A NOVEL EDMAN-TYPE DEGRADATION: DIRECT FORMATION OF THE THIOHYDANTOIN RING IN ALKALINE SOLUTION BY REACTION OF EDMAN-TYPE REAGENTS WITH *N*-MONOMETHYL AMINO ACIDS

Jui Yoa CHANG

Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, D-1000 Berlin-Dahlem, Germany

Received 3 May 1978

1. Introduction

The stepwise degradation of peptide and protein by the Edman-type reagents [1,2] comprises two separate reactions:

1. The coupling of isothiocyanate with the N-terminal amino acid to form a thiocarbamyl derivative in alkaline solution.

2. The cleavage of the thiocarbamyl amino acid to form a thiazolinone derivative in a strong acidic solution.

As the yields from the coupling and cleavage reactions do not normally reach 100% at every degradation cycle, overlapping by the preceding amino acid residues thus occurs during the extended sequence determination.

In the recent work on sequence determination of ribosomal proteins with *N*-monomethyl amino acids at the N-termini [4–6], an unusual pre-overlapping (i.e., overlapping by the succeeding amino acid residue instead of the preceding amino acid residue) has been observed. To explain this result, an isopeptide bond at the N-terminal sequence of ribosomal protein S11 has been proposed [6].

Abbreviations: PITC, phenylisothiocyanate; PTC, phenylthiocarbamyl; PTH, phenylthiohydantoin; DABITC, 4-*NN*-dimethylaminoazobenzene 4′-isothiocyanate; DABTC, 4-*NN*-dimethylaminoazobenzene 4′-thiocarbamyl; DABTH, 4-*NN*-dimethylaminoazobenzene 4′-thiohydantoin; NM, *N*-monomethyl

Studies of the reactions of several *N*-monomethyl amino acids (both as free amino acids and as N-terminal amino acids of ribosomal proteins S11, L16, L33 and IF-3 factor) with PITC and the new Edman-type reagent DABITC [3,8] are presented here. A direct formation of the thiohydantoin ring by reaction of DABITC or PITC with *N*-monomethyl amino acids (either as free amino acids or N-termini of proteins) in alkaline solution has been found. This novel Edman-type degradation accounts for the unusual pre-overlapping observed during the sequence determination of ribosomal proteins S11, L16, L33 and IF-3 factor with *N*-monomethyl amino acids as the N-terminal residues.

2. Materials and methods

Ribosomal proteins S11, L16, L33 and initiation factor IF-3 from *E. coli* used in this report were provided by Professor H. G. Wittmann. Their identity and purity were checked by two-dimensional polyacrylamide gel electrophoresis [7]. *N*-Monomethyl alanine (NMAla), *N*-monomethyl leucine (NMLeu), *N*-monomethyl valine (NMVal) and *N*-monomethyl glutamic acid (NMGlu) were obtained from Serva, Heidelberg, or Sigma, MO and 4-*NN*-dimethylaminoazobenzene-4′-isothiocyanate was prepared as in [3,8].

2.1. *Coupling of DABITC with* N-*monomethyl amino acids*

DABITC (100 nmol) was allowed to react with 2 μmol NM-amino acid in 150 μl 67% aqueous

pyridine or triethylamine/acetic acid buffer, pH 10.3 [9], at 52°C for 1 h. The mixture was dried in vacuo and the coupling product was extracted into 300 μl butyl acetate from 400 μl of 0.2 M $NaHCO_3$. The butyl acetate extract was dried and redissolved in suitable volumes of absolute ethanol for separation on a polyamide sheet [9] and silica gel plate [10], ultraviolet absorption measurement (in a Gilford spectrophotometer 250) and for identification by mass spectrum (in a Varian CH7 mass spectrometer). In a separate reaction, the dried butyl acetate extract was treated with 40 μl water and 80 μl acetic acid/saturated with HCl at 52°C for 50 min. The sample was dried, redissolved in absolute ethanol and characterized by thin-layer chromatography and ultraviolet absorption in order to compare it with the original coupling product.

2.2. *Measurement of the change in ultraviolet absorption during the coupling reaction*

Two methods were applied:

1. NM-amino acid (1 μmol), dissolved in 0.75 ml ethanol and 0.75 ml 0.2 M $NaHCO_3$, was left to react with 50 nmol DABITC in an optical cuvette at room temperature. The change in ultraviolet absorption was recorded against a blank solution without addition of DABITC.

