JOURNAL OF BACTERIOLOGY, Oct. 1991, p. 6118–6123
0021-9193/91/196118-06$02.00/0

Vol. 173, No. 19

Molecular Cloning and DNA Sequencing of the *Escherichia coli* K-12 *ald* Gene Encoding Aldehyde Dehydrogenase

ELENA HIDALGO,[1] YU-MEI CHEN,[2] E. C. C. LIN,[2] AND JUAN AGUILAR[1]*

Department of Biochemistry, School of Pharmacy, University of Barcelona, Diagonal 643, 08028 Barcelona, Spain,[1] and Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115[2]

Received 16 April 1991/Accepted 20 June 1991

The gene *ald*, encoding aldehyde dehydrogenase, has been cloned from a genomic library of *Escherichia coli* K-12 constructed with plasmid pBR322 by complementing an aldehyde dehydrogenase-deficient mutant. The *ald* region was sequenced, and a single open reading frame of 479 codons specifying the subunit of the aldehyde dehydrogenase enzyme complex was identified. Determination of the N-terminal amino acid sequence of the enzyme protein unambiguously established the identity and the start codon of the *ald* gene. Analysis of the 5′- and 3′-flanking sequences indicated that the *ald* gene is an operon. The deduced amino acid sequence of the *ald* gene displayed homology with sequences of several aldehyde dehydrogenases of eukaryotic origin but not with microbial glyceraldehyde-3-phosphate dehydrogenase.

In *Escherichia coli*, the metabolism of L-fucose and L-rhamnose yields L-lactaldehyde as an intermediate metabolite (14). An NAD-dependent enzyme that oxidizes this intermediate to L-lactate was initially named lactaldehyde dehydrogenase (26). Mutants selected for growth on L-1,2-propanediol express an NAD-dependent oxidoreductase constitutively. This enzyme oxidizes L-1,2-propanediol to L-lactaldehyde, which is then converted to L-lactate by the same dehydrogenase involved in L-fucose and L-rhamnose metabolism (14, 27). Such propanediol-positive mutants, however, lost the ability to grow on L-fucose (11, 14). A propanediol-negative mutant that lost the ability to grow on propanediol was then isolated. This secondary mutant (ECL40) lacked lactaldehyde dehydrogenase activity (26) and the immunologically cross-reacting material (5). The enzyme was later shown to be able to oxidize other α-hydroxyaldehydes such as glycoaldehyde, which is generated in the metabolism of L-arabinose (20). More recently, it was found that this enzyme was induced not only by growth on fucose, rhamnose, or arabinose, but also by growth in the presence of amino acids such as glutamate (5). It is uncertain what role the enzyme plays in the metabolism of this amino acid.

Since the dehydrogenase was found to act on a number of aldehydes, it was renamed simply aldehyde dehydrogenase. The mutation abolishing the enzyme activity was designated as *ald*, which mapped at 31.2 min (13). The native enzyme is a homotetramer with a subunit of 55,000 Da (5).

MATERIALS AND METHODS

Bacterial strains. The strains used are listed in Table 1. Strain JA111 was constructed by transducing the *recA1* mutation from strain JA120 (3) to strain ECL40 by selecting for the closely linked marker *srl*::Tn*10*. Tetracycline-resistant transductants were scored for UV sensitivity.

Growth media and preparation of cell extracts. Cells were grown aerobically as described previously (7) on Luria broth (LB) or minimal medium. For growth on minimal medium, carbon sources were added to a basal inorganic medium (6) in the following concentrations: DL-1,2-propanediol, 40 mM; casein acid hydrolysate, 0.5% (wt/vol). Ampicillin and tetracycline, when used, were routinely added at final concentrations of 100 and 15 μg/ml, respectively. For screening of the gene library, propanediol-ampicillin minimal agar plates were used that contained 40 mM DL-1,2-propanediol and 20 μg of ampicillin per ml. MacConkey-propanediol agar contained 1% DL-1,2-propanediol.

For the enzyme assay, the cells were harvested at the end of the exponential phase and the cell extract was prepared as described previously (6) with 10 mM Tris-HCl buffer (pH 7.0). When the extracts were used for enzyme purification, the buffer consisted of 10 mM phosphate buffer (pH 7.3) containing 10 mM β-mercaptoethanol and 1 mM EDTA (10).

Enzyme assay. Aldehyde dehydrogenase activity was assayed spectrophotometrically (at 340 nm) by monitoring the increased rate of NADH formation at 25°C. Since this enzyme was also reported to be responsible for the conversion of glycoaldehyde to glycolate (10), glycoaldehyde was used instead of lactaldehyde because of the commercial unavailability of lactaldehyde. The assay mixture (1 ml) consisted of 1 mM lactaldehyde or glycoaldehyde, 100 mM sodium glycine buffer (pH 9.5), and 2.5 mM NAD.

