Cloning, Expression, and Sequence Analysis of the Bacillus methanolicus C1 Methanol Dehydrogenase Gene

de Vries, Gerhard; Arfman, Nico; Terpstra, Peter; Dijkhuizen, Lubbert

Published in:
Journal of Bacteriology

Publication date:
1992

JOURNAL OF BACTERIOLOGY, Aug. 1992, p. 5346–5353
0021-9193/92/165346-08$02.00/0

Vol. 174, No. 16

Cloning, Expression, and Sequence Analysis of the *Bacillus methanolicus* C1 Methanol Dehydrogenase Gene

G. E. DE VRIES,[1] N. ARFMAN,[1] P. TERPSTRA,[2] AND L. DIJKHUIZEN[1]*

Department of Microbiology, University of Groningen, NL-9751 NN Haren,[1] and Department of Chemistry, University of Groningen, NL-9747 AG Groningen,[2] The Netherlands

Received 21 May 1992/Accepted 16 June 1992

The gene (*mdh*) coding for methanol dehydrogenase (MDH) of thermotolerant, methylotroph *Bacillus methanolicus* C1 has been cloned and sequenced. The deduced amino acid sequence of the *mdh* gene exhibited similarity to those of five other alcohol dehydrogenase (type III) enzymes, which are distinct from the long-chain zinc-containing (type I) or short-chain zinc-lacking (type II) enzymes. Highly efficient expression of the *mdh* gene in *Escherichia coli* was probably driven from its own promoter sequence. After purification of MDH from *E. coli*, the kinetic and biochemical properties of the enzyme were investigated. The physiological effect of MDH synthesis in *E. coli* and the role of conserved sequence patterns in type III alcohol dehydrogenases have been analyzed and are discussed.

Bacillus methanolicus C1 is a representative of the thermotolerant methanol-utilizing *Bacillus* spp. which oxidize methanol by use of an NAD-dependent methanol dehydrogenase (MDH) (3, 5, 9). The previously reported N-terminal amino acid sequence (27) has shown that the enzyme belongs to a novel family of NAD-dependent alcohol dehydrogenases (ADHs) which includes *Zymomonas mobilis* ADH2 (8), *Saccharomyces cerevisiae* ADH4 (29), *Escherichia coli* L-1,2-propanediol oxidoreductase (POR) (7), and the *Clostridium acetobutylicum* NADPH-dependent butanol dehydrogenase (ADH1) (31). This family of enzymes is different from the horse liver-type ADHs (long chain, type I) or *Drosophila*-type enzymes (short chain, type II) and will therefore be referred to as type III ADHs.

The *B. methanolicus* MDH is a decameric enzyme with a subunit M_r of 43,000. Each subunit contains a noncovalently but tightly bound NAD(H) cofactor molecule, one zinc atom, and one or two magnesium atoms (27). The *S. cerevisiae* ADH4 enzyme also has been reported to contain zinc (29), while iron has been detected in the *Z. mobilis* ADH2 protein (8). The *E. coli* POR enzyme is stimulated by ferrous ions (25). Further data about the presence of metals in type III ADHs is lacking, and it remains to be decided whether the presence of magnesium in the *Bacillus* MDH is unique.

The recently reported involvement of an activator protein, which stimulates NAD-dependent ADH activity of the purified MDH enzyme, certainly seems unique (4). In this respect, it is significant that in the other members of the type III ADH enzymes, the oxidative reaction does not play a metabolic role: all support a fermentative metabolic function in the respective wild-type organisms. Metabolic dependence on methanol requires a high flux of methanol oxidation: approximately twice as much carbon must be processed by this reaction compared with ethanol utilization. The activator protein has been suggested to play a dedicated role in the release of reducing equivalents from the bound NAD(H) cofactor (1). The occurrence of the activator protein in all methylotrophic *Bacillus* species investigated (4, 5) may thus form an interesting adaptation to increase the reaction rate of NAD-dependent ADHs in the oxidative direction.

