

# Expression in *Escherichia coli* of a sub-gene encoding the lipoyl domain of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus*

Frédéric Dardel\*, Leonard C. Packman and Richard N. Perham

Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK

Received 20 March 1990

A sub-gene encoding the lipoyl domain (residues 1–85) of the lipoate acetyltransferase chain of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus* was over-expressed in *Escherichia coli*. Approx. 80% of the domain was unlipoylated but most of the remainder was correctly lipoylated on Lys-42 and could be reductively acetylated by the *B. stearothermophilus* enzyme complex. A small proportion (approx. 4%) of the domain carried an aberrant substituent, possibly an octanoyl group, on Lys-42. The 400 MHz  $^1\text{H}$  NMR spectra of the lipoylated and unlipoylated domains were essentially identical and closely resembled that of the native lipoyl domain.

Lipoyl domain; Pyruvate dehydrogenase complex; Lipoylation; NMR spectroscopy

## 1. INTRODUCTION

The 2-oxo acid dehydrogenase complexes are among the most extensively studied multienzyme systems, in terms of both structure and mechanism (for recent reviews, see [1–3]). The lipoate acyltransferase (E2) component is of particular importance in that it is the middle of 3 enzymes that must act sequentially in the overall reaction and also serves as the structural core (octahedral or icosahedral) of the complexes [4].

Thus far, no 2-oxo acid dehydrogenase complex or E2 component has been crystallized but the elaborate domain-and-linker structure [1–3] of the E2 chain opens up another possibility for detailed structural analysis: determination of the structure of the separate domains and of the intervening segments of the E2 polypeptide chain. We have already made progress, by means of NMR spectroscopy, with the linker regions of the E2p chain of the *Escherichia coli* pyruvate dehydrogenase (PDH) complex [5]. Similarly, we have demonstrated that the N-terminal lipoyl domain of the E2p chain of the *Bacillus stearothermophilus* PDH complex, which can be released by limited proteolysis, is a small folded protein with a robust tertiary structure [6,7].

*Correspondence address:* R.N. Perham, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK

\* *Present address:* Laboratoire de Biochimie, Ecole Polytechnique, 91128 Palaiseau Cedex, France

*Abbreviations:* DTT, dithiothreitol; PDH, pyruvate dehydrogenase; PTH, phenyl thiohydantoin; SDS, sodium dodecylsulphate; TFA, trifluoroacetic acid

To facilitate further NMR work on the *B. stearothermophilus* lipoyl domain, we have now created a sub-gene encoding it and succeeded in over-expressing it in *E. coli*. A sub-gene encoding a lipoyl domain of the *E. coli* E2p chain has also been expressed in *E. coli* [8]. We demonstrate that *E. coli* can lipoylate the correct lysine residue of the foreign domain, at least in part, but that a small fraction (approx. 4%) of the purified domain carries another substituent on the lysine residue, possibly an octanoyl group. The 400 MHz  $^1\text{H}$  NMR spectrum of the domain closely resembles that of the domain derived by proteolysis of the native *B. stearothermophilus* PDH complex.

## 2. MATERIALS AND METHODS

### 2.1. Bacterial growth, genetic manipulation and DNA sequencing, reductive acetylation

*E. coli* strain TG1 *recO* [9] was used as host for the various plasmids. Plasmid pBst42 carries the E1 $\alpha$  and E1 $\beta$  genes and the beginning of the E2p gene from *B. stearothermophilus* [9]. Cells were grown on 2  $\times$  TY medium [10], supplemented with 10  $\mu\text{g}/\text{ml}$  lipoic acid and 100  $\mu\text{g}/\text{ml}$  ampicillin, where applicable. Restriction enzymes were used as recommended by the suppliers (Amersham, Boehringer Mannheim, BRL). DNA sequence was determined by the dideoxy chain termination method [11,12] using the Sequenase kit (US Biochemical) and [ $\alpha$ - $^{35}\text{S}$ ]dATP (Amersham). The mutagenic oligonucleotide was synthesized by Mr M. Weldon on a Milligen BioSearch Cyclone synthesizer and purified by ion-exchange chromatography [13]. Mutagenesis was performed by the phosphorothioate method [14,15] using the Amersham kit. The ability of the free lipoyl domain to become reductively acetylated by the PDH complex in the presence of [2- $^{14}\text{C}$ ]pyruvate was assayed according to Packman et al. [6].