Propanediol oxidoreductase activity was assayed at 25°C in the direction of L-lactaldehyde reduction by monitoring the decrease in the concentration of NADH. Since the oxidoreductase reduces not only lactaldehyde but also glycoaldehyde to ethylene glycol (8), glycoaldehyde was routinely used instead of lactaldehyde. The assay mixture (1 ml) consisted of 0.5 mM glycoaldehyde, 200 mM sodium phosphate buffer (pH 7.0), and 0.125 mM NADH.

The protein concentration in cell extracts was determined by the method of Lowry et al. (21) with bovine serum albumin as the standard.

DNA manipulation. Plasmid DNA was routinely prepared by the boiling method (22). For large-scale preparation, a crude DNA sample was subjected to purification by cesium chloride-ethidium bromide density gradient centrifugation or on a column (Qiagen GmbH, Düsseldorf, Federal Republic of Germany). Single-stranded DNA was isolated by using a method outlined by Stratagene Corp., La Jolla, Calif. DNA manipulations were performed essentially as described by Maniatis et al. (22). The DNA sequence was determined by using the dideoxy-chain termination procedure of Sanger et

* Corresponding author.

TABLE 1. *E. coli* strains used in this work

Strain	Genotype	Source or reference
ECL3	HfrC *fucO*(Con) *fucA*(Con) *fucPIK*(Non) *phoA8 relA1 tonA22* T2^r(lambda)	27
ECL40	HfrC *ald-1 fucO*(Con) *fucA*(Con) *fucPIK*(Non) *phoA8 relA1 tonA22* T2^r(lambda)	26
JA120	F$^-$ *thi-1 thr-1 leuB6 lacY1 tonA21 supE44* λ^- *recA1 srlA*::Tn*10*	3
JA111	HfrC *ald-1 fucO*(Con) *fucA*(Con) *fucPIK*(Non) *phoA8 relA1 tonA22* T2^r(lambda) *recA1 srlA*::Tn*10*	This study

al. (24). Double-stranded plasmid DNA was used as a template. Plasmid purified by CsCl gradient centrifugation was used for the construction of ordered deletions by using the Erase-a-Base system (Promega Biotec, Madison, Wis.). To resolve the numerous sequencing gel compressions, we used 7-deaza-dGTP and 7-deaza-dITP instead of dGTP. In some cases, treatment of the sequencing reaction with methoxyamine and bisulfite, which eliminate the secondary-structure effects in gels by modifying the cytosine residues, was required (1).

Primer extension analysis. Primer extension analysis was performed by using the procedure of Hu and Davidson (18). The procedure involved hybridization of mRNA to a single-stranded DNA template and annealing of a radiolabeled DNA primer to the template at a site upstream from the 5′ end of the mRNA. Extension of the primer by T4 DNA polymerase should stop at the 5′ end of the hybridized mRNA; therefore, the 3′ end of the growing DNA chain marks its position. Total RNA from strain ECL1 was obtained as described previously (6). T4 DNA polymerase was purchased from Promega.

Purification and amino-terminal determination of aldehyde dehydrogenase. Aldehyde dehydrogenase was purified from cells of strain ECL3 grown on casein hydrolysate (9). The fraction of higher activity (1 ml) obtained from the agarose-NAD column was exhaustively dialyzed against 0.5 liter of 0.1 M ammonium hydrogen carbonate. The process involved 10 buffer changes with a 15-min dialysis each time. The dialyzed enzyme solution was concentrated in a vacuum chamber, and the N-terminal sequence of the subunit was determined by automated Edman degradation with an Applied Biosystems 470A gas-phase sequencer.

RESULTS

Cloning of the *ald* gene. As host strain, we used strain JA111 (aldehyde dehydrogenase deficient, propanediol oxidoreductase constitutive, and fucose negative), which failed to grow on propanediol. A previously constructed *E. coli* genomic library (12) was used for transformation. The transformants were plated on propanediol-ampicillin minimal agar. Nineteen propanediol-positive clones were purified. All were fucose negative, which was a trait of the host strain. The plasmid DNA of six clones displaying more rapid growth on propanediol was prepared and reintroduced into strain JA111. All six of these plasmids were shown to complement the *ald* mutation. Three of the transformants, carrying plasmids pALD1, pALD2, and pALD3, were subjected to determination of aldehyde dehydrogenase and propanediol oxidoreductase activities; all three showed restored aldehyde dehydrogenase activity and constitutive propanediol oxidoreductase activity (Table 2). The aldehyde dehydrogenase and the propanediol oxidoreductase activities were also determined in strain JA111 transformed with the plasmid pBR322 as a negative control. The transformant remained deficient in aldehyde dehydrogenase and constitutive in propanediol oxidoreductase.