In this report, the cloning and nucleotide sequence analysis of the *B. methanolicus* C1 *mdh* gene are presented. The deduced amino acid sequence was found to share significant identity with the other known members of type III ADHs. The C-terminal part of the bifunctional *E. coli* fermentative ADH (10) could also be incorporated in this cluster. Sequence patterns, which may be related to the distinctive properties of these enzymes, such as NAD(P) and metal binding, have been analyzed and are discussed. Expression and properties of the *B. methanolicus* MDH in the *E. coli* host were investigated.

MATERIALS AND METHODS

Bacterial strains, plasmids, and culture media. *B. methanolicus* C1 cells, grown in liquid medium containing 1% tryptone, 0.5% yeast extract (autoclaved separately), and 1 mM $CaCO_3$, were used as the source of genomic DNA for the cloning procedures. *E. coli* MC1061 (18) served as host strain and was grown in Luria broth (LB) or M9 medium (18) supplemented with the carbon sources indicated. Ampicillin (100 $\mu g\ ml^{-1}$) was added if appropriate. Plasmid pBS^+ (Stratagene, La Jolla, Calif.) was used as the cloning vector.

DNA manipulations, sequencing, and computer programs. Methods for DNA handling, modification, cloning, and transformation of *E. coli* were used as previously documented (18). All DNA-modifying enzymes were used according to the recommendations of the suppliers (Biolabs, Bishop's Stortford, Hertfordshire, United Kingdom; Boehringer, Mannheim, Germany; Pharmacia, Uppsala, Sweden). DNA sequencing was carried out by the chain termination technique (23) on supercoiled plasmid DNA (32). The necessary sequence primers were synthesized according to the information obtained in prior sequence runs. The sequence determination was started from the plasmid fusion sites by using the commercially available universal and reverse primers. Custom primers were synthesized by Eurosequence (Groningen, The Netherlands). DNA and protein sequences were analyzed by a range of computer programs included in the PCGENE (University of Geneva, Geneva, Switzerland) and University of Wisconsin Genetics

* Corresponding author.

Computer Group (Madison, Wis.) packages in addition to the separate programs CLUSTAL (13), PATSCAN (Leunissen, CAOS, Nijmegen, The Netherlands), ISEARCH (National Biomedical Research Foundation, PIR), and CLONEMANAGER (SES, State Line, Pennsylvania).

B. methanolicus C1 DNA was isolated from 250-ml cultures. The cells were harvested and incubated with 25 mg of lysozyme in 25 ml of 10 mM Tris hydrochloride (pH 8.0)–0.4 M sucrose–1 mM $MgCl_2$ until spheroplasts were formed. Following the addition of 2.5 ml of 0.2% Triton X-100 and 1 mg of proteinase K, cells were lysed during incubation for 30 min at 37°C. Genomic DNA was purified by phenol and chloroform extractions, followed by ethanol precipitation. The genomic clone bank was constructed from partially *Sau*3AI-digested *B. methanolicus* C1 DNA, which was sized on a sucrose gradient. The 3- to 5-kb fractions were pooled, ligated to *Bam*HI-digested pBS$^+$ plasmid DNA, and transformed to *E. coli* MC1061. A *B. methanolicus* DNA clone bank in plasmid pBS$^+$ was prepared, with cesium chloride-gradient purification, from 30,000 *E. coli* MC1061 clones grown in batch culture after their initial selection on plates. Transformation of *E. coli* MC1061 with the clone bank DNA was followed by gene selection experiments by immunological screening.