### 2.2. Protein chemical methods

Samples were reduced with 13 mM DTT (60 min) and alkylated with vinyl pyridine (44 mM) in 0.5 M Tris-HCl, pH 8.5, containing

6 M guanidine-HCl under Ar at 22°C. After 3 h, the alkylated proteins were purified by a gradient (0–64% over 30 min) of CH<sub>3</sub>CN in 0.1% TFA on an Aquapore RP300 column (4.6 mm × 30 mm, 0.3 ml/min) using a Hewlett Packard 1090M liquid chromatograph. Pyridylethylated protein was collected, dried and digested with *Staphylococcus aureus* V8 proteinase (1%, w/w) at 37°C for 18 h in 1% NH<sub>4</sub>HCO<sub>3</sub>. Peptides were separated on a Spherisorb ODS2 column (3 mm × 100 mm) using the conditions described above. Diode array analysis was used to detect pyridylethylated peptides at 245–260 nm. Peptides were characterized by amino acid analysis [7] and sequenced on an Applied Biosystems 477A Protein Sequencer. Ion-exchange chromatography of the domains was performed on a Polypore DEAE column (2.1 mm × 30 mm) equilibrated with 10 mM NH<sub>4</sub>HCO<sub>3</sub> at 0.2 ml/min; elution of protein was by a gradient of 0–0.2 M NH<sub>4</sub>HCO<sub>3</sub> (5 min), then 0.2–0.4 M NH<sub>4</sub>HCO<sub>3</sub> (15 min). Samples of domain which were reduced (5 mM DTT, 2 h) and alkylated (25 mM vinyl pyridine, 2 h) in 10 mM Tris-HCl, pH 8.3, under Ar, at 22°C, were applied directly to the column. Intact domains were analysed on an Applied Biosystems 270A capillary electrophoresis apparatus (monitoring A<sub>200</sub>) in 0.1 M H<sub>3</sub>PO<sub>4</sub> at 15–20 kV and 30°C.

### 2.3. Mass spectrometry and <sup>1</sup>H NMR spectroscopy

FAB-mass spectrometric analysis of peptides was performed in a thioglycerol matrix. Samples of intact domain (250 pmol) were analysed by electrospray mass spectrometry on a prototype VG Bio Q instrument in 10 μl of 2% acetic acid in 50% methanol. <sup>1</sup>H NMR spectra were recorded at 400 MHz using a Bruker AM400 WB spectrometer. The lipoyl domains were freeze-dried and dissolved in <sup>2</sup>H<sub>2</sub>O twice. Apparent pD was adjusted to 5.1 by adding <sup>2</sup>HCl, and NaN<sub>3</sub> was added to 0.02% (w/v). Final concentration of the samples (0.5 ml) was 10 mM. Spectra were recorded with quadrature detection, a 5.5 kHz spectral width, pulse intervals of 1.5 s and presaturation of the HDO line. Before Fourier transformation, a Gaussian function multiplication was applied, leading to a line broadening of ~1 Hz.

## 3. RESULTS AND DISCUSSION

### 3.1. Construction and expression of a sub-gene encoding the lipoyl domain

A 378 bp *Dra*I fragment from pBst49 [9] which comprised the translation start and the first 86 codons of the *B. stearothermophilus* E2p gene was purified and ligated into the *Sma*I site of the pBSTNAV vector [13]. The resulting plasmid, pBlip, carries the truncated E2 gene under the control of the strong constitutive *lpp* promoter from *E. coli*. A stop codon (UAA) was introduced after codon 86 of the E2 sub-gene by directed mutagenesis, using the oligonucleotide 5'-CAGCAGCTGTCAAAACGTCATG-3' and single-stranded pBlip DNA. The resulting mutant plasmid, pBlipS, was characterized by completely sequencing its insert. It encodes exactly the lipoyl domain of 85 residues released by chymotryptic digestion of *B. stearothermophilus* PDH complex [7]. Crude cell extracts of *E. coli* strain TG1recO carrying pBlipS showed an over-produced polypeptide band of *M<sub>r</sub>* approx. 9500 on SDS polyacrylamide gel electrophoresis [6], as expected for the lipoyl domain.

