Protein Science (1996), *5*:1719–1736. Cambridge University Press. Printed in the USA.

Sequence and organization of genes encoding enzymes involved in pyruvate metabolism in *Mycoplasma capricolum*

PENG-PENG ZHU AND ALAN PETERKOFSKY

Laboratory of Biochemical Genetics, National Heart, Lung and Blood Institute, Bethesda, Maryland 20892

(RECEIVED March 6, 1996; ACCEPTED May 29, 1996)

Abstract

The region of the genome of *Mycoplasma capricolum* upstream of the portion encompassing the genes for Enzymes I and IIA^{glc} of the phosphoenolpyruvate:sugar phosphotransferase system (PTS) was cloned and sequenced. Examination of the sequence revealed open reading frames corresponding to numerous genes involved with the oxidation of pyruvate. The deduced gene organization is *naox* (encoding NADH oxidase)-*lplA* (encoding lipoate-protein ligase)-*odpA* (encoding pyruvate dehydrogenase EIα)-*odpB* (encoding pyruvate dehydrogenase EIβ)-*odp2* (encoding pyruvate dehydrogenase EII)-*dldH* (encoding dihydrolipoamide dehydrogenase)-*pta* (encoding phosphotransacetylase)-*ack* (encoding acetate kinase)-*orfA* (an unknown open reading frame)-*kdtB-ptsI-crr*. Analysis of the DNA sequence suggests that the *naox* and *lplA* genes are part of a single operon, *odpA* and *odpB* constitute an additional operon, *odp2* and *dldH* a third operon, and *pta* and *ack* an additional transcription unit. Phylogenetic analyses of the protein products of the *odpA* and *odpB* genes indicate that they are most similar to the corresponding proteins from *Mycoplasma genitalium*, *Acholeplasma laidlawii*, and Gram-positive organisms. The product of the *odp2* gene contains a single lipoyl domain, as is the case with the corresponding proteins from *M. genitalium* and numerous other organisms. An evolutionary tree places the *M. capricolum odp2* gene product in close relationship to the corresponding proteins from *A. laidlawii* and *M. genitalium*. The *dldH* gene encodes an unusual form of dihydrolipoamide dehydrogenase that contains an aminoterminal extension corresponding to a lipoyl domain, a property shared by the corresponding proteins from *Alcaligenes eutrophus* and *Clostridium magnum*. Aside from that feature, the protein is related phylogenetically to the corresponding proteins from *A. laidlawii* and *M. genitalium*. The phosphotransacetylase from *M. capricolum* is related most closely to the corresponding protein from *M. genitalium* and is distinguished easily from the enzymes from *Escherichia coli* and *Haemophilus influenzae* by the absence of the characteristic amino-terminal extension. The acetate kinase from *M. capricolum* is related evolutionarily to the homologous enzyme from *M. genitalium*. Map position comparisons of genes encoding proteins involved with pyruvate metabolism show that, whereas all the genes are clustered in *M. capricolum*, they are scattered in *M. genitalium*.

Keywords: acetate kinase; dihydrolipoamide dehydrogenase; lipoate-protein ligase; NADH oxidase; PTS; phosphotransacetylase; pyruvate dehydrogenase

Mycoplasmas, the smallest free-living organisms, contain the least complex genomes (Razin, 1985, 1992). They appear to have evolved from Gram-positive bacteria by selective elimination of nonessential genes (Maniloff, 1983). Growth of these organisms requires complex media, reflecting the loss of many genes encoding anabolic enzymes, but retention of genes encoding catabolic pathways (Miles, 1992). *Mycoplasma capricolum* has a genome of 1,155 kb, sufficient to encode approximately 350 genes (Miyata et al., 1991).

It was shown previously (Cirillo, 1979) that *M. capricolum* can metabolize carbohydrates by the ubiquitous phosphoenolpyruvate:sugar phosphotransferase system (PTS) (for review, see Cirillo, 1979). This system promotes phosphotransfer from phosphoenolpyruvate to the heat-stable phosphocarrier protein HPr in a reaction proposed to be catalyzed by the homodimeric form of Enzyme I. Phosphorylated HPr then transfers a phosphoryl group to the sugar-specific acceptor proteins, referred to as Enzymes II. Each Enzyme II complex consists of one or two membrane embedded proteins or domains (IIC and IID), as well

Reprint requests to: Dr. Alan Peterkofsky, National Institutes of Health, Building 36, Room 4C-11, Bethesda, Maryland 20892; e-mail: alan@codon.nih.gov.

[1] Nucleotide sequence(s) reported in this paper has (have) been submitted to the GenBank™/EMBL Data Bank with accession number (U62057).

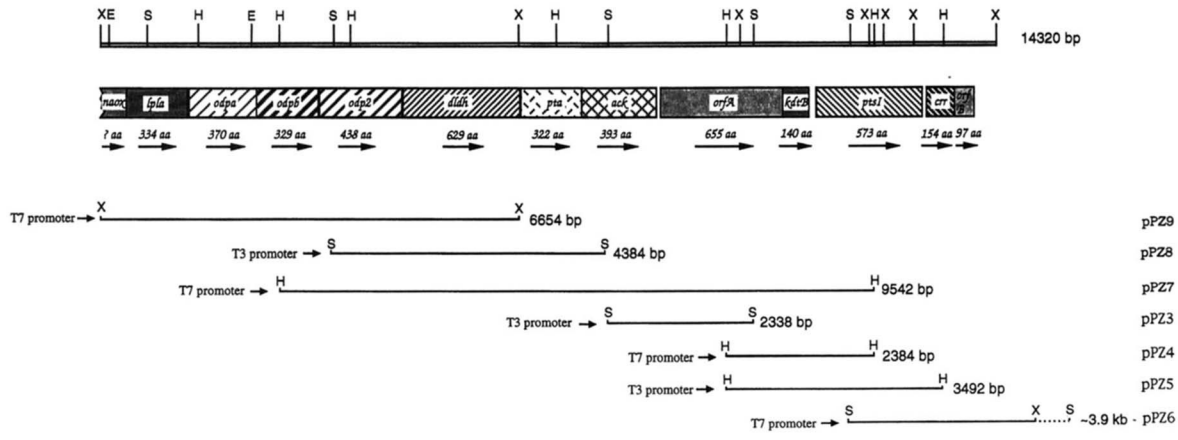


Fig. 1. Restriction and genetic maps of the region of the *M. capricolum* genome upstream of the ptsI-crr operon. The three clones that were isolated and sequenced are (**A**) pPZ7, the 9,542-bp clone derived from a partial *Hin*d III digest, (**B**) pPZ8, the 4,384-bp *Spe* I *clone*, (**C**) pPZ9, the 6,654-bp *Xba* I clone. The internal *Xba* I site is located at base 10209 (see Fig. 2). The internal *Eco*R I sites are located at bases 141 and 2396; The internal *Spe* I sites are located at bases 720, 3710, and 8094; the internal *Hin*d III sites are at bases 1551, 2834, 3966, 7266, and 9992. Horizontal arrows indicate direction of transcription of the indicated genes. --- aa, the number of amino acid residues of the deduced open reading frame. *naox*, the gene encoding NADH oxidase; *lplA*, the gene encoding lipoate-protein ligase; *odpA*, the gene encoding the EIα subunit of pyruvate dehydrogenase; *odpB*, the gene encoding the EIβ subunit of pyruvate dehydrogenase; *odp2*, the gene encoding the EII subunit of pyruvate dehydrogenase; *dldh*, the gene encoding dihydrolipoamide dehydrogenase; *pta*, the gene encoding phosphotransacetylase; *ack*, the gene encoding acetate kinase; *kdtB*, an open reading frame homologous to the *E. coli* KDTB protein product; *ptsI*, the gene encoding Enzyme I of the PTS; *crr*, the gene encoding enzyme IIA^{glc} of the PTS.

as two cytoplasmic proteins or domains (IIA and IIB) that pass the phosphoryl group from HPr to a specific sugar substrate during its translocation across the membrane (Postma et al., 1993).

In previous studies from this laboratory (Zhu et al., 1993, 1994), we reported the sequences of cloned *pts* genes from *M. capricolum*. It was shown that the *ptsH* gene, encoding HPr, constituted a unique monocistronic operon (the *ptsH* operon) in contrast to the typical occurrence in all *pts* operons sequenced to date of the *ptsH* gene immediately upstream of the *ptsI* gene, encoding Enzyme I. It was also established that *ptsI* is located in a dicistronic operon containing the gene encoding the glucose-specific Enzyme IIA (*crr*). In the present study, the region of the *M. capricolum* genome upstream of the *ptsI-crr* operon is characterized. It is shown that this region includes essentially all the genes necessary for the use of pyruvate.

Results

Cloning of the genes required for pyruvate metabolism

Previously we sequenced the region of the *M. capricolum* genome encoding the *ptsI-crr* operon (Zhu et al., 1994). Clones that were isolated in the course of that sequencing project were designated pPZ3, pPZ4, pPZ5, and pPZ6 (see Fig. 1). With the intention of investigating the organization of genes located upstream of the *ptsI-crr* operon, further clones were isolated. A partial digest with *Hin*d III provided DNA for the isolation of a 9,542-bp clone, designated pPZ7, which was screened using an oligonucleotide based on the pPZ4 sequence (from bases 3466 to 3422 of the sequence published previously) (Zhu et al., 1994) (see Materials and methods). Other DNA clones were isolated by a similar approach. pPZ8 was isolated from an *Spe* I digest and identified using an oligonucleotide sequence from pPZ7 corresponding to base numbers 8018–8062. The clone pPZ9 (from an *Xba* I digest) was also detected with a pPZ7 probe corresponding to bases 2882–2926 (see Fig. 1).

Sequence analysis of the genes required for pyruvate metabolism

Clones pPZ7, pPZ8, and pPZ9 were sequenced on both strands and then analyzed for open reading frames to reveal the gene organization illustrated in Figure 1. It was surprising to observe that a block of eight genes in this upstream region was in some way involved with the metabolism of pyruvate. The deduced gene organization, starting from the 5′-end of the sequence, is *naox* (encoding NADH oxidase), *lplA* (encoding lipoate-protein ligase), *odpA* (encoding the EIα subunit of pyruvate dehydrogenase), *odpB* (encoding the EIβ subunit of pyruvate dehydrogenase), *odp2* (encoding the EII subunit of pyruvate dehydrogenase), *dldH* (encoding dihydrolipoamide dehydrogenase), *pta* (encoding phosphotransacetylase), *ack* (encoding acetate kinase). The region downstream of *ack* (*orfA*, *kdtB*, *ptsI*, *crr*, and *orfB*) has been described previously (Zhu et al., 1994).

The combination of the three sequenced clones (pPZ7, pPZ8, and pPZ9) allowed us to deduce a total of 10,214 bp of DNA sequence (see Fig. 2). Bases 1–280 coded for an open reading frame that is homologous to the carboxyl-terminal region of the amino acid sequence of NADH oxidase from *Enterobacter faecalis* (Genbank accession no. P37061) and *M. genitalium* (MG275 from the TIGR database) (Fig. 3). Comparison (by BESTFIT analysis) of the *M. capricolum* and *E. faecalis* sequences indicated 64% similarity and 39% identity of the compared sequences over a length of 88 residues. When the *M. capricolum* sequence was compared with the *M. genitalium* se-

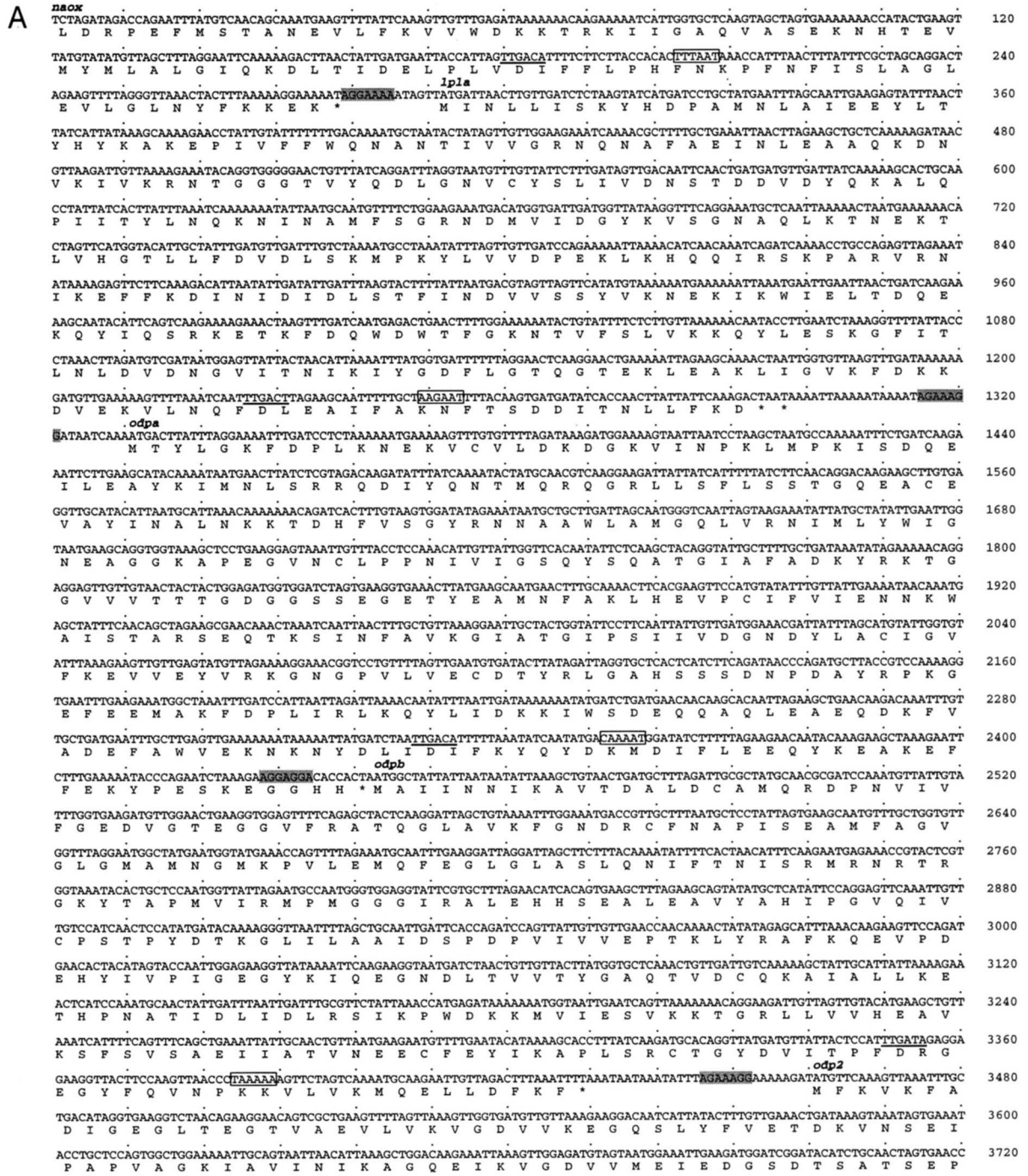


Fig. 2A. (*Figure continues on next two pages.*)