Immunological screening. Fresh *E. coli* MC1061 transformants were replica plated onto nylon membranes (Zeta Probe; Bio-Rad Laboratories, Richmond, Calif.) and incubated on fresh LB agar medium with isopropyl-β-D-thiogalactoside (IPTG, 0.1 mM) and ampicillin for 3 h. Cells were lysed by incubation for 60 min in chloroform vapor and subsequent incubation on Whatman 3MM paper, which was soaked with 50 mM Tris hydrochloride (pH 7.0)–150 mM NaCl–5 mM $MgSO_4$–0.1 mg of lysozyme ml^{-1}–5 μg of DNase I ml^{-1}–1% bovine serum albumin (BSA) (16 h at 37°C). The nylon membranes were intensively washed with tap water, and blocking was performed with 10% skim milk in 10 mM Tris hydrochloride (pH 8.0)–150 mM NaCl–25 mM EDTA at 4°C. Subsequent immunodetection of *E. coli* clones which produced *B. methanolicus* MDH was performed according to standard techniques (12). *B. methanolicus* MDH rabbit antiserum was prepared by using a previously purified MDH preparation (5). An alkaline phosphatase-coupled goat immunoglobulin G (IgG) fraction directed against rabbit IgG (Promega, Madison, Wis.) was utilized according to the instructions of the distributor.

MDH purification. Purification of MDH from *B. methanolicus* was performed as described previously (5). MDH was also purified from the *E. coli* MC1061/pDV72 and was termed cMDH. Although a different order of purification steps was used, the purification scheme for cMDH was basically similar to the procedure described for MDH. A 1 M ammonium sulphate precipitation step, a phenyl-Sepharose hydrophobic interaction chromatography step, and a second (95% saturation) ammonium sulphate precipitation were performed. A dialysis step for 3 h (buffer B [5]) preceded the further purification on a Q-Sepharose ion-exchange column. cMDH-containing fractions were pooled, and protein was concentrated in a dialysis bag which was covered with solid polyethylene glycol (PEG 6000). As a final step, 200-μl samples of the concentrated enzyme preparation (2.3 mg of protein ml^{-1}) were applied onto a Superose-12 gel filtration column, equilibrated with 100 mM Tris hydrochloride (pH 7.5)–5 mM $MgSO_4$–5 mM 2-mercaptoethanol (0.5 ml min^{-1}). The molecular weight of native cMDH was estimated by using gel filtration standards (1,350 to 670,000 range; Bio-Rad) and was compared with that of purified *B. methanolicus* MDH, which was applied in a separate run (approximately 0.5 mg). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (16) was performed with the following calibration proteins with the indicated molecular weights as references: phosphorylase A, 94,000; human transferrin, 80,000; BSA, 68,000; catalase, 58,000; fumarase, 50,000; citrate synthase, 46,000; and carbonic anhydrase, 31,000. Gels were stained with Coomassie brilliant blue G-250. MDH activator protein was purified as described previously (4).

Analytical methods. Spectrophotometric assays were performed with a Hitachi model 100-60 spectrophotometer. All enzyme assays were performed at 50°C, with prewarmed buffer solutions. ADH activities and formaldehyde reductase activities of *Bacillus* MDH and *E. coli* cMDH were measured according to previously published methods (5). Stimulation of cMDH and *Bacillus* MDH activity by purified activator protein was assayed as described previously (4).

The metal compositions of purified cMDH and MDH were determined on a Perkin-Elmer 5100 oven atomic absorption spectrophotometer. The MDH preparations were dialyzed extensively against 10 mM Tris hydrochloride (pH 7.0) containing 1 mM EDTA and subsequently against the same buffer without EDTA. The following elements were analyzed in duplicate: zinc, magnesium, iron, and copper. UV/VIS absorption spectra were measured with a Hewlett-Packard 8452 A photodiode array spectrophotometer.

The N-terminal amino acid sequence of purified cMDH was analyzed as described by Vonck et al. (27), with approximately 1 nmol of cMDH protein subunits. Protein was determined by the method of Bradford (6), with the Bio-Rad protein assay kit with BSA as a standard.

Nucleotide sequence accession number. The sequence data presented in this paper have been submitted to GenBank and assigned accession number M65004.

RESULTS

***B. methanolicus* C1 *mdh* gene cloning.** The plasmid pBS$^+$ (Stratagene), which was used in the construction of the *B. methanolicus* clone bank, has not been designed to serve as a vector for high-level expression of heterologous genes. Still, the *lac* promoter, when induced, may direct gene expression either by a fortuitous in-frame gene fusion or by a similarly fortuitous generation of an α*lacZ*-gene X operon structure. Moreover, sufficient amounts of heterologous gene product will be produced only when the gene of interest is cloned together with an expression signal that functions in *E. coli*.

