Journal of Supramolecular Structure 3:426–447 (1975)

STUDIES ON THE PRIMARY STRUCTURE OF 14 PROTEINS FROM THE LARGE SUBUNIT OF ESCHERICHIA COLI RIBOSOMES WITH AN IMPROVED PROTEIN SEQUENATOR AND WITH MASS SPECTROMETRY

B. Wittmann-Liebold, A. W. Geissler, and E. Marzinzig

Max Planck Institut für Molekulare Genetik, Berlin-Dahlem, Germany

Fourteen proteins from the large subunit of Escherichia coli ribosomes were analyzed in an improved sequenator. In addition to our previously described modifications of a Beckman sequenator, new valves which work free of a dead volume were constructed. By this and the previous improvements (e.g., a new vacuum system with a recorder, cool traps, automatic conversion) much better results were obtained than before. It was even possible to use (in addition to the standard methods, e.g., thin-layer chromatography and amino acid analysis) mass spectrometry without preceding gas chromatography for identification of the released PTH amino acids. Our experience with the various methods, especially mass spectrometry, is described and the techniques are compared. The results obtained by the described methods on the amino acid sequences of the 14 ribosomal proteins are summarized.

INTRODUCTION

In preceding papers (1, 2) our studies on the primary structure of 20 proteins from E. coli ribosomes with an improved protein sequenator have been reported. The N-terminal regions of proteins from the small subunit were compared up to positions 32–60. These investigations have now been extended to 14 proteins isolated from the large subunit of E. coli ribosomes and their N-terminal regions up to positions 32–65 were sequenced.

All data were obtained by the automatic degradation method (3) in a Beckman sequencer which was considerably improved and equipped with a device for an automatic conversion reaction. In addition to the modifications already mentioned (1), a new system of delivery and effluent valves was incorporated in the machine, together with further improvements which considerably facilitated an unambiguous identification of all amino acid derivatives. The PTH amino acids obtained after automatic conversion in the sequenator were identified by thin-layer chromatography, mass spectrometry, or back hydrolysis followed by amino acid analysis. The results are described in this paper.

MATERIAL AND METHODS

Sequenator

An apparatus (model 890) from Beckman Instruments, Palo Alto, Calif. was used throughout this study. It was considerably modified as already described (1). In addition to

Manuscript received July 29, 1975; revision accepted August 27, 1975

these modifications a completely new delivery system and effluent valve system made in our workshop has been used for more than 2 yr. These valves, which are described in full detail elsewhere (4), are arranged in series and have only one delivery line. They work free of a dead volume because the delivery capillary is completely emptied by nitrogen after each delivery step.

The advantage of this arrangement is that lesser amounts of solvents can be used than before: only 7–10 ml of benzene and 10–12 ml of ethylacetate are needed even for double coupling in a program with Quadrol buffer. Furthermore, the periods for the drying steps are shorter. Since a slow flow of nitrogen empties the delivery line instead of the restricted vacuum, the protein film can be dried more gently and without moving towards the top of the cup with the reagents which flow into the cup from the delivery line under the usual conditions of a restricted vacuum. Therefore, we can still use the original cup without the undercut groove preventing the protein film from rising too much during the degradation cycles. In the case of our delivery system the protein film stays at its original height even after 60–70 cycles. On the other hand, undercut grooved cups have the tendency to accumulate the protein near that groove, resulting in increasing thickness, so that the efficiency of the Edman degradation reaction decreases. The same types of valves were also used as effluent valves and were incorporated in the automatic conversion system.

In addition to the improved valves, a double-jacketed glass bell jar for the cup was installed. This heats the cup with water from a thermostat (53°C). This bell jar has a plug of Kel-F in which the vertical position of the scoop is fixed. It is brought into the right acceptance position by a short turn.

Automatic conversion of the thiazolinones into the PTH amino acids was performed with 20% trifluoroacetic acid plus 0.02% dithioerythritol for 15 min at 52°C. The heating water for the cup bell jar is directed into the double-jacketed conversion flask, thus maintaining a constant temperature of 52°C for the conversion reaction. The details of all newly constructed parts will be published elsewhere (4).

Reagents and Solvents

The purification of the reagents and solvents was done as previously described (1). Only Quadrol was bought from Beckman. All substances were tested for purity by infrared spectroscopy, mass spectromety, and the chemical tests previously described (1, 3).

Program

The program has two coupling steps (of 20 min and 10 min, respectively) which are necessary for a complete coupling if repetitive glutamic acid and aspartic acid residues occur. Furthermore, there are two rather short cleavage steps (150 and 80 sec, respectively).

The same programmer that regulates the main program in the sequenator is also used for the reaction in the conversion flask. This side program consists of seven different functions: numbers 36–40 for nitrogen, waste 2, vacuum, collect into the conversion flask and into the fraction collector. Furthermore, functions numbers 21 and 22 are used for delivery of R4 (conversion medium) and of R5 (solvent for PTH amino acids). Function 41 of the programmer is used for regulation of the nitrogen valve that empties the delivery line. The side program runs parallel to the main program starting with drying of the two chlorobutane washes followed by addition of the conversion medium. It times the conversion reaction parallel to the coupling time in the main program. During the benzene and ethylacetate washes in the main program, delivery of R5 (three times) into

	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
L1	Ala-Lys-Leu-Thr-Lys-Arg-Met-Arg-Val-Ile-Arg-Glu-Lys-Val-Asp-Ala-Thr-Lys-Gln-Tyr-Asp-Ile-Asn-Glu-Ala-
L2	Ala-Val-Val-Lys-Ser-Lys-Pro-Thr-Ser-Pro-Gly-Arg-Arg-His-Val-Val-Lys-Val-Val-Asn-Pro-Glu-Leu-His-Lys-
L3	Met-Ile-Gly-Leu-Val-Gly-Lys-Lys-Val-Gly-Met-Thr-Arg-Ile-Phe-Thr-Glu-Asp-Gly-Val-Ser-Ile-Pro-Val-Thr-
L5	Ala-Lys-Leu-His-Asp-Tyr-Tyr-Lys-Asp-Glu-Val-Val-Lys-Lys-Leu-Met-Thr-Glu-Phe-Asn-Tyr-Asn-Ser-Val-Met-
L12	Ser-Ile-Thr-Lys-Asp-Gln-Ile-Ile-Glu-Ala-Val-Ala-Ala-Met-Ser-Val-Met-Asp-Val-Val-Glu-Leu-Ile-Ser-Ala-
L13	Met-Lys-Thr-Phe-Thr-Ala-Lys-Pro-Glu-Thr-Val-Lys-Arg-Asp- X -Tyr-Val-Val-Asp-Ala-Thr-Gly-Lys-Thr-Leu-
L17	Met-Arg-His-Arg-Lys-Ser-Gly-Arg-Gln-Leu-Asn-Arg-Asn-Ser-Ser-His-Arg-Gln-Ala-Met-Phe-Arg-Asn-Met-Ala-
L20	Ala-Arg-Val-Lys-Arg-Gly-Val-Ile-Ala-Arg-Ala-Arg-His-Lys-Lys-Ile-Leu-Lys-Gln-Ala-Lys-Gly-Tyr-Tyr-Gly-
L22	Met-Glu-Thr-Ile-Ala-Lys-His-Arg-His-Ala-Arg-Ser-Ser-Ala-Gln-Lys-Val-Arg-Leu-Val-Ala-Asp-Leu-Ile-Arg-
L23	Met-Ile-Arg-Glu-Glu-Arg-Leu-Leu-Lys-Val-Leu-Arg-Ala-Pro-His-Val-Ser-Glu-Lys-Ala-Ser-Thr-Ala-Met-Glu-
L24	Ala-Ala-Lys-Ile-Arg-Arg-Asp-Asp-Glu-Val-Ile-Val-Leu-Thr-Gly-Tyr-Lys-Asp-Lys-Gly-Lys-Arg-Gly-Lys-Val-Lys-
L25	Met-Phe-Thr-Ile-Asn-Ala-Glu-Val-Arg-Lys-Glu-Gln-Gly-Lys-Gly-Ala-Ser-Arg-Arg-Leu-Arg-Ala-Ala-Asn-Lys-
L27	Ala-His-Lys-Lys-Ala-Gly-Gly-Ser-Thr-Arg-Asn-Gly-Arg-Asp-Ser-Glu-Ala-Lys-(Arg)-Leu-Gly-Val-Lys-(Arg)-Phe-
L29	Met-Lys-Ala-Lys-Glu-Leu-Arg-Glu-Lys-Ser-Val-Glu-Glu-Leu-Asn-Thr-Glu-Leu-Asn-Leu-Leu-Arg-Glu-Gln-

```
        26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50
L1   Ile-Ala-Leu-LoI-Lys-Glu-Leu-Ala-Thr-Ala-Lys-Phe-Val-(Lys)-Ser-Val-(Tyr)-Val-Ala-Val-Asn-LoI-
L2   Gly-Lys-Pro-Phe-Ala-Pro-Leu-Leu-Glu-Lys-Asn-Ser-Lys-Ser-Gly-
L3   Val-Ile-(Arg)-Val-(Glu)-Ala-Asn-
L5   Gln-Val-Pro-Arg-Val-Glu-Lys-Ile-Thr-Leu-Asn-Met-Gly-Val-Gly-
L12  Met-Glu-Glu-Lys-Phe-Gly-Val-Ser-Ala-Ala-Ala-Ala-Val-Ala-Val-Ala-Ala-Gly-Pro-Val-Glu-Ala-Ala-Glu-Glu-
L13  Gly-Arg-Leu-Ala-Thr-Glu-Leu-Ala-Ser-(Arg)-Leu-Ser-Gly-Lys-His-Lys-Ala-Gly-Tyr-Thr-Pro-
L17  Gly-Ser-Leu-Val-Arg-His-Glu-Ile-LoI-Lys-
L20  Ala-Arg-Ser-Arg-Val-Tyr-Arg-Val-Ala-Phe-Gln-Ala-Val-Ile-Lys-Ala-Gly-Gln-Tyr-Ala-Tyr-Arg-Asp-Arg-(Arg)-
L22  Gly-Lys-Lys-Val-Ser-Gln-Ala-Leu-Asp-Ile-Leu-Thr-Tyr-Thr-Asn-Lys-Lys-Ala-Ala-Val-LoI-Val-Lys-Lys-Val-
L23  Lys-Ser-Asn-Thr-Ile-Val-Leu-Lys-Val-Ala-Lys-Asp-Ala-Thr-Lys-Ala-(Lys)-LoI-Lys-Ala-Ala-Val-(Gln)-Lys-LoI-
L24  Asn-Val-Leu-Ser-Ser-Gly-Lys-Val-Ile-Val-Glu-Gly-Ile-Asn-Leu-Val-Lys-Lys-His-Gln-(Lys)-Pro-Val- X -Ala-
L25  Phe-Pro-Ala-Ile-Ile-Tyr-Gly-Gly-Lys-Glu-Ala-Pro-Leu-Ala-Ile-Glu-Leu-Asp-His-Asp-Lys-Val-Met-Asn-Met-
L27  Gly-Gly- X -Ser-Val-Leu-Ala-Gly-Ser-LoI-Ile-Val-(Arg)-Gln-Arg-
L29  Phe-Asn-Leu-Arg-Met-Gln-Ala-Ala-Ser-Gly-Gln-Leu-Gln-Gln-Ser-His-Leu-Leu-Lys-Leu-Gln-Leu-Asn-Thr-Lys-

        51  52  53  54  55  56  57  58  59  60  61  62  63  64  65
L22  LoI-Lys-Ser-(Pro)-LoI-Ala-Asn-Ala-Lys-Glu-
L24  Leu-Asn-Gln-Pro-Gly-
L25  Gln-Ala-Lys-Ala-Glu-Phe-Tyr-Ser-Glu-Val-Leu-Thr-Ile-Val-Val-
```

