Amino Acid Sequence around Lipoic Acid Residues in the Pyruvate Dehydrogenase Multienzyme Complex of *Escherichia coli*

Geoffrey HALE and Richard N. PERHAM

Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, U.K.

(Received 25 February 1980)

Amino-acid sequences around two lipoic acid residues in the lipoate acetyltransferase component of the pyruvate dehydrogenase complex of *Escherichia coli* were investigated. A single amino acid sequence of 13 residues was found. A repeated amino acid sequence in the lipoate acetyltransferase chain might explain this result.

The pyruvate dehydrogenase complex of *Escherichia coli* consists of multiple copies of three types of polypeptide chain; pyruvate decarboxylase (E1) (EC 1.2.4.1), lipoate acetyltransferase (E2) (EC 2.3.1.12) and lipoamide dehydrogenase (E3) (EC 1.6.4.3) (Reed, 1974; Perham, 1975). Lipoate acetyltransferase (E2), which comprises 24 polypeptide chains of apparent M_r 83000, forms the structural core of the complex to which the other components are bound (Reed & Oliver, 1968; Vogel, 1977; Danson *et al.*, 1979). Each E2 chain contains two covalently-bound lipoic acid residues, which become reductively acetylated during the course of the normal enzymic reaction (Danson & Perham, 1976; Bates *et al.*, 1977; Speckhard *et al.*, 1977; Collins & Reed, 1977).

The native complex is readily cleaved by endogenous proteinases (Perham & Thomas, 1971; Vogel *et al.*, 1971; Gebhardt *et al.*, 1978) and by trypsin (Hale & Perham, 1979a), the E2 component being particularly sensitive despite its core position. We have already shown that limited treatment of the complex with trypsin excises fragments of the E2 chain that contain the lipoic acid residues without significantly disrupting the quaternary structure of the complex (Hale & Perham, 1979a,b). This has been confirmed by Bleile *et al.* (1979). However, digestion of the excised fragments with *Staphylococcus aureus* proteinase and isolation of the lipoic-acid-containing peptides by diagonal paper electrophoresis revealed only a single amino-acid sequence of seven residues (Hale & Perham, 1979b). This sequence included the tetrapeptide sequence originally obtained by Daigo & Reed (1962). It is conceivable that these independent investigations of lipoyl-lysine-containing peptides failed to disclose the expected second sequence for unidentified technical reasons (Hale & Perham, 1979b). However, it is also possible that this sequence appears twice in each polypeptide chain.

To try to distinguish these alternatives, we have analysed the E2 chain for lipoyl-lysine-containing peptides by yet another method. We have again found only one amino-acid sequence which is longer than, but includes, the heptapeptide we have previously described (Hale & Perham, 1979b).

Materials and Methods

Pyruvate dehydrogenase complex was isolated from a mutant of *E. coli* K12 constitutive for synthesis of the complex (Danson *et al.*, 1979). It was mixed with a small amount of complex in which the lipoic acid residues were labelled by treatment with *N*-ethyl[2,3- ^{14}C]maleimide in the presence of pyruvate (Brown & Perham, 1976) and subjected to limited proteolysis with trypsin. The lipoyl-lysine-containing fragments of E2 chains were separated from the rest of the complex by gel filtration on Sepharose 6B (Hale & Perham, 1979b). They were oxidized with performic acid (Hirs, 1956) and then further digested in 0.5% (w/v) NH_4HCO_3 , pH 8.0, for 4 h with 2% (w/w) thermolysin (Calbiochem, San Diego, CA, U.S.A.) at 45°C or with 2% (w/w) *S. aureus* proteinase (Miles Laboratories, Stoke Poges, Bucks., U.K.) at 37°C.

Paper electrophoresis, amino acid analysis, amide assignments from electrophoretic mobilities, and dansyl-Edman degradation of peptides were carried out as described by Perham (1978).