2. NM-amino acid (2 μmol) was left to react with 200 nmol DABITC in 600 μl 67% aqueous pyridine at 52°C. Equal samples (50 μl) were removed at time intervals, dried thoroughly in vacuo and redissolved in 1.5 ml absolute ethanol for the ultraviolet measurements.

2.3. *Sequence determination*

The N-terminal sequences of proteins S11, L16, L33 and IF-3 (2–3 nmol) were determined by a combined DABITC–PITC double-coupling method. This method includes:

(i) A first coupling with DABITC and a second coupling with PITC to complete the reaction.
(ii) The cleavage with anhydrous trifluoroacetic acid.
(iii) Conversion with water/acetic acid saturated with HCl.

(A more detailed report of this manual microsequencing method will appear elsewhere.) The excess reagent and by-products were extracted with *n*-heptane/ethyl acetate (2:1, v/v) after the coupling reaction. The thiazolinone derivatives released were extracted into butyl acetate. They were then converted to the thiohydantoin derivatives in aqueous acid. Identification of DABTH–amino acids was carried out on 2.5 × 2.5 cm polyamide sheets (Schleicher und Schüll) [3,9]. A blue synthetic marker, DABTC–diethylamine (10 pmol), was used in each identification to give a quick discrimination between the DABTH–amino acids. The DABTH–amino acids could be separated on a silica-gel plate [10] and extracted into absolute ethanol for quantitative analysis.

3. Results

3.1. *Characterization of the coupling products of DABITC with NM-amino acid in alkaline solution*

DABITC is a modified Edman reagent [8]. The major advantage of this reagent is that it forms DABTC–amino acids and DABTH–amino acids with different colours on thin-layer chromatograms after exposure to HCl vapour. Accordingly, the quantitative conversion of DABITC (purple) to DABTC derivative (blue) and then to DABTH derivative (red) could be followed by observing the coloured spots appearing on the thin-layer plate. For common amino acids, the coupling resulted in a blue DABTC–amino acid on the thin-layer plate. The coupling reaction between DABITC and the NM-amino acid, however, gave a red-coloured product. To make sure that the red colour does not result from the existence of *N*-methyl group in a thiocarbamyl structure, the coupling product was subjected to the acid treatment (see section 2) which was normally used to convert a DABTC–amino acid to a DABTH–amino acid. This acid treatment did not change the chemical structure of the coupling product, as judged by thin-layer chromatography (fig.1) and ultraviolet absorption (fig.2), which indicated that a thiohydantoin ring had already formed in the coupling buffer. Mass spectral identification also shows the peaks of DABTH–NMAla (m/e=367), DABTH–NMLeu (m/e=409), DABTH–NMVal (m/e=395) and DABTH–NMGlu (m/e=425) for the coupling products.