Fig. 1. N-terminal amino acid sequences of 14 proteins from the large ribosomal subunit of E. coli, strain K. (LoI, leucine or isoleucine; X, amino acid is not known).

the conversion flask and transfer into the fraction collector are directed by the side program. (For details see reference 4.)

For all sequenator runs described in this paper, 0.5 M Quadrol buffer, pH 9.0 (1.0 M; diluted under nitrogen with bidistilled water) was used. In typical runs 3–4 mg of lyophilized proteins (water content 30–40%) dissolved in 0.5 ml of water plus 100 μg of dithioerythritol are transferred into the cup and dried under nitrogen followed by an automatic change of nitrogen and vacuum and continued by rough and fine vacuum (1). The run starts with one chlorobutane wash which is at the end of the normal program. This first chlorobutane wash, which does not contain any thiazolinones, is collected in the conversion flask and is subjected to the normal "conversion process" (the so called wash cycle "LC"). It is used as a blank in mass spectrometry, namely, for subtraction of the background from the spectra obtained from the degradation cycles (see below). It is further used as a check of the sequenator: 20% of the total amount from this wash cycle should not give any spot in thin-layer chromatography nor a peak when injected in the gas chromatograph.

Proteins

All investigated protein samples were obtained from Dr. H. G. Wittmann. They were isolated from ribosomes of Escherichia coli, strain K, according to our standard procedure (5) consisting of separation of ribosomal subunits by zonal centrifugation, group fractionation of proteins by treatment of the ribosomes with LiCl, extraction of proteins as well as their isolation, and purification by column chromatography on carboxymethyl cellulose followed by gel filtration on Sephadex. The purity of the protein samples was tested by two-dimensional polyacrylamide gel electrophoresis (6) and by a sequenator run with a small amount of protein (1–2 mg). These test runs give information not only about the purity of the sample but also determine if a given protein is blocked at the N-terminus, and in addition they elucidate the sequence of the first 20–30 amino acids. For the main run 3–4 mg lyophilized protein were used, an amount which corresponds, after subtraction of bound water, to about 200 nmoles. For each of the proteins listed in Fig. 1 two to four sequenator runs were performed.

Thin-Layer Chromatography

PTH amino acids were spotted on silica gel plates (DC-alufoil silica gel 60 F_{254}, 0.25 mm, from Merck, Darmstadt) and chromatographed with three different solvent systems one after the other in the same direction. The solvent mixtures (v/v) were: first, chloroform-1-propanol-2-propanol, 98:1:1 (system VII); second, chloroform-methanol 90:10 (system II); third, heptane-propionic acid-1,2-dichloroethane 58:17:25 (system IV). This combination resolves all PTH amino acids on one plate (Fig. 2) and therefore saves material for other identification methods. The following amounts were used (starting with 200 nmoles of protein): for cycles 1–10, 10–20% of the total amount; 11–20, 20–25%; 21–30, 30%; 31–40, 40%; 41–50, 40–50%; 51–70, 50–80%.

Hydrolysis and Analysis

Cleavage of the PTH amino acids to the free amino acids was made with 5.7 N HCl + 0.05% mercaptoethanol for 24 hr at 130°C. Amino acid analyses were carried out in the one-column system of LKB analyzers. The following portions of the sample were used for this purpose: cycles 1–15, 10%; 16–30, 20–40%; 31–40, 40–50%.

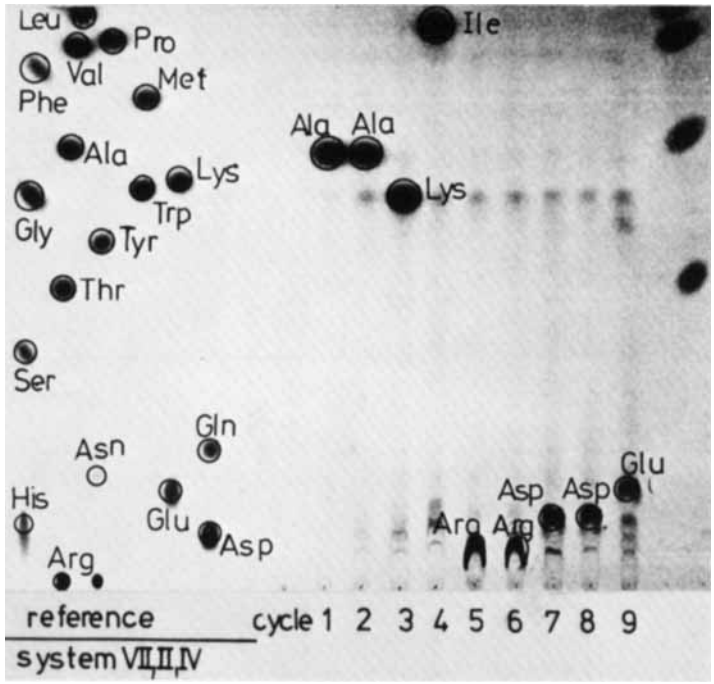


Fig. 2. Thin-layer chromatography of PTH amino acids released in cycles 1–9 of protein L24. The photograph was taken after the thin-layer plate had been developed in three solvent systems (see Methods).

Mass Spectrometry

Instruments. The identification of the PTH amino acids was performed with a Varian CH-7 instrument which was equipped with turbomolecular pumps (Balzers, Vaduz) and with a temperature program (from 25°C to 190°C with scale division 20). This instrument was connected to the Varian Data System MS 100 with a Houston plotter, a display unit, and the one magnetic tape system.