Results and Discussion

The lipoic-acid-containing fractions obtained by gel filtration of the trypsin-treated complex (Hale & Perham, 1979b) were pooled and freeze-dried. A sample (10 mg) was oxidized with performic acid to convert the lipoyl groups to 6,8-disulpho-octanoyl groups (Daigo & Reed, 1962) and then digested with thermolysin. The digest was submitted to paper

electrophoresis at pH 2. Under these conditions, only peptides containing at least two sulphonic acid groups are likely to move towards the positive electrode. When the paper was stained with fluorescamine (Perham, 1978) a single such acidic peptide (TH1) was observed, though it stained as a rather broad band. This suggested that peptide TH1 might actually consist of a mixture of peptides of very similar mobilities. However, attempts to resolve the putative mixture by further paper electrophoresis at

Table 1. *Properties of peptide TH1 and the three peptides obtained from it after digestion with S. aureus proteinase*
Peptides were hydrolysed for 24 h and analysed on a Rank-Hilger Chromaspek amino-acid analyser. Electrophoretic mobility at pH 6.5 was measured with respect to aspartic acid (defined as -1.00).

Amino acid	Amino acid composition (mol/mol of peptide)			
	TH1*	TH1A2	TH1A3a	TH1A3b
Aspartic acid	} 2.2†	} 2.2†		
Methionine sulphone				
Threonine	0.8		1.0	1.0
Serine	1.1	1.1		
Glutamic acid	2.1	1.2	1.0	1.0
Proline	2.7	1.7		
Glycine	1.4	1.0‡		
Alanine	1.7	1.8		
Valine	1.8	1.1	1.1	1.1
Isoleucine	0.7‡		1.0	0.9‡
Leucine	0.2‡		0.6‡	
Lysine	0.9	1.0		
Electrophoretic mobility at pH 6.5	-0.67	-0.83	-0.34	-0.40

* Values are the mean of several determinations.
† Aspartic acid and methionine sulphone were not resolved on the Rank Chromaspek amino-acid analyser.
‡ *N*-Terminal residue, established by dansylation. The yield of these amino acids may be low since the peptides had been stained with fluorescamine.

Fig. 1. *Amino acid sequences of peptide TH1 and peptides derived from it by digestion with S. aureus proteinase*
→, Residue established by dansyl-Edman degradation. The lipoyl group in these peptides was present as the disulpho-octanoyl derivative. Met(O₂), methionine sulphone.

pH 3.5 or 6.5 were unsuccessful. Similarly, the amino-acid composition of peptide TH1 was not integral for all residues (Table 1) and two *N*-terminal residues, leucine and isoleucine, were established by dansylation. No *N*⁶-dansyl was detected, consistent with any *N*⁶-amino group being acylated. However, the results of a dansyl-Edman degradation of peptide TH1 (in which two amino acids were identified at each step) were consistent with the presence of two identical amino acid sequences staggered by one additional residue (leucine) at the *N*-terminus (Fig. 1).

A sample of peptide TH1 was treated with *S. aureus* proteinase and three major peptides were isolated from the digest by paper electrophoresis at pH 6.5 and 2.0 (Table 1). The amino-acid sequences of peptides TH1A3a and TH1A3b were found to be identical by dansyl-Edman degradation except that peptide TH1A3a had an additional residue of leucine at its *N*-terminus (Fig. 1). The first nine residues of peptide TH1A2 were also established by means of dansyl-Edman degradation (Fig. 1) but the results were not clear-cut beyond that point. By comparing this sequence with the composition of peptide TH1A2 (Table 1), we infer that there is probably one residue of alanine and one of proline unaccounted for.

These results are all consistent with the proposition that peptide TH1 is actually a mixture of at least two overlapping peptides. One carries an extra residue of leucine at its *N*-terminus. Leucine and isoleucine are among the expected *N*-terminal residues of peptides generated by thermolytic digestion. There may also be heterogeneity at the *C*-terminus. For our present purposes, the most interesting conclusion is that a single sequence of 13 amino acids contains the lipoylated lysine residue, *viz*:

This sequence includes that of peptide TA1:

which we identified previously in digests with *S. aureus* proteinase alone (Hale & Perham, 1979b). In the present experiments, the Glu-Val bond (positions 11-12) was not cleaved to a significant extent by the treatment with *S. aureus* proteinase, possibly because the adjacent methionine residue was this time present as the sulphone.