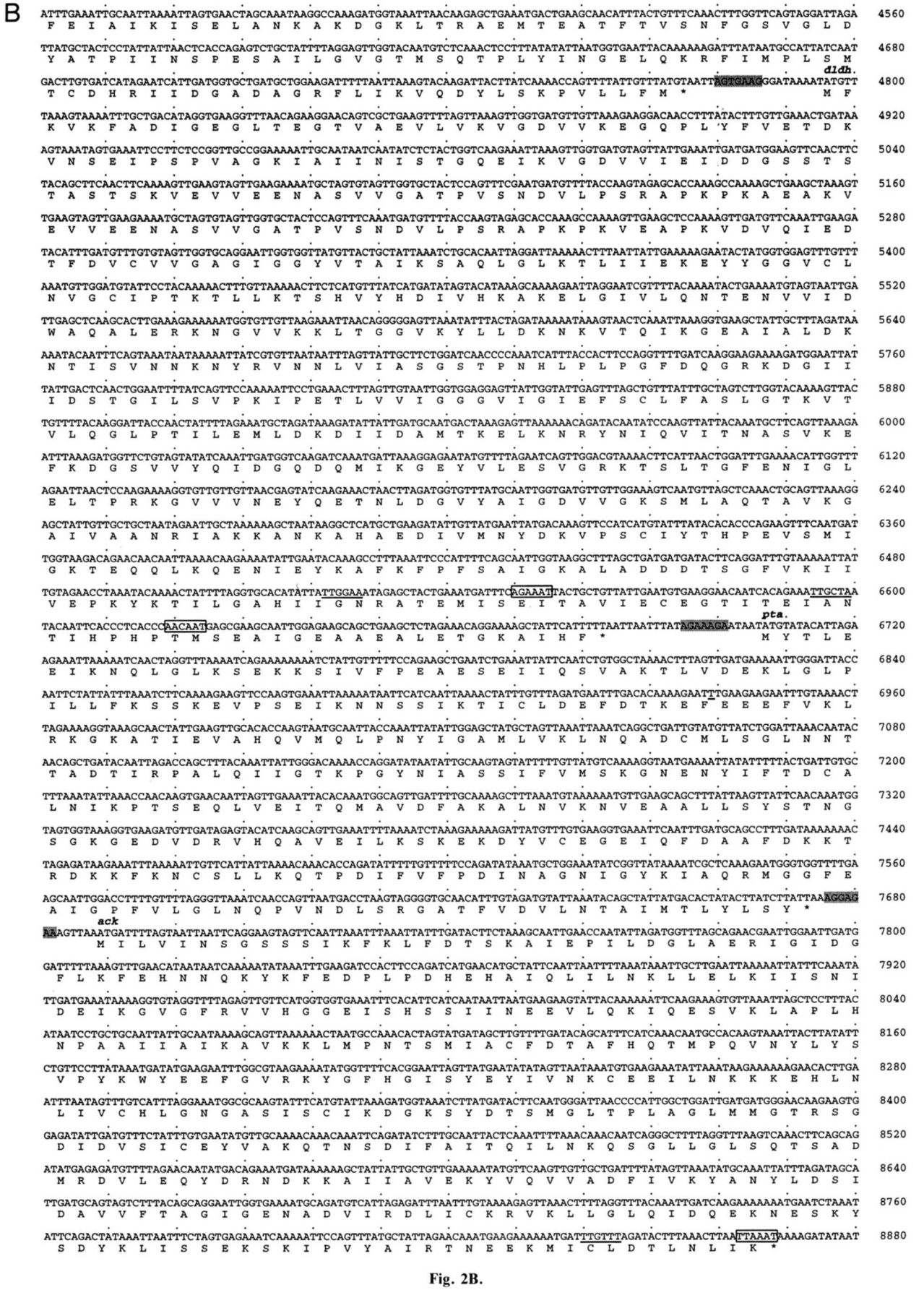


Fig. 2B.

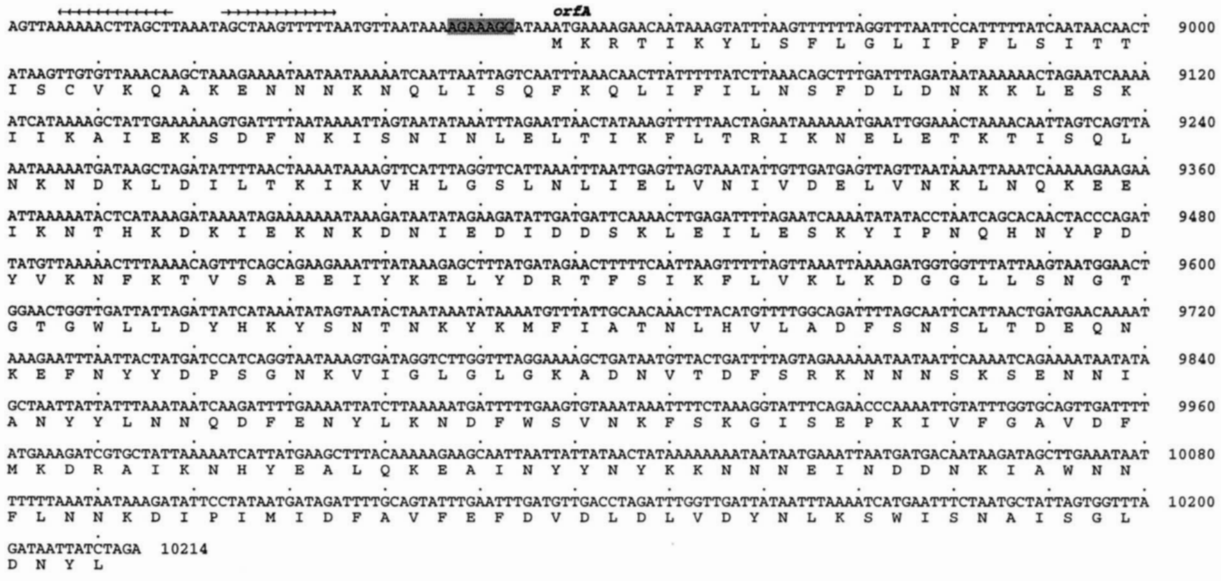


Fig. 2. Nucleotide sequence of the region of the *Mycoplasma* genome encoding genes involved with pyruvate metabolism.[1] A partial sequence of *naox* and the entire nucleotide sequences of *lplA*, *odpA*, *odpB*, *odp2*, *dldH*, *pta*, *ack*, and an ORF designated *orfA* (Zhu et al., 1994), as well as the deduced amino acid sequences, are shown. The putative promoter regions [i.e., 35 regions (underlined) and 10 regions (boxed)] are shown for the *lplA*, *odpA*, *odpB*, *odp2*, *pta*, and *orfA* genes. A sequence corresponding to a likely transcription termination site downstream of *ack* is marked with diverging arrows. Shine–Dalgarno ribosome binding sites for *lplA*, *odpA*, *odpB*, *odp2*, *dldH*, *pta*, *ack*, and *orfA* are highlighted with shaded boxes. Asterisks denote translation stop codons.

quence, it showed 77% similarity and 52% identity over a length of 92 residues.

After a gap of 11 bases, an open reading frame (from bases 292 to 1312) was identified. The end of this sequence was characterized by the presence of four in-frame stop codons (all TAA). Use of FASTA analysis of the deduced protein product led to its identification as the gene encoding lipoate–protein ligase. Alignment of the sequence with the lipoate–protein ligases from *M. genitalium* (MG270 from the TIGR database[2]) and *Escherichia coli* (Genbank accession no. P32099) is shown in Figure 4. BESTFIT analysis of the *Mycoplasma* sequence (334 amino acid residues) with the lipoate–protein ligase from *E. coli* (337 amino acid residues) showed 55% similarity and 32% identity over a length of 335 residues (data not shown). A similar analysis with the enzyme from *M. genitalium* (336 residues) showed 62% similarity and 36% identity over a length of 335 residues.

After a gap of 19 bases, an open reading frame (from bases 1331 to 2443) was observed. FASTA analysis showed the protein product to be the α-subunit of pyruvate dehydrogenase Enzyme I. Figure 5 shows an alignment of the sequence from *M. capricolum* with those from other sources recovered from the Genbank and TIGR databases. It is clear that the regions of total amino acid conservation (shown in reverse shading) are

[2] The *M. genitalium* sequence MG279 is listed in the TIGR database as unknown in function. The FASTA analysis reported here identifies the sequence as the gene encoding lipoate–protein ligase.

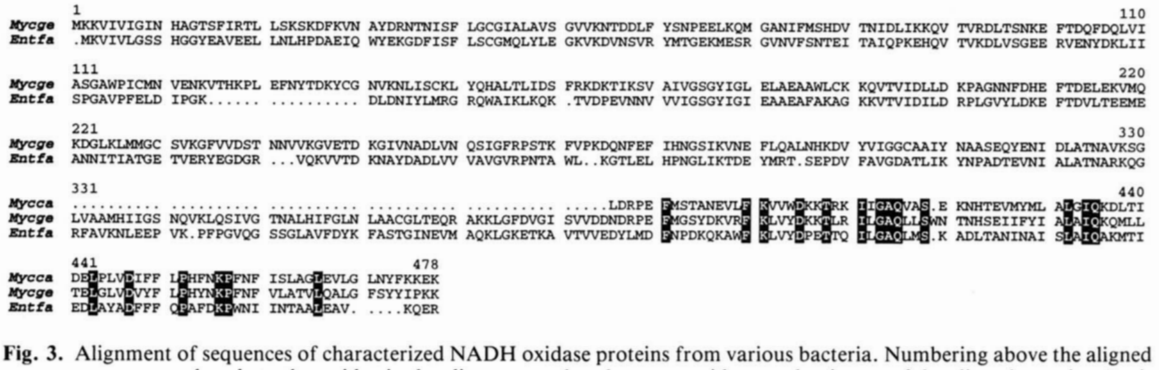


Fig. 3. Alignment of sequences of characterized NADH oxidase proteins from various bacteria. Numbering above the aligned sequences corresponds only to the residue in the alignment rather than to a residue number in any of the aligned proteins. Residues that are conserved in the listed NADH oxidase proteins are shown in reverse shading. Abbreviations used and references to published sequences are: *M. capricolum* (*Mcapr*) (this work); *M. genitalium* (*Mgeni*) (MG275 from the TIGR database); *E. faecalis* (*Entfa*) (Ross & Claiborne, 1992).

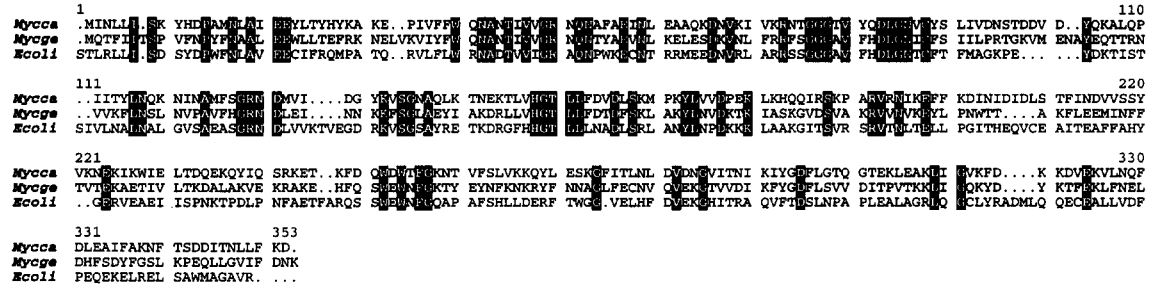


Fig. 4. Alignment of sequences of characterized lipoate-protein ligase proteins from various bacteria. Numbering above the aligned sequences corresponds only to the residue in the alignment rather than to a residue number in any of the aligned proteins. Residues that are conserved in the listed lipoate-protein ligase proteins are shown in reverse shading. Abbreviations used and references to published sequences are: *M. capricolum* (*Mcapr*) (this work); *M. genitalium* (*Mgeni*) (MG270 from the TIGR database); *E. coli* (*Ecoli*) (Morris et al., 1994).