### 3.2. Purification of the over-expressed lipoyl domain

Strain TG1recO/pBlipS was grown in a 15 litre

fermenter in 2 × TY medium supplemented with lipoic acid (10 mg/l) and glucose (0.2% w/v). Cells harvested in late log phase (*A*<sub>650</sub> 14.0, 227 g wet weight) were resuspended in 250 ml buffer A (20 mM Na-phosphate, pH 7.0; 2.7 mM EDTA; 0.02% NaN<sub>3</sub>). In all subsequent steps, the presence of the lipoyl domain was monitored by SDS polyacrylamide gel electrophoresis [6]. The cells were disrupted with a French press and cell debris were removed by centrifugation (3 h, 16000 × *g*). Nucleic acids were precipitated with streptomycin (3% w/v). The extract was then fractionally precipitated with ammonium sulphate and the 40–70% cut was resuspended in a minimal amount of buffer A (final vol. 220 ml). This extract was gel filtered, in batches of 40 ml, on a G50 Superfine column (75 cm × 2 cm, 24 ml/h), equilibrated with buffer A. The lipoyl domain eluted after the bulk of the proteins, as judged by the reductive acetylation assay. This indicated that the lipoyl domain from *B. stearothermophilus* was at least partly lipoylated by the *E. coli* lipoylation system.

The samples pooled from the G50 columns (250 ml) were further fractionated by chromatography on a DE52 column (7 × 2 cm, 50 mM to 500 mM Na-phosphate, pH 7.0, 100 ml, 60 ml/h). The ability to become reductively acetylated did not coincide exactly with the protein peak, which exhibited a shoulder. This suggested that at least two overlapping forms of the lipoyl domain were present, only one of which could become reductively acetylated. The separation was improved by repeating the DE52 step, using a shallower gradient (200 mM to 450 mM Na-phosphate, pH 7.0, 100 ml). The amino acid compositions and N-terminal sequences (10 residues) matched those of the lipoyl domain isolated from the PDH complex [6,7]. The final yield was 290 mg of the non-acetyltable form and 60 mg of the acetyltable form. The *A*<sub>280</sub><sup>0.1%</sup> of each was 0.54.

### 3.3. Characterization of the purified domains

When examined by capillary electrophoresis, the acetyltable form of the lipoyl domain was found to contain less than 3% of the non-acetyltable form. No contamination of the non-acetyltable form was evident (data not shown). Both forms of the domain were reduced and alkylated with vinyl pyridine under denaturing conditions, digested with *S. aureus* V8 proteinase, and the products separated by HPLC (Fig. 1). The peaks highlighted in profiles I (non-acetyltable domain) and II (acetyltable domain) were identified by amino acid composition and sequence analysis. Peak A of profile I was VQNDKAVVE, derived from the usual site of lipoylation (Lys-42) in the lipoyl domain; this peak is clearly absent from profile II. Profile II, in turn, has a peak (B) which is absent from profile I. Peak B was VQNDXAVVE where X, based on the amino acid composition, must represent a modified derivative of lysine; its PTH-derivative eluted well after

