The pyruvate dehydrogenase complex of the chemolithoautotrophic bacterium *Thiobacillus ferrooxidans* has an unusual E2–E3 subunit fusion

Rosamund Powles and Douglas Rawlings

Author for correspondence: Douglas Rawlings. Tel: +27 21 650 3261. Fax: +27 21 689 7573. e-mail: doug@molbiol.uct.ac.za

**Department of Microbiology, University of Cape Town, Rondebosch 7700, South Africa**

**The genes encoding pyruvate dehydrogenase (PDH) of *Thiobacillus ferrooxidans* were previously located by cloning and sequence analysis of the region upstream of the genes encoding the citrate synthase and γ-glutamylcysteine synthetase genes. The pdh genes of *T. ferrooxidans* were able to complement an *Escherichia coli* aroA pdD mutant for growth on minimal medium lacking acetate, indicating that the *T. ferrooxidans* PDH complex was functional in *E. coli*. The predicted amino acid sequence of the *T. ferrooxidans* PDH complex contained three ORFs. The first ORF encoded a 36.7 kDa homologue of the PDH complex El subunit, the second ORF a 374 kDa El/β subunit and the third ORF an unusual 102 kDa fusion of the E2 and E3 subunits. In spite of *T. ferrooxidans* being a Gram-negative bacterium, its PDH complex had more features in common with Gram-positive bacteria and eukaryotes.**

**Keywords: Thiobacillus ferrooxidans, pyruvate dehydrogenase, molecular cloning**

**INTRODUCTION**

*Thiobacillus ferrooxidans* is a Gram-negative, chemolithoautotrophic, rod-shaped bacterium, that plays an important role in industrial mineral processing. *T. ferrooxidans* is able to attack sulphide-containing minerals and convert the insoluble sulphides of metals such as copper, lead, zinc or nickel to their soluble metal sulfates, thus contributing to the recovery of these metals (Rawlings & Silver, 1995).

Pyruvate dehydrogenase complexes consist of three components: a pyruvate decarboxylase component (E1); a dihydrolipoamide acetyltransferase component (E2) and a dihydrolipoamide dehydrogenase component (E3). The E2 component forms the structural core of the complex, to which the peripheral subunits E1 and E3 bind (Oliver & Reed, 1982). Pyruvate dehydrogenases from Gram-negative bacteria have a core of 24 E2 subunits arranged with octahedral symmetry, whereas those from Gram-positive bacteria and eukaryotes have a larger E2 core of 60 subunits arranged with icosahedral symmetry (Mattevi et al., 1992b). The E1 component can be in one of two forms. In the octahedral complex (Gram-negative bacteria), E1 exists as an $\alpha_2$ homodimer, with a subunit molecular mass of approximately 100 kDa, whereas in the icosahedral complex (Gram-positive bacteria and eukaryotes), it is found as an $\alpha_2\beta_2$ heterotetramer, with subunit molecular masses of approximately 41 and 36 kDa, respectively (Mattevi et al., 1992a).

The pyruvate dehydrogenase (PDH) complex catalyses the oxidative decarboxylation of pyruvate generating acetyl-CoA and NADH, and releasing CO$_2$, thereby providing a link between glycolysis and the Kreb's cycle. The reaction takes place in several steps. In the first, pyruvate decarboxylase converts pyruvate to hydroxyethylthiamin pyrophosphate. Next the transacetylase shifts the two-carbon fragment from thiamin pyrophosphate to lipoic acid and from there to CoA, forming acetyl-CoA. Lastly, the flavoprotein dihydrolipoamide dehydrogenase reoxidizes the reduced form of lipoic acid. Citrate synthase then catalyses the Claisen condensation of oxaloacetate and acetyl-CoA to form citrate (Rawn, 1983).