Fig. 5 (*on next page*). Alignment of sequences of characterized Enzyme I-α proteins of the pyruvate and α-ketoacid dehydrogenases and acetoin catabolism complexes from various bacteria. Numbering above the aligned sequences corresponds only to the residue in the alignment rather than to a residue number in any of the aligned proteins. Residues that are conserved in all listed Enzyme I-α proteins are shown in reverse shading. Abbreviations used and references to published sequences are: *M. capricolum* (odpa-mycca) (this work); *M. genitalium* (odpa-mycge) (MG274 from the TIGR database); *A. laidlawii* (odpa-achla) (Wallbrandt et al., 1992); *B. subtilis* (odpa-bacsu) (Hemila et al., 1990) and (odba-bacsu) (Wang et al., 1993); *B. stearothermophilus* (odpa-bacst) (Borges et al., 1990); *A. thaliana* (odpa-arath) (Luethy et al., 1995); yeast (odpa-yeast) (Behal et al., 1989); human (odba-human) (McKean et al., 1992), (odpa-human) (Ho et al., 1989), and (odpt-human) (Dahl et al., 1987); pig (odpa-pig) (Sermon et al., 1990); mouse (odpa-mouse) (Fitzgerald et al., 1992) and (odpt-mouse) (Fitzgerald et al., 1992); rat (odba-rat) (Zhang et al., 1987), (odpt-rat) (Cullingford et al., 1993), and (odpa-rat) (Cullingford et al., 1994); *Ascarus suum* (odpt-ascu) (Johnson et al., 1992) and (odpa-ascu) (Johnson et al., 1992); *C. magnum* (acoa-cloma) (Kruger et al., 1994); *A. eutrophus* (acoa-alceu) (Priefert et al., 1991); *Pelobacter carbinolicus* (acoa-pelca) (Oppermann & Steinbuchel, 1994); *K. pneumoniae* (acoa-klepn) (Deng et al., 1994); bovine (odba-bovin) (Hu et al., 1988); *Pseudomonas putida* (odba-psepu) (Burns et al., 1988). odpa and odpt designations correspond to pyruvate dehydrogenase, whereas odba designations correspond to α-keto acid dehydrogenase complex gene products, respectively, and the acoa designation corresponds to acetoin catabolism complexes.

also shared by the sequence from *M. capricolum*. The signature sequence (shown as residues 250–280) G(D/E) (G/A) (X26)NN, characteristic of thiamine diphosphate-dependent enzymes, is also found in the *M. capricolum* sequence. A phylogenetic tree (Fig. 6) of the sequences of the α-subunit family shows that the *M. capricolum* protein is related closely to the corresponding proteins from *M. genitalium*, *Acholeplasma laidlawii*, and those from Gram-positive organisms (*Bacillus subtilis* and *B. stearothermophilus*).

The last base of the termination codon (TAA) of the sequence encoding the α-subunit of pyruvate dehydrogenase Enzyme I is also the first base of an open reading frame (from bases 2443 to 3432) that codes for the β-subunit of pyruvate dehydrogenase Enzyme I. Figure 7 shows an alignment of the sequence from *M. capricolum* with those from other sources recovered from the Genbank and TIGR databases. The alignment shows that residues that are totally conserved (shown as reverse shading) are also identical in the sequence from *M. capricolum*. It is noteworthy that all the bacterial sequences are from 40 to 70 residues shorter at the amino-terminal end than those sequences from eukaryotes. A phylogenetic tree (Fig. 8) shows a similar pattern to that observed for ODPA; the protein from *M. capricolum* is most closely related to those from *M. genitalium* (MG273, TIGR, database) and *A. laidlawii*.

After a gap of 28 bases, a new open reading frame (from bases 3461 to 4877), corresponding to the Enzyme II subunit of pyruvate dehydrogenase, was identified. An alignment of the *M. capricolum* sequence with other Enzyme II sequences (Fig. 9) shows clearly that the conserved residues (reverse shading) are found in the *M. capricolum* sequence. Boxes shaded in grey cor-

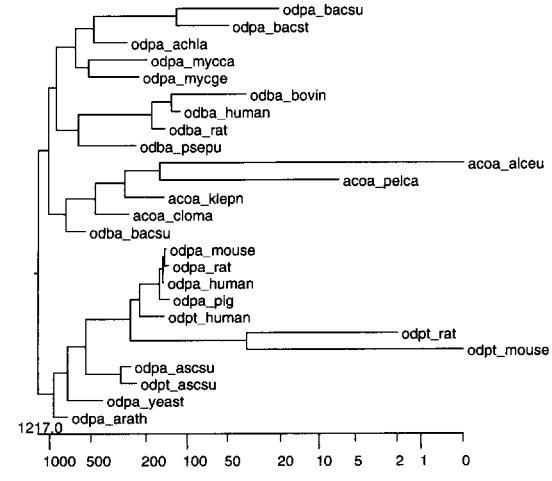
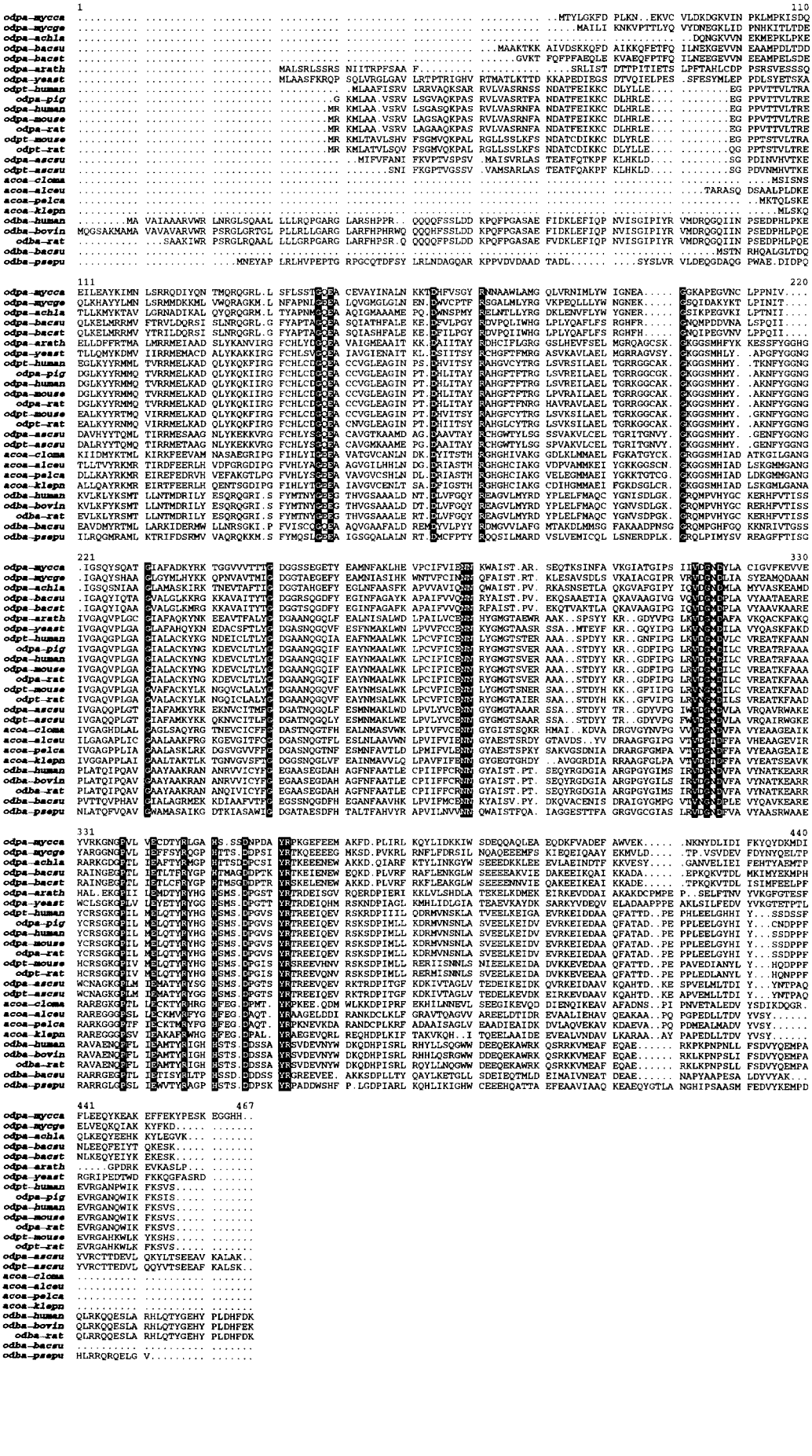


Fig. 6. Phylogenetic tree of sequenced proteins of the Enzyme I-α subunit family. Relative evolutionary distances are shown on the horizontal scale. Abbreviations are as in the legend to Figure 5.



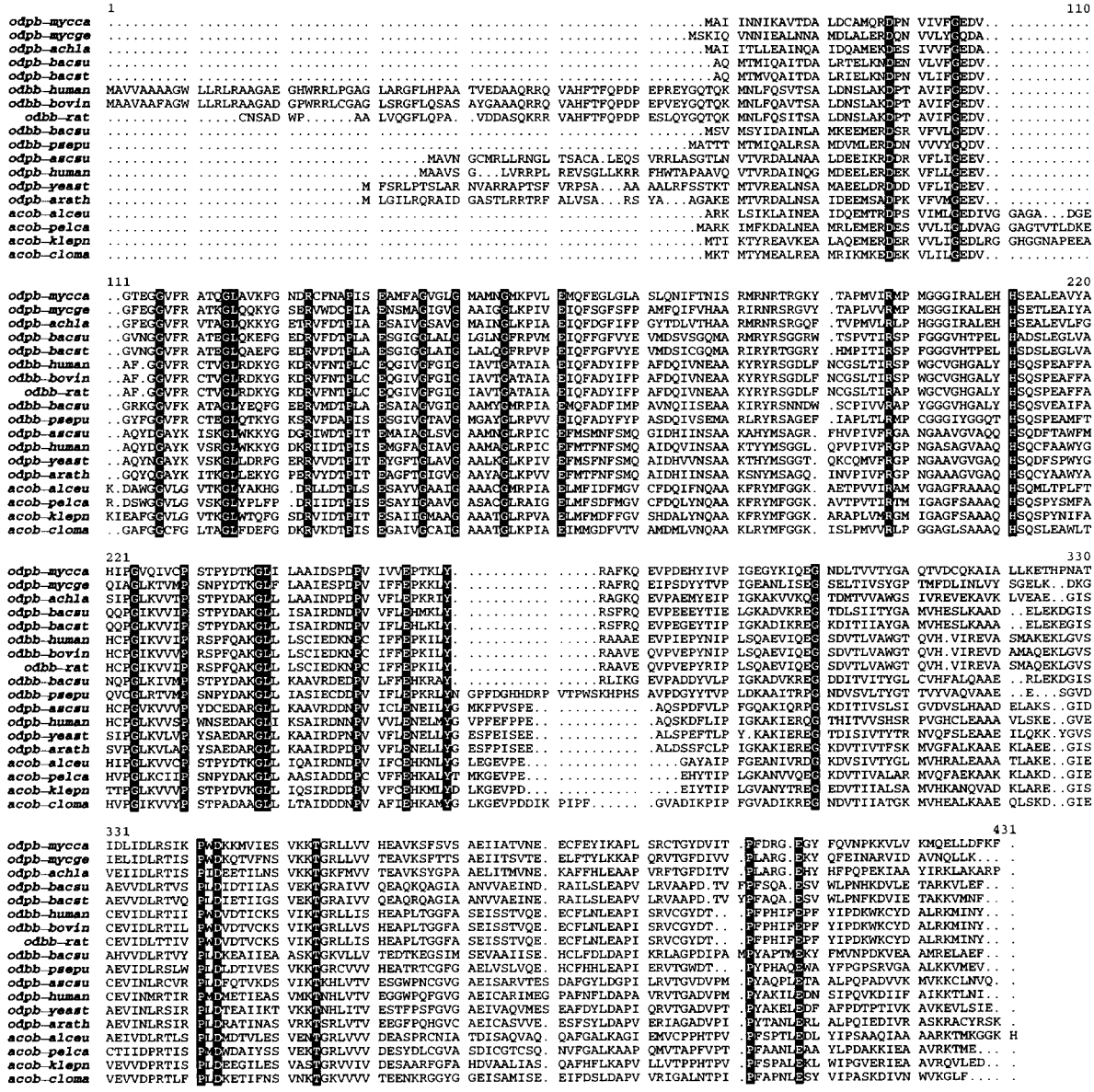


Fig. 7. Alignment of sequences of members of the Enzyme I-β subunit family. Numbering above the aligned sequences corresponds only to the residue in the alignment rather than to a residue number in any of the aligned proteins. Residues that are conserved in the listed Enzyme I-β proteins are shown in reverse shading. Abbreviations used and references to published sequences are: *M. capricolum* (odpb-mycca) (this work); *M. genitalium* (odpb-mycge) (MG273 from the TIGR database); *A. laidlawii* (odpb-achla) (Wallbrandt et al., 1992); *B. subtilis* (odpb-bacsu) (Hemila et al., 1990) and (odbb-bacsu) (Wang et al., 1993); *B. stearothermophilus* (odpb-bacst) (Borges et al., 1990), human (odbb-human) (Nobukuni et al., 1990a) and (odpb-human) (Ho & Patel, 1990); bovine (odbb-bovin) (Nobukuni et al., 1990b); rat (odbb-rat) (Zhao et al., 1992); *P. putida* (odbb-psepu) (Burns et al., 1988); *A. suum* (odpb-ascu) (Wheelock et al., 1991); yeast (odpb-yeast) (Miran et al., 1993); *A. thaliana* (odpb-arath) (Luethy et al., 1994); *A. eutrophus* (acob-alceu) (Priefert et al., 1991); *P. carbinolicus* (acob-pelca) (Oppermann & Steinbuchel, 1994); *K. pneumoniae* (acob-klepn) (Deng et al., 1994); *C. magnum* (acob-cloma) (Kruger et al., 1994). odpb, odbb, and acob refer to the genes encoding the β-subunits of Enzyme I of the pyruvate dehydrogenase, 2-oxoisovalerate dehydrogenase, and acetoin catabolism complexes.