TABLE 2. Enzyme activities of *E. coli* JA111 containing plasmids complementing the ALDH-negative phenotype

Strain[a]	Sp act (U/mg of protein) of:	
	Aldehyde dehydrogenase	Propanediol oxidoreductase
JA111(pBR322)[b]	0.01	0.410
JA111(pALD1)	1.74	0.390
JA111(pALD2)	1.59	0.400
JA111(pALD3)	1.67	0.370

[a] Cells were grown in casein hydrolysate aerobically.
[b] Control plasmid.

Determination of *ald* coding region and direction of transcription. One of the cloned plasmids, pALD1, having approximately 9.2 kb of insert DNA, was chosen for further

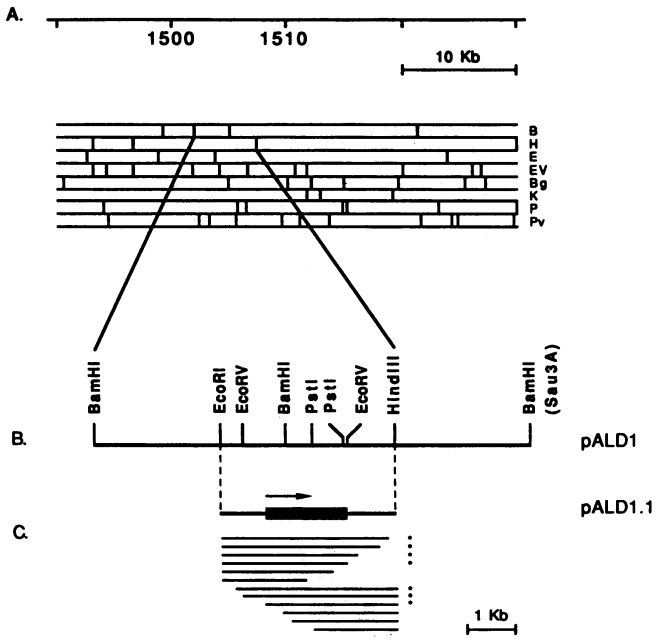

FIG. 1. (A) Fragment of the restriction map of Kohara et al. (19) encompassing the *ald* gene. The restriction sites for several cleaving enzymes are indicated in the horizontal open bars, which are labeled as follows: B, *Bam*HI; H, *Hin*dIII; E, *Eco*RI; EV, *Eco*RV; Bg, *Bgl*I; K, *Kpn*I; P, *Pst*I; Pv, *Pvu*II. (B) Restriction map of the insert of plasmid pALD1 and its correspondence to the map of Kohara et al. (19). The insert was cloned from a *Sau*3A library, which generated the restriction site *Bam*HI marked *Sau*3A. (C) Deletion analysis of the 3.7-kb *Eco*RI-*Hin*dIII insert of pALD1.1. The solid line represents the DNA fragment of pALD1.1. Thinner lines below pALD1.1 represent the fragments of the subclones constructed by deletion of pALD1.1. The fragments complementing the *ald* mutation are marked by an asterisk. The solid bar represents the *ald* coding region. The arrow represents the direction of transcription.