The Kreb's cycle has a twofold metabolic function: (i) to oxidize metabolites and generate NADH which on conversion to NAD$^+$ may be coupled to ATP formation,
and (ii) to generate metabolites required for biosynthesis, such as the amino acids of the glutamate family. T. ferrooxidans is an obligately chemolithoautotrophic bacterium that obtains its energy by oxidizing Fe$^{3+}$ to Fe$^{2+}$ or reduced sulfur compounds to sulfuic acid (Kelly & Harrison, 1989). It would, therefore, not require the Kreb’s cycle for energy generation, but would require intermediates of the cycle to generate metabolites for biosynthesis. In this paper we report the identification and analysis of the genes for T. ferrooxidans PDH complex which have novel features.

**METHODS**

**Bacterial strains and plasmids.** Genotypes of the strains used are: Escherichia coli JM109, endA1 recA1 gyrA96 thi hsdR17 (rK m4); relA1 supE44 Δ(lac–pro–AB) (F’ traD36 proAB lacF2ΔM15); E. coli JRG1342 recA rpsL [Δ18(aroP–lpd)]; metB thyA azi pps ton. Professor John Guest (University of Sheffield, UK) kindly donated E. coli strain JRG1342 and plasmid pGS87, which is a derivative of pBR322 containing the aceEF–lpd region of the E. coli chromosome (Guest & Lewis, 1985). Plasmids pBluescript SK, pBluescript KS (Stratagene), pUCMB20 and pUCBM21 (Boehringer Mannheim) were used for subcloning and sequencing.

**Media.** Tetrathionate medium was made from mineral salts solution (g l$^{-1}$): (NH$_4$)$_2$SO$_4$, 3.0; KCl, 0.1; K$_2$HPO$_4$, 0.5 and Ca(NO$_3$)$_2$, 0.01 adjusted to pH 2.5 with H$_2$SO$_4$ and autoclaved. Trace elements solution (mg l$^{-1}$): FeCl$_3$, 6H$_2$O, 11.0; CuSO$_4$.5H$_2$O, 0.5; HBO$_3$, 20; Na$_2$MoO$_4$.2H$_2$O, 0.8; CoCl$_2$.6H$_2$O, 0.6 and ZnSO$_4$.7H$_2$O was filter-sterilized. One millilitre of trace elements solution was added to 100 ml mineral salts solution and to this was added either 50 mM K$_3$PO$_4$ or 100 mM FeSO$_4$, and pH adjusted such that the final pH of the medium was 2.5. Luria–Bertani medium was used for growth of E. coli on complex medium, and M9 minimal medium was used for growth of E. coli on selective medium (Sambrook et al., 1989).

**DNA techniques, sequencing and analysis.** Standard methods were used for plasmid preparation, restriction endonuclease digests, gel electrophoresis, ligations and Southern blot hybridizations (Sambrook et al., 1989). Labelling of probes, hybridization and detection were carried out according to the digoxigenin-dUTP non-radioactive DNA labelling and detection kit (Boehringer Mannheim). DNA sequencing was by the dideoxy chain-termination method (Sanger et al., 1977) using the Sequenase version 2.0 from USB. The entire sequence presented in Fig. 2 was determined from both strands. The Genetic Computer Group (GCG) software package was used for sequence analysis (Deveraux et al., 1984). Multiple alignments of sequences and the phylogenetic trees were calculated using the clustalw (Thompson et al., 1994). The TreeView software package was used to draw the phylogenetic trees (Page, 1993).

**Complementation of the PDH complex.** Constructs to be tested were transformed into the E. coli JRG1342 aroP–aceEF–lpd deletion mutant. Expression mixes were plated on Luria–Bertani agar plates with appropriate antibiotic selection (ampicillin 100 µg ml$^{-1}$) and placed at 37 °C for approximately 36 h or until colonies were visible. Individual colonies from each plate were then streaked onto M9

