JOURNAL OF BACTERIOLOGY, Feb. 1994, p. 871–885
0021-9193/94/$04.00+0

Vol. 176, No. 3

Molecular Characterization of an Aldehyde/Alcohol Dehydrogenase Gene from *Clostridium acetobutylicum* ATCC 824

RAMESH V. NAIR,[1] GEORGE N. BENNETT,[2] AND ELEFTHERIOS T. PAPOUTSAKIS[1]*

Department of Chemical Engineering, Northwestern University, Evanston, Illinois 60208,[1] and Department of Biochemistry and Cell Biology, Rice University, Houston, Texas 77251[2]

Received 17 June 1993/Accepted 2 December 1993

A gene (*aad*) coding for an aldehyde/alcohol dehydrogenase (AAD) was identified immediately upstream of the previously cloned *ctfA* (J. W. Cary, D. J. Petersen, E. T. Papoutsakis, and G. N. Bennett, Appl. Environ. Microbiol. 56:1576–1583, 1990) of *Clostridium acetobutylicum* ATCC 824 and sequenced. The 2,619-bp *aad* codes for a 96,517-Da protein. Primer extension analysis identified two transcriptional start sites 83 and 243 bp upstream of the *aad* start codon. The N-terminal section of AAD shows homology to aldehyde dehydrogenases of bacterial, fungal, mammalian, and plant origin, while the C-terminal section shows homology to alcohol dehydrogenases of bacterial (which includes three clostridial alcohol dehydrogenases) and yeast origin. AAD exhibits considerable amino acid homology (56% identity) over its entire sequence to the trifunctional protein encoded by *adhE* from *Escherichia coli*. Expression of *aad* from a plasmid in *C. acetobutylicum* showed that AAD, which appears as a ~96-kDa band in denaturing protein gels, provides elevated activities of NADH-dependent butanol dehydrogenase, NAD-dependent acetaldehyde dehydrogenase and butyraldehyde dehydrogenase, and a small increase in NADH-dependent ethanol dehydrogenase. A 957-bp open reading frame that could potentially encode a 36,704-Da protein was identified upstream of *aad*.

The strict anaerobe *Clostridium acetobutylicum* is a gram-positive, spore-forming, saccharolytic bacterium capable of fermenting a wide variety of sugars, oligosaccharides, and polysaccharides to butanol, acetone, ethanol, butyrate, and acetate. Advances in the science of metabolic engineering coupled with today's better understanding of the regulation of key acidogenic and solventogenic genes may potentially lead to the development of superior solvent-producing industrial strains.

Butyraldehyde dehydrogenase (BYDH) catalyzes the conversion of butyryl coenzyme A (butyryl-CoA), a key intermediate in the formation of butanol, to butyraldehyde, accompanied by the oxidation of NAD(P)H and the release of CoA from butyryl-CoA. Previously, a BYDH with a subunit molecular mass of 56 kDa was purified from *C. acetobutylicum* B643 (35). However, no genes encoding this or any other BYDH have yet been cloned from *C. acetobutylicum*.

Acetoacetyl-CoA:acetate/butyrate:CoA transferase (CoA-transferase) is the first enzyme in the acetone formation pathway. This enzyme catalyzes the transfer of the CoA moiety from acetoacetyl-CoA to either butyrate or acetate and thus is responsible for the uptake of the two acids during the non-growth-associated solventogenic phase of the fermentation. The two genes (*ctfA* and *ctfB*) encoding the CoA-transferase have recently been cloned and sequenced (4, 38). The end of an open reading frame (ORF) was identified 32 bp upstream of *ctfA* (38). Northern (RNA) analysis revealed that the *ctf* genes were contained within a 4.1-kb transcript in *C. acetobutylicum* DSM 792 (12). A terminator has been identified immediately downstream of *ctfB*, and the coding region of the *ctfA* and *ctfB* genes spans only ~1.4 kb. It is thus believed that at least one other gene is encoded on the remaining ~2.7 kb of transcript upstream of *ctfA*. Enzymatic studies had indicated that BYDH has an induction pattern similar to that of CoA-transferase (19), which plays a critical role in both acetone formation and acid uptake leading to butanol and ethanol formation. It thus appears likely that the gene encoding BYDH is cotranscribed with the *ctf* genes.

Here we report the identification, sequencing, and characterization of a gene, designated *aad*, upstream of *ctfA*. This gene encodes a protein that demonstrates at least two different activities (aldehyde and alcohol dehydrogenase). Both of these activities are necessary for the formation of the two alcohols, butanol and ethanol, from butyryl-CoA and acetyl-CoA, respectively.

(Preliminary results of this study have been presented elsewhere [32a].)

MATERIALS AND METHODS

Bacterial strains. *C. acetobutylicum* ATCC 824 was obtained from the American Type Culture Collection (Rockville, Md.). *Escherichia coli* JM109 [*endA1 recA1 gyrA96 thi hsdR17* (r_K^- m_K^+) *relA1 supE44* λ^- Δ(*lac-proAB*) (F' *traD36 proAB lacI*^{q}ZΔM15)] was obtained from Promega Corp. (Madison, Wis.) and used in all cloning steps. *E. coli* BL21(DE3) (48) [F^- *ompT* (r_B^- m_B^-) (DE3)] was purchased from Novagen, Inc. (Madison, Wis.) and used as the host in T7 expression studies. *E. coli* ER2275 (*recA mcrBC*) was purchased from New England Biolabs (Tozer, Mass.) and used as the methylation host.

Growth conditions and maintenance. All *E. coli* strains were grown aerobically, unless otherwise stated, at 37°C in Luria-Bertani medium supplemented when necessary with ampicillin [50 μg/ml for JM109 cultures and 200 μg/ml for BL21(DE3) cultures]. Both recombinant and wild-type strains were stored at −85°C in 15% (vol/vol) glycerol (41). For [^{35}S]methionine

* Corresponding author. Mailing address: Department of Chemical Engineering, Northwestern University, 2145 Sheridan Rd., Evanston, IL 60208-3120. Phone: (708) 491-7455. Fax: (708) 491-3728.

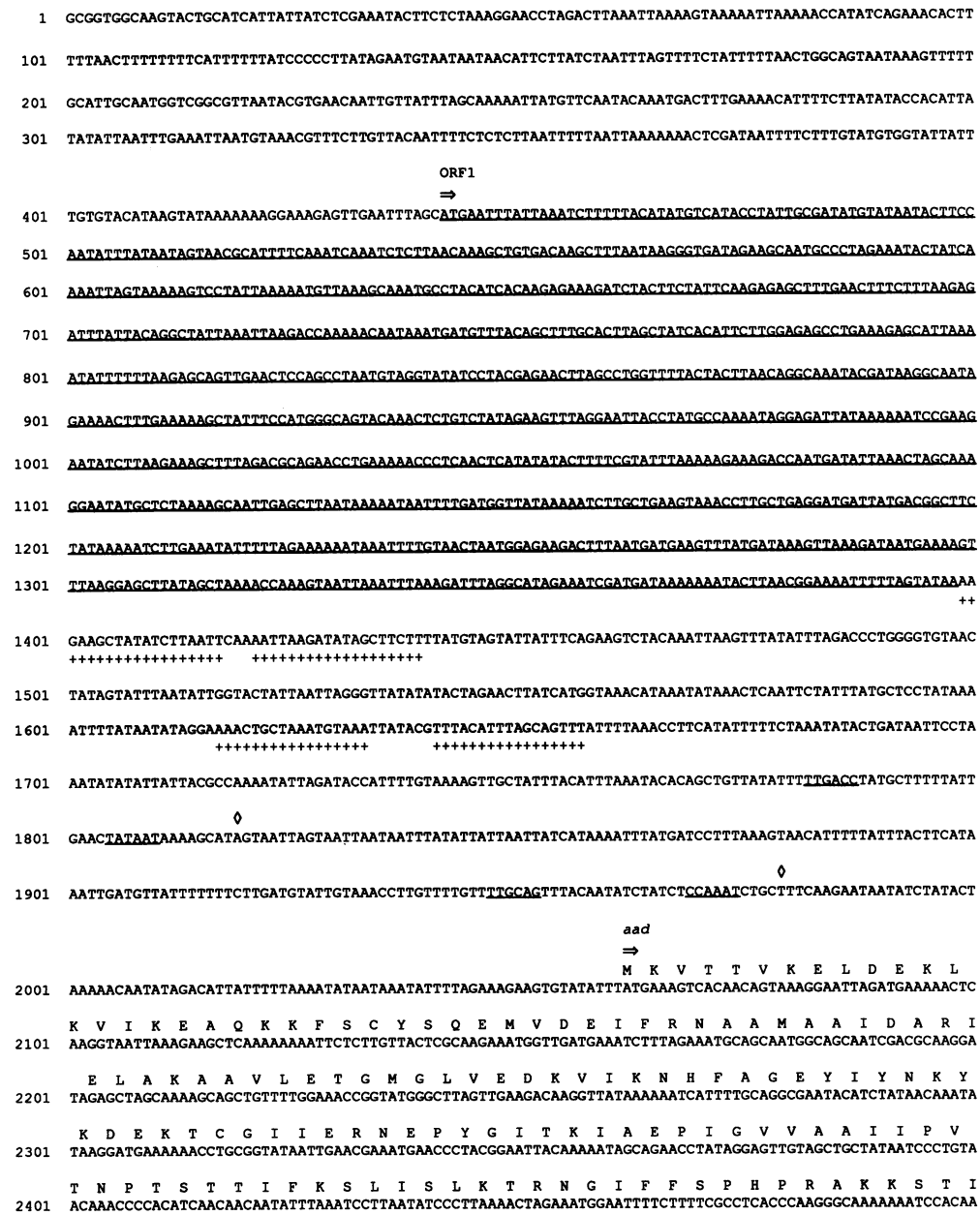

FIG. 1. Nucleotide sequence of *aad*. The translation in the single-letter amino acid representation is placed above the first nucleotide of each codon. The transcriptional start sites of *aad* at nucleotide positions 1819 (S_2) and 1979 (S_1) are indicated (◇). Putative promoter −10 and −35 regions corresponding to these start sites are underlined, as are the nucleotide sequence of the upstream ORF1 and a portion of the downstream *ctfA*. Two inverted repeat segments downstream of ORF1 are underscored (+).

labeling studies, both recombinant and wild-type strains of BL21(DE3) were grown aerobically or anaerobically in M9 minimal medium (41) supplemented with 1% (wt/vol) glucose, 20 μg of thiamine hydrochloride per ml, and 0.01% (wt/vol) 18 amino acids (minus Cys and Met). For recombinant strains, the medium was supplemented with carbenicillin (50 μg/ml). *C. acetobutylicum* was grown in clostridial growth medium (56), supplemented when necessary with 100 μg of erythromycin per ml, and maintained as previously described (32).