The extended amino acid sequence around the lipoyl-lysine residues that we have established (Hale & Perham, 1979b; the present paper) includes the original tetrapeptide sequence of Daigo & Reed

(1962) and its later extension by two residues (Horney, 1967). It also includes the sequence:

of Bleile *et al.* [unpublished work quoted in Bleile *et al.* (1979)]. The proteolytic enzymes used and the method of peptide purification employed have differed in all these experiments. We cannot discount the possibility that we, and others, have failed to purify one of the two sequences to be expected (Danson & Perham, 1976; Bates *et al.*, 1977; Speckhard *et al.*, 1977; Collins & Reed, 1977). However, an alternative explanation is that this sequence is repeated in the E2 polypeptide chain (Hale & Perham, 1979b). The lipoic acid residues are localized in a region of the E2 chain with an apparent *M_r* of 45 000 (Bleile *et al.*, 1979; G. Hale & R. N. Perham, unpublished work). It remains to be seen whether the primary structure of the E2 chain is divisible into highly homologous halves, as proposed by Gebhardt *et al.* (1978), or whether any repeating sequence is confined to the lipoic-acid-containing region, as indicated by the experiments of Bleile *et al.* (1979).

We thank the Science Research Council for a research grant (to R. N. P.) and the Wellcome Trust for funds to purchase equipment. We are grateful to Mr. G. Parr for skilled assistance with amino-acid analysis.

References

- Bates, D. L., Danson, M. J., Hale, G., Hooper, E. A. & Perham, R. N. (1977) *Nature (London)* **268**, 313-316
- Bleile, D. M., Munk, P., Oliver, R. M. & Reed, L. J. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4385-4389
- Brown, J. P. & Perham, R. N. (1976) *Biochem. J.* **155**, 419-427
- Collins, J. H. & Reed, L. J. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 4223-4227
- Daigo, K. & Reed, L. J. (1962) *J. Am. Chem. Soc.* **84**, 666-671
- Danson, M. J. & Perham, R. N. (1976) *Biochem. J.* **159**, 677-682
- Danson, M. J., Hale, G., Johnson, P., Perham, R. N., Smith, J. & Spragg, P. (1979) *J. Mol. Biol.* **129**, 603-617
- Gebhardt, C., Mecke, D. & Bisswanger, H. (1978) *Biochem. Biophys. Res. Commun.* **84**, 508-514
- Hale, G. & Perham, R. N. (1979a) *Eur. J. Biochem.* **94**, 119-126
- Hale, G. & Perham, R. N. (1979b) *FEBS Lett.* **105**, 263-266
- Hirs, C. W. (1956) *J. Biol. Chem.* **219**, 611-621
- Horney, D. L. (1967) Ph.D. Thesis, University of Texas
- Perham, R. N. (1975) *Philos. Trans. R. Soc. London. Ser. B* **272**, 123-136

- Perham, R. N. (1978) in *Techniques in Protein and Enzyme Biochemistry* (Kornberg, H. L., Metcalfe, J. C., Northcote, D. H., Pogson, C. I. & Tipton, K. F., eds.), **B110**, pp. 1–39, Elsevier/North Holland, Amsterdam and New York
- Perham, R. N. & Thomas, J. O. (1971) *FEBS Lett.* **15**, 8–12
- Reed, L. J. (1974) *Accounts Chem. Res.* **7**, 40–46
- Reed, L. J. & Oliver, R. M. (1968) *Brookhaven Symp. Biol.* **21**, 397–411
- Speckhard, D. C., Ikeda, B. H., Wong, S. S. & Frey, P. A. (1977) *Biochem. Biophys. Res. Commun.* **77**, 708–713
- Vogel, O. (1977) *Biochem. Biophys. Res. Commun.* **74**, 1235–1241
- Vogel, O., Beikirch, H., Muller, H. & Henning, U. (1971) *Eur. J. Biochem.* **20**, 169–178