With these conditions, 50,000 colonies were replica plated on nylon membranes, lysed, and immunologically screened with polyclonal MDH rabbit antiserum. Positive clones were detected by the tetrazolium-dependent enzyme reaction of alkaline phosphatase coupled to mouse anti-rabbit IgG antibodies. A single positive clone (pDV72) was recovered among several initial candidates, all of which showed unstable immunogenic activity during subsequent colony purification steps. Although IPTG was used in the selection procedure to induce the *lacZ* promoter of the cloning vector pBS$^+$, its inclusion in the medium appeared not to be required for the induction of immunogenic activity in the single clone that had been retained. The restriction map obtained after sequencing of the entire 2,500-bp insert corroborated this observation (Fig. 1). A large open reading frame (ORF) is directed in reverse orientation with respect to the *lacZ* promoter. SDS-PAGE of cell extracts from strain

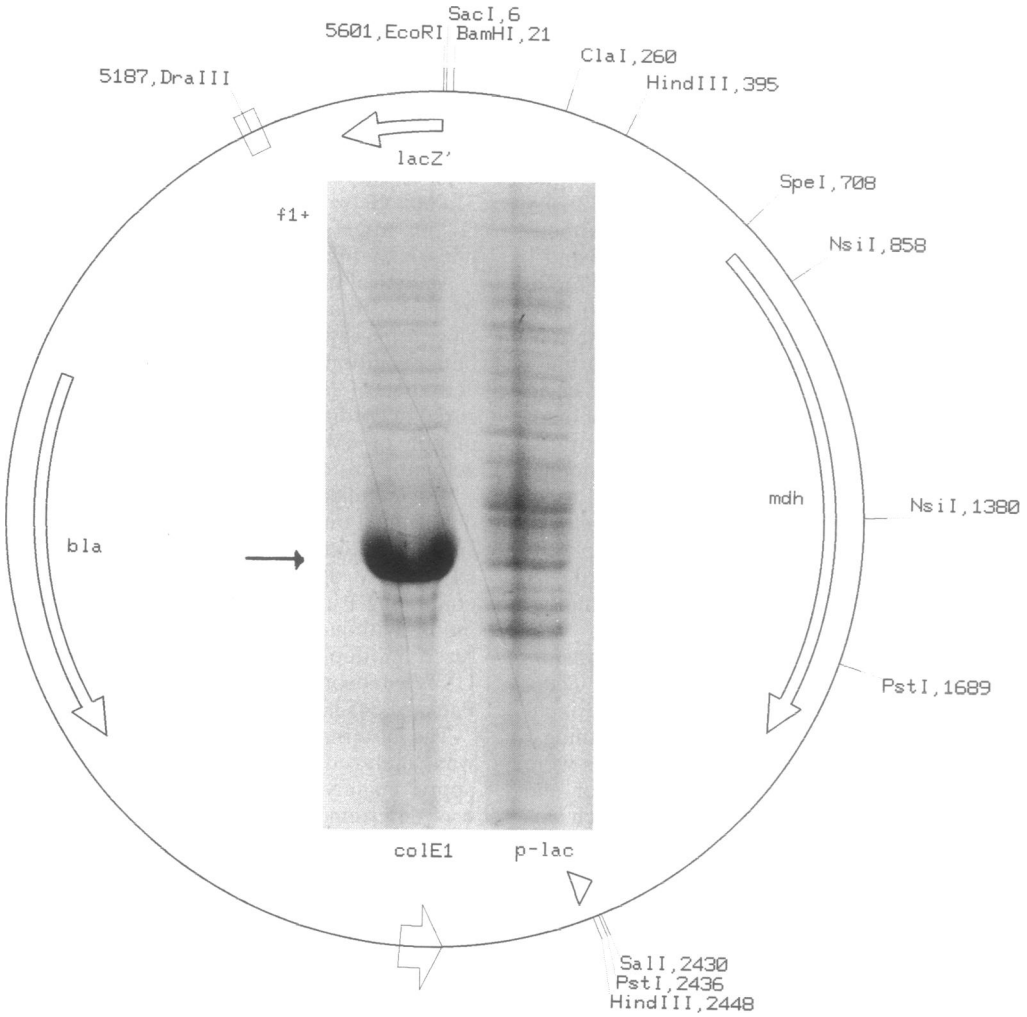

FIG. 1. Restriction map of plasmid pDV72. The SDS-PAGE insert shows the protein profiles of strain MC1061 pDV72 (left lane) and strain MC1061 (right lane) extracts. The black arrow indicates a molecular weight of 43,000.

MC1061 pDV72 showed an extraordinary prominent protein band with an estimated molecular weight that corresponded to the subunit size of *B. methanolicus* MDH protein (Fig. 1, insert). The protein displayed strong immunogenic activity towards the antiserum preparation that had been used in the selection procedure (data not shown).