**Table 1. Genetic complementation of the E. coli pdhR–lpd mutant JRG1342**

| Plasmid/cosmid | Carbon source | G + A | S + A | S
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>+ + + +</td>
<td>+ +</td>
<td>+ +</td>
<td>+ +</td>
</tr>
<tr>
<td>pSK</td>
<td>+ + + +</td>
<td>+ +</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pGS87</td>
<td>+ + + +</td>
<td>+ +</td>
<td>+ +</td>
<td>+ + +</td>
</tr>
<tr>
<td>Cosmid 5.1</td>
<td>+ + + +</td>
<td>+ +</td>
<td>+</td>
<td>+ *</td>
</tr>
<tr>
<td>pTHIO81</td>
<td>+ + + +</td>
<td>+ +</td>
<td>+</td>
<td>+ *</td>
</tr>
<tr>
<td>pTHIO251</td>
<td>+ + +</td>
<td>+ +</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Carbon sources: G + A + S, 10 mM glucose + 2 mM acetate + 2 mM succinate; G + A, 10 mM glucose + 2 mM acetate; G, 10 mM glucose; S + A, 50 mM succinate + 4 mM acetate; S, 50 mM succinate. + + + +, Normal growth; + +, poor growth; −, no growth; + + + *, growth of orange-pigmented colonies after 6–7 d; + + + + *, growth of orange-pigmented colonies after 2–3 d.
minimal media agar plates, containing L-methionine (20 µg L⁻¹) and thiamin (5 µg ml⁻¹), with either glucose (10 mM), potassium succinate (50 mM) or potassium acetate (50 mM) as carbon sources. Glucose minimal medium was supplemented with 2 mM acetate and/or succinate, succinate minimal medium with 4 mM acetate and acetate minimal medium with 2 mM succinate as indicated.

In vitro synthesis of polypeptides. The in vitro synthesis of polypeptides from cosmid 5.1, plasmids pTHIO81, pTHIO191, pTHIO51, pTHIO251, pPD23 and pBlueScript SK was determined using the E. coli S30 DNA-directed transcription/translation system from Promega. Reactions were performed according to the manufacturer’s specifications. The proteins were separated by SDS-PAGE using a 5–15% gradient gel.

RESULTS

Isolation, complementation and localization of the genes for the PDH complex

Cosmid 5.1 contained a T. ferrooxidans chromosomal DNA insert of approximately 42 kb cloned into the vector pHC79. This cosmid was isolated by its ability to complement an E. coli γ-glutamylcysteine synthetase (gshA) mutant for growth on minimal medium lacking glutathione as described previously (Powles et al., 1996). Partial nucleotide sequencing of the region upstream of the gshA gene had revealed the presence of the genes for the PDH complex. A region of cosmid 5.1 between the SacII and BglII restriction sites, was subcloned into pBlueScript SK to give the pTHIO81 (Fig. 1). A restriction map of pTHIO81 was constructed and smaller fragments subcloned into the vectors pBlueScript SK and KS as illustrated.

E. coli strain JRG1342 is an aceEF-lpd deletion mutant which lacks pyruvate decarboxylase, dihydrolipoamide transacylase and dihydrolipoamide dehydrogenase activities and is unable to grow on minimal medium lacking acetate whereas PDH positive strains can (Guest et al., 1983). Cosmid 5.1, plasmids pTHIO81 and pTHIO251 were transformed into E. coli JRG1342 and tested for their ability to enable growth of the host on supplemented minimal medium. Plasmid pGS87 contained the cloned E. coli pdhR-lpd region and mutants transformed with pGS87 served as a positive control, while pBlueScript SK transformants were the negative control.

The results of the supplementation of the E. coli pdhR-lpd mutant with the cloned T. ferrooxidans genes are shown in Table 1. Cosmid 5.1 and plasmid pTHIO81 enabled growth on glucose and succinate minimal medium whereas plasmids pBlueScript SK and pTHIO251 did not. This indicated that the genes for the T. ferrooxidans PDH complex were able to complement the E. coli pdhR-lpd mutant and that they were situated between the SacII (adjacent to EcoRV 0·0) and BglII (6·5) sites of cosmid 5.1 (Fig. 1). The orange-pigmented colonies obtained on growing E. coli JRG1342 containing either pGS87, cosmid 5.1 or pTHIO81 on succinate minimal media were shown to be E. coli by testing for lysis by T4 phage. The experiment was repeated three times and on each occasion colonies grew with the same orange appearance. The reason for the orange pigmentation is unknown.