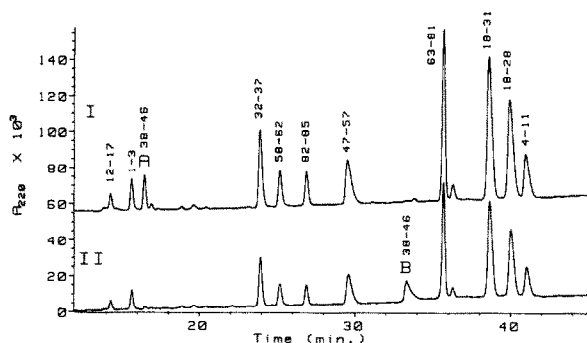


Fig. 1. Reverse-phase HPLC of *S. aureus* V8 proteinase digests of alkylated lipoyl domain on Spherisorb ODS2. Profile I, digest of non-acetyltable domain; Profile II, digest of acetyltable domain. Numbers above each peak denote the sequence position of the peptide [7] determined from amino acid and/or sequence analysis.

PTH-Leu on an extended gradient. Sequence analysis of a reduced and alkylated model peptide TVEGD(lipoyl)KASME [16] showed the PTH-derivative of residue X in the *Bacillus* sequence to co-chromatograph with that of the alkylated lipoyl-lysine of the model compound. FAB-mass spectrometric analysis of the model peptide confirmed the lipoyl group to be  $S^6, S^8$ -(bis)pyridylethylated. Surprisingly, the peptide VQND(lipoyl)KAVVE was only detectable in its bis-pyridylethylated form; it could not be located in any part of the HPLC effluent unless the lipoyl group was alkylated. This property was not shared by the model peptide TVEGD(lipoyl)KASME, which is readily isolated on this reverse-phase support. The remaining peaks in profiles I and II were identified as shown in Fig. 1. The acetyltable form of the domain evidently carries a lipoyl group at position Lys-42 and Lys-42 appears fully substituted. Similarly, Lys-42 in the non-acetyltable form of the domain is unsubstituted.

When examined by high-performance ion-exchange chromatography, the unlipoylated domain gave a single peak and its retention time was unaffected by reduction and alkylation with vinyl pyridine, indicating that its single cysteine residue, Cys-36, is not available for reaction in the folded protein. However, the reduced lipoylated domain gave two principal peaks. After subsequent alkylation with vinyl pyridine, the major component (80%) of the doublet moved to a higher position in the gradient. This was puzzling because the modified domain has acquired 2 extra positive charges and would therefore be expected to elute earlier in the gradient. Other interactions must increase the retention time of the domain. The minor component (18%) was unaffected (Fig. 2) and diode array analysis showed it not to be pyridylethylated. Given that the lipoylated domain preparation possessed no unsubstituted Lys-42 (see above), the minor component must be carrying a substituent different from lipoic acid.

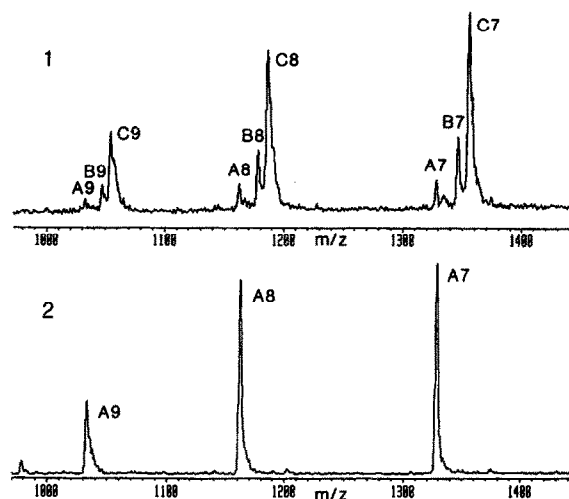


Fig. 2. Ion-exchange chromatography of lipoylated and unlipoylated domain preparations on Polypore DEAE. Lipoylated domain before (A) and after (B) alkylation with vinyl pyridine. Unlipoylated domain before (C) and after (D) alkylation with vinyl pyridine.