Fermentor experiments. For enzyme assay experiments, batch fermentations of *C. acetobutylicum* were performed in a BioFlo II fermentor (New Brunswick Scientific, Edison, N.J.) with a culture volume of 5 liters (clostridial growth medium with glucose at 80 g/liter instead of 50 g/liter). For fermentations of recombinant *C. acetobutylicum*, the medium was supplemented with 75 μg of clarithromycin (Abbott Laboratories, North Chicago, Ill.) per ml (31). The reactor medium (initial pH of 6.2) was inoculated with 500 ml of preculture at an optical density at 600 nm (OD_{600}) of ~0.2. The pH dropped with culture progression until it reached a value of 4.5, at which point it was maintained through additions of 6 M NH_4OH or 2 M HCl. Growth was monitored through OD_{600} measure-

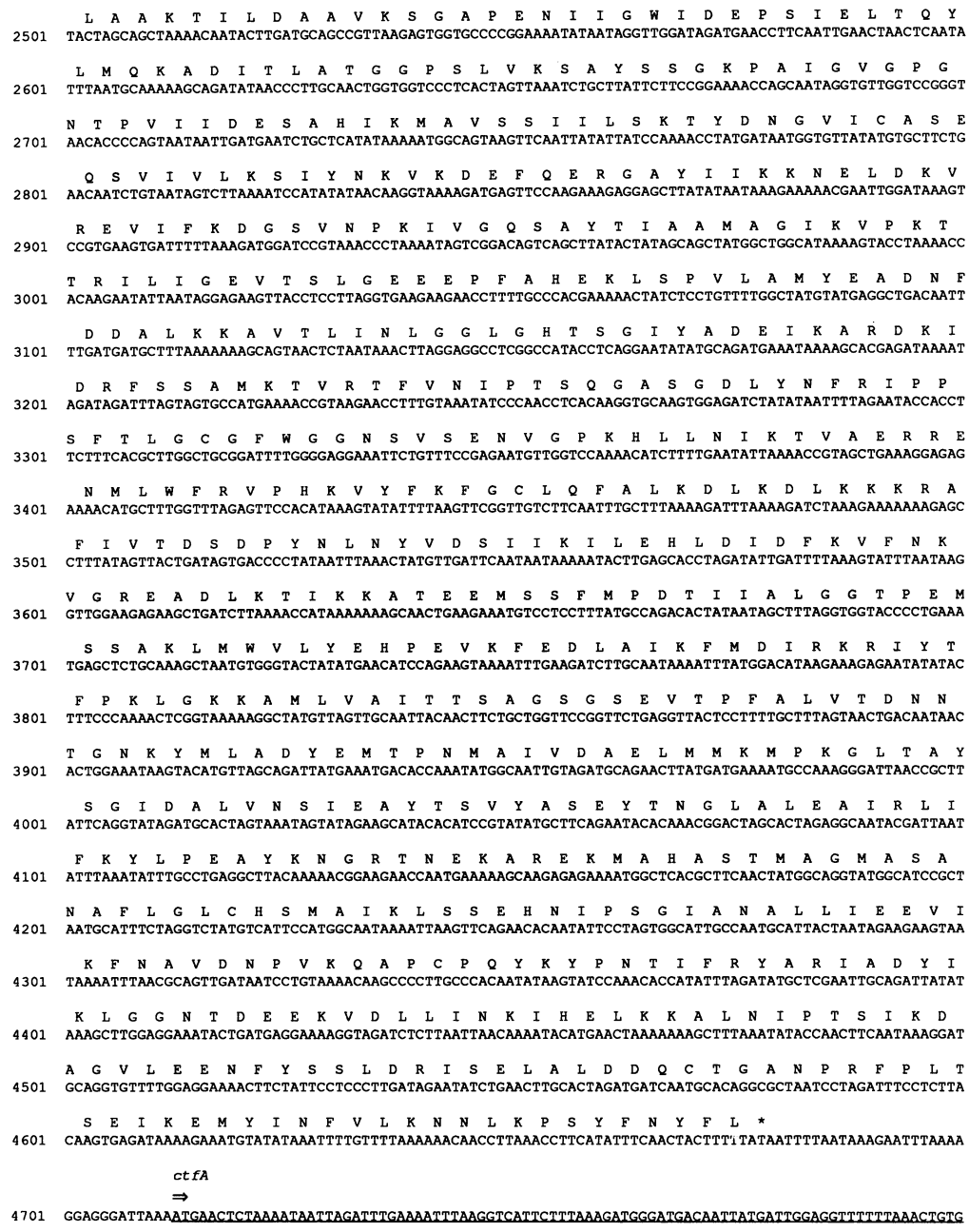

FIG. 1—*Continued.*

ments with a Beckman DU-65 spectrophotometer (Beckman Instruments, Fullerton, Calif.).

DNA isolation, transformation, and manipulation. Isolation of *C. acetobutylicum* chromosomal DNA was performed as described previously (32). Plasmid isolation from *E. coli* was done by the method of Birnboim and Doly (2), with additional steps of the procedure of Wu and Welker (59) when the DNA was to be sequenced. Previously published methods were used for electrotransformation of *E. coli* (11) and *C. acetobutylicum* (30, 32), using a Bio-Rad Gene Pulser (Bio-Rad, Hercules, Calif.). Restriction endonucleases and T4 DNA ligase were purchased from New England Biolabs and used according to the manufacturers' specifications.

Southern blots. Appropriately digested DNA was transferred from agarose gels to nitrocellulose filters (Schleicher & Schuell, Inc., Keene, N.H.) as previously described (47). DNA probes were labeled with [α-^{32}P]dATP (Du Pont Co., NEN Research Products, Wilmington, Del.), using the a nick translation system (Bethesda Research Laboratories [BRL], Life Technologies, Inc., Gaithersburg, Md.). Unincorporated radionucleotides were removed by using Select-D, G-25 spin columns (5 Prime→3 Prime, Inc., Boulder, Colo.).

DNA sequencing. Both DNA strands were sequenced by the dideoxy-chain termination method (42), using synthetic oligonucleotide (20-mer) primers. Double-stranded plasmid DNA was prepared for use as a template, and sequencing reactions

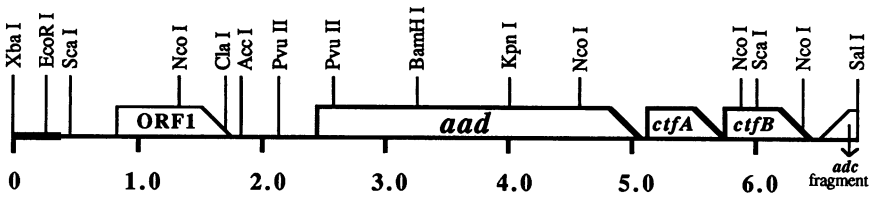

FIG. 2. Transcriptional organization of *aad* and the acetone formation genes. *ctfA* and *ctfB* encode CoA-transferase, and *adc* encodes acetoacetate decarboxylase. The restriction map of the 6.7-kb *SalI-XbaI* fragment of *C. acetobutylicum* DNA in pHXS5 containing the genes is shown. The orientation of transcription of *aad*, *ctfA*, and *ctfB* is indicated by arrowed boxes. Intergenic spans are 663 bp between ORF1 and *aad*, 32 bp between *aad* and *ctfA*, 27 bp between *ctfA* and *ctfB*, and 65 bp between *ctfB* and *adc*. Thin baselines indicate regions where the DNA has been sequenced. Thick baselines indicate regions which have not been sequenced, and consequently, restriction sites within these regions were physically mapped.

were performed by using the Sequenase version 2.0 kit (U.S. Biochemical Corp., Cleveland, Ohio) as specified by the manufacturer.

RNA isolation. Total RNA was isolated from *C. acetobutylicum* as previously described (52). RNA for primer extension studies was isolated from cells that had been producing butanol for 5 h, at which point the butanol concentration was 30% of the final culture concentration.

Primer extension. Primer extension reactions were performed as previously described (12), using Moloney murine

```
aad    1 MKVTTVKELDEKLKVIKEAQKKFSCYSQEMVDEIFRNAAMAAIDARIELAKAAVL
adhE   1 MAVTNVAELNALVERVKKAQREYASFTQEQVDKIFRAAALAAADARIPLAKMAVA

aad   56 ETGMGLVEDKVIKNHFAGEYIYNKYKDEKTCGIIERNEPYGITKIAEPIGVVAAI
adhE  56 ESGMGIVEDKVIKNHFASEYIYNAYKDEKTCGVLSEDDTFGTITIAEPIGIICGI

aad  111 IPVTNPTSTTIFKSLISLKTRNGIFFSPHPRAKKSTILAAKTILDAAVKSGAPEN
adhE 111 VPTTNPTSTAIFKSLISLKTRNAIIFSPHPRAKDATNKAADIVLQAAIAAGAPKD

aad  166 IIGWIDEPSIELTQYLMQKADI--TLATGGPSLVKSAYSSGKPAIGVGPGNTPVI
adhE 166 LIGWIDQPSVELSNALMHHPDINLILATGGPGMVKAAYSSGKPAIGVGAGNTPVV

aad  219 IDESAHIKMAVSSIILSKTYDNGVICASEQSVIVLKSIYNKVKDEFQERGAYIIK
adhE 221 IDETADIKRAVASVLMSKTFDNGVICASEQSVVVVDSVYDAVRERFATHGGYLLQ

aad  274 KNELDKVREVIFKDGSVNPKIVGQSAYTIAAMAGIKVPKTTRILIGEVTSLGEEE
adhE 276 GKELKAVQDVILKNGALNAAIVGQPAYKIAELAGFSVPENTKILIGEVTVVDESE

aad  329 PFAHEKLSPVLAMYEADNFDDALKKAVTLINLGGLGHTSGIYADEIKARDKIDRF
adhE 331 PFAHEKLSPTLAMYRAKDFEDAVEKAEKLVAMGGIGHTSCLYTDQDNQPARVSYF

aad  384 SSAMKTVRTFVNIPTSQGASGDLYNFRIPPSFTLGCGFWGGNSVSENVGPKHLLN
adhE 386 GQKMKTARILINTPASQGGIGDLYNFKLAPSLTLGCGSWGGNSISENVGPKHLIN

aad  439 IKTVAERRENMLWFRVPHKVYFKFGCLQFALKDLKDLKKKRAFIVTDSDPYNLNY
adhE 441 KKTVAKRAENMLWHKLPKSIYFRRGSLPIALDEVITDGHKRALIVTDRFLFNNGY

aad  494 VDSIIKILEHLDIDFKVFNKVGREADLKTIKKATEEMSSFMPDTIIALGGTPEMS
adhE 496 ADQITSVLKAAGVETEVFFEVEADPTLSIVRKGAELANSFKPDVIIALGGGSPMD

aad  549 SAKLMWVLYEHPEVKFEDLAIKFMDIRKRIYTFPKLGKKAMLVAITTSAGSGSEV
adhE 551 AAKIMWVMYEHPETHFEELALRFMDIRKRIYKFPKMGVKAKMIAVTTTSGTGSEV

aad  604 TPFALVTDNNTGNKYMLADYEMTPNMAIVDAELMMKMPKGLTAYSGIDALVNSIE
adhE 606 TPFAVVTDDATGQKYPLADYALTPDMAIVDANLVMDMPKSLCAFGGLDAVTHAME

aad  659 AYTSVYASEYTNGLALEAIRLIFKYLPEAYKNGRTNEKARENMAHASTMAGMASA
adhE 661 AYVSVLASEFSDGQALQALKLLKEYLPASYHEGSKNPVARERVHSAATIAGIAFA

aad  714 NAFLGLCHSMAIKLSSEHNIPSGIANALLIEEVIKFNAVDNPVKQAPCPQYKYPN
adhE 716 NAFLGVCHSMAHKLGSQFHIPHGLANALLICNVIRYNANDNPTKQTAFSQYDRPQ

aad  769 TIFRYARIADYIKLGGNTD--EEKVDLLINKIHELKKALNIPTSIKDAGVLEENF
adhE 771 ARRRYAEIADHLGLSAPGDRTAAKIEKLLAWLETLKAELGIPKSIREAGVQEADF

aad  822 YSSLDRISELALDDQCTGANPRFPLTSEIKEMYINFVLKNNLKPSYFNYFL
adhE 826 LANVDKLSEDAFDDQCTGANPRYPLISELKQILLDTYYGRDYVEGETAAKKEAAP
```

FIG. 3. Amino acid alignment of *aad* and *adhE* gene products. The two proteins share a very high identity of 56% (75% similarity). Boxed positions indicate strict conservation. Abbreviations: aad, *C. acetobutylicum* AAD (this study); adhE, *E. coli* ADH/ACDH/PFL deactivase (13, 21).

