enzymatic methylation of protein fractions from calf thymus nuclei. Biochem Biophys Res Commun 29: 14

Palmiter RD (1977) Prevention of NH_2-terminal acetylation of proteins synthesized in cell-free system. J Biol Chem 252: 8781

Sterner R, Vivali G, Allfrey VG (1979) Studies of acetylation and deacetylation in high mobility group proteins. Identification of the sites of acetylation in HMG-1. J Biol Chem 254: 11577

Sterner R, Vivali G, Allfrey VG (1981) Studies of acetylation and deacetylation in high mobility group proteins. Identification of the sites of acetylation in high mobility group proteins 14 and 17. J Biol Chem 256: 8892

Trayer IP, Harris CI, Perry SV (1968) 3-methyl histidine and adult and foetal forms of skeletal muscle myosin. Nature (London) 217: 452

Williams KR, Stone KL, Lopresti MB, Mervill, Planck SR (1985) Amino acid sequence of the UP1 calf thymus helix-destabilizing protein and its homology to an analogous protein from mouse myeloma. Proc Natl Acad Sci U.S.A. 82: 5666

Wold F (1981) *In vivo* chemical modification of proteins (post-translational modification). Annu Rev Biochem 50: 783

Authors' address: Dr. T. Choli-Papadopoulou, School of Chemistry, Laboratory of Biochemistry, Aristotle University of Thessaloniki, GR-54006 Thessaloniki, Greece.

Received October 5, 1996