Sequence determination. The complete 2,500-bp insert in plasmid pDV72 was sequenced in both directions. A single large ORF was found, potentially spanning 385 amino acids. The N-terminal amino acid sequence of the *B. methanolicus* MDH (27) could be located in the determined nucleotide sequence, although some discrepancies were noted. The original data of the N-terminal amino acid sequence determination were therefore reexamined. It was concluded that the inconsistent residues (no. 17, 41, 49, and 58) had been previously misinterpreted, mainly because of decreasing signal-to-noise ratios in the course of the original amino acid sequence analysis (26a). Sequence similarity with the *Z. mobilis* ADH2 protein (8, 30) indicated that the probable start codon is located at the nucleotide which is numbered 1 in Fig. 2, consistent with the N-terminal amino acid sequence of the MDH protein. The amino acid composition of MDH, as determined biochemically (27), compared favorably with the composition which now could be determined from the *mdh* DNA sequence (data not shown).

Several *E. coli* σ^{70} or *Bacillus* σ^{43} consensuslike promoter sequences could be detected in a 400-bp region preceding the *mdh* ORF. A possible promoter with the highest score, calculated according to the method of Harley and Reynolds (11), has its putative −35 region at position −68 with respect to the apparent translational start. The exact start site of mRNA synthesis remains to be determined, however. A sequence that could be the ribosome binding site is located 12 bp upstream of the ATG start codon (Fig. 2). It should be noted that the 3′ end of the *B. methanolicus* 16S rRNA structure was found to be completely identical to the corresponding *Bacillus subtilis* sequence (3). The 5′ upstream region has a very high AT content and displays repetitive A_n and T_n tracts. The existence of such sequences has been implicated to influence curvature of the DNA helix structure (15, 17, 19).