Measurements. Experimental conditions included: 300 μA of emitter wire current, 800 resolution, 3 kV accelerating potential, and 5s/decade scanning rate. Electron impact spectra were obtained with the direct probe technique. Since the various PTH amino acids (for mass spectra of PTH amino acids, see reference 7) are volatile at different temperatures, 20 spectra were taken from every degradation cycle in 4 min with a scan rate of 21 to ensure that no characteristic molecular or fragment ion was lost during the temperature program. Care was taken to use the same conditions for the measurement of all degradation cycles. The following portions of the sample were applied: for cycles 1–20, 1–2%; 21–40, 4%; 41–60, 10% of the total amount.

LM spectra. After the conversion into mass spectra by the computer, the main spectrum of every degradation cycle (the so-called LM spectrum) was selected by the display unit or from the printed mass lists (mass numbers 190–300) and plotted by the Houston plotter. In Figs. 3a, 4a, and 5a examples of LM spectra are shown.

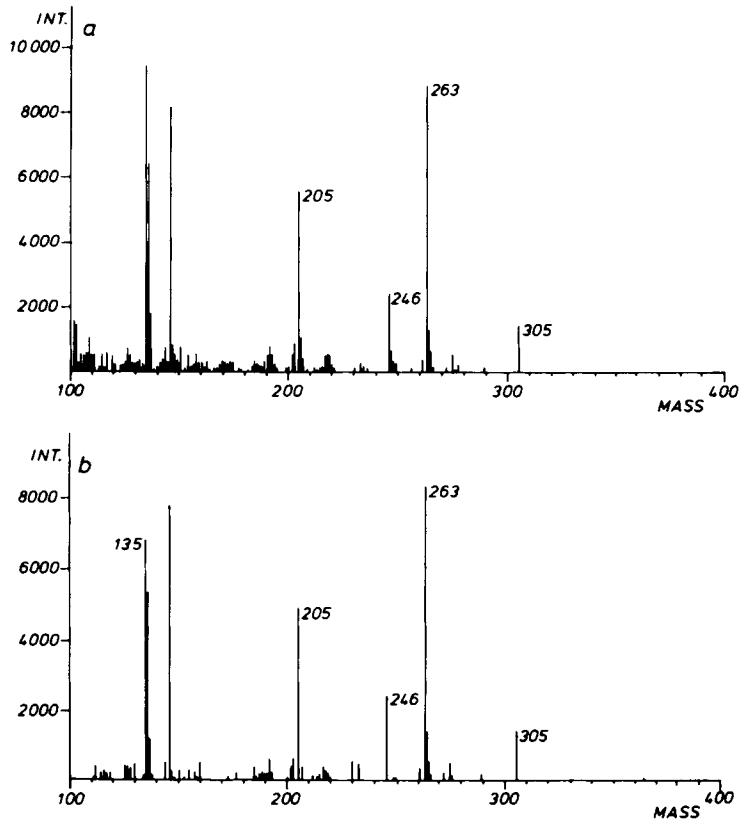


Fig. 3. Mass spectrometry of cycle 9 of protein L29. The main LM spectrum (number 19) of the PTH amino acid in this cycle (PTH-Lys) is given without (Fig. 3a) and with (Fig. 3b) subtraction of the background determined in the "wash cycle" (see text). The fragment ions characteristic for PTH-Lys are marked by their mass numbers.

LS spectra. From these LM spectra the background of the sequenator caused by the reagents and solvents was subtracted as follows: Aliquots of the wash cycle (see Methods) were used as a blank in the mass spectrometry. 20 spectra were taken from this blank cycle under the same conditions as for normal degradation cycles and the spectrum with the highest intensities was selected for subtractions (for illustration of the LS spectra see Figs. 3b, 4b, and 5b).

In addition to the described procedure another method for subtraction was used. From the main spectra of every degradation cycle the corresponding spectrum of the previous cycle was subtracted, for instance, spectrum no. 9 of the 37th degradation cycle minus spectrum no. 9 of the 36th degradation cycle (see Fig. 6). The use of the computer for this kind of subtractions is essential for degradation cycles in which overlap occurs especially at high cycles. For a detailed description of these studies see reference 8.

Mass chromatograms. From all sequenator cycles selected mass chromatograms were plotted by the Houston plotter. The intensities of every molecular or characteristic fragment ion of the PTH amino acids are displayed in the so-called D plot. This was done

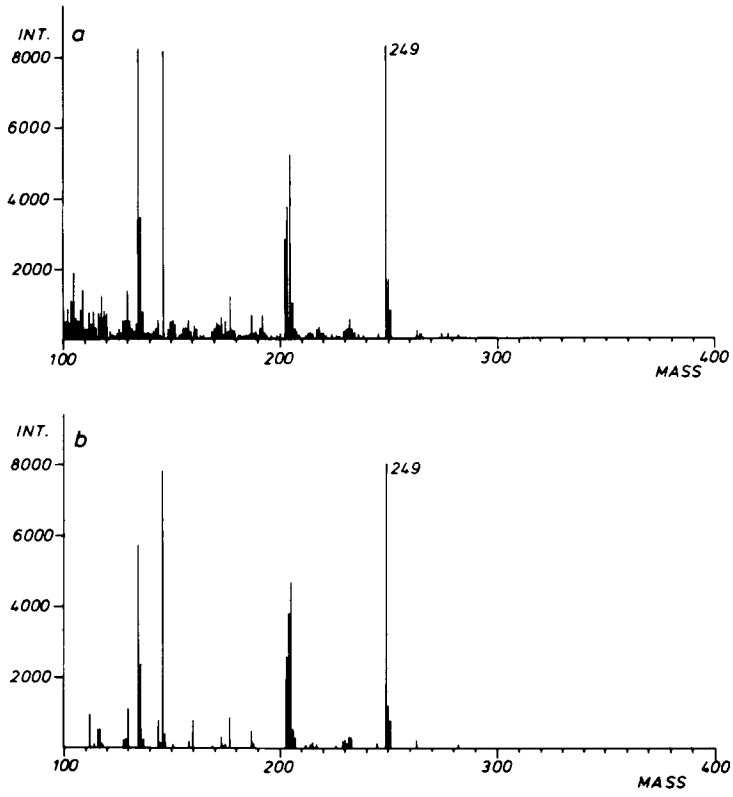


Fig. 4. Mass spectrometry of cycle 20 of protein L29. The main LM spectrum (number 15) of the PTH amino acid in this cycle (PTH-Asn) is given without (Fig. 4a) and with (Fig.4b) subtraction of the background. The molecular and fragment ions characteristic for PTH-Asn are marked by their mass numbers.

for all spectra taken from the degradation cycles. Examples are given in Figs. 7a–c and 8a–c and in more detail elsewhere (8).

RESULTS AND DISCUSSION

The use of a sequenator for the determination of the amino acid sequences within the N-terminal regions is the first and a very efficient step towards the complete elucidation of the primary structure of the proteins from E. coli ribosomes. In a previous publication (2) our sequenator studies on 20 E. coli ribosomal proteins, mainly from the small subunit, were presented.

In this paper the results on 14 proteins isolated from the large subunit are summarized in Fig. 1. Depending on the proteins, between 32 and 65 amino acid residues were determined. This corresponds to 15–80% of the total length of the ribosomal proteins which differ very much in size. The sequences shown in Fig. 1 are confirmed by the isolation and analysis of tryptic peptides for many of these proteins.

Complete primary structures are already known for proteins S4 (9), S6 (10), S8 (11), S9 (12), L7 and L12 (13), L25 (14), and L29 (15). Other proteins are still under investigation. Protein L26 was also studied in the sequenator: it has the same N-terminal region

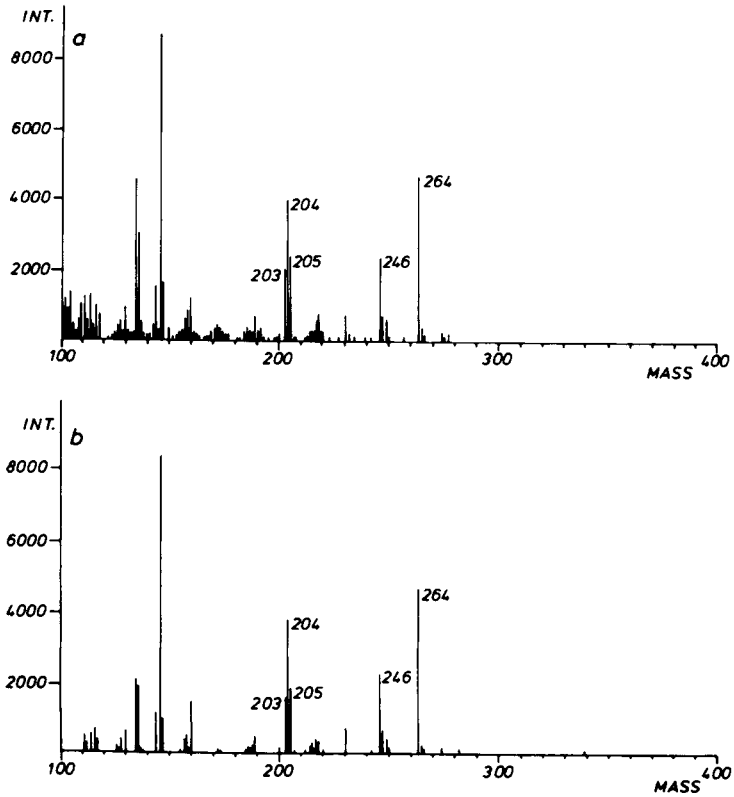


Fig. 5. Mass spectrometry of cycle 24 of protein L29. The main LM spectrum (number 16) of the PTH amino acid in this cycle (PTH-Glu) is given without (Fig. 5a) and with (Fig. 5b) subtraction of the background. The molecular and fragment ions characteristic for PTH-Glu are marked by their mass numbers.