### 3.4. An alternative to lipoylation

Electrospray-mass spectrometry (Fig. 3) of the lipoylated domain preparation gave a major component (78%) of  $9457.5 \pm 1.5$  Da, as expected of the lipoylated domain (calc. mass 9457.6 Da); a second component (17%) of  $9395.9 \pm 1.9$  Da; and a minor component (5%) of  $9269.1 \pm 1.3$  Da, corresponding to the low level of unlipoylated domain (calc. mass 9269.6 Da) already detected in the preparation. The unlipoylated domain preparation gave a single component of  $9269.5 \pm 1.8$  Da, consistent with its known purity. The second component in the lipoylated domain preparation must be carrying a substituent on Lys-42 that is uncharged and released by acid hydrolysis. The estimated mass of the substituent is 126.8 Da; an obvious possibility is the octanoyl group (calc. mass

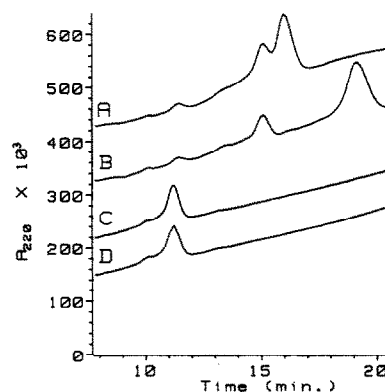


Fig. 3. Electrospray-mass spectrometry of (1) lipoylated domain and (2) unlipoylated domain preparations. A-series, 9269.5 Da (no lipoyl group); B-series, 9395.9 Da (putative octanoyl domain); C-series, 9457.5 Da (lipoyl domain). Numbers in each series denote the positive-charge state of each ion.

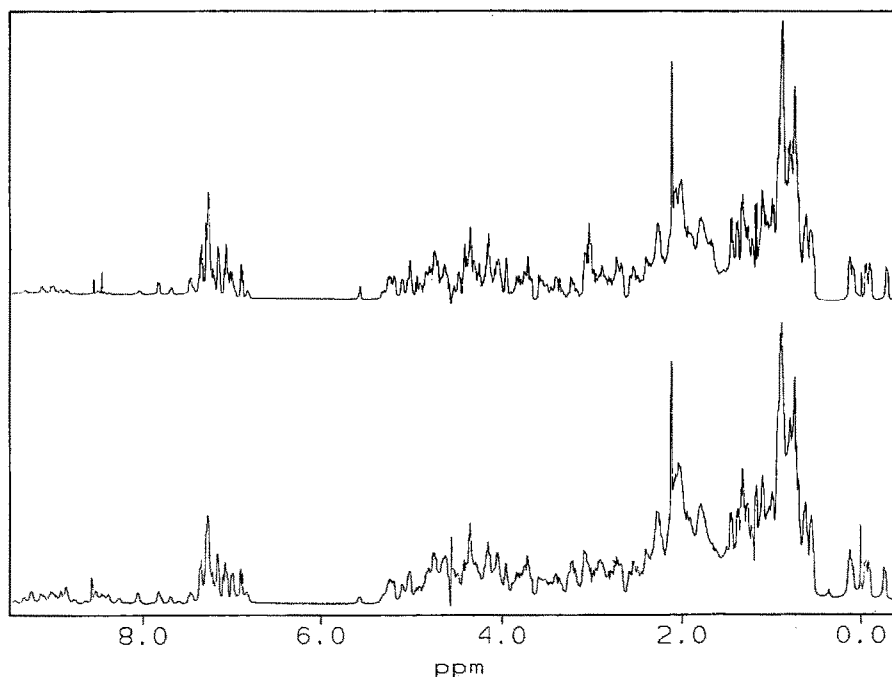


Fig. 4. 400 MHz  $^1\text{H}$  NMR spectra of the isolated domains. The spectra were recorded at 45°C and are the sum of 128 transients. Top trace, unlipoylated lipoyl domain; bottom trace, lipoylated lipoyl domain. The chemical shifts are expressed relative to the internal trimethylsilyl propionate.