respond to the lipoyl domains (approximately 70 residues). The proteins from *E. coli* and *Azotobacter vinelandii* contain three such domains; those from *E. faecalis*, *A. laidlawii*, *Alcaligenes eutrophus*, *Haemophilus influenzae*, *Pseudomonas aeruginosa*, *Dictyostelium discoideum*, *Arabidopsis thaliana*, and human contain two; and all the others shown, including the *Mycoplasma* sequences, contain one lipoyl domain. Each one of the lipoyl domains contains a conserved lysine residue (the site of lipoylation), generally preceded by an aspartyl residue. In the case of *A. vinelandii*, the aspartate of the first lipoyl domain is replaced by alanine. For *Klebsiella pneumoniae*, the conserved aspartate is replaced by serine. All the lipoyl domains are also character-

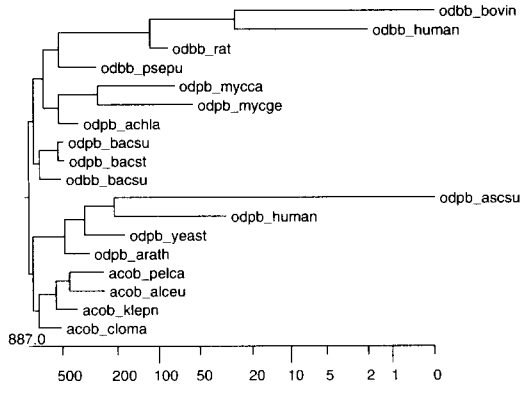


Fig. 8. Phylogenetic tree of sequenced proteins of the Enzyme I-β subunit family. Relative evolutionary distances are shown on the numerical scale. Abbreviations are as in the legend to Figure 7.

ized by the presence of a conserved GD pair. In the case of odp2 from *B. subtilis*, the G is replaced by N. For odo2 of *A. vinelandii*, odp2 of rat, and odp2 of human, the D is replaced by E. In approximately half of the proteins (odb2-psepu, odo2-azovi, odp2-rat, odo2-rat, odo2-human, odp2-entfa, odpa-achla, odp2-alceu, odp2-haein, odp2-pseae, odp2-dicdi, odp2-arath, odp2-human, odp2-ecoli, and odp2-azovi), the lipoyl domain is separated from the remainder of the protein by an A,P-rich linker. It is worth noting that the linker in odp2-dicdi is enriched in serine. The central portions of the proteins contain an E3 binding domain of approximately 50 residues and the carboxyl-terminal 250 residues corresponds to the catalytic domain, which effects the acyl transfer reaction. The conserved histidine (at residue 753) acts as the general base catalyst (Mattevi et al., 1992). A phylogenetic tree (Fig. 10) shows the preservation of the evolutionary relationship of the proteins from *M. capricolum*, *M. genitalium*, and *A. laidlawii*.

Following a gap of 18 bases, an open reading frame (from bases 4796 to 6686) was found; it corresponds to the gene encoding dihydrolipoamide dehydrogenase. The alignment shown in Figure 11 emphasizes a unique feature of the dihydrolipoamide dehydrogenase from *M. capricolum*. This protein contains an aminoterminal lipoyl domain (shown as a boxed grey area), a feature observed previously only in the DLDH proteins from *A. eutrophus* (Hein & Steinbuchel, 1994) and *Clostridium magnum* (Kruger et al., 1994). These lipoyl domains contains the characteristic conserved lysine residue (the site of lipoylation) (Mattevi et al., 1992), preceded by an aspartyl residue (residues 42 and 43). There are a number of totally conserved residues (shown in reverse shading) characteristic of dihydrolipoamide dehydrogenase sequences. The region from residues 171 to 180 corresponds to the motif GXGXXGYXXA, which is a possible nucleotide binding site. The conserved region from residues 206 to 218 corresponds to the signature GGXCLNXGCXP(S/T)K. In this region, the combination of the flavin ring with an adjacent disulphide bridge forms the redox center, which is involved with the transfer of the reducing equivalents from dihydrolipoamide to NAD^+. The vicinal cysteines in this region may bind lipoic acid and undergo reversible oxidation-reduction. The active species (dimeric) of E3 has four domains: FAD binding, NAD binding, central, and interface (Mattevi et al., 1992). A phylogenetic tree (Fig. 12) documents the consistency of the relationship of the *M. capricolum*, *M. genitalium*, and *A. laidlawii* sequences.

A space of 21 bases separates the coding sequence for dihydrolipoamide dehydrogenase from the next open reading frame. The region from bases 6707 to 7675 corresponds to the gene encoding phosphotransacetylase. The seven known sequences for phosphotransacetylase are aligned in Figure 13. The proteins from *E. coli* and *H. influenzae* are unique in that they have amino-terminal extensions of approximately 400 residues. All of the sequences show characteristic regions of total conservation, highlighted by reverse shading. A phylogenetic tree (Fig. 14) shows clearly a distinct branch for the proteins from *H. influenzae* and *E. coli* and a separate branch for the proteins from *M. genitalium* and *M. capricolum*.

A 13-base spacer separates the *pta* gene from the next open reading frame, encoding acetate kinase (from bases 7689 to 8866). The alignment shown in Figure 15 compares the six known acetate kinase sequences. The amino-terminal regions are characterized by a highly conserved sequence (residues 12–17, shown in reverse shading), GSSS(I/L)K, that may be involved in nucleotide binding. The phylogenetic tree shown in Figure 16 demonstrates the sequence relatedness of the *M. capricolum* and *M. genitalium* proteins, as well as the *E. coli*, *H. influenzae* similarities. Seventy-one bases separate the end of the sequence encoding acetate kinase from the one encoding *orfA*.

The sequence shown in Figure 2 was examined for regulatory features. Upstream of the *lplA* gene, within the coding sequence of *naox*, there is a possible −35 (TTGACA, bases 180–185, underlined) and −10 (TTTAAT, bases 204–209, boxed) promoter sequence. Upstream of the *odpA* gene, within the *lplA* coding sequence, there is a possible −35 (TTGACT, bases 1226–1231, underlined) and −10 (AAGAAT, bases 1249–1254, boxed) sequence. Upstream of *odpB*, within the *odpA* coding sequence, is found the potential −35 (TTGACA, bases 2328–2333, underlined) and −10 (CAAAAT, bases 2353–2358, boxed) sequence. Within the *odpB* coding sequence, upstream of *odp2*, potential −35 (TTGATA, bases 3350–3355, underlined) and −10 (TAAA AA, bases 3384–3389, boxed) sequences are located. Within the *dldH* coding region, upstream of *pta*, there are two potential −35 sequences (TTGGAA, bases 6523–6528 and TTGCTA, bases 6594–6599, underlined) and two potential −10 (AGAAAT, bases 6550–6555 and AACAAT, bases 6619–6624, boxed) sequences. Near the end of the *ack* coding sequence is located a potential −35 (TTGTTT, bases 8841–8846, underlined) and −10 (TTAAAT, bases 8864–8869, boxed) sequence. The spacer between the end of the *ack* gene and the beginning of *orfA* (71 bases) is AT-rich (marked with diverging arrows) and may form a stem-loop structure that plays a role in transcription termination of the *ack*-containing message.

Probable ribosome-binding sites are recognizable at appropriate positions before coding sequences. AGGAAAA (bases 270–276, shaded) is found before *lplA*; AGAAAGG (bases 1315–1321, shaded) is found before *odpA*; AGGAGGA (bases 2425–2431, shaded) is localized preceding *odpB*; AGAAAGG (bases 3447–3453, shaded) precedes *odp2*; AGTGAAG (bases 4781–4787, shaded) is localized in front of *dldH*; AGAAAGA (bases 6696–6702, shaded) precedes *pta*; AGGAGAA (bases 7675–7681, shaded) is found in front of *ack*; AGAAAGC (bases 8927–8933, shaded) is found in front of *orfA*.

A

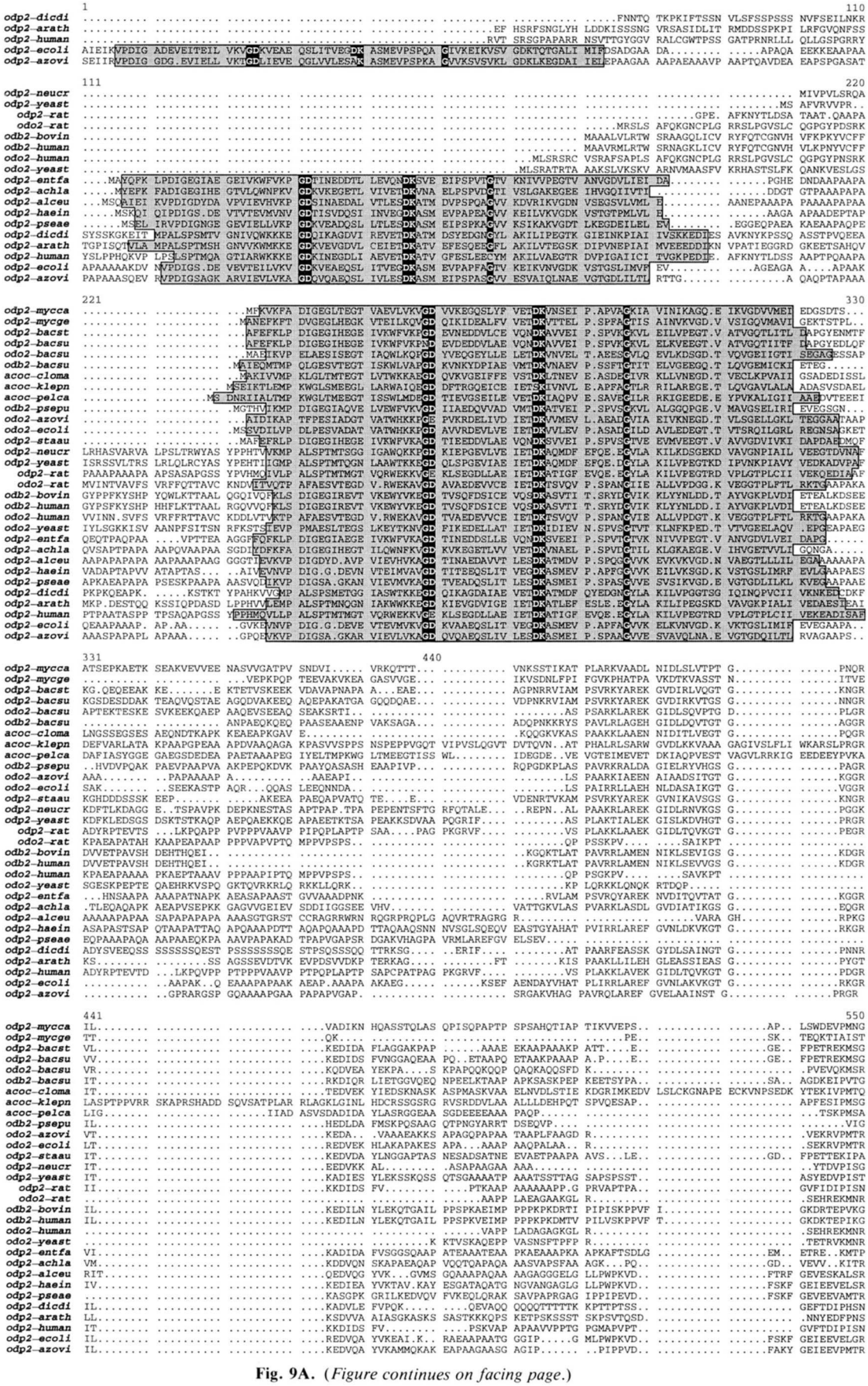


Fig. 9A. (*Figure continues on facing page.*)