TABLE 1. ALDHs with high identities to AAD

Organism and protein	Identity (%)	Size (aa)	Reference
Mus musculus cytosolic ALDH (AHD-2)	23	500	40
Vibrio cholerae ALDH	23	506	37
Aspergillus nidulans ALDH	22	497	39
Escherichia coli ALDH	22	495	17
E. coli succinate semialdehyde dehydrogenase	22	482	34
E. coli lactaldehyde dehydrogenase	21	478	18
Bos taurus mitochondrial ALDH	20	520	15
E. coli betaine ALDH	20	491	3
Pisum sativum ALDH-related turgor-responsive protein	20	508	16
Pseudomonas oleovorans ALDH	17	483	23

TABLE 2. ADHs with high identities to AAD

Organism and protein	Identity (%)	Size (aa)	Reference
Clostridium acetobutylicum P262 ADH1	42	388	62
Zymomonas mobilis ADH2	34	382	8
Bacillus methanolicus methanol dehydrogenase	33	382	10
Saccharomyces cerevisiae ADH4	32	382	57
Escherichia coli 1,2-propanediol oxidoreductase	30	383	5
Clostridium acetobutylicum ATCC 824 BDH I	25	390	52
C. acetobutylicum ATCC 824 BDH II	22	391	52

leukemia virus reverse transcriptase (U.S. Biochemical Corp.). Two different end-labeled oligonucleotides, PE-1 (5′-TTT ACTGTTGTGACTTTCAT-3′) and PE-2 (5′-TTACCTTG AGTTTTTCATCT-3′), complementary to the N-terminal region of *aad* were used in separate primer extension reactions. To map the exact transcriptional start sites, sequencing reactions were performed on the corresponding DNA by using the same primers that were used for the primer extension reactions.

PCR. DNA fragments were amplified by PCR in a reaction mixture which included 300 μM each dATP and dTTP, 200 μM each dGTP and dCTP, 4 mM $MgSO_4$, 0.01 μg of each primer per μl, 100 μg of nonacetylated bovine serum albumin per ml, template DNA, and 0.05 U of Vent DNA polymerase (New England Biolabs) per μl. A total of 15 cycles were executed in a DNA Thermal Cycler (Perkin-Elmer Cetus, Norwalk, Conn.), with each cycle consisting of annealing at 40°C for 1 min, extension at 72°C for 3.5 min, and denaturation at 94°C for 1 min.

Construction of plasmid pCADEX1. *aad*, together with its two putative promoters, was amplified by PCR using plasmid pHXS5 DNA (which contains *aad*; see Results) as a template. The upstream primer, 5′-AAATATACTGAGAATTCCTAA ATATATA-3′ was generated by substituting a G for a T at nucleotide position 1692 (Fig. 1) to provide an internal *Eco*RI site (underlined). The downstream primer, 5′-AATCTAATT ATTCTAGAGTTCATTTTAA-3′, resulted from a substitution of a C for a T at nucleotide position 4723 (Fig. 1) on the complementary strand to provide an internal *Xba*I site (underlined). The amplified product was sequentially digested with *Eco*RI and *Xba*I and ligated into the similarly digested vector pT7/T3α-18 to yield plasmid pCADEX1 (see Fig. 7). The precise orientation of *aad* with respect to the T7 promoter was verified by sequencing the region upstream of *aad* in the construct pCADEX1 by using primer PE-1.

Construction of plasmids pCAAD and pCCL. *Acc*I-digested pCADEX1 was ligated to *Cla*I-digested pIM13 (28) to yield the 8.1-kb plasmid pCAAD (see Fig. 8a). The *Bacillus subtilis* plasmid pIM13 provides MLS^r and a gram-positive origin of replication (28). Plasmid pCAAD was digested sequentially with *Bam*HI (position 2923; Fig. 1) and *Bgl*II (position 4434; Fig. 1) and religated to generate the control plasmid pCCL (see Fig. 8b), which contains a truncated *aad* (58% of the structural gene deleted).

T7 expression. For T7 expression studies, *aad* was inserted into the vector pT7/T3α-18 (BRL). For assay experiments, 500 ml of *E. coli* BL21(DE3) cultures were grown to an OD_{600} of ~0.7. The T7 RNA polymerase under the control of the isopropylthiogalactopyranoside (IPTG)-inducible *lac*UV5 promoter was then induced for 2 h with 0.4 mM IPTG (49) at 37°C under either aerobic or anaerobic (Forma Scientific Anaerobic System, Forma Scientific, Inc., Marietta, Ohio) conditions with another addition of 200 μg of ampicillin per ml. For the rifampin studies, rifampin (200 μg/ml; Sigma Chemical Co., St. Louis, Mo.) was added to cultures 30 min after the start of the induction. The cells were harvested by centrifugation at 5,000 × *g* for 5 min at 4°C. For [^{35}S]Met (Du Pont) labeling experiments, BL21(DE3) cells were grown to an OD_{600} of 0.4 and then subjected to IPTG induction for 20 min and a further 30-min incubation subsequent to rifampin addition.

Enzyme assays. Cells from 500 ml of *C. acetobutylicum* culture were harvested, and the pellets were stored at −85°C until used (always within 3 weeks). Cells were suspended in 15 mM potassium phosphate buffer (pH 7.0) containing 1 mM dithiothreitol and 0.1 mM $ZnSO_4$ (1 g of frozen cells per 5 ml of cell lysis buffer). The cell suspension was sonicated on ice with an ultrasonic cell disruptor (model W-225R; Heat Systems-Ultrasonics, Farmingdale, N.Y.) at level 10 for 18 min with a pulse setting of 90%. The cell debris was removed by centrifugation at 35,000 × *g* for 20 min at 4°C. Crude extracts were prepared from *E. coli* in a similar manner, with the following modifications: the cell lysis buffer contained 1 mg of lysozyme per ml, cells were incubated for 1 h on ice in cell lysis buffer prior to sonication, and cells were sonicated at level 5 for 5 min at a pulse setting of 90%. Butyraldehyde dehydrogenase (BYDH) and acetaldehyde dehydrogenase (ACDH) assays were performed in the reverse physiological direction as previously described (19). One unit is defined as the amount of enzyme necessary to reduce 1 μmol of NAD(P) per min. Butanol dehydrogenase (BDH) assays were performed in the physiological direction as previously described (54), with butyraldehyde at a final assay concentration of 25 mM. An identical procedure was adopted for ethanol dehydrogenase (EDH) assays, with acetaldehyde at a final assay concentration of 50 mM. One unit is defined as the amount of enzyme necessary to oxidize 1 μmol of NAD(P)H per min.

[^{35}S]methionine labeling. Five milliliters of *E. coli* cell samples was withdrawn prior to induction, at the end of the induction period, and after incubation with rifampin addition. Five milliliters of *C. acetobutylicum* cultures was withdrawn at the early exponential (OD_{600} of about 0.6) and late exponential (onset of solventogenesis) stages of growth. These samples were then transferred to prewarmed tubes containing [^{35}S]Met (final concentrations of 50 μCi/ml for *E. coli* cultures and 100

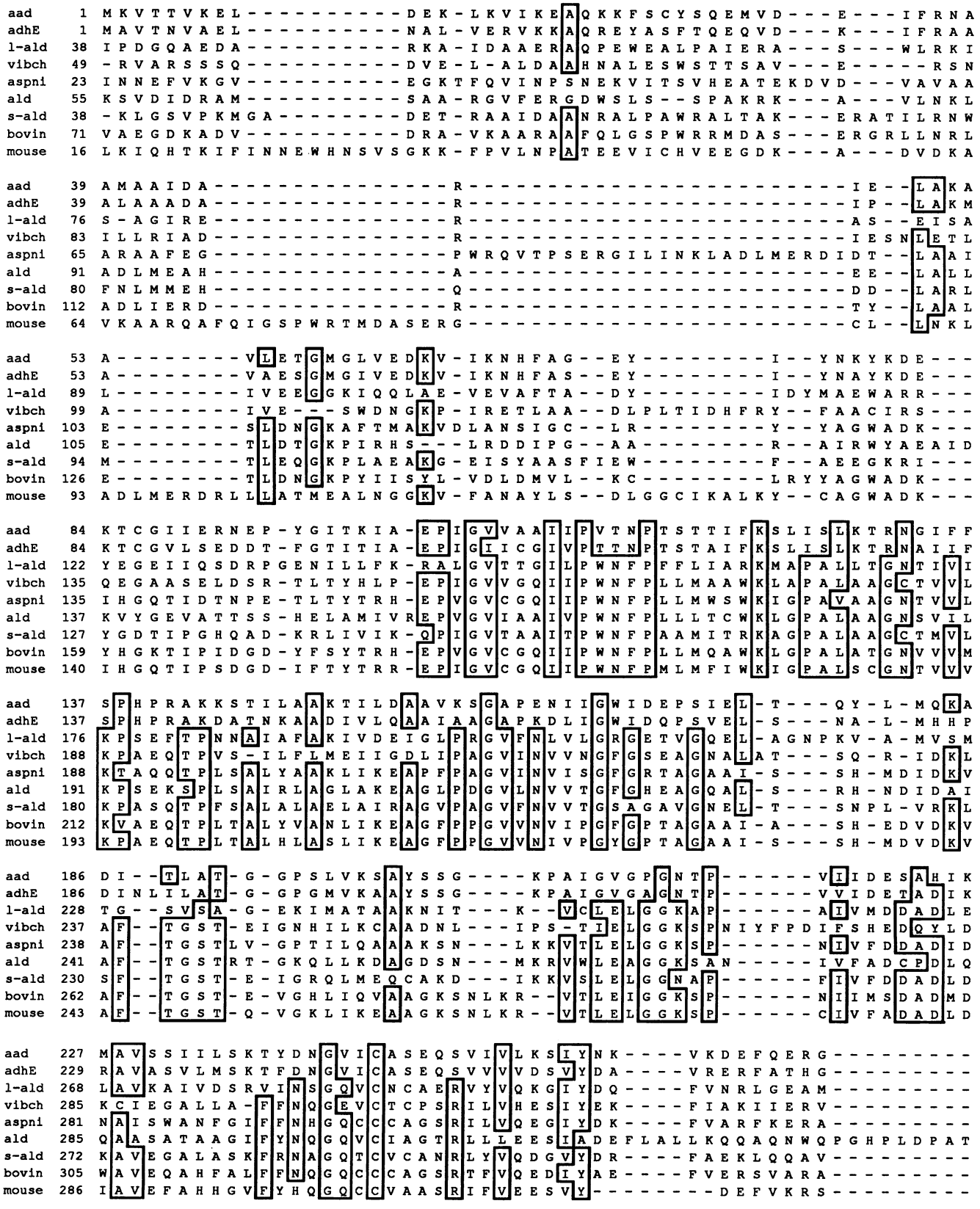

FIG. 4. Amino acid alignment of nine ALDHs. Boxed positions indicate conservation in at least six of the nine proteins. Abbreviations: aad, *C. acetobutylicum* AAD (this study); adhE, *E. coli* ADH/ACDH/PFL deactivase (13, 21); l-ald, *E. coli* lactaldehyde dehydrogenase (18); vibch, *Vibrio cholerae* ALDH (37); aspni, *Aspergillus nidulans* ALDH (39); ald, *E. coli* ALDH (17); s-ald, *E. coli* succinate semialdehyde dehydrogenase (34); bovin, *Bos taurus* mitochondrial ALDH (15); mouse, *Mus musculus* cytosolic ALDH (40).