127.2 Da), given that octanoic acid is a biosynthetic precursor of lipoic acid [17,18], but this remains to be proven.

### 3.5. $^1\text{H}$ NMR spectroscopy of the purified domains

The 400 MHz  $^1\text{H}$  NMR spectra of the lipoylated and un-lipoylated domains are essentially identical to each other (Fig. 4) and to that of the lipoyl domain released from the intact PDH complex by proteolytic cleavage [6]. The high field methyl and the aromatic proton resonances are particularly well conserved and stable over a wide temperature range (10–65°C), indicating a well-defined tertiary structure. 2D-COSY spectra of the two forms are identical (data not shown), apart from a few additional cross-peaks in the spectrum of the lipoylated form which can be assigned to the lipoamide moiety.

### 3.6. Conclusions

The lipoyl domain of the *B. stearotherophilus* E2p chain can now be isolated in large amounts, sufficient for 2D-NMR spectroscopy, from *E. coli* cells transformed with pBlipS. It is essentially identical to the domain obtained by limited proteolysis of the native PDH complex [6]. It is of particular interest that *E. coli* cells can correctly lipoylate the *B. stearotherophilus* domain, at least in part. Presumably the non-lipoylated form arises from an inability of the host *E. coli* cells to process the large amounts of domain emanating from the highly expressed sub-gene. The

small amount of lipoyl domain carrying another substituent on the target lysine residue, very likely an octanoyl group, may arise similarly. Whether any incorrect substitution of the target lysine residue of native *E. coli* lipoyl domains takes place remains to be investigated.

**Acknowledgements:** This work was supported by the SERC (Grant GR/D/84542) and its Cambridge Centre for Molecular Recognition and by The Wellcome Trust. We thank Ms Jennifer Nutkins for FAB-ms and Stuart Jarvis and Brian Green at VG Masslab for electrospray-ms, Mrs Z.J. Jacoby for technical assistance with the amino acid analysis, and ICI-France for a Visiting Fellowship at Churchill College, Cambridge (for F.D.).

### REFERENCES

- [1] Perham, R.N., Packman, L.C. and Radford, S.E. (1987) Biochem. Soc. Symp. 54, 67–81.
- [2] Miles, J.S. and Guest, J.R. (1987) Biochem. Soc. Symp. 54, 45–65.
- [3] Perham, R.N. and Packman, L.C. (1989) Ann. NY Acad. Sci. 573, 1–20.
- [4] Reed, L.J. (1974) Acc. Chem. Res. 7, 40–46.
- [5] Radford, S.E., Laue, E.D., Perham, R.N., Martin, S.R. and Appella, E. (1989) J. Biol. Chem. 264, 767–775.
- [6] Packman, L.C., Perham, R.N. and Roberts, G.C.K. (1984) Biochem. J. 217, 219–227.
- [7] Packman, L.C., Borges, A. and Perham, R.N. (1988) Biochem. J. 252, 79–86.
- [8] Miles, J.S. and Guest, J.R. (1987) Biochem. J. 245, 869–874.
- [9] Hawkins, C.F., Borges, A. and Perham, R.N. (1990) Eur. J. Biochem., in press.

- [10] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- [11] Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) *J. Mol. Biol.* 143, 161–178.
- [12] Biggin, M.D., Gibson, T.G. and Hong, G.F. (1983) *Proc. Natl. Acad. Sci. USA* 80, 1405–1414.
- [13] Meinnel, T., Mechulam, Y. and Fayat, G. (1988) *Nucleic Acids Res.* 16, 8095–8096.
- [14] Taylor, J.W., Schmidt, W., Cosstick, R., Okruszek, A. and Eckstein, F. (1985) *Nucleic Acids Res.* 13, 8749–8764.
- [15] Taylor, J.W., Ott, J. and Eckstein, F. (1985) *Nucleic Acids Res.* 13, 8765–8785.
- [16] Graham, L.D., Packman, L.C. and Perham, R.N. (1989) *Biochemistry* 8, 1574–1581.
- [17] White, R.H. (1980) *Biochemistry* 19, 15–19.
- [18] Parry, R.J. (1983) *Tetrahedron* 39, 1215–1238.