B

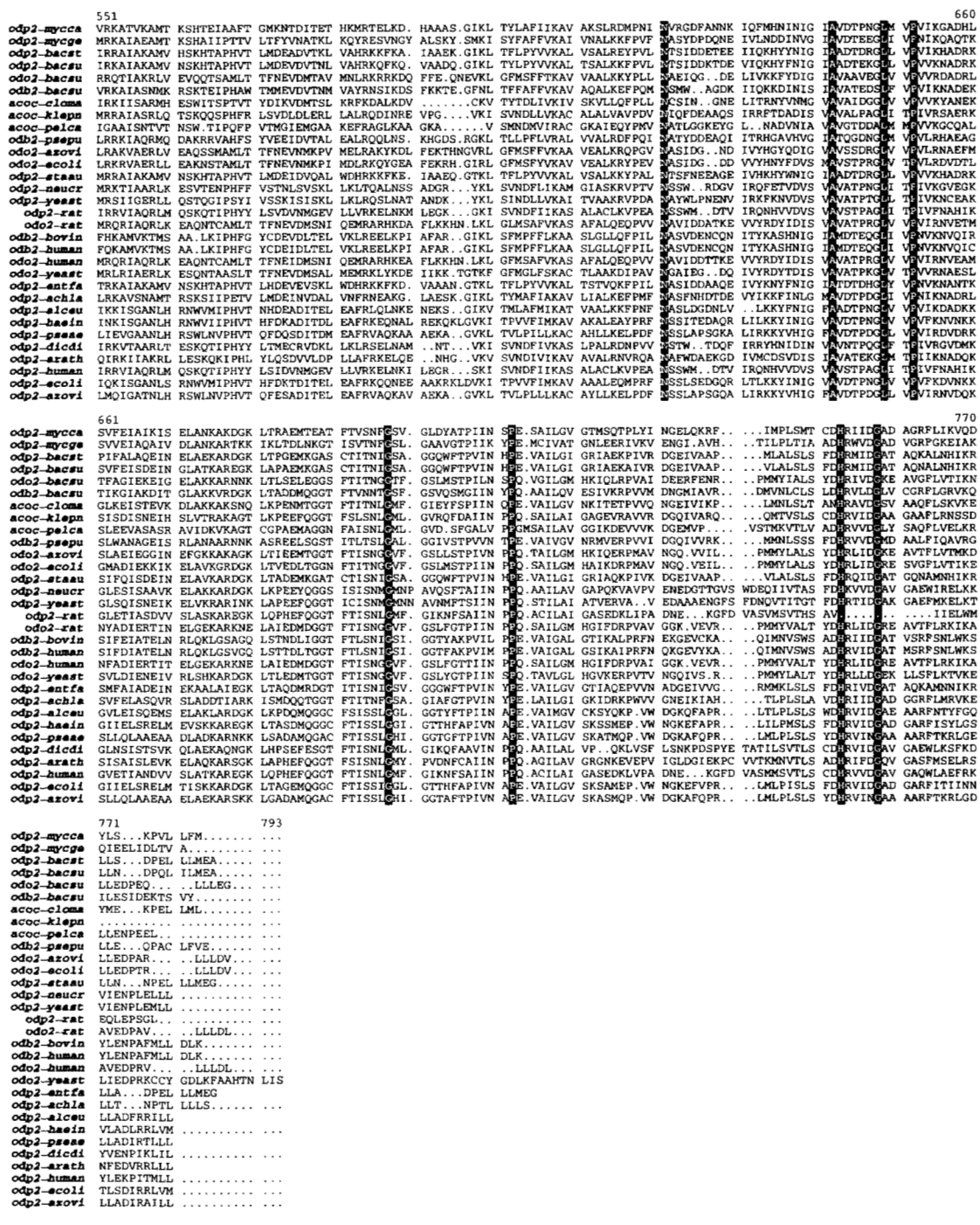


Fig. 9. Alignment of sequences of members of the Enzyme II family. Numbering above the aligned sequences corresponds only to the residue in the alignment rather than to a residue number in any of the aligned proteins. Residues that are conserved in the listed Enzyme II proteins are shown in reverse shading. Boxed shaded regions correspond to lipoyl domains. Abbreviations used and references to published sequences are: *M. capricolum* (odp2-mycca) (this work); *M. genitalium* (odp2-mycge) (MG272 from the TIGR database); *B. stearothermophilus* (odp2-bacst) (Borges et al., 1990); *B. subtilis* (odp2-bacsu) (Wang et al., 1993), (odb2-bacsu) (Wang et al., 1993), and (odo2-bacsu) (Carlsson & Hederstedt, 1989); *C. magnum* (acoc-cloma) (Kruger et al., 1994); *K. pneumoniae* (acoc-klepn) (Deg et al., 1994); *P. carbinolicus* (acoc-pelca) (Oppermann & Steinbuchel, 1994); *P. putida* (odb2-psepu) (Burns et al., 1988); *A. vinelandii* (odo2-azovi) (Westphal & de Kok, 1990) and (odp2-azovi) (Hanemaaijer et al., 1988); *E. coli* (odo2-ecoli) (Spencer et al., 1984) and (odp2-ecoli) (Guest, 1987); *Staphylococcus aureus* (odp2-staur) (Hemila, 1991); *Neurospora crassa* (odp2-neucr) (Kreader et al., 1989); yeast (odp2-yeast) (Niu et al., 1988) and (odo2-yeast) (Repetto & Tzagoloff, 1990); rat (odp2-rat) (Gershwin et al., 1987) and (odo2-rat) (Nakano et al., 1991); bovine (odb2-bovin) (Lau et al., 1988); human (odb2-human) (Lau et al., 1992), (odp2-human) (Thekkumkara et al., 1988), and (odo2-human) (Nakano et al., 1993); *E. faecalis* (odp2-entfa) (Allen & Perham, 1991); *A. laidlawii* (odp2-achla) (Wallbrandt et al., 1992); *A. eutrophus* (odp2-alceu) (Hein & Steinbuchel, 1994); *H. influenzae* (odp2-haein) (HI232 from the TIGR database); *P. aeruginosa* (odp2-pseae) (Genbank accession no. U47920); *D. discoideum* (odp2-dicdi) (Genbank accession no. u06634); *A. thaliana* (odp2-arath) (Guan et al., 1995). odp2, odb2, and odo2 refer to the genes encoding the Enzyme IIs of the pyruvate dehydrogenase; α-oxo acid dehydrogenase and 2-oxoglutarate dehydrogenase complexes, respectively; and acoc refers to acetoin dehydrogenase EII.

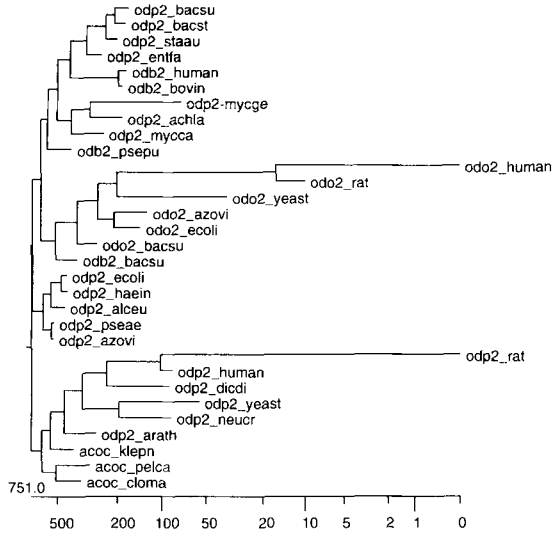


Fig. 10. Phylogenetic tree of sequenced proteins of the Enzyme II family. Relative evolutionary distances are shown on the numerical scale. Abbreviations are as in the legend to Figure 9.

Gene transcription analyses

The cloned sequence encodes 13 genes that might be transcribed in the same direction (see Fig. 1). The only significant gap between sequences occurs between the *ack* and *orfA* open reading frames. The program TERMINATOR was used to search the cloned sequence for potential transcription termination sequences (designated by diverging arrows in Fig. 2). Nineteen bases downstream of the translation termination codon for *ack*, there is a possible transcription termination region (from nucleotides 8886–8914). Using the program FOLDRNA, the stability of this stem-loop structure was calculated to be −11.8 kcal/mol. The stem-loop structure formed contains a perfectly matched stem 12 bases long. It is therefore possible that the region from *naox* to *ack* is transcribed as a single mRNA. Evidence was presented previously (Zhu et al., 1994) that the *ptsI* and *crr* genes constituted an operon. Consequently, it may be the case that the *orfA* and *kdtB* genes are cotranscribed.

Discussion

The sequence analysis presented here (see Figs. 1, 2) demonstrates the presence in *M. capricolum* of a unique arrangement of genes involved in the metabolism of pyruvate. In other bacterial species where gene mapping has been performed (*E. coli*, *H. influenzae*, *M. genitalium*), the genes are somewhat scattered throughout the genome. For example, in *E. coli*, the genes encoding the pyruvate dehydrogenase complex are located at 3 min on the genetic map, whereas the *ack*, *pta* and *ptsH*, *ptsI* and *crr* genes are in the 50–52-min region. In contrast, in *M. capricolum*, all the relevant genes are clustered in a single region. The positioning of these genes close to the *ptsI-crr* operon (whose products use PEP to form pyruvate) may also be of some regulatory significance. The scheme shown in Figure 17 indicates that all the enzymes required for the conversion of PEP to acetate and ATP, including those activities used for lipoylation of the pyruvate dehydrogenase and regeneration of NAD from NADH, are accounted for in the region of the *M. capricolum* genome sequenced in this study (see Fig. 1).

We reported previously (Zhu et al., 1993) that the HPr protein from *M. capricolum* has an unusually high isoelectric point. Using the computer program PEPTIDESORT, we calculated the pIs of the protein sequences deduced in this study. The pIs of the EI and EII proteins of the pyruvate dehydrogenase complex were in the range 5.3–6.75 and were similar for both *M. capricolum* and *M. genitalium*. However, several of the proteins from *M. genitalium* showed significantly higher pIs than the corresponding proteins from *M. capricolum*. For DLDH, the pIs for *M. capricolum* and *M. genitalium* were 5.35 and 7.08, respectively. For lipoate-protein ligase, the pIs for *M. capricolum* and *M. genitalium* were 6.97 and 9.29, respectively. For phosphotransacetylase, the pIs for *M. capricolum* and *M. genitalium* were 5.24 and 7.48, respectively. For acetate kinase, the pIs for *M. capricolum* and *M. genitalium* were 6.67 and 9.09, respectively. The significance of the widespread occurrence of proteins with high isoelectric points in *M. genitalium* remains to be clarified.

In order to evaluate the possibility that the genes encoding the enzymes involved in pyruvate metabolism are part of a polycistronic operon(s), northern blotting and primer extension experiments using probes derived from the plasmids described in Figure 1 were performed. No detectable mRNA species were found. Similar lack of success in detecting *ack*- or *pta*-specific mRNA species was reported for *Methanosarcina thermophila* (Latimer & Ferry, 1993) and it was suggested that the mRNAs may have a short half-life or be degraded rapidly during the RNA preparation. Consequently, the nature of the transcripts encoding the enzymes for pyruvate metabolism remains to be established.

The dihydrolipoamide dehydrogenase from *M. capricolum* is atypical, containing an amino-terminal lipoyl domain (see Fig. 11). This structure is also shared by the enzymes from *A. eutrophus* and *C. magnum*. This observation opens the possibility that effective function of the pyruvate dehydrogenase complex may be possible with the association of lipoate residues with either the E2 or E3 components. It is interesting to note that a recent description of an outer membrane protein from *Neisseria meningitidis* (de la Sierra et al., 1994) indicated that it contained an amino-terminal lipoyl domain and was otherwise homologous to lipoamide dehydrogenases.

The *M. capricolum* pyruvate dehydrogenase and *C. magnum* acetoin dehydrogenase complexes show a unique similarity. Each of these complexes contain an E2 with a single lipoyl domain, as well as an E3 with a single lipoyl domain. Comparison of the sequences of these lipoyl domains (see Figs. 9, 11) show that they are almost exact duplicates. This suggests that the lipoyl domains in the E3 proteins of these organisms arose by a duplication of the preexisting domain in the genes encoding the E2s.

In contrast, the pyruvate dehydrogenase complex of *A. eutrophus* is characterized by an E2 with two lipoyl domains, as well as an E3 with a single lipoyl domain. In this case, the two lipoyl domains of E2 are essentially identical, consistent with the idea that they arose by a duplication mechanism. However, the single lipoyl domain in the E3 diverges considerably from the sequences in the E2. Therefore, it seems unlikely that the lipoyl domain of the E2 in this organism arose by a simple duplication of the sequences in the gene encoding the E2.