as protein S20. The identity of both proteins was already concluded from immunological studies (16). With the exception of proteins L7/L12 and S20/L26 there are no extended homologous structures among E. coli ribosomal proteins as revealed by immunological methods (17). Therefore, the amino acid sequence determinations of the unrelated ribosomal proteins are much more difficult than in the many studies on homologous proteins of similar primary structure.

Since the isolation and purification of ribosomal proteins are very laborious and time consuming, and since the highly purified proteins are also needed for other studies (e.g., for production of antibodies, for functional studies, and for those on protein-RNA interaction), only relatively small amounts of proteins were available for our sequence studies. Furthermore, ribosomal proteins contain a rather high number of amino acids which are easily destroyed during degradation under usual sequenator conditions or are released in poor yields by the Edman degradations.

For these reasons (no homologous structures; small amount; unfavorable composition) it was essential that the sequenator and the methods for identification of the PTH amino acids were most reliable and efficient. Therefore and because our original Beckman sequenator did not give satisfactory results with ribosomal proteins and many runs were

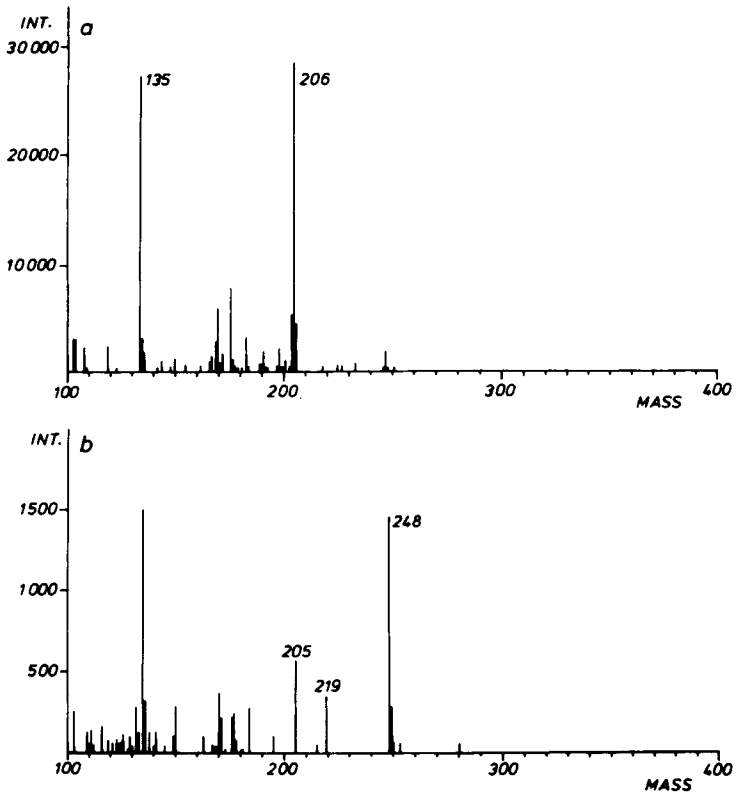


Fig. 6. LS spectra of protein L29: cycle 33 (PTH-Ala) in Fig. 6a and cycle 37 (PTH-Leu) in Fig. 6b. Subtractions were made as follows. The main LM spectrum (number 5) of cycle 33 minus LM spectrum (number 5) of cycle 31 (Fig. 6a). The main LM spectrum (number 9) of cycle 37 minus LM spectrum (number 9) of cycle 36 (Fig. 6b).

lost by failure of the apparatus (1), we began to modify the instrument and incorporated a system for automatic conversion of the thiazolinones into the PTH amino acids. In addition to the modifications previously reported (1) we constructed new valves which work completely free of a dead volume. Further, we used direct mass spectrometry as a new means for identification of the PTH amino acids. This enabled us to obtain much better and more unambiguous results with less amounts of protein.

In the following we describe our experience with the recent modification of the sequenator as well as the results obtained by our methods for identification, especially by a mass spectrometer combined with a data system. Finally, the amino acid sequences obtained for 14 proteins from the large subunit of E. coli ribosomes are compared to each other and to those of other ribosomal proteins.

Modification of the Sequenator

The high purity of the released PTH amino acids is due to a newly designed vacuum system (1, 4) which differs very much from that originally installed in the instrument. It allows us to completely omit the ethylacetate extraction of the converted PTH amino acids which is otherwise absolutely necessary.

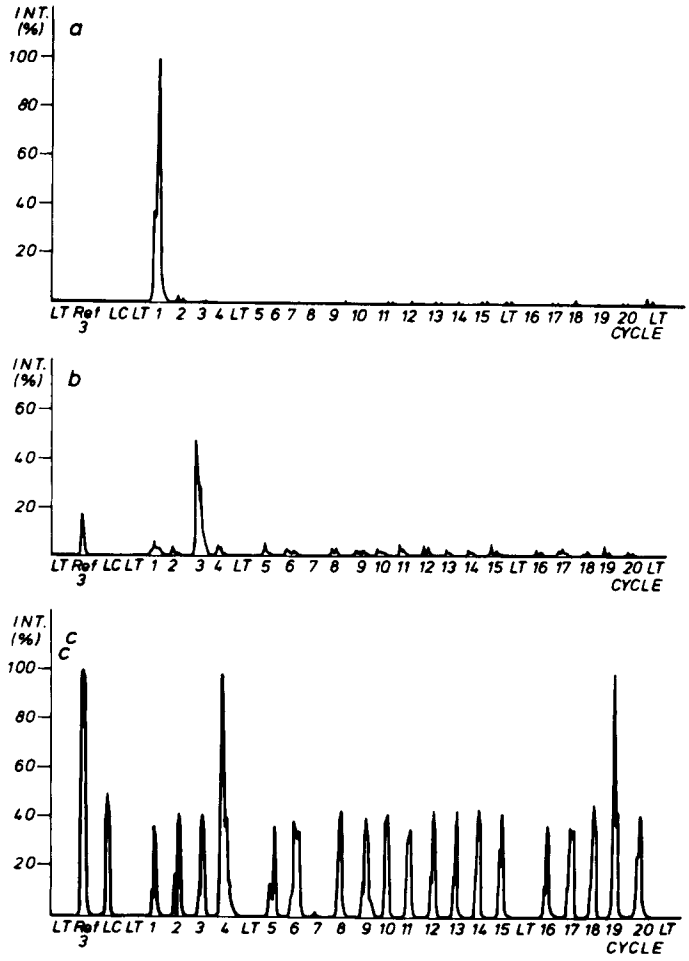


Fig. 7. Mass chromatogram plots of cycles 1–20 of protein L22 for the following characteristic ions: m/e = 266 (PTH-Met) in Fig. 7a; m/e = 218 (PTH-Thr minus H_2O) in Fig. 7b and m/e = 248 (PTH-Leu and PTH-Ile) in Fig. 7c.

Furthermore, the automatic conversion reaction incorporated in our sequenator (1) results in better yields of several PTH amino acids, since not only are both chlorobutane washes collected and used for the conversion but also less destruction can take place. This is due to oxygen-free conditions during the conversion reaction and because all released thiazolinones are immediately converted to the more stable PTH amino acids. In this way serine, threonine, histidine, and arginine which normally become at least partially destroyed can be easily detected. The same is true for glutamine and asparagine even up to step 70. Each of these two amino acids gives only one spot on thin-layer plates without detectable amounts of glutamic or aspartic acid, respectively.