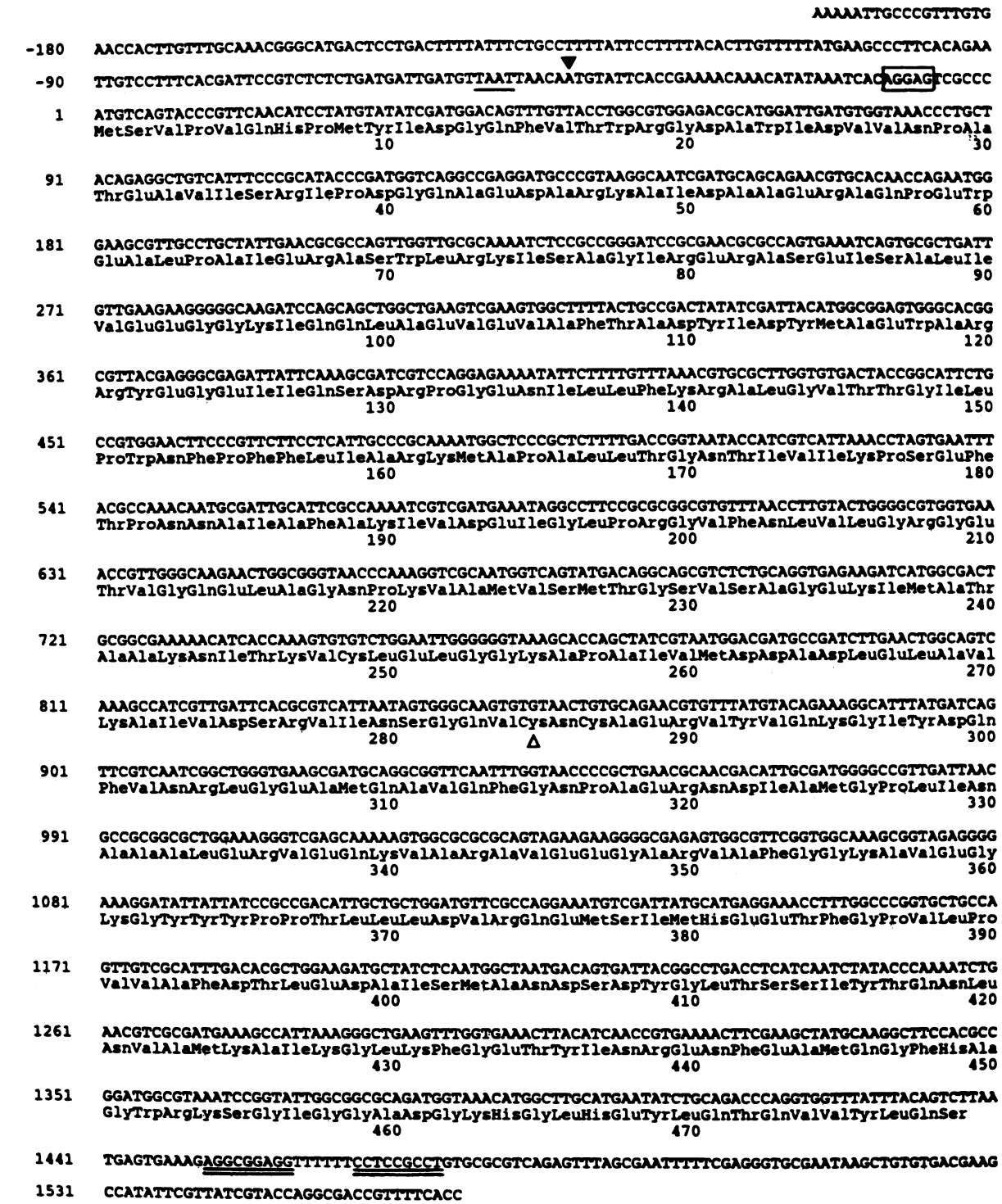

FIG. 2. Nucleotide and deduced amino acid sequences of the *ald* gene. The deduced amino acid sequence is shown below the nucleotide sequence. The putative Shine-Dalgarano sequence is boxed. The transcriptional start point is marked by a solid triangle. The −10 region (Pribnow box) is underlined. The putative *rho*-dependent terminator is underlined twice. The cysteine of the active center is marked by an open triangle.

physical mapping with various restriction enzymes. Several subclones were then constructed and tested for complementation in strain JA111. The restriction map of the *ald* region is shown in Fig. 1A. One of the subclones, plasmid pALD1.1 containing the *Eco*RI-*Hin*dIII fragment of 3.7 kb in the vector Bluescript plasmid (Stratagene), was introduced into JA111 and showed complementation of the *ald* mutation by restoring the enzyme activity.

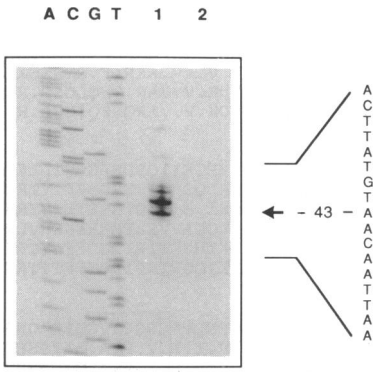

FIG. 3. Identification of the 5′ end of the ald transcript. The primer-extended products (lane 1) were coelectrophoresed with a sequencing ladder (lanes A, C, G, and T) generated by using the same template and primer. A control reaction without RNA was run in lane 2. A portion of the nucleotide sequence deduced from the sequencing lanes is shown on the right, with the transcriptional start point indicated by an arrow at position −43.

Deletion analysis of the 3.7-kb *Eco*RI-*Hin*dIII fragment by use of exonuclease III delimited the region that complemented the *ald* mutation to 1.8 kb (Fig. 1B). By comparing the restriction map of the cloned *ald* region with the corresponding region of the physical map of Kohara et al. (19), we found that the *ald* gene is located between 1504 and 1506 kb on the physical map and around 31.8 min on the chromosomal map (2). Sequencing data of the *ald* region (see below) have shown that the start codon of the open reading frame (ORF) of the *ald* gene is proximate to the *Eco*RI site and distal to the *Hin*dIII site on plasmid pALD1 (Fig. 1). Accordingly, the *ald* gene is transcribed clockwise on the *E. coli* chromosome.