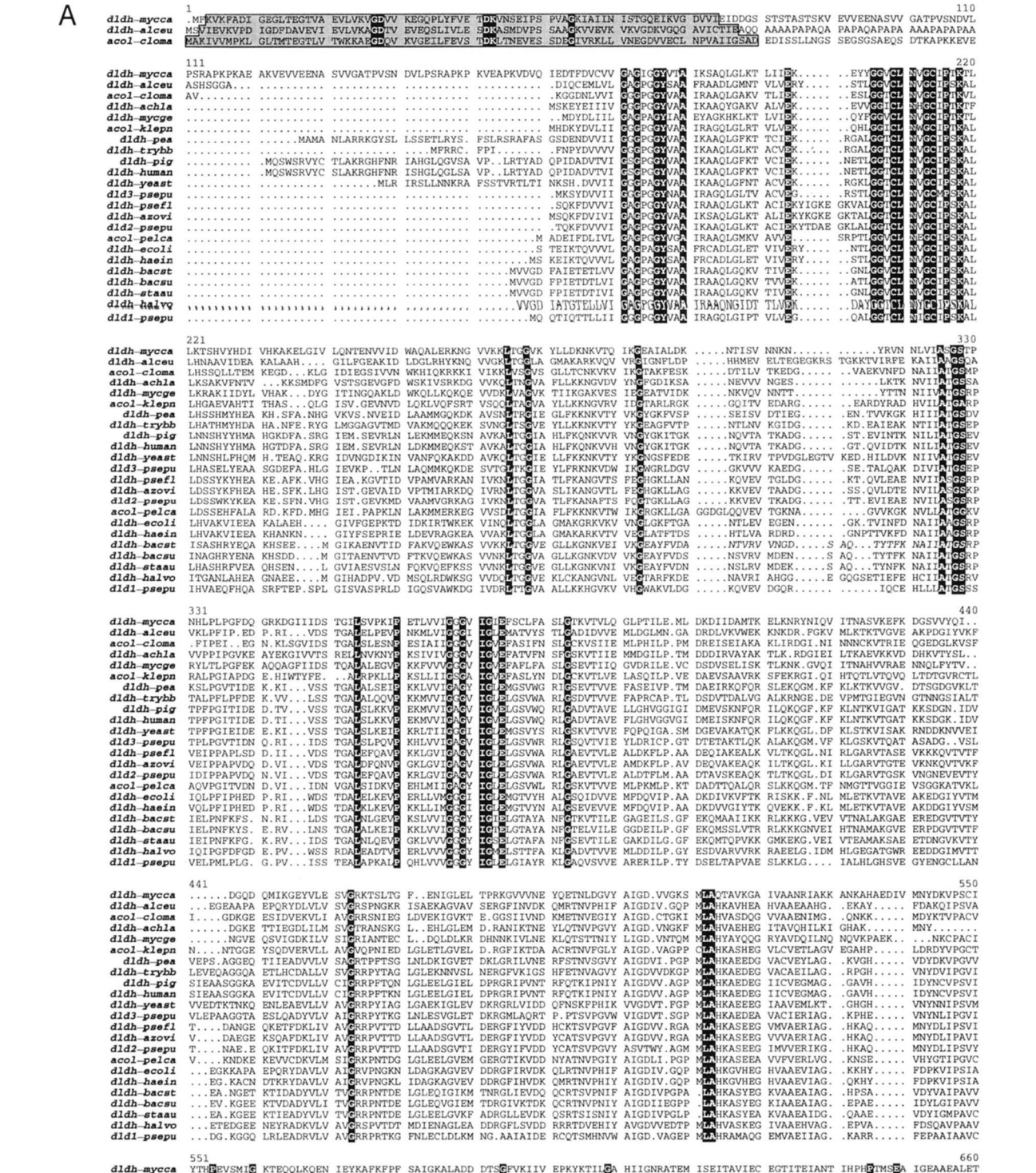


Fig. 11A. (*Figure continues on following page.*)

B

```
                661              678
dldh-mycca      GKAIHF.... ........
dldh-alceu      GTCTDVPPPR KR......
acol-cloma      NQAIHMMNK. ........
dldh-achla      .......... ........
dldh-mycge      DYPS..... ........
acol-klepn      DQPLHQ.... ........
  dldh-pea      DKPIHI.... ........
dldh-trybb      AKTINF.... ........
  dldh-pig      GKAINF.... ........
dldh-human      GKSINF.... ........
dldh-yeast      DKAIHC.... ........
dld3-psepu      GMAMQI.... ........
dldh-psefl      GHAIHIANRK KR......
dldh-azovi      GHAIHVANRK K.......
dld2-psepu      GGAIHVANRK KR......
acol-pelca      GAAVHC.... ........
dldh-ecoli      GSITDLPNPK AKKK....
dldh-haein      GSITDLPNAK AKEKIISI
dldh-bacst      GTPIHIITK. ........
dldh-bacsu      GSPIHIVK.. ........
dldh-staau      GYPIHTM... ........
dldh-halvo      GQAIHTLNR. ........
dld1-psepu      GHALHI.... ........
```

Fig. 11. Alignment of sequences of members of the dihydrolipoamide dehydrogenase family. Numbering above the aligned sequences corresponds only to the residue in the alignment rather than to a residue number in any of the aligned proteins. Residues that are conserved in the listed Enzyme II proteins are shown in reverse shading. Boxed shaded region corresponds to a lipoyl domain. Abbreviations used and references to published sequences are: *M. capricolum* (dldh-mycca) (this work); *A. eutrophus* (dldh-alceu) (Hein & Steinbuchel, 1994); *C. magnum* (acol-cloma) (Kruger et al., 1994); *A. laidlawii* (dldh-achla) (Wallbrandt et al., 1992); *M. genitalium* (dldh-mycge) (MG271 from the TIGR database); *K. pneumoniae* (acol-klepn) (Genbank accession no. U30887); pea (dldh-pea) (Bourguignon et al., 1992); *Trypanosoma brucei* (dldh-trybb) (Else et al., 1993); pig (dldh-pig) (Otulakowski & Robinson, 1987); human (dldh-human) (Pons et al., 1988); yeast (dldh-yeast) (Browning et al., 1988); *P. putida* (dld2-psepu) (Palmer et al., 1991a), (dld1-psepu) (Burns et al., 1989), and (dld3-psepu) (Palmer et al., 1991b); *Pseudomonas fluorescens* (dldh-psefl) (Benen et al., 1989); *A. vinelandii* (dldh-azovi) (Westphal & de Kok, 1988); *P. carbinolicus* (acol-pelca) (Oppermann & Steinbuchel, 1994); *E. coli* (dldh-ecoli) (Guest, 1987); *H. influenzae* (dldh-haein) (HI1231 frrom the TIGR database); *B. stearothermophilus* (dldh-bacst) (Borges et al., 1990); *B. subtilis* (dldh-bacsu) (Hemila et al., 1990); *S. aureus* (dldh-staau) (Hemila, 1991); *Haloferax volcanii*, (dldh-halvo) (Vettakkorumakankav & Stevenson, 1992). dldh, dld1, and dld2 refer to the genes encoding dihydrolipoamide dehydrogenases from the pyruvate, 2-oxoglutarate, and branched chain α-ketoacid dehydrogenase complexes, respectively. dld3 refers to the third dehydrogenase isolated from *P. putida* and acol refers to the dihydrolipoamide dehydrogenase of the acetoin dehydrogenase complex.

Mycoplasmas are generally believed to be descendants of Gram-positive bacteria. All the *M. capricolum* proteins involved with the metabolism of pyruvate described here show phylogenetic relatedness to the homologous proteins from Gram-positive bacteria and *M. genitalium*.

The complete genome of *M. genitalium* has been reported recently (Fraser et al., 1995). Because it might be expected that these two species are closely related, the question was posed concerning the genomic locations of the genes encoding enzymes of pyruvate metabolism. Figure 18 shows a comparison of the location of the genes of interest in *M. capricolum* and *M. genitalium*. It is most surprising to see that, whereas all the genes involved with pyruvate metabolism are clustered in *M. capricolum*, this is not the case in *M. genitalium*. In both organisms, the *ptsI* genes are located approximately 500 kb from the replication origin. Further, the *gyrA,B* complex is separated from the *naox*, *odpA*, *odpB*, *odp2*, *dldH*, *lplA* complex by approximately 300 kb in both organisms. Clearly, there have been extensive rearrangements in the genomes of these organisms during evolution, resulting in resolution of *gyrA,B* from the replication origin in the case of *M. capricolum* and scattering of genes involved with pyruvate metabolism in *M. genitalium*.

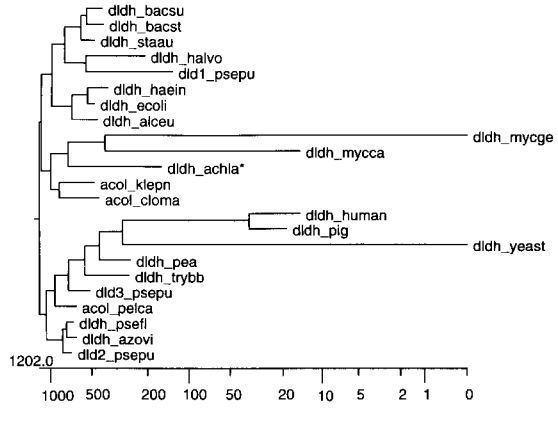


Fig. 12. Phylogenetic tree of sequenced proteins of the dihydrolipoamide dehydrogenase family. Relative evolutionary distances are shown on the numerical scale. Abbreviations are as in the legend to Figure 11. *, corresponds to a partial amino acid sequence (see Fig. 11).

In summary, the present work has demonstrated the unique arrangement of the genes encoding enzymes involved with pyruvate metabolism in *M. capricolum* and described some unique properties of the products of these genes.

Materials and methods

Growth of cells

M. capricolum (kid strain) were grown in modified Edwards medium at pH 8, as described previously (Mugharbil & Cirillo, 1978). After harvesting, cells were stored as frozen pellets for future use.

Nucleic acid preparations

DNA was prepared from frozen cells as previously described (Ausubel et al., 1990). RNA was prepared by suspension of cells (100 mg) in buffer (1 mL) containing 50 mM Tris-HCl, pH 6.8, 2 mM EDTA, and 1% SDS; the suspension was mixed with 5 mL of 4 M guanidium thiocyanate homogenization buffer

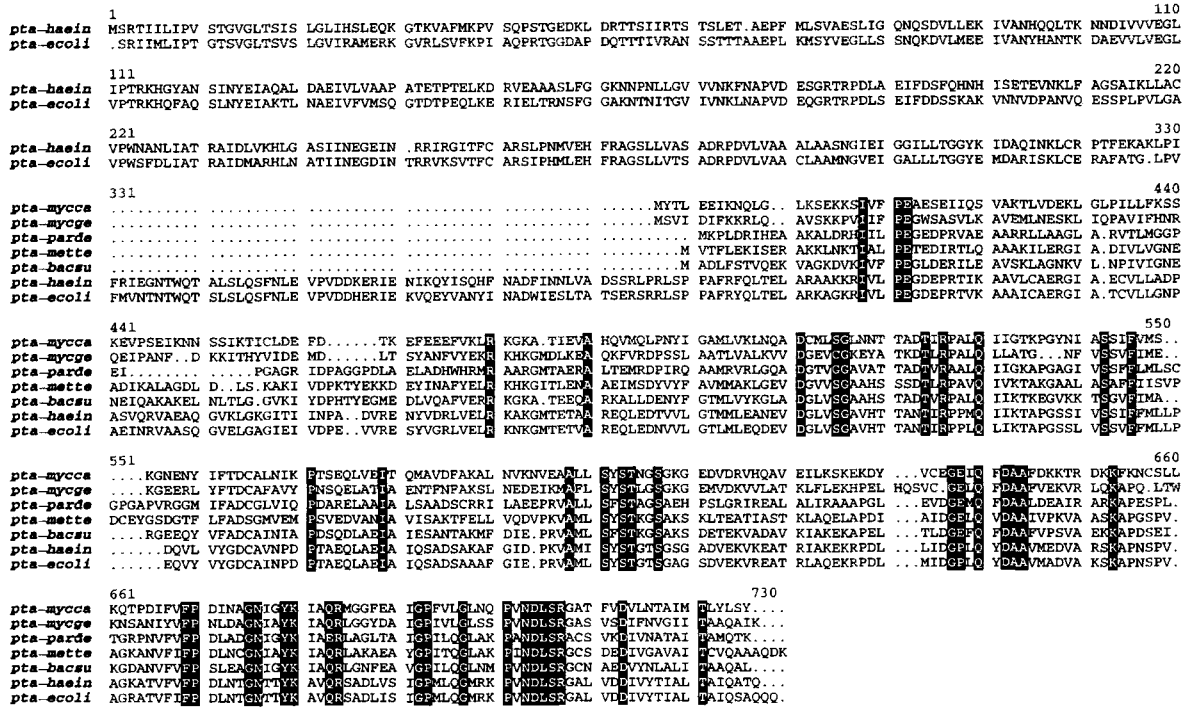


Fig. 13. Alignment of sequences of members of the phosphotransacetylase family. Numbering above the aligned sequences corresponds only to the residue in the alignment rather than to a residue number in any of the aligned proteins. Residues that are conserved in the listed phosphotransacetylase proteins are shown in reverse shading. Abbreviations used and references to published sequences are: *M. capricolum* (pta-mycca) (this work); *M. genitalium* (pta-mycge) (MG299 from the TIGR database); *Paracoccus denitrificans* (pta-parde) (Van Spanning et al., 1995); *M. thermophila* (pta-mette) (Latimer & Ferry, 1993); *B. subtilis* (pta-bacsu) (Glaser et al., 1993); *H. influenzae* (pta-haein) (H11203 from the TIGR database); *E. coli* (pta-ecoli) (Kakuda et al., 1994).

(Sambrook et al., 1989). The cell suspension was frozen in dry ice-ethanol, then warmed briefly in a 64 °C water bath (Salser et al., 1967). Total RNA was purified by ultracentrifugation through 5.7 M CsCl/10 mM EDTA (Sambrook et al., 1989).

Cloning and screening

Genomic DNA fragments produced by digestion of *M. capricolum* DNA with *Hind* III (9,542-bp fragment), *Xba* I (6,654-bp fragment), or *Spe* I (4,384-bp fragment) were cloned into the vector pBluescript II KS$^+$ (pKSII$^+$). Recombinant plasmids were used to transform Epicurean coli XLI-Blue Competent cells (Stratagene). Colonies were lifted onto nylon membranes (NEN Research Products, NEF-978). [^{32}P]5′-end-labeled oligonucleotide probes (1×10^6 cpm/mL of hybridization solution) were used for selecting positive clones. Oligonucleotide probes, synthesized as trityl-off derivatives on an Applied Biosystems 380B DNA synthesizer, were labeled with [γ^{32}P]ATP by the DNA 5′-end-labeling method (Sambrook et al., 1989). Prehybridization was performed at 40 °C for 4 h in 6× SSPE/0.1% SDS/10× Denhardt's solution containing 20 mg/mL tRNA and 50 mg/mL of denatured heterologous DNA. Hybridization was performed at 42 °C for 16 h in 6× SSPE/10% SDS solution containing 1.4×10^6 cpm of [^{32}P]-labeled oligonucleotide/mL. The membrane was finally washed in 0.5× SSPE/1% SDS solution at 40 °C for 30 min. Positive clones were detected by autoradiography.