Our new delivery valves allow us to use lesser amounts of proteins or peptides than were necessary before. For typical runs, 200 nmoles of proteins and 100 nmoles of peptides were used. The advantage of the delivery valve system (4) is that the delivery line is emptied by nitrogen after every delivery step, thus resulting in dead volume free

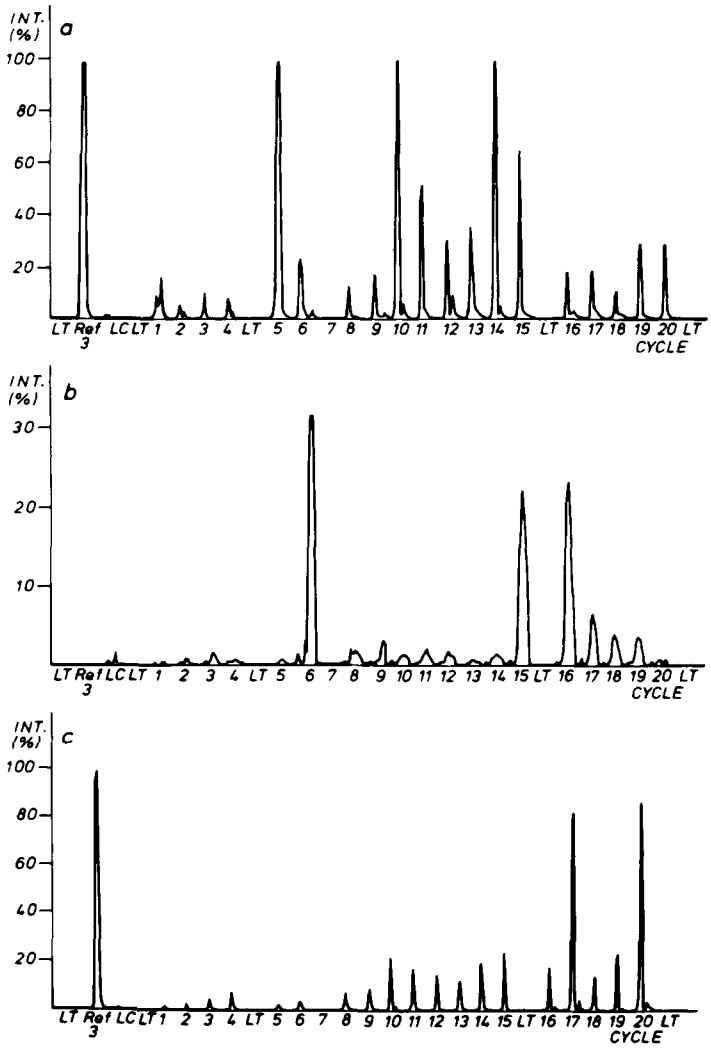


Fig. 8. Mass chromatogram plots of cycles 1–20 of protein L22 for the following characteristic ions: m/e = 206 (PTH-Ala) in Fig. 8a; m/e = 263 (PTH-Gln and fragment ion of PTH-Lys) in Fig. 8b and m/e = 234 (PTH-Val) in Fig. 8c.

delivery of reagents and solvents. The protein film in the cup keeps its height and homogeneous thickness during all degradation cycles. Furthermore, a lesser amount of solvents than is normally necessary in the Quadrol program can be used even with double couplings which we routinely use, and shorter drying periods can be applied.

With the modifications performed in the sequenator the efficiency of the machine is much improved. Overlaps occur later than before, yields and purity of the released PTH amino acids are much better, and the apparatus works very reliably and gives unambiguous and completely reproducible results. This is especially useful when proteins have to be sequenced which are very difficult to isolate and are therefore available in small amounts only, as is the case with ribosomal proteins (see above).

Identification of PTH Amino Acids

Three methods were used for identification: thin-layer chromatography, hydrolysis of PTH amino acids, and mass spectrometry (see Table I).

Thin-layer chromatography. This was routinely performed up to 50–70 cycles. This technique allows identification of all PTH amino acids under our conditions. Only at high degradation cycles and with the low protein amounts which we have to use, the yields for the derivatives of arginine, threonine, and serine become too low for a clear identification by this method. On the other hand, clear results are obtained for glutamine and asparagine up to very high steps. Increasing the number of cycles necessitates an increasing portion of the samples to be applied to thin-layer plates. This raises the background from the sequenator, since rather large aliquots (about 50–75%, see Methods) have to be taken at the high degradation cycles.

Thin-layer chromatography is very useful and was therefore routinely used for the identification of the PTH amino acids. Figure 9 shows the thin-layer chromatograms for cycles 1–34 of protein L29, and Fig. 2 those for cycles 1–9 of protein L24. In the case of protein L29 two different plates were used for each degradation cycle with solvent mixtures previously described (1). For protein L24 spots were chromatographed with three different solvent mixtures used one after the other on the same plate, as described under Methods. Both figures are given here to illustrate the clear distinction between PTH-Glu, PTH-Gln, PTH-Asp, PTH-Asn, and PTH-Arg under our sequenator conditions.

In the case of protein L29 the sequenator run was not optimal because of (1) a quicker wash-out rate than normal caused by the short protein length of 63 amino acids; (2) repetitive sequences, e.g., Leu-Leu-Asn-Leu-Leu in positions 18–22; (3) accumulated Glu sequences (positions 5, 8, 12, 13, 17, and 24) which made the Edman degradation less effective as compared to normal sequences. The same was observed for protein L23 where a Glu-Glu sequence occurs in positions 4 and 5 and in the case of another protein rich in glutamic acid whose sequence is not given in this paper. Double coupling used as a routine method for all degradation cycles helped to overcome the difficulties with such unfavorable sequences.

TABLE I. Identification Procedures for PTH Amino Acids

Protein	Published positions	Thin-layer chromatography	Hydrolysis and analysis	Mass spectrometry
L1	1–47	1–60	1–45	–
L2	1–40	1–50	1–36	1–50
L3	1–32	1–49	1–32	1–40
L5	1–40	1–50	1–40	1–38
L12	1–50	1–60	–	–
L13	1–46	1–58	1–37	–
L17	1–35	1–50	1–35	1–30
L20	1–50	1–70	1–50	1–26
L22	1–60	1–60	1–47	1–42
L23	1–50	1–58	1–37	1–15
L24	1–55	1–60	–	–
L25	1–65	1–65	1–50	1–20
L27	1–40	1–53	1–40	–
L29	1–50	1–50	1–40	1–45

More examples for thin-layer chromatography of PTH amino acids from the studied ribosomal proteins are presented elsewhere, e.g., cycles 1–58 for protein L25 (14) and cycles 1–36 for protein L5 (8).

Hydrolysis of PTH amino acids followed by amino acid analysis was done for the first 35–50 cycles. Since a great portion of the available material was used for thin-layer chromatography (see above), too little was left for hydrolysis and amino acid analysis beyond 35–50 cycles.