Sequencing of the region coding for aldehyde dehydrogenase. When the nucleotide sequence of plasmid pALD1.1 was completely sequenced for both strands of DNA, a single long ORF was found. This ORF, presented in Fig. 2, encodes 479 amino acids with a predicted molecular mass of 52,278 Da, which is close to that of the enzyme subunit (about 55,000 Da), as indicated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (5). The amino acid composition deduced from the nucleotide sequence was in close agreement with the experimentally determined composition of the enzyme protein reported previously (5).

The ORF starts from the ATG initiation codon at position 1 of the sequenced region, with a potential Shine-Dalgarno sequence at position −11. There is a set of 9-bp inverted repeats 10 nucleotides after the TAA translational stop codon that could form a stable stem-loop structure with a calculated free energy of stabilization of −25.8 kcal (ca. −107.9 kJ) (underlined twice in Fig. 2). This structure is likely to correspond to a *rho*-dependent transcriptional termination signal.

A possible consensus sequence of 7 amino acids for coenzyme specificity was found starting at amino acid 207. This sequence, G-X-G-X-X-X-G, is highly conserved among aldehyde dehydrogenases (see Fig. 4). A cysteine at position 285 most probably corresponds to the highly conserved cysteine of the active center of the aldehyde dehydrogenases. A decapeptide of sequence V-T-L-E-L-G-G-K-S-P, which is highly conserved among other aldehyde dehydrogenases (28), was also found in the *E. coli* enzyme: starting at position 248, the sequence was V-C-L-E-L-G-G-K-A-P.

N-terminal amino acid sequencing of aldehyde dehydrogenase. Aldehyde dehydrogenase was purified from *E. coli* K-12 strain ECL3, and the amino acid of the N terminus was sequenced by automated Edman degradation. The sequence yielded was S-V-P-V-Q-H-P-M-Y-I. This corresponded exactly to the deduced amino-terminal sequence of the ORF, except that the initial Met was lost in the mature protein. This correlation unambiguously showed that the cloned and sequenced gene in fact encoded the aldehyde dehydrogenase protein.

Transcriptional start site of the *ald* gene. The site of transcription initiation of the *ald* gene was determined by primer extension analysis. Total mRNA was prepared from strain ECL1 grown aerobically on rhamnose. The DNA template used in these experiments was derived from plasmid pALD1.1 with an insert containing the *ald* gene and a 200-bp DNA fragment upstream from the nucleotide 1. Single-stranded DNA of this template was hybridized with the total mRNA preparation. After annealing with ^{32}P-labeled sequencing primer that was complementary to the Bluescript polylinker region, the primer was extended with T4 DNA polymerase. The samples were subjected to polyacrylamide gel electrophoresis in parallel with sequencing mixtures prepared with the same template and primer. Two major products differing in 2 bp were observed (Fig. 3, lane 1) that were absent when the primer extension reaction was carried out without mRNA (lane 2). Of the two observed products, the first termination of the primer extension would be the most likely to represent the transcriptional start point, and it corresponded to the adenine at position −43 (Fig. 3). A putative TATA box appears at −10 of the transcriptional start point (underlined in Fig. 2).

Amino acid sequence homology with other aldehyde dehydrogenases. The availability of the amino acid sequence of the *E. coli* aldehyde dehydrogenase gave us an opportunity to determine the amino acid sequence homology between this enzyme and eukaryotic and prokaryotic aldehyde dehydrogenases already determined. Our sequence was introduced and compared with the EMBL data bank. Scores of high homology appeared for all eukaryotic aldehyde dehydrogenases in the bank. Comparison of the sequences of the enzymes with the highest scores was performed by using the Microgenic (Beckman) computer program (Fig. 4). Overall amino acid identities between the entire *E. coli* aldehyde dehydrogenase protein and various aldehyde dehydrogenases showed similar percentages of homology: 35.0% for the horse cytoplasmic enzyme (4); 34.5% for the *Aspergillus niger* enzyme (23); 34.1% for a plant betaine enzyme (28); and 33.5% for the human liver cytoplasmic (15) and mitochondrial (16) enzymes.