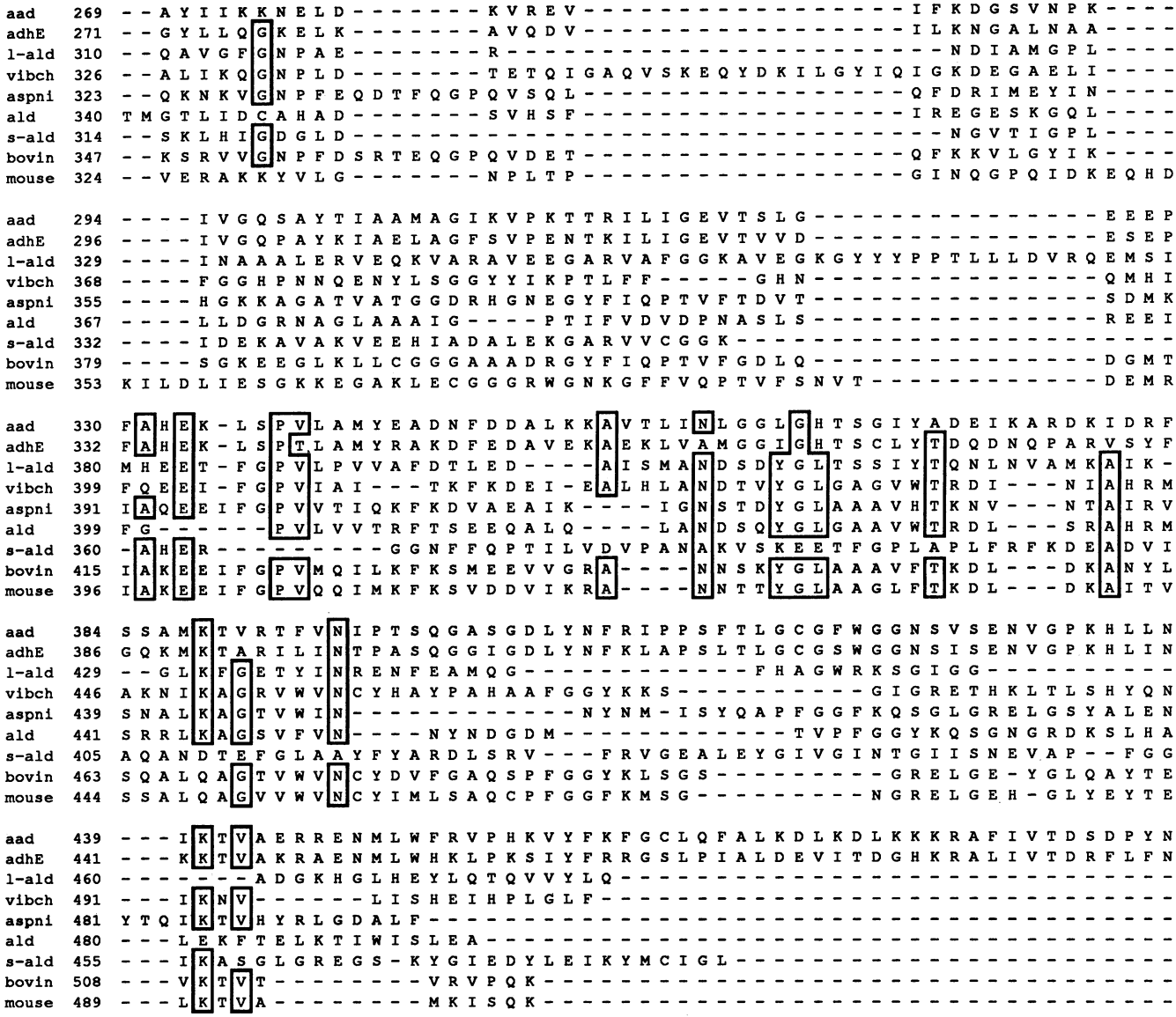

FIG. 4—*Continued.*

μCi/ml for *C. acetobutylicum* cultures) and mixed. Labeling by incubation at 37°C for 10 min was followed by a chase of unlabeled Met (300 μg/ml), and the tubes were then chilled on ice. The cells were washed thrice with 1.5 ml of cold buffer (50 mM Tris-HCl [pH 8.0]–2 mM EDTA for *E. coli* samples and 25 mM Tris [pH 5.5] for *C. acetobutylicum* samples), and cell pellets were stored at −85°C until used.

SDS-polyacrylamide gel electrophoresis (PAGE). Cells from [^{35}S]Met experiments were suspended in sample buffer (41) and placed in a 100°C water bath for 5 min. Cell debris was spun out twice, and an aliquot of the resulting supernatant was subjected to liquid scintillation counting (Model 1900CA Tri-Carb liquid scintillation analyzer; Packard Instrument Co., Meriden, Conn.). Samples were loaded at 500,000 cpm per lane in a sodium dodecyl sulfate (SDS)–8% polyacrylamide gel (25) alongside ^{14}C-labeled high-molecular-weight protein standards (BRL). Upon electrophoresis, gels were dried on Whatman 3MM paper, and imaging plates (exposed for at least 24 h) were scanned in a model 400S PhosphorImager (Molecular Dynamics, Sunnyvale, Calif.). *C. acetobutylicum* samples from bioreactor runs (early exponential, late exponential, and stationary stages of growth) were suspended in sample buffer, boiled, and loaded as described above. High-range (BRL) and mid-range (Promega) protein molecular weight markers were loaded alongside. Gels were stained with 0.25% (wt/vol) Coomassie brilliant blue R-250 (Sigma).

Analysis of fermentation products. The concentrations of butanol, acetone, ethanol, butyrate, and acetate were determined by gas chromatography as described previously (29).

Computer programs. Sequence comparisons and homology searches of the GenBank (release 75.0, February 1993) and SwissProt (release 24.0, January 1993) data bases were done by using the Wisconsin Genetics Computer Group (9) sequence analysis software package (version 7.2). The programs used included FastA, TFastA, BestFit, StemLoop, Fold, Terminator, Composition, PeptideStructure, and PlotStructure. PhosphorImager scans were performed by using ImageQuant version 3.15 (Molecular Dynamics).

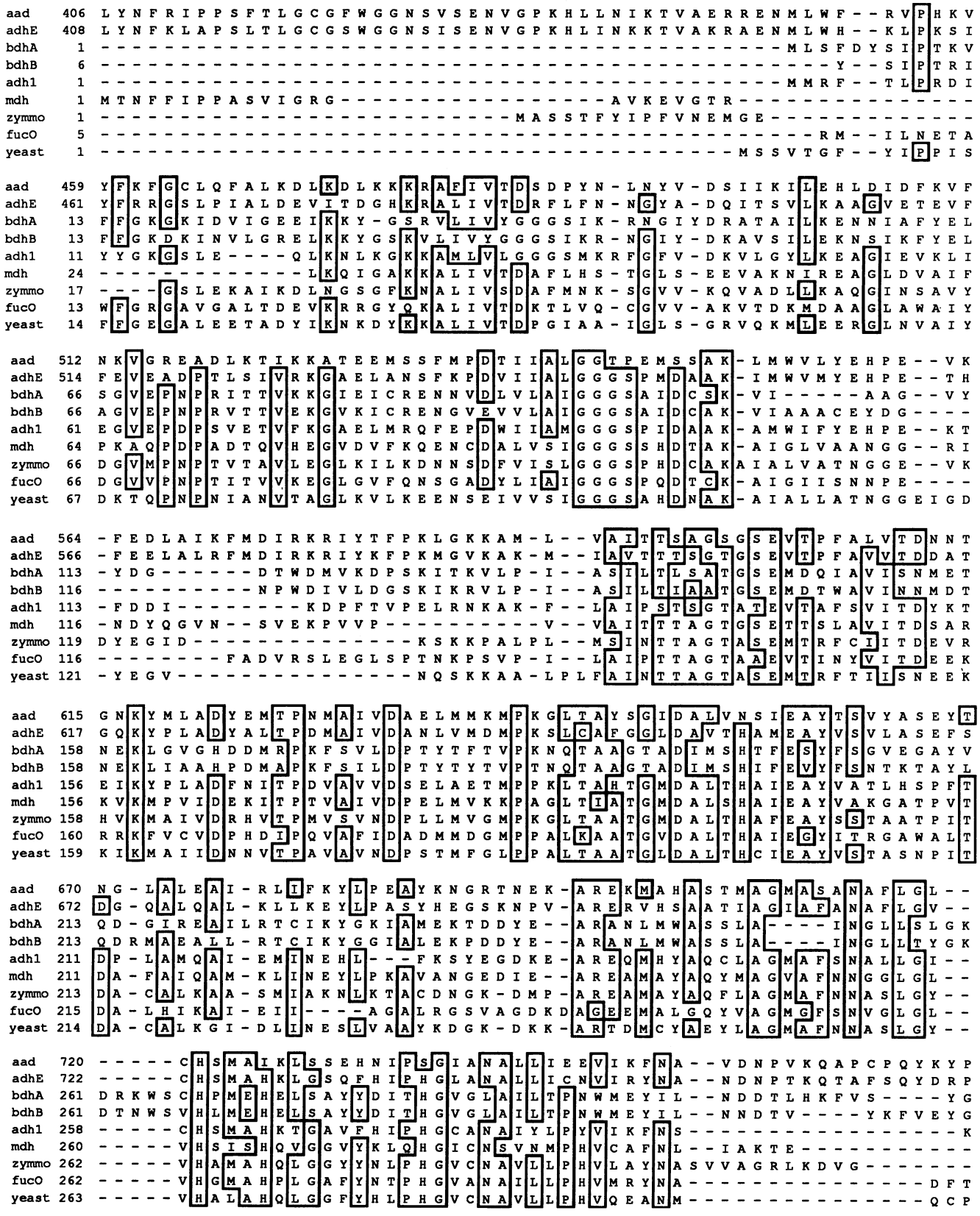

FIG. 5. Amino acid alignment of nine ADHs. Boxed positions indicate conservation in at least six of the nine proteins. Abbreviations: aad, *C. acetobutylicum* AAD (this study); adhE, *E. coli* ADH/ACDH/PFL deactivase (13, 21); bdhA, *C. acetobutylicum* BDH I (52); bdhB, *C. acetobutylicum* BDH II (52); adh1, *C. acetobutylicum* ADH1 (62); mdh, *Bacillus methanolicus* methanol dehydrogenase (10); zymmo, *Zymomonas mobilis* ADH2 (8); fucO, *E. coli* 1,2-propanediol oxidoreductase (5); yeast, *Saccharomyces cerevisiae* ADH4 (57).