The source of the genes encoding the PDH complex was confirmed by hybridization of the labelled plasmid pTHIO51 to T. ferrooxidans ATCC 33020 chromosomal DNA, cosmid 5.1 and various subclones (data not shown).

Sequence analysis

Analysis of the sequence from the EcoRV site (0·0) to the ClaI site (6·3) (see Fig. 1) revealed three complete ORFs.
The sequence from 540 to 6300 bp is shown in Fig. 2. A putative promoter region (TTGACG-N_{15-17}CATCAT) which corresponds closely to a consensus σ^{70} promoter (TTGACA-N_{15-17}TATAAT) was located 88 bp upstream of the first ORF. A region with features of a rho-independent transcription terminator, consisting of a 10 bp inverted repeat (ΔG = -17 kcal mol^{-1}/-71 kJ mol^{-1}) followed by a poly(T) region was present 528 bp downstream of the last ORF.

The first ORF was preceded by a strong RBS and encoded a polypeptide of 327 amino acids, corresponding to a protein of 367 kDa. The predicted amino acid sequence was closely related to the α subunit of the heteromeric El subunits of the acetoin dehydrogenase of *Clostridium magnum* (42% identity, GenBank/EMBL accession no. L31844), the PDH complexes of *Homo sapiens* (40% identity, accession no. M86808) and *Arabidopsis thaliana* (38% identity, accession no. U21214) and contained the highly conserved thiaminpyrophosphate-binding site (Fig. 2) (Hawkins et al., 1989). The second ORF was also preceded by a strong RBS and overlapped the first ORF by 5 bp. This ORF encoded a polypeptide of 343 amino acids, corresponding to a protein of 374 kDa, with highest sequence identity to the β subunits of the El heteromeric subunits of the PDH complexes of *Arabidopsis thaliana* (46% identity, accession no. U09137), *H. sapiens* (42% identity, accession no. M34056) and *Bacillus subtilis* (40% identity, accession no. M57435). Both the first and the second ORFs had limited sequence identity to the tetrameric E1 components which are composed of a single type of subunit (e.g. 18% sequence identity to the *E. coli* E1 subunit after excluding the region where the two proteins do not correspond due to the larger size of the *E. coli* subunit).

The third ORF was separated from the second ORF by 4 bp and encoded a polypeptide of 978 amino acids, corresponding to a protein of 102 kDa. This predicted polypeptide was considerably larger than previously reported for the E2 and E3 components of the PDH complex and appeared to be a fusion of these two subunits. Analysis of the amino acid sequence revealed two repeating units which exhibited features associated with lipoyl domains (highlighted in Fig. 2) (Stephens et al., 1983). Lys residues which are lipoylated in the E2 components of the 2-oxo-acid dehydrogenase complexes, were identified at the centre of each domain. Both repeating units are followed by regions rich in Ala and Pro residues, resembling the interdomain linker segments of other dihydrolipoyl transacetylase enzymes. These regions were followed by a subunit-binding domain (Fig. 2). The central region of the *T. ferrooxidans* protein exhibited strong similarity to the catalytic domains of the E2 component of other PDH complexes and contained the highly conserved putative active site sequence Ser-(X)_{55}-His-(X)_{36}-Asp-Gly (Fig. 3) (Guest, 1987). The E2 residues Ser, His and Asn are thought to be involved in transition state stabilization, proton transfer and activation of proton transfer, respectively (Hendle et al., 1995). The presence of an active site Asn residue has previously been recorded only in *Azotobacter vinelandii* E2, as all other dihydrolipoamide acetyltransferase sequences contain an Asp residue at this position (Hendle et al., 1995). The N-terminal region of the predicted amino acid sequence of this ORF had highest sequence identity to the dihydrolipoyl transacetylase (E2) component of the PDH complexes of *Neurospora crassa* (36% identity, accession no. P20285) and *H. sapiens* (33% identity, accession no. J03866).