DISCUSSION

Identification of the cloned ORF as the structural gene encoding aldehyde dehydrogenase has been established by several criteria. First, the cloned region complements strain JA111 deficient in aldehyde dehydrogenase. This strain was made *recA* negative to avoid misleading recombination in the complementation experiments. The complementation was assessed by analyzing the restoration of aldehyde dehydrogenase activity and the retention of the constitutivity of propanediol oxidoreductase as a control trait of the host. Second, the N-terminal sequence deduced from the ORF matched that of the experimentally determined sequence of the purified protein. Third, the molecular weight of the subunit and the amino acid composition of the protein

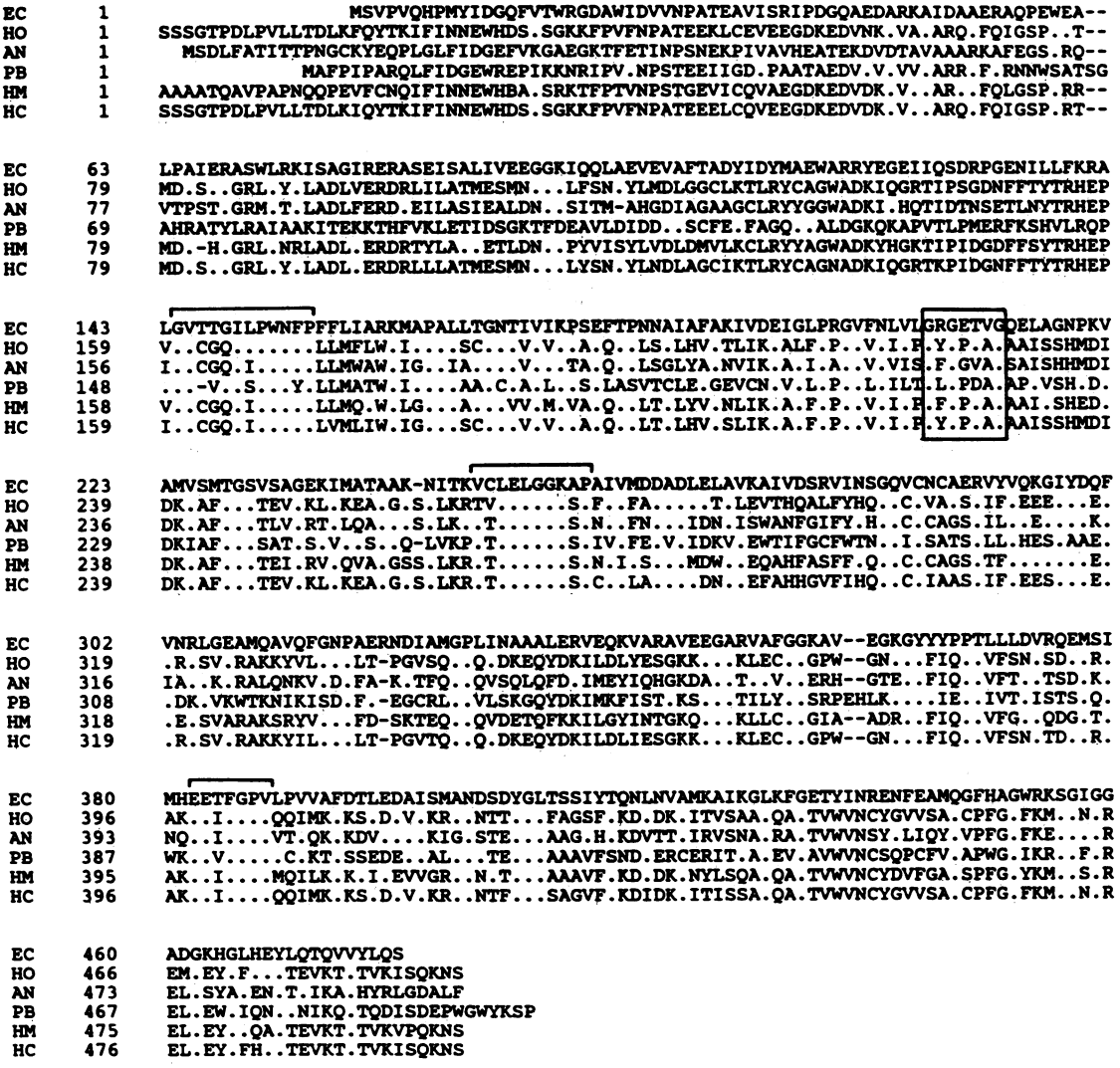

FIG. 4. Comparison of amino acid sequences of several aldehyde dehydrogenases. From top to bottom: *E. coli* K-12, horse cytoplasm, *Aspergillus niger*, betaine-aldehyde dehydrogenase of *Spinacia oleracea*, human liver mitochondria, and human liver cytoplasm. Numbers refer to the amino acid of the *E. coli* protein. Dots indicate amino acid identity with the corresponding amino acid in *E. coli* aldehyde dehydrogenase. Dashes indicate spaces inserted in sequences to give best alignments. Conserved peptides reported in the text are indicated by overlines. The consensus sequence for coenzyme binding is boxed.

purified by Baldoma and Aguilar (5) are in accord with those predicted by the sequence of the structural gene. Finally, the deduced amino acid sequence displays important homologies with other aldehyde dehydrogenases that have highly conserved regions, such as the coenzyme-binding domain and other regions of unknown function.