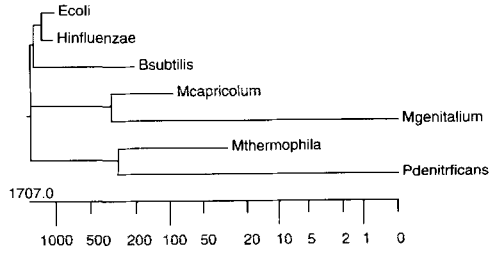


Fig. 14. Phylogenetic tree of sequenced proteins of the phosphotransacetylase family. Relative evolutionary distances are shown on the numerical scale. Abbreviations are as in the legend to Figure 13.

DNA sequencing

DNA sequencing on both strands of the DNA was performed by the dideoxy chain termination method of Sanger et al. (1977), with [α^{35}S]dATP, using Sequenase 2.0 (United States Biochemicals) DNA sequencing kits. M13 forward or reverse primers or

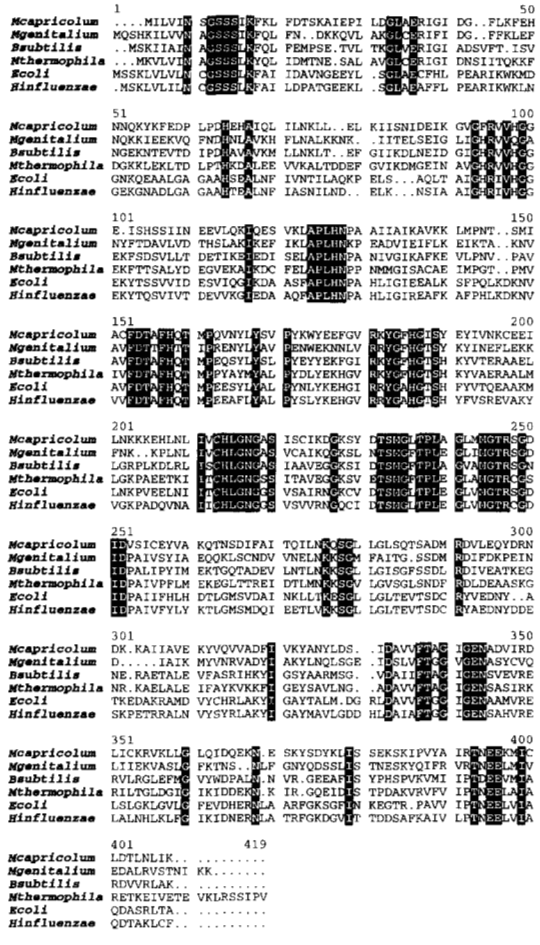


Fig. 15. Alignment of sequences of members of the acetate kinase family. Numbering above the aligned sequences corresponds only to the residue in the alignment rather than to a residue number in any of the aligned proteins. Residues that are conserved in the listed phosphotransacetylase proteins are shown in reverse shading. Abbreviations used and references to published sequences are: *M. capricolum* (Mcapricolum) (this work); *M. genitalium* (Mgenitalium) (MG357 from the TIGR database); *B. subtilis* (Bsubtilis) (Grundy et al., 1993); *M. thermophila* (Mthermophila) (Latimer & Ferry, 1993); *E. coli* (Matsuyama et al., 1989); *H. influenzae* (Hinfluenzae) (H11204 from the TIGR database).

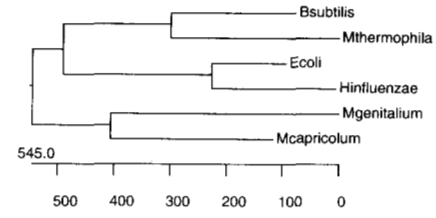


Fig. 16. Phylogenetic tree of sequenced proteins of the acetate kinase family. Relative evolutionary distances are shown on the numerical scale. Abbreviations are as in the legend to Figure 15.

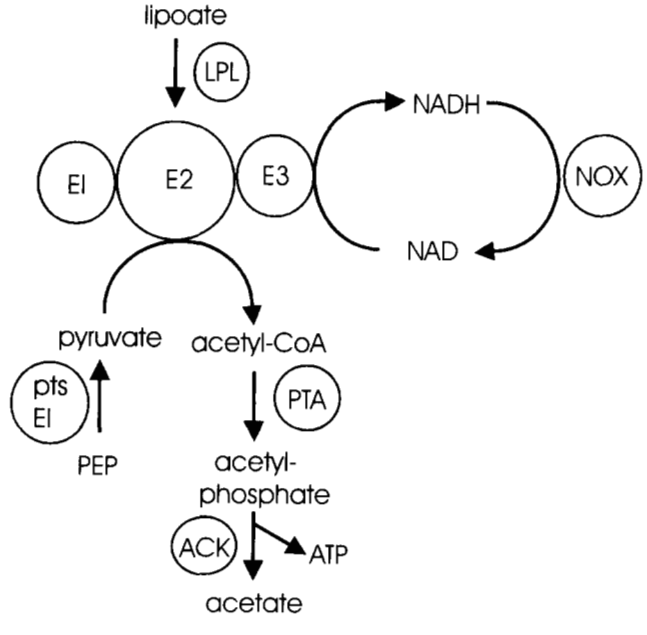


Fig. 17. Metabolic scheme for conversion of phosphoenolpyruvate to acetate in *M. capricolum*.

specific primers complementary to previously determined sequences were used.

Computer analyses

Analyses of DNA and protein sequence were performed using the GCG programs, version 7.2 (Devereux et al., 1984). Isoelectric points were calculated using the PEPTIDESORT program. A search for transcription termination sites used the TERMINATOR program. Stem-loop structures were analyzed using FOLDRNA. Translation frames were detected using the MAP program. Phylogenetic trees were constructed using the MEGALIGN module of the LaserGene program (DNAStar, Madison, Wisconsin) by the method of Hein (1990).

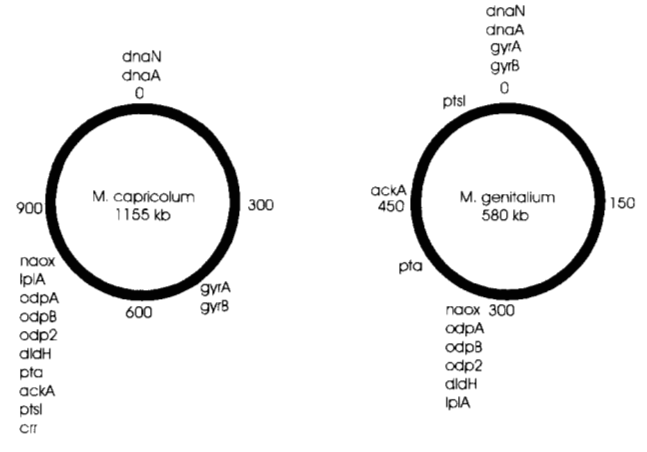


Fig. 18. Comparison of positions on the genetic map of various genes in *M. capricolum* and *M. genitalium*.

References

Allen AG, Perham RN. 1991. Two lipoyl domains in the dihydrolipoamide acetyltransferase chain of the pyruvate dehydrogenase multienzyme complex of *Streptococcus faecalis*. *FEBS Lett 287*:206–210.

Ausubel FA, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K, eds. 1990. *Current protocols in molecular biology*. New York: Greene Publishing and Wiley-Interscience. pp 2.4.1–2.4.5.

Behal RH, Browning KS, Reed LJ. 1989. Nucleotide and deduced amino acid sequence of the alpha subunit of yeast pyruvate dehydrogenase. *Biochem Biophys Res Commun 164*:941–946.

Benen JAE, van Berkel WJ, van Dongen WMAM, Mueller F, de Kok A. 1989. Molecular cloning and sequence determination of the lpd gene encoding lipoamide dehydrogenase from *Pseudomonas fluorescens*. *J Gen Microbiol 135*:1787–1797.

Borges A, Hawkins CF, Packman LC, Perham RN. 1990. Cloning and sequence analysis of the genes encoding the dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase components of the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. *Eur J Biochem 194*:95–102.

Bourguignon J, Macherel D, Neuberger M, Douce R. 1992. Isolation, characterization, and sequence analysis of a cDNA clone encoding L-protein, the dihydrolipoamide dehydrogenase component of the glycine cleavage system from pea-leaf mitochondria. *Eur J Biochem 204*:865–873.

Browning KS, Uhlinger DJ, Reed LJ. 1988. Nucleotide sequence for yeast dihydrolipoamide dehydrogenase. *Proc Natl Acad Sci USA 85*:1831–1834.

Burns G, Brown T, Hatter K, Sokatch JR. 1988. Comparison of the amino acid sequences of the transacylase components of branched chain oxoacid dehydrogenase of *Pseudomonas putida*, and the pyruvate and 2-oxoglutarate dehydrogenases of *Escherichia coli*. *Eur J Biochem 176*:165–169.

Burns G, Brown T, Hatter K, Sokatch JR. 1989. Sequence analysis of the lpdV gene for lipoamide dehydrogenase of branched-chain-oxoacid dehydrogenase of *Pseudomonas putida*. *Eur J Biochem 179*:61–69.

Carlsson P, Hederstedt L. 1989. Genetic characterization of *Bacillus subtilis* odhA and odhB, encoding 2-oxoglutarate dehydrogenase and dihydrolipoamide transsuccinylase, respectively. *J Bacteriol 171*:3667–3672.

Cirillo VP. 1979. Transport systems. In: Barile MF, Razin S, eds. *The mycoplasmas*. New York: Academic Press. pp 323–350.

Cullingford TE, Clark JB, Phillips IR. 1993. Characterization of cDNAs encoding the rat testis-specific E1 alpha subunit of the pyruvate dehydrogenase complex: Comparison of expression of the corresponding mRNA with that of the somatic E1 alpha subunit. *Biochim Biophys Acta 1216*: 149–153.

Cullingford TE, Clark JB, Phillips IR. 1994. The pyruvate dehydrogenase complex: Cloning of the rat somatic E1 alpha subunit and its coordinate expression with the mRNAs for the E1 beta, E2, and E3 catalytic subunits in developing rat brain. *J Neurochem 62*:1682–1690.

Dahl HH, Hunt SM, Hutchison WM, Brown GK. 1987. The human pyruvate dehydrogenase complex. Isolation of cDNA clones for the E1 alpha subunit, sequence analysis, and characterization of the mRNA. *J Biol Chem 262*:7398–7403.

de la Sierra IL, Prange T, Fourme R, Padron G, Fuentes P, Musacchio A, Madrazo J. 1994. Crystallization and preliminary X-ray investigation of a recombinant outer membrane protein from *Neisseria meningitidis*. *J Mol Biol 235*:1154–1155.

Deng WL, Chang HY, Peng HL. 1994. Acetoin catabolic system of *Klebsiella pneumoniae* CG43: Sequence, expression, and organization of the aco operon. *J Bacteriol 176*:3527–3535.

Devereux J, Haeberli P, Smithies O. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res 12*:387–395.

Else AJ, Hough DW, Danson MJ. 1993. Cloning, sequencing, and expression of *Trypanosoma bruceii* dihydrolipoamide dehydrogenase. *Eur J Biochem 212*:423–429.

Fitzgerald J, Hutchison WM, Dahl HH. 1992. Isolation and characterization of the mouse pyruvate dehydrogenase E1 alpha genes. *Biochim Biophys Acta 1131*:83–90.

Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, Bult CJ, Kerlavage AR, Sutton G, Kelley JM, et al. 1995. The minimal gene complement of *Mycoplasma genitalium*. *Science 270*:397–403.

Gershwin ME, MacKay IR, Sturgess A, Coppel RL. 1987. Identification and specificity of a cDNA encoding the 70 kd mitochondrial antigen recognized in primary biliary cirrhosis. *J Immunol 138*:3525–3531.

Glaser P, Kunst F, Arnaud M, Coudart MP, Gonzales W, Hullo MF, Ionescu M, Lubochinsky B, Marcelino L, Moszer I, Presecan E, Santana M, Schneider E, Schweizer J, Vertes A, Rapoport G, Danchin A. 1993. *Bacillus subtilis* genome project: Cloning and sequencing of the 97 kb region from 325 degrees to 333 degrees. *Mol Microbiol 10*:371–384.

Grundy FJ, Waters DA, Allen SH, Henkin TM. 1993. Regulation of the *Bacillus subtilis* acetate kinase gene by CcpA. *J Bacteriol 175*:7348–7355.

Guan Y, Rawsthorne S, Scofield G, Shaw P, Doonan J. 1995. Cloning and characterization of a dihydrolipoamide acetyltransferase (E2) subunit of the pyruvate dehydrogenase complex from *Arabidopsis thaliana*. *J Biol Chem 270*:5412–5417.

Guest JR. 1987. Functional implications of structural homologies between chloramphenicol acetyltransferase and dihydrolipoamide acetyltransferase. *FEMS Microbiol Lett 44*:417–422.

Hanemaaijer R, Janssen A, de Kok A, Veeger C. 1988. The dihydrolipoyltransacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Molecular cloning and sequence analysis. *Eur J Biochem 174*:593–599.

Hein J. 1990. Unified approach to alignment and phylogenies. *Methods Enzymol 183*:626–644.