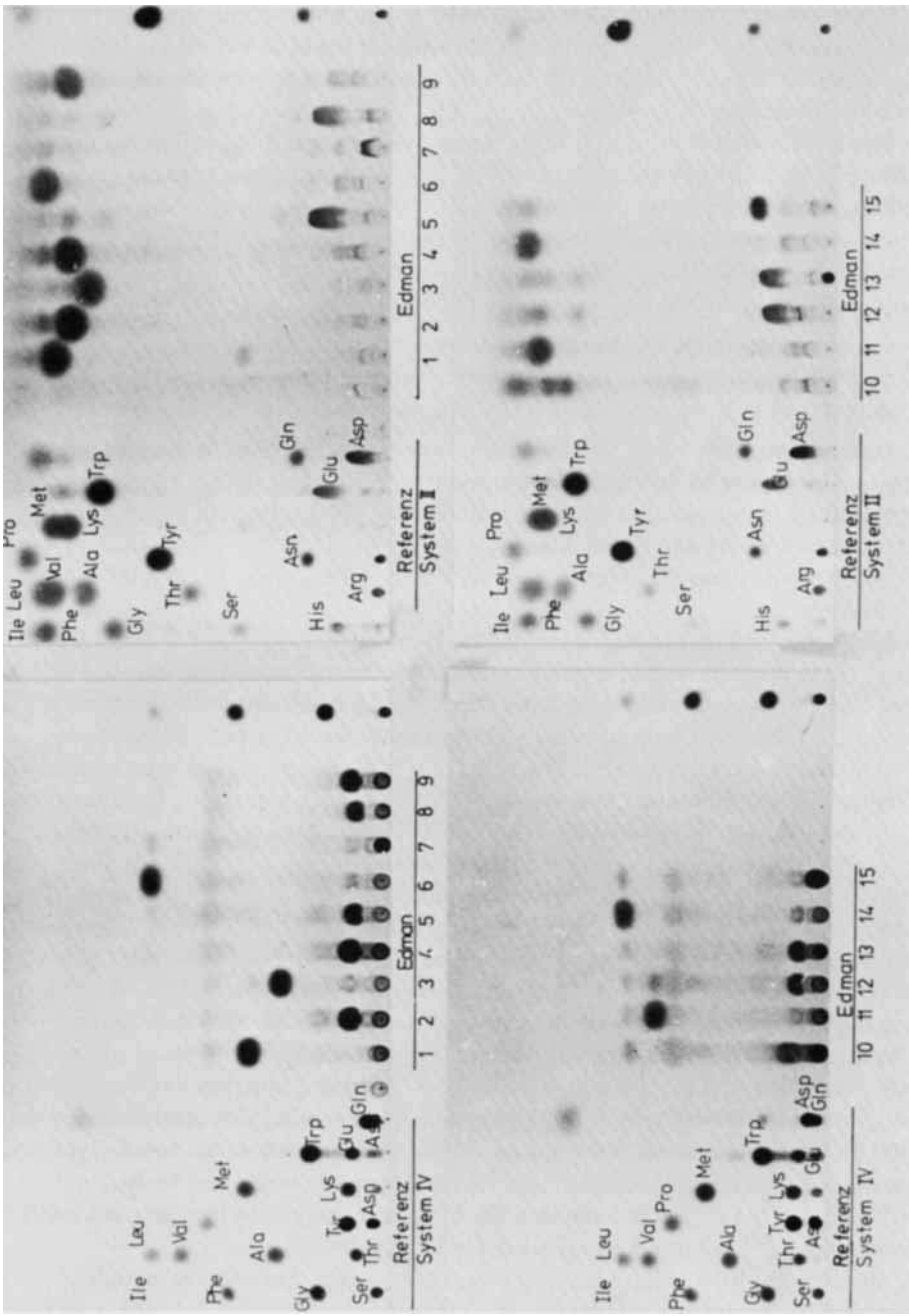
The information from amino acid analysis was very useful as confirmation for identification by thin-layer chromatography, although no distinction between glutamine and glutamic acid or asparagine and aspartic acid is possible. Negative results were obtained for serine and threonine as well as for histidine, proline, tyrosine, and methionine beyond approximately step 20. This was due mainly to the low amount of proteins available for the sequenator runs and the bad recovery of these amino acids after hydrolysis.

Gas chromatography was not used as a routine method for identification since ribosomal proteins are rather rich in polar and basic amino acids whose PTH derivatives can only badly or not at all be determined by gas chromatography.

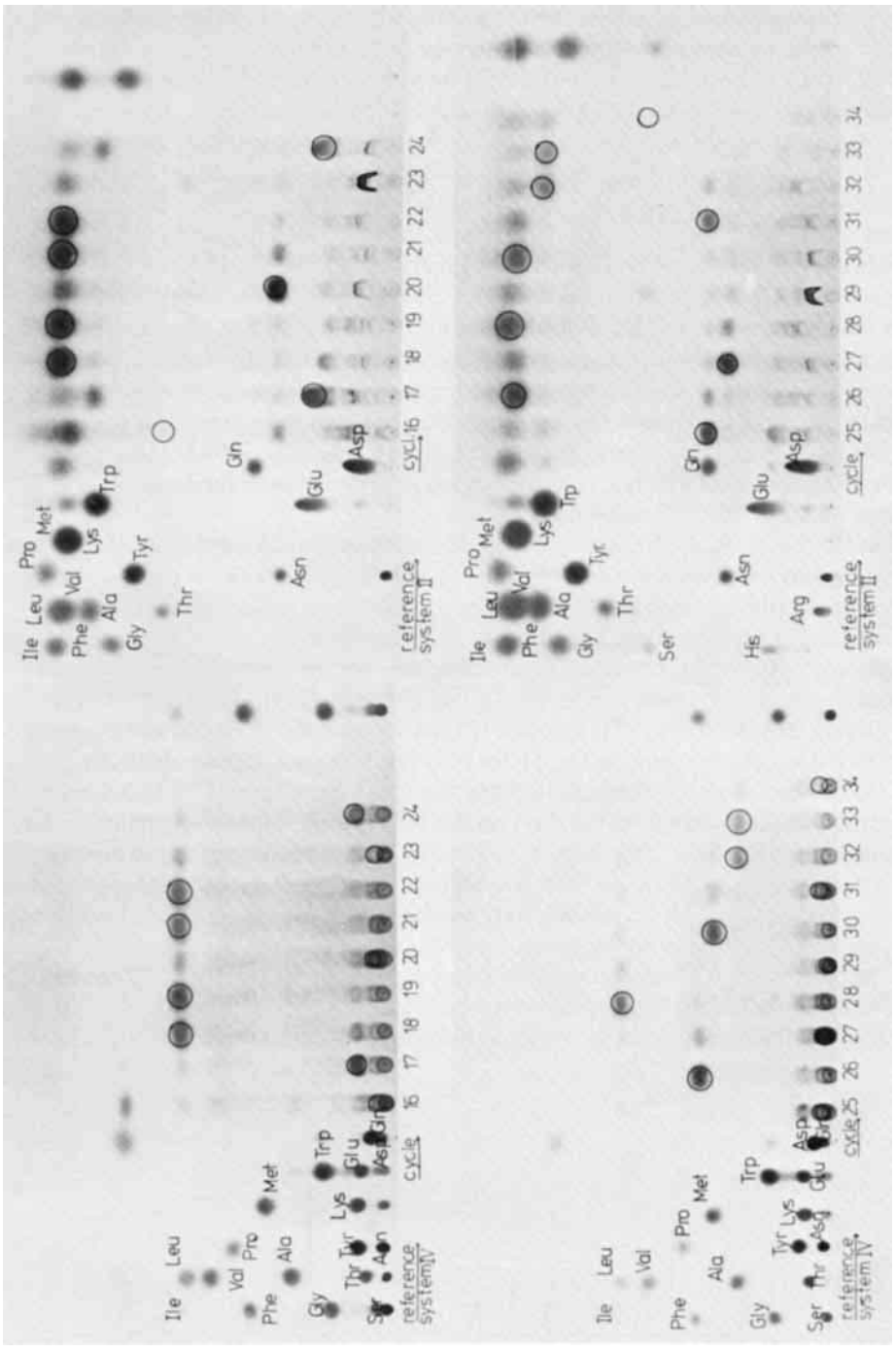
Mass spectrometry could be used without the need to subject the samples to gas chromatography prior to their analysis in the mass spectrometer. This is an independent proof that the PTH amino acids obtained from our sequenator without the usual extraction procedure are pure enough, and that the background from the sequenator is low. Otherwise it would make identification by mass spectrometry very difficult if not impossible.

Mass spectrometry is a very useful method for identification for the following reasons: (a) In contrast to thin-layer chromatography or gas chromatography it does not depend on Rf values or on retention times which vary considerably with the concentration of the substances subjected to chromatography. Since the yield of the various PTH amino acids decreases with the increasing length of the run, it is sometimes difficult to differentiate between some amino acids by thin-layer chromatography (e.g., between PTH-His and PTH-Glu or between PTH-Asp and PTH-Arg). Mass spectrometry, on the other hand, gives absolute values for the molecular and fragment ions and is a direct proof for the PTH amino acids to be identified. Therefore this identification is less subjective and more reliable than other methods, especially at high degradation cycles. (b) Mass spectrometry is more sensitive than other methods. Under our conditions (see below) 0.02 μg of a PTH amino acid can be easily detected. Only 2–6% of the sample is sufficient even at high degradation cycles, whereas 50–70% is necessary for thin-layer chromatography. (c) The mass spectrometer combined with a data system allows two kinds of subtractions, namely that of the background caused by solvents and reagents (determined in the "wash-cycle"), and subtraction of the overlap due to incomplete degradation during the cycle(s) before the one under investigation. In this way, and in contrast to other methods, a quantitative and objective measurement for the degree of overlap and background, i.e., the efficiency of the sequenator run, can be obtained. Examples to illustrate this point are given below and published elsewhere (8) in more detail.