Lipoamide dehydrogenases belong to the family of flavin-containing, homodimeric, disulfide reductases (Williams, 1992). The C-terminal region of the predicted amino acid sequence of ORF 3 contained the characteristic sequence motifs conserved among disulfide reductases such as a FAD-binding region, a redox-active disulfide active site (Cys-530 and Cys-535), the NAD(H)-binding region and an interface region (Fig. 2) (Carothers et al., 1989). The C-terminal region of the *T. ferro-
Pyruvate dehydrogenase of *Thiobacillus ferrooxidans*

---

**In vitro protein analysis**

To establish a correlation between an ORF derived from sequence analysis and the polypeptide produced, and to confirm that the E2–E3 subunits were fused, protein products from various subclones of cosmid 5.1 were examined using an *E. coli*–derived *in vitro* transcription-translation system (Fig. 3). As the sequence-derived sizes of the E1α and E1β subunits are so similar (36.7 and 37.4 kDa, respectively), the two proteins would be likely to give a single signal on an autoradiograph. A protein of approximately 37 kDa was produced by cosmid 5.1 and plasmids pTHIO1, pTHIO1, pTHIO191 (Fig. 3a, lanes 1–4), but not by pBluescript SK (lane 5). pTHIO191 contains the ORF encoding the E1α subunit, but not the E1β subunit, whereas plasmid pTHIO51 contains the ORF encoding the E1β subunit, and the 5′ end of the gene for the E2–E3 fusion but is missing the 3′ end of the gene encoding the E1β subunit. Both plasmids produced a protein of approximately 37 kDa, however the band produced by the plasmid containing the E1α-subunit gene (lane 4) was much darker than that from the plasmid with an intact gene for only the E1β subunit (lane 3). The darker bands for the 37 kDa polypeptide(s) produced by pTHIO191 and pTHIO81 was probably because in these constructs expression of the 37 kDa proteins was in the same direction as the vector *lacZ* promoter, whereas in pTHIO51 the expression of the E1β subunit was in the reverse direction. Both subunits must be produced by cosmid 5.1 and plasmid pTHIO81 since both of these clones complement the *E. coli* JRG1342 aceEF–lpd deletion mutant for growth on minimal medium plus glucose (Table 1).

A protein of approximately 100 kDa, corresponding in size to the 102 kDa fusion of the E2–E3 subunits of the PDH complex, was produced from cosmid 5.1, plasmid pTHIO81 and plasmid pPD23 (Fig. 3a, lanes 1 and 2, subunits. The GenBank accession numbers of the sequences used were as follows: (a) *Sinthopsis macroura*, L20774; pig, P29804; mouse, P35486; *Arabidopsis thaliana*, Z62630; *Escherichia coli*, J01579; *Acholeplasma laidlawi*, M57435; *Bacillus subtilis*, M57435; *Bacillus stearothermophilus*, X53560; *Clostridium magnum*, L31844; *Alcaligenes eutrophus*, P27745; *Pseudomonas putida*, L35343; *Ascaris suum*, P26257; and human, M34056. (b) *P. putida*, L35343; *E. coli*, J01579; *Acholeplasma laidlawi*, M57435; *Bacillus stearothermophilus*, X53560; *Neurospora crassa*, P20285; *Saccharomyces cerevisiae*, J03645; *Azotobacter vinelandii*, X12455; *E. coli*, J01579; *Acholeplasma laidlawi*, M57435; *Staphylococcus aureus*, M57435; and *Dictyostelium discoideum*, U06634.