The presence of a transcriptional start point (determined by primer extension) at the −43 position and the stem-loop structure found at the 3′ end of the *ald* gene are in accordance with an individually transcribed gene. The specification of this transcriptional start point permitted us to determine the RNA polymerase-binding site. The Pribnow box was partially conserved (GTTAAT), whereas no consensus site was found in the −35 region.

In addition to the typical decapeptide (28) occurring in aldehyde dehydrogenase (see Results), some other sequences, such as the dodecapeptide situated between amino acids 144 and 155 and the heptapeptide between positions 382 and 388 of aldehyde dehydrogenase, also highly conserved in our case, have not been pointed out so far. Even though homologous sections are spread all over the sequences of the five representative aldehyde dehydrogenases (Fig. 4), the most conserved regions are confined to the central part of the protein (from amino acids 144 to 412). Not surprisingly, glycines are conserved not only in the consensus sequence for the coenzyme binding but also frequently throughout the protein. The N-terminal part of aldehyde dehydrogenase, on the other hand, is poorly conserved.

The coenzyme-binding consensus sequence found in aldehyde dehydrogenases contains an additional amino acid between the second and third glycines when compared with that of the other NAD(H) general dehydrogenases reported (25). Unexpectedly, glyceraldehyde-3-phosphate dehydrogenase has the general type of coenzyme-binding consensus sequence (9, 25). In this context, it is worth pointing out that no homology with other prokaryotic or eukaryotic aldehyde

dehydrogenases appears in the complete sequence of glyceraldehyde-3-phosphate dehydrogenase.

The general dehydrogenases, such as alcohol dehydrogenases, with a coenzyme-binding consensus sequence G-X-G-X-X-G catalyze reactions that are normally reversible, whereas aldehyde dehydrogenases with a coenzyme-binding consensus sequence G-X-G-X-X-X-G catalyze reactions that are normally irreversible. Interestingly, glyceraldehyde-3-phosphate dehydrogenase, although having an aldehyde as the substrate, catalyzes a reversible reaction and has the consensus sequence of the enzymes catalyzing reversible reactions.

Many sequences of glyceraldehyde-3-phosphate dehydrogenase of different species have been shown to be highly conserved and have been used for phylogenetic comparison (17). Sequences of aldehyde dehydrogenases also seem to be promising for this purpose. Unfortunately, no sequences are yet available for other microbial aldehyde dehydrogenases with a wide spectrum of metabolic functions.

ACKNOWLEDGMENTS

We thank Alison Ulrich for editorial assistance.

This work was supported by grant GM39693 from the National Institute of General Medical Sciences to Y.-M.C. and E.C.C.L. and by grant PB85-0084 from the Comisión Asesora para la Investigación Científica y Técnica to E.H. and J.A. E.H. was the recipient of a predoctoral fellowship (FPI) from the Ministerio de Educación y Ciencia of Spain.