By using the mass spectrometer, it soon became clear that without connection to a data system the characteristic molecular and fragment ions cannot be detected in sufficient intensities for all PTH amino acids since not enough spectra could be registered before the ion cloud disappeared or too much fragmentation occurred. When the experiments were repeated with the data system it turned out that at least 20 spectra had to be taken


Ile Leu Val Phe Met Ala Gly Trp Tyr Lys Ser Thr Asp Glu Gln
Referenz
System IV
Edman
1 2 3 4 5 6 7 8 9
10 11 12 13 14 15
Ile Leu Pro Val Met Phe Ala Gly Thr Tyr Ser Asn His Gln Glu Asp Arg Lys Trp
Referenz
System II
Edman
1 2 3 4 5 6 7 8 9
10 11 12 13 14 15

9a



9b

Fig. 9. Thin-layer chromatograms of cycles 1–34 of proteins L29 developed in the two solvent systems II and IV (see reference1).

from each degradation cycle with a rapid scan rate. A temperature program was used to obtain all characteristic ions in sufficient intensities without using too big an aliquot of the sample. The number of spectra depends on the speed at which the temperature is raised (25°C to 190°C in about 4 min) and on the scan rate. When low scan rates were used some molecular ions of the PTH amino acids could not be detected, and when the temperature was raised too quickly a distinction between glutamic acid or serine and glutamine or lysine was not possible. However, under our conditions for measurement (see Methods) all PTH amino acids can be identified, and an unambiguous distinction between glutamic acid and serine whose fragment ion (m/e:204) appears in good intensities at different times is possible. Even small quantities of the molecular ion of glutamic acid (m/e:264) can be found (see Figs. 5a, b). Furthermore a distinction between lysine and glutamine is easy: the molecular ion for PTH-Gln (m/e:263) appears about four to five spectra earlier than the characteristic fragment ion of PTH-Lys (see Fig. 3 and reference 8). Identification of PTH-Gly in the sequenator cycles is also possible since its molecular ion (m/e:192) appears earlier during the temperature program than the fragment ion (m/e:192) which derives, for instance, from PTH-Leu or PTH-Ile. Distinction between PTH-Leu and PTH-Ile is made by the relative intensities of their specific fragment ions 205 and 219, respectively.

In the following, some examples for the identification of degradation cycles by mass chromatograms are given. More details will be published elsewhere (8). A mass chromatogram (the so-called D plot) for the molecular ion of proline (m/e:232) is presented in Fig. 10 for protein L2, cycles 7–26. It can clearly be seen that proline occurs at positions 7, 10, and 21. Other mass chromatograms are given in Figs. 7 and 8 for protein L22 (cycles 1–20): mass numbers 234 (PTH-Val), 206 (PTH-Ala), 218 (PTH-Thr minus H_2O), 248 (PTH-Leu, PTH-Ile), 266 (PTH-Met), and 263 (PTH-Gln and fragment ion of PTH-Lys); furthermore, in Fig. 11 for PTH-Thr in protein L25 (cycles 1–9).

As mentioned above, the background in the "wash cycle" caused by reagents and solvents can be subtracted from the main spectra (LM spectra) of the degradation cycles resulting in the LS spectra. This is illustrated in Figs. 3–5 for several cycles of protein L29, namely, cycles 9 (Lys), 20(Asn), and 24 (Glu). In these figures the LM spectra are plotted together with their LS spectra. As is evident from these examples, the background by reagents and solvents in our sequenator is rather low.

Another method of subtraction is the following: From the spectra of a given degradation cycle (e.g., the 45th) the corresponding spectra of the cycle before (i.e., the 44th) is subtracted by the data system. The resulting difference spectra are shown in Fig. 12 for

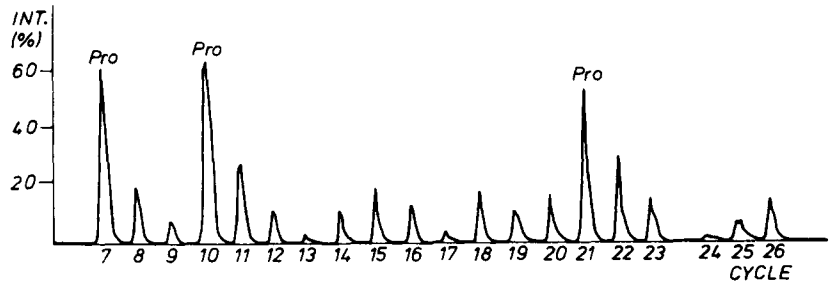


Fig. 10. Mass chromatogram plots of cycles 7–26 of protein L2 for the molecular ion m/e = 232 (PTH-Pro).

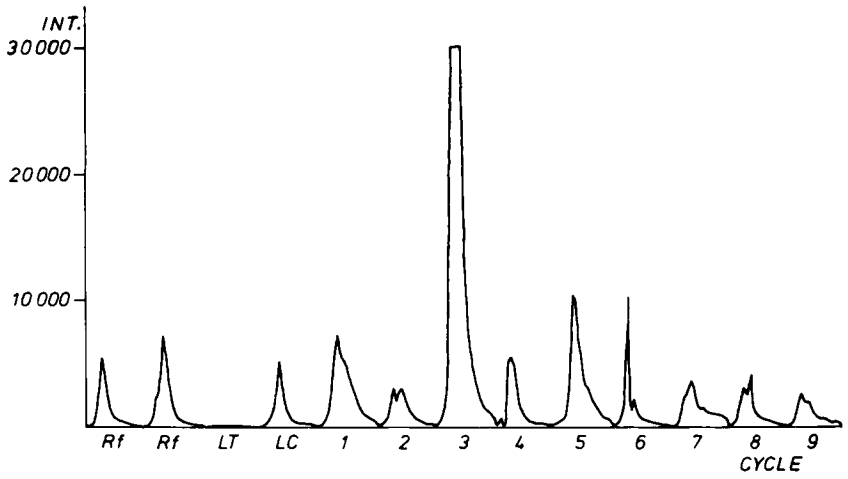


Fig. 11. Mass chromatogram plots of cycles 1–9 of protein L25 for the molecular ion m/e = 218 (PTH-Thr minus H_2O).

cycles 28–30 (Pro, Phe, Ala) of protein L2, and in Fig. 13 for cycle 14 (Ser) of protein L17. The LS spectra are especially useful to eliminate overlaps at high cycles by subtraction. In Fig. 6 examples of LS spectra from higher degradation cycles of protein L29, namely, cycles 33 (Ala) and 37 (Leu), are presented. In this way it was possible to determine amino acid sequences beyond the regions elucidated by thin-layer chromatography as well as by hydrolysis of PTH amino acids and amino acid analysis.

Comparison of Sequences

It is noteworthy that 13 out of the 14 proteins listed in Fig. 1 begin with methionine or alanine (seven and six times, respectively) and that two proteins (L1 and L5) have an identical sequence (Ala-Lys-Leu) for the first three positions. The sequence Met-Ile at the N-terminus is present in two proteins (L3 and L23) and Met-Lys in proteins L13 and L29. In some proteins there are regions with a frequent occurrence of basic (but no acidic) amino acids, e.g., positions 2–11 in protein L1; 4–17 in L2; 2–31 in L17; 2–47 in L20; 6–18 in L22; and 3–13 in L27. On the other hand, regions with acidic amino acids, e.g., positions 5–50 in L12 and 12–17 in L29, are much less frequent (see the distribution of acidic and basic amino acids in Fig. 14).

Comparing the amino acid sequences so far known for the proteins of the 50S subunit (Fig. 1) with one another and with sequences of the 30S proteins (2), one does not find extensive homologues among them. There are only rather short regions (two–four residues) with identical sequences in different proteins. Some examples of homologous regions are listed in Table II and in reference 18. A somewhat more extensive degree of homology exists between proteins L12 and S6. This is presented elsewhere (10) in more detail. A comparison of the primary structures of E. coli ribosomal proteins will be made when our studies on the N-terminal regions of all proteins of E. coli ribosomes are completed in the near future.

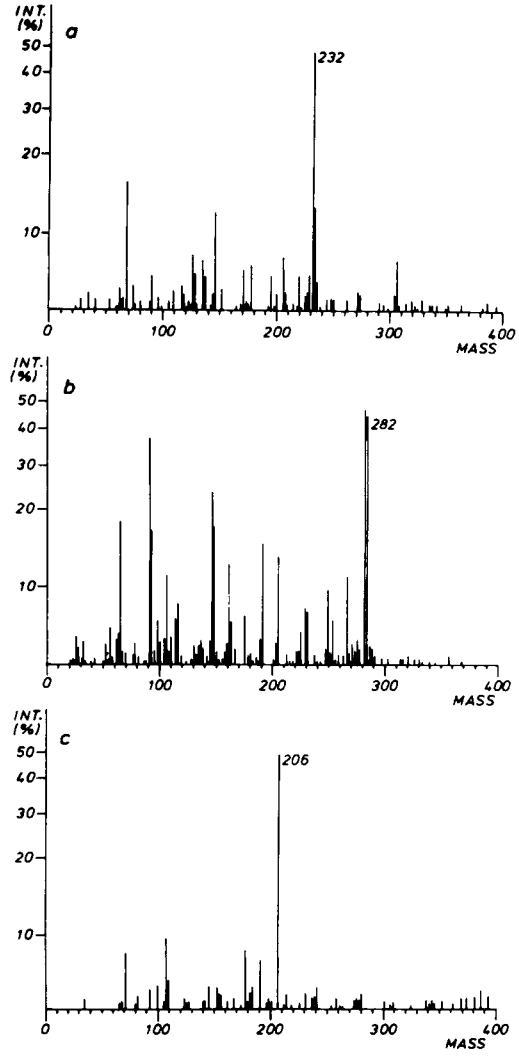


Fig. 12. LS spectra of cycle 28 (PTH-Pro), 29 (PTH-Phe), and 30 (PTH-Ala) of protein L2. For each cycle the corresponding spectrum of the preceding cycle was subtracted as described in Fig. 6. Molecular ion m/e = 232 is characteristic for PTH-Pro, m/e = 282 for PTH-Phe, and m/e = 206 for PTH-Ala.