---

**Fig. 4.** Phylograms based on the amino acid sequences of the E1α and β subunits of pyruvate and acetoin dehydrogenase complexes (acetoin dehydrogenases amino acid sequences are indicated by asterisk). The E1α subunits are shown in (a), E1β subunits in (b) and the E2 portion of the E2–E3 fusion from *T. ferrooxidans* in (c). The CLUSTALW and TreeView software packages were used to calculate and draw the phylograms. Bootstrapting values are indicated at branch points and the scale of 0.1 indicates 10% amino acid sequence divergence. Broken circles group related clusters of PDH oxidans fusion polypeptide showed relatively low amino acid sequence identity to the dihydrolipoamide dehydrogenase (E3) component of PDH complexes compared with the E1α, E1β and E2 subunits. This ranged between 27% sequence identity (*H. sapiens*, accession no. L13761) and 32% sequence identity (*Bacillus stearothermophilus*, accession no. X53560).
Fig. 3b, lane 3). Plasmids pTHIOS1 and pTHIOS11 (Fig. 3a, lane 3, and Fig. 3b, lane 1) also produced a protein of about 75 kDa which corresponded approximately to the predicted truncated form of the 102 kDa E2–E3 protein.

DISCUSSION

The primary structure of the *T. ferrooxidans* PDH genes differ from that reported for other Gram-negative bacteria. The E1 component of the *T. ferrooxidans* PDH complex consists of two subunits, resembling the heterotetrameric $\alpha_2\beta_2$ form found in Gram-positive bacteria and eukaryotes, rather than the homodimeric $\alpha_2$ form found in Gram-negative bacteria such as *E. coli*, *Alcaligenes eutrophus* (Hein & Steinbüchel, 1994), *A. vinelandii* (Hanemaaijer et al., 1988) and *Neisseria meningitidis* (accession no. X82637). The dihydrolipoil transacetylase and lipoamide dehydrogenase subunits are different from any organism previously reported in that they are translated as a single fusion polypeptide. Two pieces of evidence support this observation. Firstly, DNA sequence data indicated that there was no stop codon between the dihydrolipoil transacetylase and the lipoamide dehydrogenase portions of the fusion protein. Secondly, a protein of approximately 100 kDa was synthesized from plasmid pPD23. This plasmid contained the complete ORF for the 102 kDa fusion protein but did not contain intact genes for any other polypeptides from the cloned *T. ferrooxidans* DNA.

The unusual relationship of the *T. ferrooxidans* PDH complex to similar enzymes from other organisms was confirmed by comparison of the amino acid sequence alignments of the subunits. Examination of the phylogram based on amino acid sequence alignments of the acetoin and PDH E1 $\alpha$ subunits suggested that these subunits may be divided into three clusters, a eukaryote cluster, a bacterial PDH cluster and a bacterial acetoin dehydrogenase cluster (Fig. 4a). The *T. ferrooxidans* E1 $\alpha$ subunit was positioned between the eukaryote cluster and the bacterial acetoin dehydrogenase cluster. Although the clusters are less clear in the case of the E1 $\beta$ subunits, the *T. ferrooxidans* subunit is placed in a similar position between those of the eukaryotes and the other bacteria (Fig. 4b). A comparison of the E2 subunits with the E2 portion of the *T. ferrooxidans* E2–E3 fusion protein confirmed the placement of the *T. ferrooxidans* E2-homologous region in a position between the bacterial and eukaryote clusters (Fig. 4c). The phylogram of the E3 subunits was less clear because relatively low sequence similarities resulted in bootstrap values with much lower confidence levels than the other phylograms (data not shown).

ACKNOWLEDGEMENTS

We thank Professor John Guest for the *E. coli* strain JRG1342 and the plasmid pGS87 and Di James for assistance with sequencing. This work was supported by grants from the Foundation for Research Development and Gencor (South Africa).

REFERENCES


Pyruvate dehydrogenase of *Thiobacillus ferrooxidans*


Received 31 December 1996; revised 20 March 1997; accepted 1 April 1997.