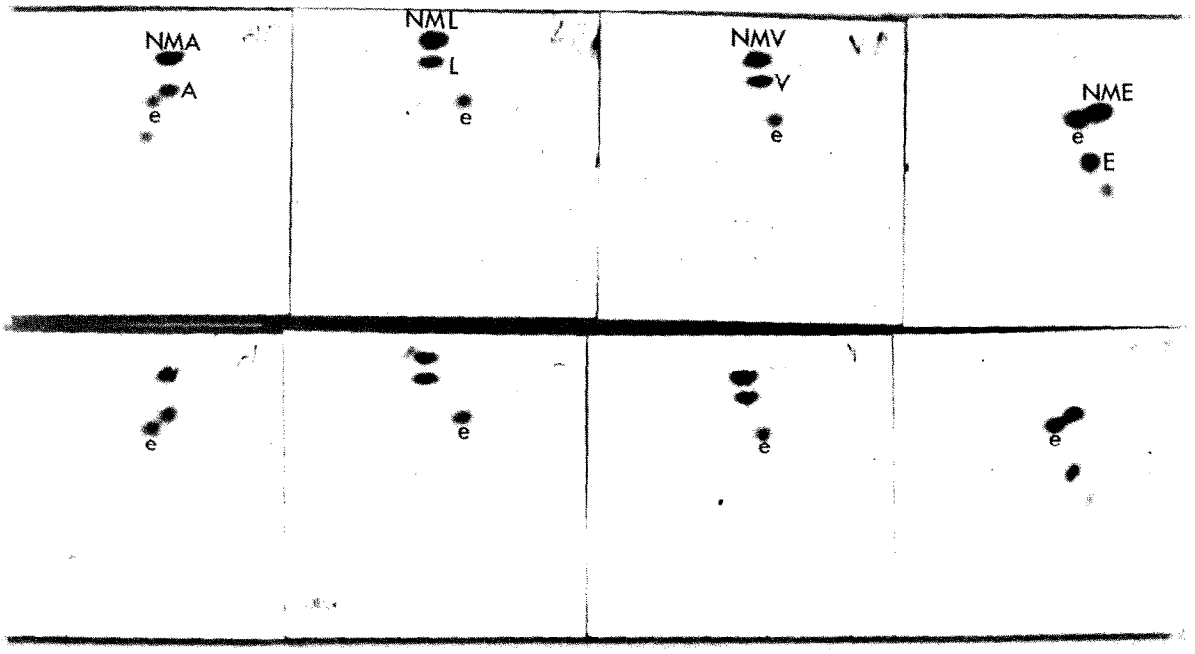

Fig.1. Characterization of the coupling products (first row) and the acid-treated coupling products (second row) of DABITC with NMAla (NMA), NMLeu (NML), NMVal (NMV) and NMGlu (NME) on polyamide sheets. The DABTH derivatives of alanine (A), leucine (L), valine (V) and glutamic acid (E) were chromatographed with their methylated derivatives to show their relative positions. The acid treatment does not change the chemical nature of the coupling products as judged by their R_F values and colour (both coupling products and acid-treated coupling products appeared as red coloured spots). It is concluded that the thiohydantoin structure had already been formed in the coupling products. 'e' is an authentic blue marker, DABTC–diethylamine. The other minor spots (also red colour), which could arise from the impurities of the samples, were not identified. Solvents used were 33% acetic acid (first dimension) and toluene/*n*-hexane/acetic acid (2:1:1, v/v/v) (second dimension).

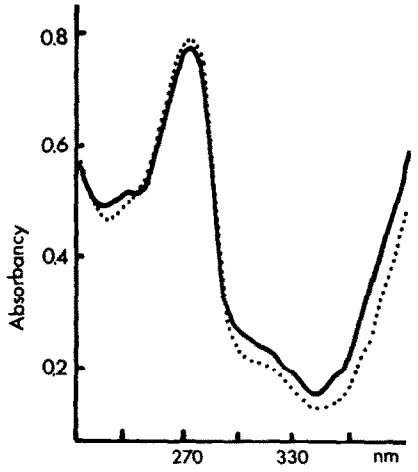

Fig.2. Ultraviolet absorption characteristic of the coupling product (solid) and the acid-treated coupling product (dotted) of DABITC and NMLeu. The absorption peak at 270 nm indicates that a thiohydantoin ring exists in both, coupling product and acid-treated coupling product.

3.2. *Measurement of the change in ultraviolet absorption during the coupling of DABITC and NM-amino acid in alkaline solution*

In order to give further evidence that a thiohydantoin ring did indeed form in the coupling buffer, experiments were designed to measure the ultraviolet absorption change in the region 265–270 nm (a characteristic absorption peak of thiohydantoin structure) during the coupling reaction. Figure 3A shows the ultraviolet absorption change at 264 nm

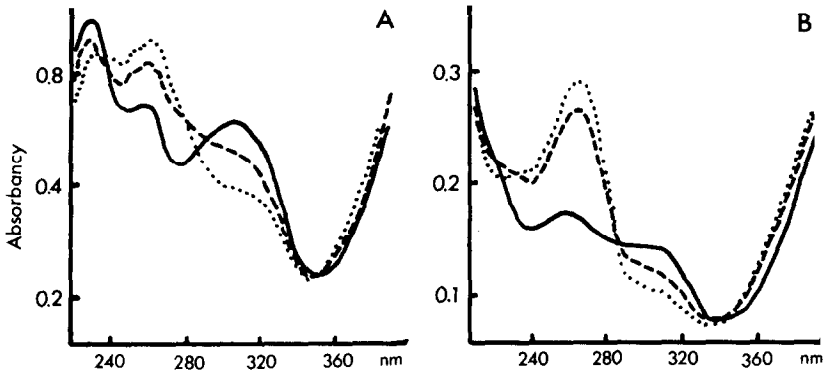

Fig.3. The change in ultraviolet absorption spectra during the coupling of DABITC with NMAla in sodium bicarbonate buffer (A) for 2 min (solid), 15 min (broken) and 40 min (dotted), and in aqueous pyridine (B) for 5 min (solid), 20 min (broken) and 40 min (dotted). For the experimental details see section 2.