ACKNOWLEDGMENTS

We thank Mr. H. Graffunder and Mr. H. Kohls for their excellent technical help in this work, which was supported by a grant from the Deutsche Forschungsgemeinschaft.

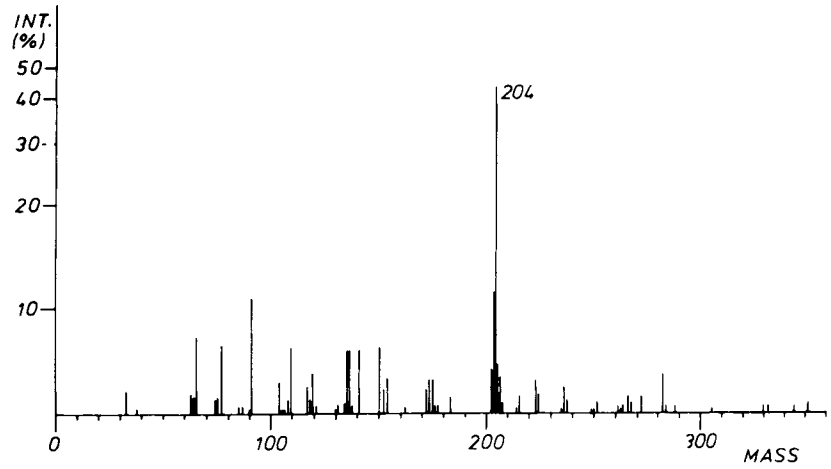


Fig. 13. LS spectrum of cycle 14 (PTH-Ser) of protein L17. The main LM spectrum (number 12) of cycle 14 minus LM spectrum (number 12) of cycle 13 results in the molecular ion m/e = 204 characteristic for PTH-Ser minus H_2O.

TABLE II. Some Homologous Regions in E. Coli 50S Ribosomal Proteins

Protein	Positions	Sequence
L5	9–12	Asp-Glu-Val-Val
L24	8–11	Asp-Glu-Val-Ile
L5	30–35	Val-Glu-Lys-Ile-Thr-Leu
L24	35–40	Val-Glu-Gly-Ile-Asn-Leu
L22	46–49	Lol-Val-Lys-Lys
L24	40–43	Leu-Val-Lys-Lys

REFERENCES

1. Wittmann–Liebold, B., Hoppe–Seyler's Z. Physiol. Chem. 354:1415–1431 (1973).
2. Wittmann–Liebold, B., FEBS Lett. 36:247–249 (1973).
3. Edman, P., and Begg, G., Eur. J. Biochem. 1:80–91 (1967).
4. Graffunder, H., Kohls, H., and Wittmann–Liebold, B., manuscript in preparation.
5. Hindennach, I., Kaltschmidt, E., and Wittmann, H.G., Biochemistry 23:12–16 (1971).
6. Kaltschmidt, E., and Wittmann, H. G., Anal. Biochem. 36:401–412 (1970).
7. Hagenmaier, H., Ebbighausen, W., Nicholson, G., and Vötsch, W., Z. Naturforschung 25b:681–689 (1970).
8. Wittmann–Liebold, B., Graffunder, H., and Schäfer, W., submitted for publication.
9. Reinbolt, J., and Schiltz, E., FEBS Lett. 36:250–252 (1973).
10. Hitz, H., Schäfer, D., and Wittmann–Liebold, B., FEBS Lett. 56:259–262 (1975).
11. Stadler, H., FEBS Lett. 48:114–116 (1974).
12. Chen, R., and Wittmann–Liebold, B., FEBS Lett. 52:139–140 (1975).
13. Terhorst, C., Möller, W., Laursen, R., and Wittmann–Liebold, B., Eur. J. Biochem. 34:138–152 (1973).

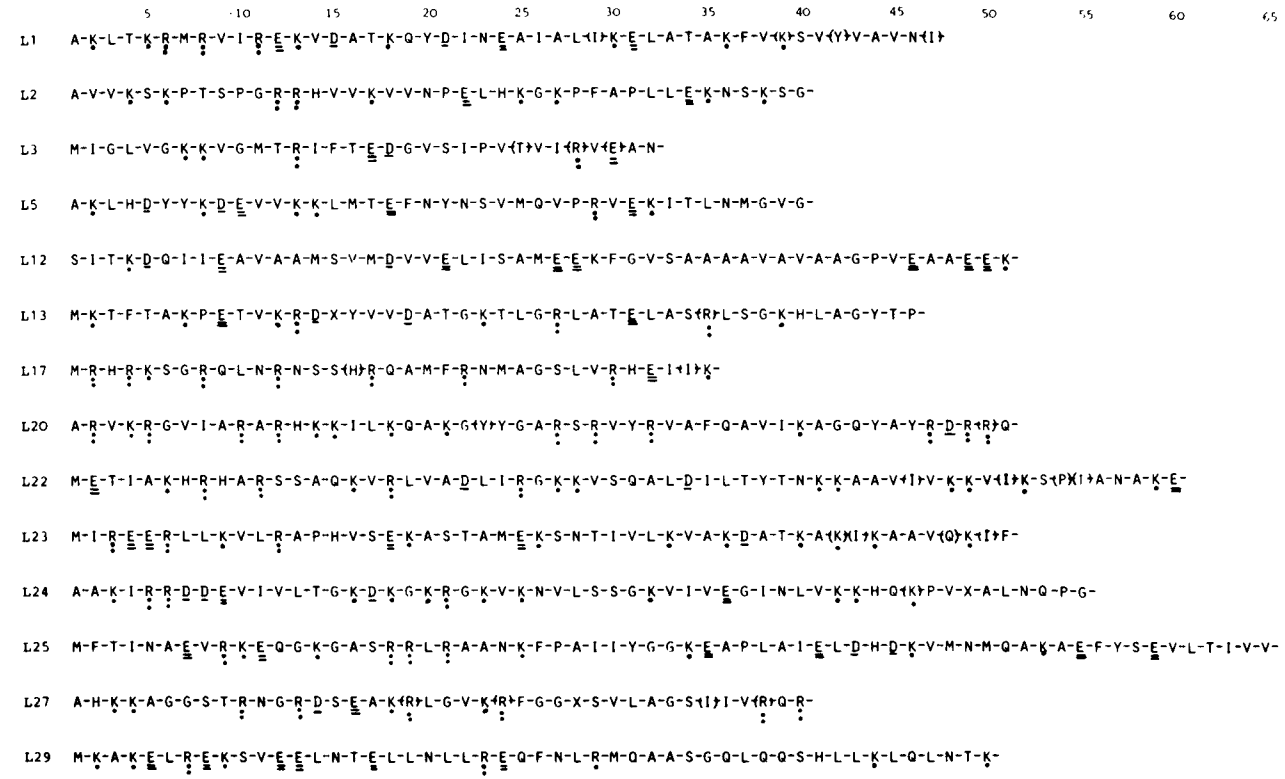


Fig. 14. Comparison of the N-terminal regions of the proteins described in this paper. Abbreviation of amino acids is according to reference 19. The acidic amino acids are underlined (D = Asp, E = Glu); lysine (=K) and arginine (R) are marked by dots.

14. Bitar, K. G., and Wittmann–Liebold, B., Hoppe–Seylers's Z. Physiol. Chem. 356:1343–1352 (1975).
15. Bitar, K. G., Biochem. Biophys. Acta 386:99–106 (1975).
16. Tischendorf, G., Stöffler, G., and Wittmann, H. G., personal communication.
17. Stöffler, G., in "The Ribosome." M. Nomura, A. Tissières, and P. Lengyel (Eds.). Cold Spring Harbor Laboratory Monograph Series 615–667 (1974).
18. Wittmann, H. G., and Wittmann–Liebold, B., in "The Ribosome" M. Nomura, A. Tissières, and P. Lengyel (Eds.). Cold Spring Harbor Laboratory Monograph Series 115–140 (1974).
19. Dayhoff, M. O., "Atlas of Protein Sequence and Structure," National Biomedical Research Foundation, Washington, D. C.