REFERENCES

1. **Ambartsumyan, N. S., and A. M. Mazo.** 1980. Elimination of the secondary structure effect in gel sequencing of nucleic acids. FEBS Lett. **114**:265–269.
2. **Bachmann, B. J.** 1990. Linkage map of *Escherichia coli* K-12, edition 8. Microbiol. Rev. **54**:130–197.
3. **Badia, J., L. Baldoma, J. Aguilar, and A. Boronat.** 1989. Identification of the *rhaA*, *rhaB* and *rhaD* gene products from *Escherichia coli* K-12. FEMS Microbiol. Lett. **65**:253–258.
4. **Bahr-Lindstrom, H., J. Hempel, and H. Jornvall.** 1984. The cytoplasmic isoenzyme of horse liver aldehyde dehydrogenase. Relationship to the corresponding human isoenzyme. Eur. J. Biochem. **141**:37–42.
5. **Baldoma, L., and J. Aguilar.** 1987. Involvement of lactaldehyde dehydrogenase in several metabolic pathways of *Escherichia coli* K-12. J. Biol. Chem. **262**:13991–13996.
6. **Belasco, J. G., J. T. Beatty, C. W. Adams, A. von Gabain, and S. N. Cohen.** 1985. Differential expression of photosynthesis genes in *R. capsulata* results from segmental differences in stability within the polycistronic rxcA transcript. Cell **40**:171–181.
7. **Boronat, A., and J. Aguilar.** 1979. Rhamnose-induced propanediol oxidoreductase in *Escherichia coli*: purification, properties, and comparison with the fucose-induced enzyme. J. Bacteriol. **140**:320–326.
8. **Boronat, A., E. Caballero, and J. Aguilar.** 1983. Experimental evolution of a metabolic pathway for ethylene glycol utilization by *Escherichia coli*. J. Bacteriol. **153**:134–139.
9. **Branlant, G., and C. Branlant.** 1985. Nucleotide sequence of the *Escherichia coli gap* gene. Different evolutionary behavior of the NAD^+-binding domain and of the catalytic domain of D-glyceraldehyde-3-phosphate dehydrogenase. Eur. J. Biochem. **150**:61–66.
10. **Caballero, E., L. Baldoma, J. Ros, A. Boronat, and J. Aguilar.** 1983. Identification of lactaldehyde dehydrogenase and glycolaldehyde dehydrogenase as functions of the same protein in *Escherichia coli*. J. Biol. Chem. **258**:7788–7792.
11. **Chen, Y.-M., Z. Lu, and E. C. C. Lin.** 1989. Constitutive activation of the *fucAO* operon and silencing of the divergently transcribed *fucPIK* operon by an IS*5* element in *Escherichia coli* mutants selected for growth on L-1,2-propanediol. J. Bacteriol. **171**:6097–6105.
12. **Chen, Y.-M., Y. Zhu, and E. C. C. Lin.** 1987. The organization of the *fuc* regulon specifying L-fucose dissimilated in *Escherichia coli* K12 as determined by gene cloning. Mol. Gen. Genet. **210**:331–337.
13. **Chen, Y.-M., Y. Zhu, and E. C. C. Lin.** 1987. NAD-linked aldehyde dehydrogenase for aerobic utilization of L-fucose and L-rhamnose by *Escherichia coli*. J. Bacteriol. **169**:3289–3294.
14. **Cocks, G. T., J. Aguilar, and E. C. C. Lin.** 1974. Evolution of the L-1,2-propanediol catabolism in *Escherichia coli* by recruitment of enzymes for L-fucose and L-lactate metabolism. J. Bacteriol. **118**:83–88.
15. **Hempel, J., H. Bahr-Lindstrom, and H. Jornvall.** 1984. Aldehyde dehydrogenase from human liver. Primary structure of the cytoplasmic isoenzyme. Eur. J. Biochem. **141**:21–35.
16. **Hempel, J., R. Kaiser, and H. Jornvall.** 1985. Mitochondrial aldehyde dehydrogenase from human liver. Primary structure, differences in relation to the cystolic enzyme, and functional correlations. Eur. J. Biochem. **153**:13–28.
17. **Hensel, R., P. Zwicki, S. Fabry, J. Lang, and P. Palm.** 1989. Sequence comparison of glyceraldehyde-3-phosphate dehydrogenases from the three urkingdoms: evolutionary implication. Can. J. Microbiol. **35**:81–85.
18. **Hu, M. C.-T., and N. Davidson.** 1986. Mapping transcriptional start points on cloned genomic DNA with T4 DNA polymerase: a precise and convenient technique. Gene **42**:21–29.
19. **Kohara, Y., K. Akiyama, and K. Isomo.** 1987. The physical map of the whole *Escherichia coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50**:495–508.
20. **Leblanc, D., and R. P. Mortlock.** 1971. Metabolism of D-arabinose: a new pathway in *Escherichia coli*. J. Bacteriol. **106**:90–96.
21. **Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall.** 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193**:265–273.
22. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **O'Connell, M. J., and J. M. Kelly.** 1989. Physical characterization of the aldehyde-dehydrogenase-encoding gene of *Aspergillus niger*. Gene **84**:173–180.
24. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **174**:5463–5467.
25. **Scrutton, N. S., A. Berry, and R. N. Perham.** 1990. Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. Nature (London) **343**:38–43.
26. **Sridhara, S., and T. T. Wu.** 1969. Purification and properties of lactaldehyde dehydrogenase from *Escherichia coli*. J. Biol. Chem. **244**:5233–5238.
27. **Sridhara, S., T. T. Wu, M. Chused, and E. C. C. Lin.** 1969. Ferrous-activated nicotinamide adenine dinucleotide-linked dehydrogenase from a mutant of *Escherichia coli* capable of growth on 1,2-propanediol. J. Bacteriol. **93**:87–95.
28. **Weretilnyk, E. A., and A. D. Hanson.** 1990. Molecular cloning of a plant betaine-aldehyde dehydrogenase, an enzyme implicated in adaptation to salinity and drought. Proc. Natl. Acad. Sci. USA **87**:2745–2749.