Hein S, Steinbuchel A. 1994. Biochemical and molecular characterization of the *Alcaligenes eutrophus* pyruvate dehydrogenase complex and identification of a new type of dihydrolipoamide dehydrogenase. *J Bacteriol 176*:4394–4408.

Hemila H, Palva A, Paulin L, Arvidson S, Palva I. 1990. Secretory S complex of *Bacillus subtilis*: Sequence analysis and identity to pyruvate dehydrogenase. *J Bacteriol 172*:5052–5063.

Hemila H. 1991. Lipoamide dehydrogenase of *Staphylococcus aureus*: Nucleotide sequence and sequence analysis. *Biochim Biophys Acta 1129*:119–123.

Ho L, Patel MS. 1990. Cloning and cDNA sequence of the beta-subunit component of human pyruvate dehydrogenase complex. *Gene 86*:297–302.

Ho L, Wexler ID, Liu TC, Thekkumkara TJ, Patel MS. 1989. Characterization of cDNAs encoding human pyruvate dehydrogenase alpha-subunit. *Proc Natl Acad Sci USA 86*:5330–5334.

Hu CWC, Lau KS, Griffin TA, Chuang JL, Fisher CW, Cox RP, Chuang DT. 1988. Isolation and sequencing of a cDNA encoding the decarboxylase (E1)-alpha precursor of bovine branched-chain alpha-keto acid dehydrogenase complex: Expression of E1-alpha mRNA and subunit in maple-syrup-urine-disease and 3T3-L1 cells. *J Biol Chem 263*:9007–9014.

Johnson RK, Komuniecki R, Sun Y, Wheelock MJ. 1992. Characterization of cDNA clones for the alpha subunit of pyruvate dehydrogenase from *Ascaris suum*. *Mol Biochem Parasitol 51*:37–48.

Kakuda H, Hosono K, Shiroishi K, Ichihara S. 1994. Identification and characterization of the *ackA* (acetate kinase A)-*pta* (phosphotransacetylase) operon and complementation analysis of acetate utilization by an *ackA-pta* deletion mutant of *Escherichia coli*. *J Biochem 116*:916–922.

Kreader CA, Langer CS, Heckman JE. 1989. A mitochondrial protein from *Neurospora crassa* detected both on ribosomes and in membrane fractions. Analysis of the gene, the message and the protein. *J Biol Chem 264*:317–327.

Kruger N, Oppermann FB, Lorenzl H, Steinbuchel A. 1994. Biochemical and molecular characterization of the *Clostridium magnum* acetoin dehydrogenase enzyme system. *J Bacteriol 176*:3614–3630.

Latimer MT, Ferry JG. 1993. Cloning, sequence analysis, and hyperexpression of the genes encoding phosphotransacetylase and acetate kinase from *Methanosarcina thermophila*. *J Bacteriol 175*:6822–6829.

Lau KS, Chuang JL, Herring WJ, Danner DJ, Cox RP, Chuang DT. 1992. The complete cDNA sequence for dihydrolipoyl transacylase (E2) of human branched-chain alpha-keto acid dehydrogenase complex. *Biochim Biophys Acta 1132*:319–321.

Lau KS, Griffin TA, Hu CWC, Chuang DT. 1988. Conservation of primary structure in the lipoyl-bearing and dihydrolipoyl dehydrogenase binding domains of mammalian branched-chain alpha-keto acid dehydrogenase complex: Molecular cloning of human and bovine transacylase (E2) cDNAs. *Biochemistry 27*:1972–1981.

Luethy MH, Miernyk JA, Randall DD. 1994. The nucleotide and deduced amino acid sequences of a cDNA encoding the E1 beta-subunit of the *Arabidopsis thaliana* mitochondrial pyruvate dehydrogenase complex. *Biochim Biophys Acta 1187*:95–98.

Luethy MH, Miernyk JA, Randall DD. 1995. The mitochondrial pyruvate dehydrogenase complex: Nucleotide and deduced amino-acid sequences of a cDNA encoding the *Arabidopsis thaliana* E1 alpha-subunit. *Gene 164*:251–254.

Maniloff J. 1983. Evolution of wall-less prokaryotes. *Annu Rev Microbiol 37*:477–499.

Matsuyama A, Yamamoto H, Nakano E. 1989. Cloning, expression, and nucleotide sequence of the *Escherichia coli* K-12 *ackA* gene. *J Bacteriol 171*:577–580.

Mattevi A, deKok A, Perham RN. 1992. The pyruvate dehydrogenase multienzyme complex. *Curr Opin Struct Biol 2*:877–887.

McKean MC, Winkeler KA, Danner DJ. 1992. Nucleotide sequence of the 5′ end including the initiation codon of cDNA for the E1 alpha subunit

of the human branched chain alpha-ketoacid dehydrogenase complex. *Biochim Biophys Acta 1171*:109–112.
Miles RJ. 1992. Catabolism in mollicutes. *J Gen Microbiol 138*:1773–1783.
Miran SG, Lawson JE, Reed LJ. 1993. Characterization of PDH beta 1, the structural gene for the pyruvate dehydrogenase beta subunit from *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA 90*:1252–1256.
Miyata M, Wang L, Fukumura T. 1991. Physical mapping of the *Mycoplasma capricolum* genome. *FEMS Microbiol Lett 79*:329.
Morris TW, Reed KE, Cronan JE Jr. 1994. Identification of the gene encoding lipoate-protein ligase A of *Escherichia coli*. Molecular cloning and characterization of the *lplA* gene and gene product. *J Biol Chem 269*:16091–16100.
Mugharbil U, Cirillo VP. 1978. *Mycoplasma* phosphoenolpyruvate-dependent sugar phosphotransferase system: Glucose-negative mutant and regulation of intracellular cyclic AMP. *J Bacteriol 133*:203–209.
Nakano K, Matsuda S, Sakamoto T, Takase C, Nakagawa S, Ohta S, Ariyama T, Inazawa J, Abe T, Miyata T. 1993. Human dihydrolipoamide succinyltransferase: cDNA cloning and localization on chromosome 14q24.2–q24.3. *Biochim Biophys Acta 1216*:360–368.
Nakano K, Matuda S, Yamanaka T, Tubouchi H, Nakagawa S, Titani K, Ohta S, Miyata T. 1991. Purification and molecular cloning of succinyltransferase of rat alpha-ketoglutarate dehydrogenase complex. Absence of a sequence motif of the putative E3 and/or E1 binding site. *J Biol Chem 266*:19013–19017.
Niu XD, Browning KS, Behal RH, Reed LJ. 1988. Cloning and nucleotide sequence of the gene for dihydrolipoamide acetyltransferase from *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA 85*:7546–7550.
Nobukuni Y, Mitsubuchi H, Endo F, Akaboshi I, Asaka J, Matsuda I. 1990a. Maple syrup urine disease. Complete primary structure of the E1 beta subunit of human branched chain alpha-ketoacid dehydrogenase complex deduced from the nucleotide sequence and a gene analysis of patients with this disease. *J Clinical Investigation 86*:242–247.
Nobukuni Y, Mitsubuchi H, Endo F, Asaka Y, Oyama R, Titani K, Matsuda I. 1990b. Isolation and characterization of a complementary cDNA clone coding for the E1-beta subunit of the bovine branched-chain alpha-ketoacid dehydrogenase complex: Complete amino acid sequence of the precursor protein and its proteolytic processing. *Biochemistry 29*:1154–1160.
Oppermann FB, Steinbuchel A. 1994. Identification and molecular characterization of the *aco* genes encoding the *Pelobacter carbinolicus* acetoin dehydrogenase enzyme system. *J Bacteriol 176*:469–485.
Otulakowski G, Robinson BH. 1987. Isolation and sequence determination of cDNA clones for porcine and human lipoamide dehydrogenase: Homology to other disulfide oxidoreductases. *J Biol Chem 262*:17313–17318.
Palmer JA, Hatter K, Sokatch JR. 1991a. Cloning and sequence analysis of the LPD-glc structural gene of *Pseudomonas putida*. *J Bacteriol 173*:3109–3116.
Palmer JA, Madhusudhan KT, Hatter K, Sokatch JR. 1991b. Cloning, sequence and transcriptional analysis of the structural gene for LPD-3, the third lipoamide dehydrogenase of *Pseudomonas putida*. *Eur J Biochem 202*:231–240.
Pons G, Raefsky-Estrin C, Carothers DJ, Pepin RA, Javed AA, Jesse BW, Ganapathi MK, Samols D, Patel MS. 1988. Cloning and cDNA sequence of the dihydrolipoamide dehydrogenase component of human alpha-ketoacid dehydrogenase complexes. *Proc Natl Acad Sci USA 85*:1422–1426.
Postma PW, Lengeler JW, Jacobson GR. 1993. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol Rev 57*:543–594.
Priefert H, Hein S, Kreuger N, Zeh K, Schmidt B, Steinbuchel A. 1991. Identification and molecular characterization of the *Alcaligenes eutrophus* H16 *aco* operon genes involved in acetoin catabolism. *J Bacteriol 173*:4056–4071.
Razin S. 1985. Molecular biology and genetics of *Mycoplasmas* (mollicutes). *Microbiol Rev 49*:419–455.
Razin S. 1992. Peculiar properties of Mycoplasmas: The smallest self-replicating prokaryotes. *FEMS Microbiol Lett 100*:423–432.
Repetto B, Tzagoloff A. 1990. Structure and regulation of KGD2, the structural gene for yeast dihydrolipoyl transsuccinylase. *MCB 10*:4221–4232.
Ross RP, Claiborne A. 1992. Molecular cloning and analysis of the gene encoding NADH oxidase from *Streptococcus faecalis* 10C1. Comparison with NADH peroxidase and the flavoprotein disulfide reductases. *J Mol Biol 227*:658–671.
Salser W, Gesteland RF, Bolle A. 1967. In vitro synthesis of bacteriophage lysozyme. *Nature 215*:588–591.
Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA 74*:5463–5467.
Sermon K, DeMeirleir L, Elpers I, Lissens W, Liebaers I. 1990. Characterization of cDNA for porcine PDHE1-alpha and comparison with the human cDNA. *Nucleic Acids Res 18*:4925
Spencer ME, Darlison MG, Stephens PE, Duckenfield IK, Guest JR. 1984. Nucleotide sequence of the sucB gene encoding the dihydrolipoamide succinyltransferase of *Escherichia coli* K12 and homology with the corresponding acetyltransferase. *Eur J Biochem 141*:361–374.
Thekkumkara TJ, Ho L, Wexler ID, Pons G, Liu TC, Patel MS. 1988. Nucleotide sequence of a cDNA for the dihydrolipoamide acetyltransferase component of human pyruvate dehydrogenase complex. *FEBS Lett 240*:45–48.
Van Spanning RJM, De Boer APN, Slotboom DJ, Reijnders WNM, Stouthamer AH. 1995. Isolation and characterization of a novel insertion sequence element, IS1248, in *Paracoccus denitrificans*. *Plasmid 34*:11–21.
Vettakkorumakankav NN, Stevenson KJ. 1992. Dihydrolipoamide dehydrogenase from *Haloferax volcanii*: Gene cloning, complete primary structure and comparison to other dihydrolipoamide dehydrogenases. *Biochemical Cellular Biol 70*:656–663.
Wallbrandt PHM, Tegman V, Jonsson BH, Wieslander A. 1992. Identification and analysis of the genes coding for the putative pyruvate dehydrogenase enzyme complex in *Acholeplasma laidlawii*. *J Bacteriol 174*:1388–1396.
Wang GF, Kuriki T, Roy KL, Kaneda T. 1993. The primary structure of branched-chain alpha-oxo acid dehydrogenase from *Bacillus subtilis* and its similarity to other alpha-oxo acid dehydrogenases. *Eur J Biochem 213*:1091–1099.
Westphal AH, de Kok A. 1988. Lipoamide dehydrogenase from *Azotobacter vinelandii*: Molecular cloning, organization and sequence analysis of the gene. *Eur J Biochem 172*:299–305.
Westphal AH, de Kok A. 1990. The 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. 2. Molecular cloning and sequence analysis of the gene encoding the succinyltransferase component. *Eur J Biochem 187*:235–239.
Wheelock MJ, Komuniecki R, Duran E, Johnson KR. 1991. Characterization of cDNA clones for the beta subunit of pyruvate dehydrogenase from *Ascaris suum*. *Mol Biochem Parasitol 45*:9–18.
Zhang B, Kuntz M, Goodwin GW, Harris RA, Crabb DW. 1987. Molecular cloning of a cDNA for the E1-alpha subunit of rat liver branched chain alpha-ketoacid dehydrogenase. *J Biol Chem 262*:15220–15224.
Zhao Y, Kuntz MJ, Harris RA, Crabb DW. 1992. Molecular cloning of the E1 beta subunit of the rat branched chain alpha-ketoacid dehydrogenase. *Biochim Biophys Acta 1132*:207–210.
Zhu PP, Reizer J, Peterkofsky A. 1994. Unique dicistronic operon (*ptsI-crr*) in *Mycoplasma capricolum* encoding enzyme I and the glucose-specific enzyme IIA of the phosphoenolpyruvate:sugar phosphotransferase system: Cloning, sequencing, promoter analysis, and protein characterization. *Protein Sci 3*:2115–2128.
Zhu PP, Reizer J, Reizer A, Peterkofsky A. 1993. Unique monocistronic operon (*ptsH*) in *Mycoplasma capricolum* encoding the phosphocarrier protein, HPr, of the phosphoenolpyruvate:sugar phosphotransferase system. *J Biol Chem 268*:26531–26540.