during the coupling of DABITC and NMAla in the sodium bicarbonate buffer according to expt 1 (see section 2.2.). Figure 3B shows the increase in ultraviolet absorption at 270 nm during the coupling of DABITC and NMAla by expt 2 (see section 2.2.).

3.3. *Sequence determination of proteins S11, L16, L33 and IF-3*

It is interesting to see whether the direct formation of the DABTH–amino acid also proceeds when the NM-amino acid is located at the N-terminus of a protein. Ribosomal proteins S11, L16, L33 and initiation factor IF-3 with NMAla (S11 and L33) and NMMet (L16 and IF-3) as the N-termini were subjected to sequence determination by the micro-sequencing DABITC–PITC double-coupling method. The N-terminal sequences of these proteins have been determined as:

S11	NMAla–Lys–Ala–Pro–Ile–	[6]
L33	NMAla–Lys–Gly–Ile–Arg–	[5]
L16	NMMet–Leu–Gln–Pro–Lys–	[11]
IF-3	NMMet–Lys–Gly–Gly–Lys–	[12]

For S11 (fig.4; first row), 92% of the N-terminal NMAla was cleaved off (determined quantitatively from the third degradation, see section 2) to form DABTH–NMAla at the coupling reaction of the first degradation and it was extracted into the *n*-heptane/ethyl acetate solution together with excess DABITC, PITC and by-products (fig.4; sheet no. 0 of S11). As a result, the second amino acid residue lysine also reacted with DABITC (and PITC) in the coupling reaction of the first degradation and appeared after the first cleavage. The expected residue therefore appeared one step earlier during the following degradations. This result is consistent with that obtained [6] during the sequence determination of the N-terminal tryptic peptide from S11. For L33 (fig.4; second row), 67% of NMAla was cleaved off (determined from the third degradation) during the coupling of the first degradation and was found in the *n*-heptane/ethyl acetate extract (sheet no. 0 of L33). Hence, during the extended degradations, the expected residues were pre-overlapped by an almost double amount of their succeeding residues. For L16 (fig.4; third row), a small amount (22%, determined from the second degradation) of the N-terminal NMMet was cleaved off during the coupling in the first degradation. The amount of pre-overlapping by the succeeding residues could be seen from the second (Leu and Gln) and third (Gln and Pro) degradations. For IF-3, an estimated 50% of the N-terminal NMMet was cleaved off at the coupling in the first degradation and found in the *n*-heptane/ethyl acetate extract. The existence of a minor form of IF-3 and the isolation of the N-terminal tryptic peptide of this minor