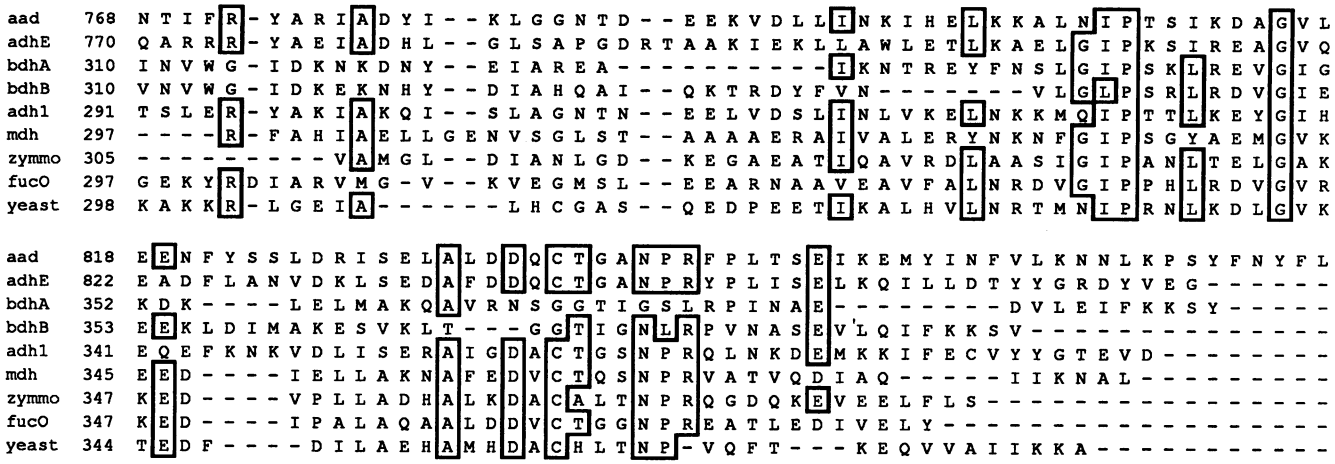

FIG. 5—*Continued.*

Nucleotide sequence accession number. The 4,800-bp sequence (Fig. 1) that includes *aad* and the upstream ORF1 has been submitted to GenBank (1) and has been assigned accession number L14817.

RESULTS

Plasmid pHXS5. Previously, a bacteriophage lambda EMBL3 library of *C. acetobutylicum* DNA was screened by using both *ctfA* and *ctfB* subunit probes to identify a positive phage isolate, H1 (4). A ~6.7-kb *SalI-XbaI* DNA fragment (Fig. 2) from this phage isolate, which was later found to contain intact *ctfA* and *ctfB* genes and a ~5.0-kb region upstream of *ctfA*, was cloned into the similarly digested pUC19 vector to yield plasmid pHXS5. The *SalI* end of this fragment corresponds to that of the 2.0-kb *SalI-BglII* DNA fragment (containing both *ctf* genes) previously cloned on plasmid pCoAT9 (4).

Southern blotting. *NcoI*-digested *C. acetobutylicum* chromosomal DNA was probed with nick-translated pHXS5. The resulting Southern blot (not shown) showed that up to 3.8 kb of DNA upstream of *ctfA* was intact. Figure 2 shows the locations of the *NcoI* sites.

Sequence analysis. Both strands of a 4,712-bp segment of *C. acetobutylicum* DNA upstream of *ctfA* were sequenced (Fig. 1) from pHXS5. Two large complete ORFs were identified on this segment (Fig. 2). The 2,619-bp ORF immediately upstream of *ctfA* was designated *aad*. On the basis of homology searches of protein and DNA data bases, this ORF was found later to encode an aldehyde/alcohol dehydrogenase (AAD). This AAD is a 873-amino-acid (aa) protein with a calculated molecular mass of 96,517 Da. A putative ribosome binding site (RBS) (44) (5′-AAGAAG-3′) was found 9 bp upstream of the *aad* start codon. The sequence and spacing of the RBS are similar to those of other *C. acetobutylicum* genes (36).

The function of the second ORF (ORF1) upstream of *aad* is unknown. If this were a gene, the encoded 319-aa protein would have a molecular mass of 36,704 Da. A putative RBS (5′-GGAAAGAG-3′), similar in sequence and spacing to those of other *C. acetobutylicum* genes (36), was found 11 bp upstream of the ORF1 start codon. Two inverted repeat segments (Fig. 1) were identified in the region of DNA between ORF1 and *aad* ($\Delta G = -20.0$ kcal [1 kcal = 4.184 kJ]/mol [63], positions 1399 to 1439; $\Delta G = -19.4$ kcal/mol, positions 1617 to 1657).

Amino acid comparison. Data base searches revealed that AAD shares a high amino acid homology (56% identity and 75% similarity) with the trifunctional 96,008-Da protein encoded by *adhE* from *E. coli* K-12 (13, 21). Identity was observed over the entire spans of these proteins except for the last 22 aa at the C-terminal end of AAD (Fig. 3).

AAD also showed homology to several aldehyde dehydrogenases (ALDHs) (Table 1) and alcohol dehydrogenases (ADHs) (Table 2). The 10 ALDHs listed in Table 1 (six bacterial, two mammalian, one fungal, and one plant) have a 17 to 23% identity (41 to 46% similarity) to AAD. Homology of these 10 ALDHs (average size = 496 aa) is restricted to within ~460 aa residues at the N-terminal end of AAD. The amino acid alignment of seven of these proteins which show the best identities to AAD and the *adhE* gene product (ADH/ACDH) is shown in Fig. 4. There are 10 amino acid positions that are conserved in all nine proteins and an additional 16 positions that are conserved in eight of nine proteins.

The seven ADHs, which include three clostridial ADHs (52, 62), that show homology to AAD are listed in Table 2 (six bacterial and one yeast). Here the identities to AAD are higher than those of the ALDHs, with the range being 22 to 42% (46 to 65% similarity). Homology of the seven ADHs in Table 2 (average size = 385 aa) is restricted to within ~420 aa residues of the C-terminal section of AAD. The amino acid alignment of these seven proteins with AAD and ADH/ACDH is shown in Fig. 5. There are 19 amino acid positions that are strictly conserved and an additional 20 positions that are conserved in eight of nine proteins.

Transcriptional start sites. It was shown previously that the *ctfA* and *ctfB* genes were contained on a 4.1-kb transcript in *C. acetobutylicum* DSM 792 (12). These genes have also been found on a similar-size transcript in strain ATCC 824 (51). These findings indicate that *aad*, *ctfA*, and *ctfB* are expressed as an operon, since the corresponding size of the transcript based on sequence analysis agrees with the Northern blot results.

A primer extension experiment was performed to identify the precise 5′ end of the *aad-ctf* operon. The results are shown in Fig. 6. Two bands were observed that corresponded to positions 83 bp (S_1; intense band) and 243 bp (S_2; faint band)

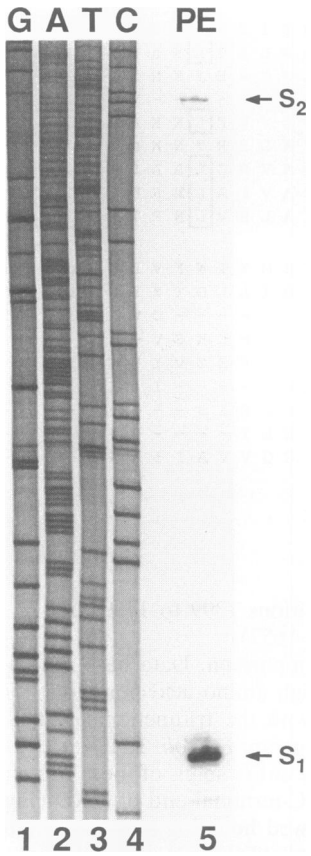

FIG. 6. Primer extension analysis. Primer extension products made by using primer PE-1 complementary to *aad* are shown in lane 5. The arrows indicate bands corresponding to transcriptional start sites S_1 and S_2 (83 and 243 bp, respectively, upstream of the *aad* start codon). RNA used in this experiment was obtained from *C. acetobutylicum* cells which had been producing butanol for 5 h. Regions of plasmid pHXS5 were sequenced by using the same primer, and the resulting DNA sequences are shown in lanes 1 to 4.

upstream of the *aad* start codon. To rule out the possibility of nonspecific hybridization of the PE-1 primer with different mRNA, the experiment was repeated (not shown) with a second primer (PE-2) complementary to a different N-terminal region of *aad*. Both experiments revealed identical dual start sites with the same relative band intensities. The −10 and −35 regions of the apparent promoters, along with those of other *C. acetobutylicum* solventogenic genes, are shown in Table 3. The distal, apparently weaker, *aad-ctf* promoter was similar to that previously identified as a consensus clostridial promoter (61).

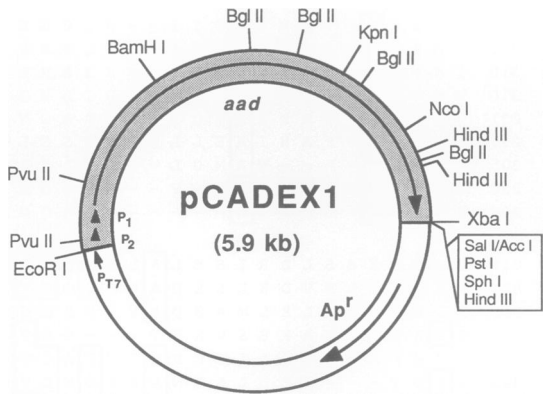

FIG. 7. Plasmid pCADEX1. The 3.0-kb insert (shaded region) containing *aad* and its two natural promoters (P_1 and P_2 upstream of transcriptional start sites S_1 and S_2, respectively) was amplified by PCR, using single-base-substituted primers to generate *Eco*RI and *Xba*I ends, and cloned into the T7 expression vector pT7/T3α-18. The T7 promoter is designated P_{T7}.

However, the other *aad-ctf* promoter differed from those of other solventogenic genes, especially in the −10 region.

As expected, no terminator structures could be located in the 32-bp intergenic region between *aad* and *ctfA* or the 27-bp intergenic region between *ctfA* and *ctfB*. The presence of a bidirectional terminator has been identified downstream of *ctfB* (38).