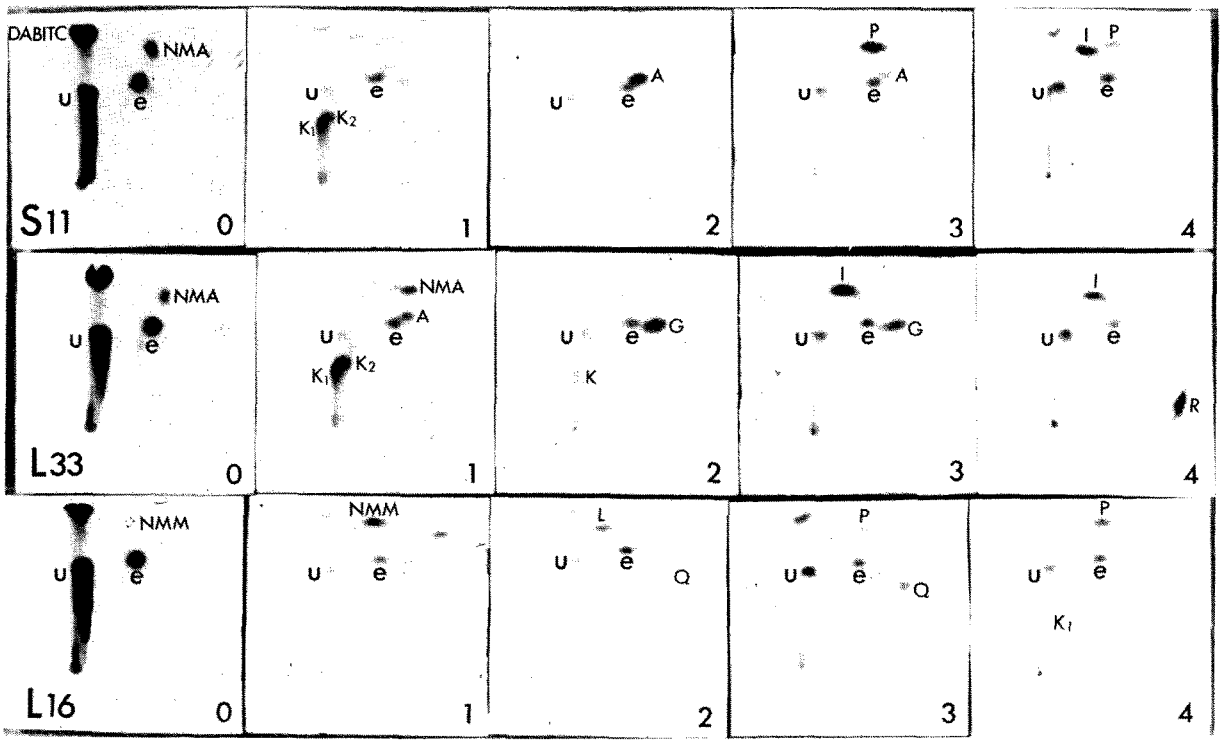

Fig.4. N-Terminal sequences determination of proteins S11, L33 and L16 by the DABITC–PITC double coupling method. The numbers indicate the number of degradation cycles. No. 0 sheets were the *n*-heptane/ethyl acetate extract of the excess reagent and by-products after the coupling reactions of the first degradations. The expected N-terminal sequences were: NMAla–Lys–Ala–Pro–Ile– (S11); NMAla–Lys–Gly–Ile–Arg– (L33); and NMMet-Leu–Gln–Pro–Lys– (L16). A = DABTH–Ala, P = DABTH–Pro, G = DABTH–Gly, I = DABTH–Ile, L = DABTH–Leu, Q = DABTH–Gln, R = DABTH–Arg, NMA = DABTH–NMAla and NMM = DABTH–NMMet, all appeared as red coloured spots. K_1 = α-DABTH–ε-DABTC–Lys (purple) and K_2 = α-PTH–ε-DABTC–Lys (blue). 'U' is an unextracted by-product (a blue coloured thiourea formed by coupling of PITC with hydrolyzed DABITC). 'e' is the marker, DABTC–diethylamine. Solvents used are given in fig.1 legend.

IF-3 with the absence of N-terminal NMMet has also been reported [12].

It appeared that for a NM-amino acid sited at the N-terminus of a protein, the completeness of the cleavage to form DABTH–NM-amino acid in the alkaline coupling solution depends not only on the NM-amino acid itself but also on the protein (from a given isolation procedure). Thus, although S11 and L33 have the same N-terminal sequence NMAla–Lys–, the extent of DABTH–NMAla formation in the coupling solution has been found to be different. Using the PITC single coupling method, a PTH–NMAla was also found for S11 and L33 in the *n*-heptane/ethyl acetate extract after the first coupling reaction (fig.5).

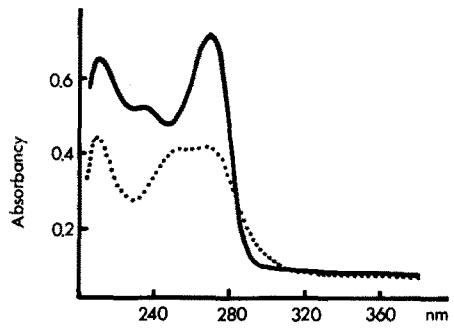

Fig.5. Ultraviolet absorption spectra of the coupling products of PITC and Ala (dotted) which resulted in PTC–Ala, and PITC and NMAla (solid) which resulted in PTH–NMAla. The coupling solution is 67% aqueous pyridine.

4. Discussion

For the first time a direct formation of the thiohydantoin ring in alkaline solution has been observed by reaction of DABITC (or PITC) with free *N*-monomethyl amino acids. Although the formation of a thiohydantoin ring must occur by way of the thiocarbamyl derivative, this intermediate has not been detected by thin-layer chromatography during the coupling reaction. This indicates that cyclization of a DABTC–NM-amino acid to form a DABTH–NM-amino acid in the coupling solution is fast and is not a rate-determining step. The formation of a DABTH–NM-amino acid in the mild alkaline solution (and in aqueous pyridine) reflects the fact that a DABTH–NM-amino acid must possess an especially stable thiohydantoin structure compared with normal DABTH–amino acids.

The direct formation of a DABTH–NM-amino acid in the coupling solution was also observed when the NM-amino acid is located at the N-terminal end of a protein. This abnormal Edman-type degradation which resulted in the unusual pre-overlapping has been observed during the sequence determinations of ribosomal proteins S11 [4,6] and L33 (Wittmann-Liebold, personal communication). A presumption of two polypeptide chains differing only in their N-terminal residue had to be ruled out when the N-terminal tryptic peptide of S11 was isolated and that Edman degradation worked similarly on this hexapeptide as on the intact S11 protein [6]. They proposed the existence of an isopeptide bond between the carboxyl group of the N-terminal NMAla and the ϵ-amino group of lysine. They concluded that the α-amino group of lysine is also reactive toward the Edman degradation (but with incomplete reaction) and pre-overlapping by the succeeding residue therefore occurred. According to their model, a PTH–NMAla and an α-PTH–Lys should be detected after the first Edman degradation. However, spots corresponding to α-PTH–ϵ-PTC–Lys in addition to PTH–NMAla were found [6]. A similar result was also observed by the DABITC–PITC double-coupling method: two major spots (α-DABTH–ϵ-DABTC–Lys and α-PTH–ϵ-DABTC–Lys) and a minor spot (DABTH–NMAla) were found after the first degradation (see fig.4; no. 1 sheet of S11). It is obvious that, owing to the direct cleavage of the N-terminal amino acid (NMAla) in the coupling reaction, both α- and ϵ-amino groups of the second amino acid (lysine) were open to react with DABITC and PITC during the coupling reaction of the first degradation cycle.

This novel Edman-type degradation could be a challenge to the well established phenylisothiocyanate (Edman) method. In addition, since the cleavage to form a thiohydantoin ring proceeds easily in the mild alkaline solution (or in aqueous pyridine), further developments in methods of protein sequence determination might be made which exploit this behaviour.

Acknowledgements

I am grateful to Drs H. G. Wittmann and B. Wittmann–Liebold for their interest and faithful encouragement during this work. I thank Dr D. Brauer and Dr P. Woolley for discussion.

References

[1] Edman, P. (1956) Acta Chem. Scand. 10, 761–768.
[2] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[3] Chang, J. Y. (1977) Biochem. J. 163, 517–520.
[4] Chen, R., Brosius, J., Wittmann-Liebold, B. and Schäfer, W. (1977) J. Mol. Biol. 111, 173–181.
[5] Wittmann-Liebold, B. and Pannenbecker, R. (1976) FEBS Lett. 68, 115–118.
[6] Chen, R. and Chen-Schmeisser, U. (1977) Proc. Natl. Acad. Sci. USA 74, 4905–4908.
[7] Kaltschmidt, E. and Wittmann, H. G. (1970) Anal. Biochem. 36, 401–412.
[8] Chang, J. Y., Creaser, E. H. and Bentley, K. W. (1976) Biochem. J. 153, 607–611.
[9] Chang, J. Y. and Creaser, E. H. (1977) J. Chromatog. 132, 303–307.
[10] Chang, J. Y., Creaser, E. H. and Hughes, G. J. (1977) J. Chromatog. 140, 125–128.
[11] Brosius, J. and Chen, R. (1976) FEBS Lett. 68, 105–109.
[12] Brauer, D. and Wittmann-Liebold, B. (1977) FEBS Lett. 79, 269–275.