ALDH and ADH activities of AAD in *E. coli* and *C. acetobutylicum*. Cell crude extracts were assayed for BYDH, ACDH, BDH, and EDH activities. Preliminary attempts to detect ALDH and ADH enzyme activities in JM109(pCADEX1) crude extracts were unsuccessful, suggesting either that the natural promoters could not function in *E. coli* or that the protein itself was not functional. To overcome this potential difficulty, pCADEX1 (Fig. 7), whereby *aad* is transcribed from the T7 promoter, was used to transform *E. coli* BL21(DE3). Under aerobic growth conditions, low activities were detected for NAD-specific BYDH, NAD- and NADP-specific ACDH, and NADPH-specific BDH (results not shown). Previously, an ALDH from *Clostridium beijerinckii* NRRL B592 was found to be sensitive to oxygen (60). Moreover, *adhE-lacZ* fusion studies have indicated a 10- to 20-fold induction of ADH/ACDH under anaerobic conditions (26). Hence, the assays were repeated with cells grown anaerobically. Rifampin was also added to these cultures to selectively inhibit the host RNA polymerase (50) and thereby reduce any background activity of the host. Under these conditions, more pronounced activities were detected for NAD-specific BYDH and NADH-specific BDH, with both vector and host controls yielding no detectable

TABLE 3. Mapped promoter regions of solventogenic *C. acetobutylicum* genes

Gene	Sequence			Reference
	−35 region	−10 region	−1	
aad-ctf	TTATATTT TTGACC	TATGCTTTTTATTGAAC TATAAT	AAAAGCAT(S_2)[a]	This study
	GTTTTGTT TTGCAG	TTTACAATATCTATCT CCAAAT	CTGC(S_1)[a]	This study
adc	TGTAAAAA TTTACT	TAAAAAAACAATATGTGT TATAAT	GTAAAT	12
Consensus clostridial sequence	-T------ TtgAcA	----T-----AAtATga TAtAAT	---T---	61
bdhA	TTAGATGC TTGTAT	TAAAATAATAAAATAG TAAAAT	ATAAGTA	52
bdhB	AAATATTA TTGTAA	TAATTTTAAGTAGGTT TAAAAT	ATATATA	52

[a] S_1 and S_2 are the transcriptional start sites referred to in Fig. 6.

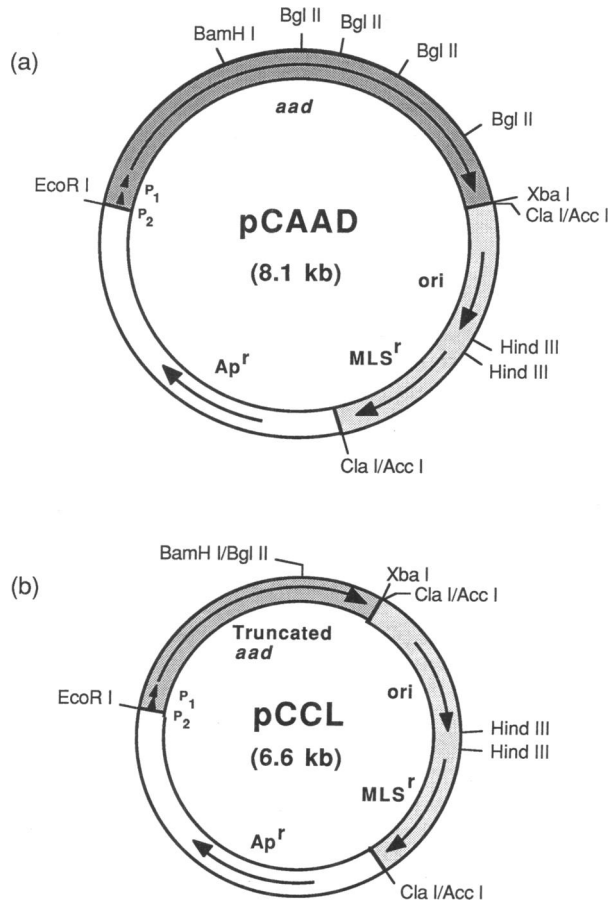

FIG. 8. (a) Plasmid pCAAD. *Acc*I-digested pCADEX1 was ligated to *Cla*I-digested *B. subtilis* plasmid pIM13 (28) (which confers MLS^r and a gram-positive origin of replication). (b) Plasmid pCCL. pCAAD was sequentially digested with *Bam*HI and *Bgl*II and religated (*Bam*HI at position 2923 [Fig. 1] and *Bgl*II at position 4434 [Fig. 1]) to yield this control plasmid carrying a truncated *aad* (58% of structural gene deleted).

activities (results not shown). Under no conditions could NADH- or NADPH-specific EDH activity be detected from plasmid pCADEX1.

When assays were performed with crude extracts of *C. acetobutylicum* samples from fermentor experiments at pH 4.5 of wild-type and recombinant strains carrying plasmid pCAAD (Fig. 8a) or pCCL (Fig. 8b), activities were detected for NAD-specific BYDH and ACDH and for NADH-specific BDH and EDH (Table 4). While BDH and ACDH activities from pCAAD were higher than those of the wild-type strain (up to 2.5- and 70-fold, respectively) and ATCC 824(pCCL) control strain (up to 2- and 60-fold, respectively), the pCCL control also reflects some degree of enhancement of these activities (up to 2- and 45-fold higher, respectively). EDH activities were lower, but the same trend can be observed. However, BYDH specific activity indicates a more selective amplification in the strain carrying the intact *aad* on pCAAD. Although the control 824(pCCL) fermentation shows higher final butanol and acetone concentrations, the ethanol-to-acetone and butanol-to-acetone ratios (Table 5) indicate that the recombinant strain carrying *aad* (on pCAAD) produces alcohol with improved selectivity compared with either of the controls. In addition, the 824(pCAAD) fermentation shows 2- and 2.7-fold-higher ethanol concentrations compared with the control fermentations 824(pCCL) and 824, respectively.

TABLE 4. ALDH and ADH specific activities in *C. acetobutylicum* from fermentor experiments at pH 4.5

C. acetobutylicum ATCC 824	Growth phase[a]	Sp act[b] (mU/mg of protein)			
		NAD specific		NADH specific	
		BYDH	ACDH	BDH	EDH
Wild type	A	<1	<1	120	<1
	B	<1	3	161	<1
	C	<1	<1	372	<1
Harboring pCCL	A	<1	1	268	4
	B	<1	138	329	1
	C	<1	16	166	<1
Harboring pCAAD	A	<1	63	252	5
	B	69	224	426	3
	C	<1	60	418	2

[a] A, early exponential growth phase; B, late exponential growth phase; C, stationary phase.

[b] NADP-specific BYDH and ACDH activities and NADPH-specific BDH and EDH activities were very low and hence are not reported. Errors in activity numbers are typically of the order of 20% (53).

Detection of AAD. Amplification of *aad* in *C. acetobutylicum* ATCC 824(pCAAD) revealed a strong induction of a ~96-kDa protein during the onset of solventogenesis (late exponential stage) in Coomassie blue-stained denaturing gels (Fig. 9, lane 7) and in gels (not shown) with [^{35}S]Met-labeled proteins. In strain 824, this band is seen faintly at the onset of solventogenesis (lane 3). Strain 824(pCCL), like strain 824, has only a single intact copy (chromosomal) of *aad*. However, plasmid (pCCL)-imposed metabolic stress (see Discussion) seems to lead to elevated expression of chromosomal *aad* (lane 12). Gels (not shown) run with [^{35}S]Met-labeled proteins from T7 expression studies in *E. coli* did not reveal overexpression of a ~96-kDa protein which would be expected from the molecular weight of the translated *aad*. Coomassie blue-stained gels of total protein also confirmed this finding. However, an intense band corresponding to ~42 kDa was observed (gel not shown) in the rifampin-treated sample of cells carrying pCADEX1. Several fainter bands corresponding to higher-molecular-weight proteins were observed for both aerobically and anaerobically grown *E. coli* cells (results not shown).

DISCUSSION

C. acetobutylicum AAD, which demonstrates both aldehyde and alcohol dehydrogenase functions, shows high identity to the trifunctional *adhE* from *E. coli*. Under anaerobic conditions, the *adhE* gene product has ACDH and ADH activities, which are responsible for ethanol formation (13), and pyruvate-formate-lyase (PFL) deactivase activity (21). PFL deactivase inactivates PFL, which is required in the nonoxidative cleavage of pyruvate to form acetyl-CoA and formate (6, 22). The anaerobically induced PFL deactivase exists as an iron-dependent homopolymeric rod-shaped protein (~40 × 96 kDa) (21). Amino acid alignment of AAD and ADH/ACDH (Fig. 3) revealed that identity is established at the very N-terminal end and the loss in identity is in the C-terminal region (22 aa residues of AAD and 37 aa residues of ADH/ACDH). *E. coli* ADH/ACDH is 19 aa residues longer than AAD. The presence of PFL deactivase activity has not been verified for AAD, since no evidence exists for the presence of PFL in *C. acetobutylicum*. PFL activity has been detected in

TABLE 5. Final product levels in *C. acetobutylicum* fermentor experiments at pH 4.5

C. acetobutylicum ATCC 824	Concn (mM)					Ratio	
	Butanol	Acetone	Ethanol	Butyrate	Acetate	Butanol/acetone	Ethanol/acetone
Wild type	74	66	15	36	53	1.12	0.23
Harboring pCCL	138	95	21	6	12	1.45	0.22
Harboring pCAAD	118	72	41	5	31	1.64	0.57

Clostridium butyricum (a species substantially different from *C. acetobutylicum*). In contrast to the PFL of *E. coli*, the *C. butyricum* PFL functions mainly to furnish formate for anabolic C_1-unit formation (58).

The location of the 957-bp ORF1 upstream of the *aad-ctf* operon suggested that it could perhaps encode a regulatory protein. However, data base searches did not reveal homology to DNA-binding proteins.

As previously observed (18), the N-terminal ends of ALDHs appear to be poorly conserved; however, ALDH domains (N-terminal regions) of AAD and ADH/ACDH show homology to several ALDHs (Fig. 4). Strictly conserved residues are found in 10 positions, half of which are Gly residues. The two strictly conserved Pro residues are present as part of a highly conserved 15-aa peptide (starting at position 102 [Fig. 4] in AAD). The dodecapeptide that includes the last 12 of these 15 residues has previously been reported to be conserved among ALDHs (18). Proteins which demonstrate only ALDH function have the highly hydrophobic residues Phe and Trp between the Pro residues. AAD and ADH/ACDH have hydrophilic (Thr) to moderately hydrophobic (Val) residues in this region of unknown function. A decapeptide of sequence VTLELGGKSP (starting at position 261 in the protein designated aspni in Fig. 4) is highly conserved among ALDHs (55) but not in AAD or ADH/ACDH. The heptapeptide EExFGPV (starting at position 382 in the protein designated l-ald in Fig. 4; x is any amino acid) is also conserved among several ALDHs (18), but the corresponding sequence in AAD and ADH/ACDH is a poor match. These comparisons clearly indicate that the N-terminal regions of AAD and ADH/ACDH, while related to ALDHs, have diverged in some ways from this group.

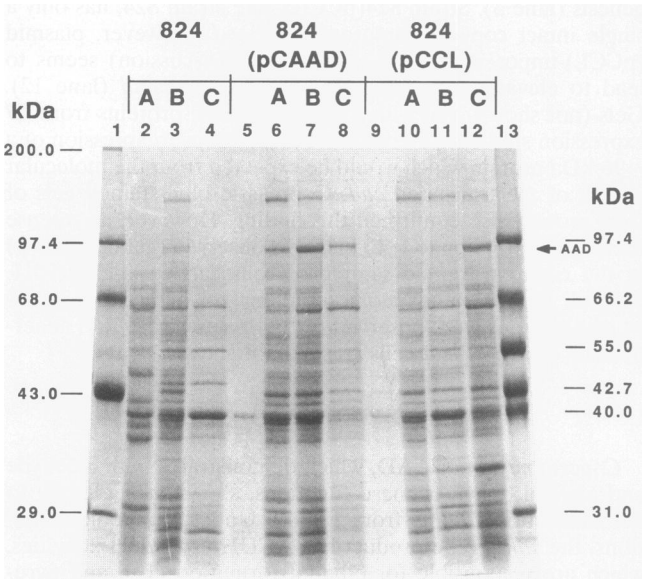

FIG. 9. SDS-PAGE. Cell samples from *C. acetobutylicum* controlled-pH (4.5) batch fermentations of both wild-type and recombinant strains were taken at the early exponential (A), late exponential (B), and stationary (C) stages of growth, and proteins were electrophoresed in SDS–8% polyacrylamide gels and Coomassie blue stained. Lanes 2 to 4 correspond to the wild-type strain, lanes 5 to 8 correspond to the recombinant strain carrying intact *aad* on plasmid pCAAD, and lanes 9 to 12 correspond to the recombinant control strain carrying a truncated *aad* on plasmid pCCL. Lanes 5 and 9 represent early acidogenic stages (prior to stage A) of the corresponding recombinant strains. The ~96-kDa band representing AAD is indicated. The prominent ~39-kDa-band protein observed in lanes 2 to 12 is probably the butyrate kinase (36). High-range (lane 1) and medium-range (lane 13) protein molecular weight markers are shown alongside.

The ADH domains (C-terminal regions) of AAD and ADH/ACDH show homology to several ADHs (Fig. 5). Of the 19 positions that are strictly conserved, there are four Gly and three Pro residues. This overrepresentation in conserved Gly residues has been suggested as typical of ADHs with largely conserved conformation (20). Cys and His residues have been implicated as metal-binding ligands in zinc-containing long-chain ADHs (20). A comparison of AAD and ADH/ACDH (Fig. 3) reveals seven His (four in the N-terminal part) and four Cys (two in the N-terminal part) residues that are conserved throughout both proteins.

The coenzyme specificity of dehydrogenases has been fairly well documented (20, 24, 43). A highly conserved glycine-rich region with a GxGxxG pattern within a βαβ Rossmann fold was found to be common to the NADH-binding domain of many ADHs (20, 43). A GxGxxGG pattern is observed in the central regions (Fig. 3) of AAD (starting at residue 418) and ADH/ACDH (starting at residue 420). With the exception of glyceraldehyde-3-phosphate dehydrogenase (43), a coenzyme-binding consensus sequence among ALDHs seems to contain an additional amino acid between the second and third glycines (18). Such a sequence is observed in several ALDHs (Fig. 4; aligned to initiating Gly-168 in AAD) but not in AAD and ADH/ACDH. However, a similar sequence is present in AAD starting at position 362 in a region that is not very well conserved in the initiating glycine. A GxGxxA sequence within a βαβ fold (with another alanine almost always present a further four residues downstream) is conserved among some NADP(H)-binding proteins (43). However, some significant differences among NADPH-binding domains (24, 43) prevents identification of such a region on the basis of sequence comparisons alone. A GxGxxVxxxA sequence exists in AAD starting at position 598 (Fig. 5) in a region highly conserved among ADHs. The ADH portion of AAD has a 42% identity (65% similarity) with the NADPH-dependent ADH1 from the industrial strain *C. acetobutylicum* P262 (62).

It appears that in strain ATCC 824, ALDH (both BYDH and ACDH) functions are present as part of a multifunctional ~96-kDa (monomer size) protein (Fig. 9 and Table 4) which is the product of *aad*. Earlier, a CoA-linked BYDH (monomer size of 56 kDa and native size of 115 kDa) was purified from *C. acetobutylicum* B643 (35), and a CoA-acylating ALDH (mono-

mer size of 55 kDa and native size of 100 kDa) was purified from *C. beijerinckii* NRRL B592 (60). It is not known whether proteins similar to these proteins exist in strain ATCC 824. However, different genetic organizations and coenzyme specificities of similar enzymes in different strains of clostridia is not an unusual phenomenon. Of the three CoA-linked ALDHs isolated from *Clostridium kluyveri*, an NAD-specific enzyme from the soluble fraction (45) has a molecular mass of 290 kDa, an enzyme (NAD and NADP specific) from the particulate fraction of strain DSM 555 has a molecular mass of 194 kDa composed of 55- and 42-kDa subunits (27), and an NADP-specific succinate semialdehyde dehydrogenase from *C. kluyveri* DSM 555 has a monomer size of 55 kDa and native size of 115 kDa (46).

Table 5 shows that AAD confers a superior ethanol formation capability compared with both the wild-type strain and the 824(pCCL) plasmid control strain. This is possibly due to elevated activities (Table 4) of the NAD-specific ACDH and to a lesser extent the NADH-specific EDH. Since no attempt was made to enhance acetone pathway genes, it is meaningful to compare ethanol-to-acetone and butanol-to-acetone ratios obtained in wild-type and recombinant strain fermentations (Table 5). The ethanol-to-acetone ratio for the recombinant strain carrying intact *aad* (on pCAAD) is more than double that for either control (Table 5). The butanol-to-acetone ratio also reflects an enhancement (Table 5) although not quite as pronounced. As observed earlier (31), host-plasmid interactions also lead to elevated final solvent levels and decreased final acid levels (Table 5) due to a possible induction of stress proteins upon plasmid-imposed metabolic stress. We have found this to be a general effect with several other plasmid constructs, as we shall report shortly (52a). We also note the appearance of a strong band at ~65 to 66 kDa, which parallels in intensity the AAD band and solvent production in all three fermentations (Fig. 9; Table 5). This band could represent DnaK, a 65.6-kDa *C. acetobutylicum* stress protein (33). This would be consistent with the proposed connection between the stress protein induction and solventogenesis (33).

The high BDH activities observed in *C. acetobutylicum* (Table 4), which are likely in large part due to BDH I and BDH II isozymes (52), coupled with the use of the same substrate (butyraldehyde) in both BYDH and BDH assays, leads to a masking of BYDH activities due to an uptake by BDH of the NAD(P)H generated in the BYDH reaction. This could explain the low BYDH activities detected (Table 4). However, detectable ACDH activities are higher, possibly because of low detectable levels of EDH (Table 4), which would minimize such a masking effect on the ACDH activities.

Expression of *aad* in *E. coli* from its natural promoters was undetectable even though one of the natural promoters (Table 3) shows conservation to the *E. coli* σ^{70} RNA polymerase promoter in 11 of 12 positions in the -10 and -35 regions. Also, the T7 expression system described in this report did not lead to the expected high ALDH and ADH activities of AAD in *E. coli*. Since no overexpressed ~96-kDa-band protein was detected in any of the *E. coli* BL21(DE3) gels, the question remains as to whether AAD exists in *E. coli* as a 96.5-kDa protein or whether it undergoes posttranslational processing into an ALDH (average size of ~55 kDa [Table 1]) and an ADH (average size of ~42 kDa [Table 2]). The idea that the 42-kDa band observed in *E. coli* protein gels is perhaps a truncated AAD could be explained by the presence of a consensus *E. coli* RBS starting at position 3395 (Fig. 1) upstream of a Met start site at position 3406 (Fig. 1), which is precisely where the ADH homology region is initiated (Fig. 5). However, this does not explain the observed ALDH activities in *E. coli*. These ambiguous results in *E. coli* prompted the *C. acetobutylicum* studies described in this report. Internal translation initiation at position 3406 could also occur in *C. acetobutylicum*, and the prominent bands of ~43 to 44 kDa appearing in lanes 3, 6 to 8, and 10 to 12 of Fig. 9 could represent such a translated product from the chromosomal *aad* [for 824 and 824(pCCL)] and the chromosomal and plasmid *aad* [for 824(pCAAD)].

The *aad-ctf* operon seems to play a very important role in solvent formation. CoA-transferase and AAD are induced and derepressed simultaneously at the onset of solventogenesis. Activity data suggest that AAD could provide both ALDH and ADH functions in butanol and ethanol formation pathways. Specific mutations in ALDH and ADH regions of AAD would be necessary to verify that these are mutually exclusive functions of the same protein. Coregulation of butanol and acetone formation genes implies that a mutation in a key regulatory region in this common locus could lead to loss of several solventogenic enzyme functions. The existence of a common regulatory factor for acetone and butanol formation has been proposed (14). Chemical mutagenesis of ATCC 824 has resulted in a mutant, M5, which lacks BYDH (NAD and NADP specific) and CoA-transferase activities but expresses both NADH- and NADPH-specific BDH activities (7). The presence of the BDH isozymes (BDH I and BDH II) in ATCC 824 (52) in a different region on the chromosome could perhaps explain the BDH activities detected in mutant M5.

ACKNOWLEDGMENTS

We thank K. Walter and L. Mermelstein for advice and technical assistance, J. Cary for isolation of the clone that was sequenced, and N. Welker and M. Heaton for helpful suggestions.

This work was supported by grants BCS-9210108 and BCS-9209627 from the National Science Foundation.

REFERENCES

1. **Bilofsky, H. S., and C. Burks.** 1988. The GenBank® genetic sequence data bank. Nucleic Acids Res. **16**:1861–1863.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
3. **Boyd, L. A., L. Adam, L. E. Pelcher, A. McHughen, R. Hirji, and G. Selvaraj.** 1991. Characterization of an *Escherichia coli* gene encoding betaine aldehyde dehydrogenase (BADH): structural similarity to mammalian ALDHs and a plant BADH. Gene **103**:45–52.
4. **Cary, J. W., D. J. Petersen, E. T. Papoutsakis, and G. N. Bennett.** 1990. Cloning and expression of *Clostridium acetobutylicum* ATCC 824 acetoacetyl-coenzyme A:acetate/butyrate:coenzyme A-transferase in *Escherichia coli*. Appl. Environ. Microbiol. **56**:1576–1583.
5. **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.
6. **Clark, D. P.** 1989. The fermentation pathways of *Escherichia coli*. FEMS Microbiol. Rev. **63**:223–234.
7. **Clark, S. W., G. N. Bennett, and F. B. Rudolph.** 1989. Isolation and characterization of mutants of *Clostridium acetobutylicum* ATCC 824 deficient in acetoacetyl-coenzyme A:acetate/butyrate:coenzyme A-transferase (EC 2.8.3.9) and in other solvent pathway enzymes. Appl. Environ. Microbiol. **55**:970–976.
8. **Conway, T., G. W. Sewell, Y. A. Osman, and L. O. Ingram.** 1987. Cloning and sequencing of the alcohol dehydrogenase II gene from *Zymomonas mobilis*. J. Bacteriol. **169**:2591–2597.
9. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.

10. **De Vries, G. E., N. Arfman, P. Terpstra, and L. Dijkhuizen.** 1992. Cloning, expression, and sequence analysis of the *Bacillus methanolicus* C1 methanol dehydrogenase gene. J. Bacteriol. **174:**5346–5353.
11. **Dower, W. J., J. F. Miller, and C. W. Ragsdale.** 1988. High efficiency transformation of *Escherichia coli* by high voltage electroporation. Nucleic Acids Res. **16:**6127–6145.
12. **Gerischer, U., and P. Dürre.** 1992. mRNA analysis of the *adc* gene region of *Clostridium acetobutylicum* during the shift to solventogenesis. J. Bacteriol. **174:**426–433.
13. **Goodlove, P. E., P. R. Cunningham, J. Parker, and D. P. Clark.** 1989. Cloning and sequence analysis of the fermentative alcohol-dehydrogenase-encoding gene of *Escherichia coli*. Gene **85:**209–214.
14. **Grupe, H., and G. Gottschalk.** 1992. Physiological events in *Clostridium acetobutylicum* during the shift from acidogenesis to solventogenesis in continuous culture and presentation of a model for shift induction. Appl. Environ. Microbiol. **58:**3896–3902.
15. **Guan, K., and H. Weiner.** 1990. Sequence of the precursor of bovine liver mitochondrial aldehyde dehydrogenase as determined from its cDNA, its gene, and its functionality. Arch. Biochem. Biophys. **277:**351–360.
16. **Guerrero, F. D., J. T. Jones, and J. E. Mullet.** 1990. Turgor-responsive gene transcription and RNA levels increase rapidly when pea shoots are wilted: sequence and expression of three inducible genes. Plant Mol. Biol. **15:**11–26.
17. **Heim, R., and E. E. Strehler.** 1991. Cloning an *Escherichia coli* gene encoding a protein remarkably similar to mammalian aldehyde dehydrogenases. Gene **99:**15–23.
18. **Hidalgo, E., Y.-M. Chen, E. C. C. Lin, and J. Aguilar.** 1991. Molecular cloning and DNA sequencing of the *Escherichia coli* K-12 *ald* gene encoding aldehyde dehydrogenase. J. Bacteriol. **173:**6118–6123.
19. **Hüsemann, M. H. W., and E. T. Papoutsakis.** 1989. Enzymes limiting butanol and acetone formation in continuous and batch cultures of *Clostridium acetobutylicum*. Appl. Microbiol. Biotechnol. **31:**435–444.
20. **Jörnvall, H., B. Persson, and J. Jeffery.** 1987. Characteristics of alcohol/polyol dehydrogenases: the zinc-containing long-chain alcohol dehydrogenases. Eur. J. Biochem. **167:**195–201.
21. **Kessler, D., I. Leibrecht, and J. Knappe.** 1991. Pyruvate-formate-lyase-deactivase and acetyl-CoA reductase activities of *Escherichia coli* reside on a polymeric protein particle encoded by *adhE*. FEBS Lett. **281:**59–63.
22. **Knappe, J., and G. Sawers.** 1990. A radical-chemical route to acetyl-CoA: the anaerobically induced pyruvate formate-lyase system of *Escherichia coli*. FEMS Microbiol. Rev. **75:**383–398.
23. **Kok, M., R. Oldenhuis, M. P. G. van der Linden, C. H. C. Meulenberg, J. Kingma, and B. Witholt.** 1989. The *Pseudomonas oleovorans alkBAC* operon encodes two structurally related rubredoxins and an aldehyde dehydrogenase. J. Biol. Chem. **264:**5442–5451.
24. **Krauth-Siegel, R. L., R. Blatterspiel, M. Saleh, E. Schiltz, R. H. Schirmer, and R. Untucht-Grau.** 1982. Glutathione reductase from human erythrocytes: the sequences of the NADPH domain and of the interface domain. Eur. J. Biochem. **121:**259–267.
25. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
26. **Leonardo, M. R., P. R. Cunningham, and D. P. Clark.** 1993. Anaerobic regulation of the *adhE* gene, encoding the fermentative alcohol dehydrogenase of *Escherichia coli*. J. Bacteriol. **175:**870–878.
27. **Lurz, R., F. Mayer, and G. Gottschalk.** 1979. Electron microscopic study on the quaternary structure of the isolated particulate alcohol-acetaldehyde dehydrogenase complex and on its identity with the polygonal bodies of *Clostridium kluyveri*. Arch. Microbiol. **120:**255–262.
28. **Mahler, I., and H. O. Halvorson.** 1980. Two erythromycin-resistance plasmids of diverse origin and their effect on sporulation in *Bacillus subtilis*. J. Gen. Microbiol. **120:**259–263.
29. **McLaughlin, J. K., C. L. Meyer, and E. T. Papoutsakis.** 1984. Gas chromatography and gateway sensors for on-line state estimation of complex fermentations (butanol/acetone fermentation). Biotechnol. Bioeng. **27:**1246–1257.
30. **Mermelstein, L. D., and E. T. Papoutsakis.** 1993. In vivo methylation in *Escherichia coli* by *Bacillus subtilis* phage ϕ3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. Appl. Environ. Microbiol. **59:**1077–1081.
31. **Mermelstein, L. D., E. T. Papoutsakis, D. J. Petersen, and G. N. Bennett.** 1993. Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for increased solvent production by enhancement of acetone formation enzyme activities using a synthetic acetone operon. Biotechnol. Bioeng. **42:**1053–1060.
32. **Mermelstein, L. D., N. E. Welker, G. N. Bennett, and E. T. Papoutsakis.** 1992. Expression of cloned homologous fermentative genes in *Clostridium acetobutylicum* ATCC 824. Bio/Technology **10:**190–195.
32a. **Nair, R. V., J. W. Cary, G. N. Bennett, and E. T. Papoutsakis.** 1993. Molecular characterization of an alcohol/aldehyde dehydrogenase gene of *Clostridium acetobutylicum* ATCC 824, abstr. O-65, p. 330. Abstr. 93rd Gen. Meet. Am. Soc. Microbiol. 1993.
33. **Narberhaus, F., K. Giebeler, and H. Bahl.** 1992. Molecular characterization of the *dnaK* gene region of *Clostridium acetobutylicum*, including *grpE*, *dnaJ*, and a new heat shock gene. J. Bacteriol. **174:**3290–3299.
34. **Niegemann, E., A. Schulz, and K. Bartsch.** 1992. SwissProt accession no. P25526.
35. **Palosaari, N. R., and P. Rogers.** 1988. Purification and properties of the inducible coenzyme A-linked butyraldehyde dehydrogenase from *Clostridium acetobutylicum*. J. Bacteriol. **170:**2971–2976.
36. **Papoutsakis, E. T., and G. N. Bennett.** 1993. Cloning, structure, and expression of acid and solvent pathway genes of *Clostridium acetobutylicum*, p. 157–199. *In* D. R. Woods (ed.), The clostridia and biotechnology. Butterworth-Heinemann, Stoneham, Mass.
37. **Parsot, C., and J. J. Mekalanos.** 1991. Expression of the *Vibrio cholerae* gene encoding aldehyde dehydrogenase is under control of ToxR, the cholera toxin transcriptional activator. J. Bacteriol. **173:**2842–2851.
38. **Petersen, D. J., J. W. Cary, J. Vanderleyden, and G. N. Bennett.** 1993. Sequence and arrangement of genes encoding enzymes of the acetone-production pathway of *Clostridium acetobutylicum* ATCC 824. Gene **123:**93–97.
39. **Pickett, M., D. I. Gwynne, F. P. Buxton, R. Elliott, R. W. Davies, R. A. Lockington, C. Scazzocchio, and H. M. Sealy-Lewis.** 1987. Cloning and characterization of the *aldA* gene of *Aspergillus nidulans*. Gene **51:**217–226.
40. **Rongnoparut, P., and S. Weaver.** 1991. Isolation and characterization of a cytosolic aldehyde dehydrogenase-encoding cDNA from mouse liver. Gene **101:**261–265.
41. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
42. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
43. **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.
44. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71:**1342–1346.
45. **Smith, L. T., and N. O. Kaplan.** 1980. Purification, properties, and kinetic mechanism of coenzyme A-linked aldehyde dehydrogenase from *Clostridium kluyveri*. Arch. Biochem. Biophys. **203:**663–675.
46. **Söhling, B., and G. Gottschalk.** 1993. Purification and characterization of a coenzyme-A-dependent succinate-semialdehyde dehydrogenase from *Clostridium kluyveri*. Eur. J. Biochem. **212:**121–127.
47. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.
48. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of

cloned genes. J. Mol. Biol. **189:**113–130.
49. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185:**60–89.
50. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82:**1074–1078.
51. **Walter, K. A. (Northwestern University).** Personal communication.
52. **Walter, K. A., G. N. Bennett, and E. T. Papoutsakis.** 1992. Molecular characterization of two *Clostridium acetobutylicum* ATCC 824 butanol dehydrogenase isozyme genes. J. Bacteriol. **174:**7149–7158.
52a. **Walter, K. A., L. D. Mermelstein, G. N. Bennett, and E. T. Papoutsakis.** Unpublished data.
53. **Welch, R. W., S. W. Clark, G. N. Bennett, and F. B. Rudolph.** 1992. Effects of rifampicin and chloramphenicol on product and enzyme levels of the acid- and solvent-producing pathways of *Clostridium acetobutylicum* (ATCC 824). Enzyme Microb. Technol. **14:**277–283.
54. **Welch, R. W., F. B. Rudolph, and E. T. Papoutsakis.** 1989. Purification and characterization of the NADH-dependent butanol dehydrogenase from *Clostridium acetobutylicum* (ATCC 824). Arch. Biochem. Biophys. **273:**309–318.
55. **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.
56. **Wiesenborn, D. P., F. B. Rudolph, and E. T. Papoutsakis.** 1988. Thiolase from *Clostridium acetobutylicum* ATCC 824 and its role in the synthesis of acids and solvents. Appl. Environ. Microbiol. **54:**2717–2722.
57. **Williamson, V. M., and C. E. Paquin.** 1987. Homology of *Saccharomyces cerevisiae* ADH4 to an iron-activated alcohol dehydrogenase from *Zymomonas mobilis*. Mol. Gen. Genet. **209:**374–381.
58. **Wood, N. P., and K. Jungermann.** 1972. Inactivation of the pyruvate formate lyase of *Clostridium butyricum*. FEBS Lett. **27:**49–52.
59. **Wu, L., and N. E. Welker.** 1991. Cloning and characterization of a glutamine transport operon of *Bacillus stearothermophilus* NUB36: effect of temperature on regulation of transcription. J. Bacteriol. **173:**4877–4888.
60. **Yan, R.-T., and J.-S. Chen.** 1990. Coenzyme A-acylating aldehyde dehydrogenase from *Clostridium beijerinckii* NRRL B592. Appl. Environ. Microbiol. **56:**2591–2599.
61. **Young, M., N. P. Minton, and W. L. Staudenbauer.** 1989. Recent advances in the genetics of the clostridia. FEMS Microbiol. Rev. **63:**301–326.
62. **Youngleson, J. S., W. A. Jones, D. T. Jones, and D. R. Woods.** 1989. Molecular analysis and nucleotide sequence of the *adh1* gene encoding an NADPH-dependent butanol dehydrogenase in the gram-positive anaerobe *Clostridium acetobutylicum*. Gene **78:**355–364.
63. **Zuker, M., and P. Stiegler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. **9:**133–148.