An 18-bp perfect inverted repeat, resembling a factor-independent transcription terminator, was found just downstream of the *mdh* ORF. No indication for the start of another ORF was observed within the following 550-bp sequence (data not shown).

```
-728 GGATCCGTCC ATACCATTCT GAGTCGGATG ATGAATAAGA ATGCTGTCAT CGTTTTCTTT

-668 TATGGAAATG TCGGGTACAT ATACATTAAA AAAGACCGCC GATGATACAA CACCAAACGG

-608 TTATGTAATA AAAGGTTCCC TTAAAAGAGG GATATCAGCT AACCAAAGAA AAGTAAATCC

-548 AGACCTTTAA CGCTTACCGG CTCAAGGATG GATTTACTTT TTTTATGTTA AAGCAATATA

-488 TCGATCTAGC TTAAAAACGA AATAGCGAAC ATCAATGCTT GAAAAACGGT CAGAAGCCGC

-428 TCTCTCTTTC GGGAATACTG CTACCGTTGA GTTTATCTAT ATTTATTGCC TATTCTATTA

-368 CTCTTTTATT ATAAAAGCTT CAAGTTGACC ACAAAAAAAC TTTCATTTTT TAAGACAATT

-308 TGATGCCTTA GGATCTAGTC GATCAGTCGT TTTATGATAT AACTTTAATC CTTTTTACAA

-248 AGTCAAATTT AAATCCTGCC CTTCCACTTT AATCCTCCAC ACTATATAAG CGTCAAAATC

-188 ACTTTTTCTT TGGAAATCTT TTATTAGGTT GTTTTTCCCT TGAGCTTTTT ATATAAAATT

-128 AATATTCAGA TCCTTCTAAA AAATATAATT TCGAAAACTA AGAACATTAC GACAACTATA

 -68 TTGACA AACATTTAGATTAGCATT TAAACT AGTTTTGTAAACAATTACATAAAT AGGAGGT

  -7 AGTAAGA

   1           ATG ACA AAC TTT TTC ATT CCA CCA GCC AGC GTA ATT GGA CGA
  43  GGT GCA GTA AAG GAA GTA GGA ACA AGA CTT AAG CAA ATT GGA GCT AAG
  91  AAA GCG CTT ATC GTT ACA GAT GCA TTT CTT CAT AGC ACA GGT TTA TCT
 139  GAA GAA GTT GCT AAA AAC ATT CGT GAA GCT GGC CTT GAT GTT GCG ATT
 187  TTC CCA AAA GCT CAA CCA GAT CCA GCA GAT ACA CAA GTT CAT GAA GGT
 235  GTA GAT GTA TTC AAA CAA GAA AAC TGT GAT GCA CTT GTT TCT ATC GGT
 283  GGA GGT AGC TCT CAC GAT ACA GCT AAA GCA ATC GGT TTA GTT GCA GCA
 331  AAC GGC GGA AGA ATC AAT GAC TAT CAA GGT GTA AAC AGT GTA GAA AAA
 379  CCA GTC GTT CCA GTA GTT GCA ATC ACT ACA ACA GCT GGT ACT GGT AGT
 427  GAA ACA ACA TCT CTT GCA GTT ATT ACA GAC TCT GCA CGT AAA GTA AAA
 475  ATG CCT GTT ATT GAT GAG AAA ATT ACT CCA ACT GTA GCA ATT GTT GAC
 523  CCA GAA TTA ATG GTG AAA AAA CCA GCT GGA TTA ACA ATC GCA ACT GGT
 571  ATG GAC GCA TTA TCA CAC GCA ATT GAA GCA TAT GTT GCA AAA GGT GCT
 619  ACA CCA GTT ACT GAT GCA TTT GCA ATT CAA GCA ATG AAA CTC ATC AAT
 667  GAA TAC TTA CCA AAA GCG GTG GCA AAC GGA GAA GAC ATC GAA GCA CGT
 715  GAA GCA ATG GCT TAT GCA CAA TAC ATG GCA GGA GTG GCA TTT AAC AAC
 763  GGT GGT TTA GGA TTA GTA CAC TCT ATT TCT CAC CAA GTA GGT GGA GTT
 811  TAC AAA TTA CAA CAC GGA ATC TGT AAC TCA GTT AAT ATG CCA CAC GTT
 859  TGC GCA TTC AAC CTA ATT GCT AAA ACT GAG CGC TTC GCA CAC ATT GCT
 907  GAG CTT TTA GGC GAG AAT GTT TCT GGC TTA AGC ACT GCA GCA GCT GCT
 955  GAG AGA GCA ATT GTA GCG CTT GAA CGC TAT AAC AAA AAC TTC GGT ATC
1003  CCA TCT GGC TAT GCA GAA ATG GGC GTG AAA GAA GAG GAT ATC GAA TTA
1051  TTA GCG AAA AAC GCA TTC GAA GAC GTA TGT ACT CAA AGC AAC CCA CGT
1099  GTT GCT ACA GTT CAA GAC ATT GCA CAA ATC ATC AAA AAC GCT CTG TAA

1147 TCATAATCTT   GGAAAAGCGTCTCTTGAA A TTCAAGAGACGCTTTTCC   GTGATGCTAG
                 >>>>>>>>>>>>>>>>>> . <<<<<<<<<<<<<<<<<<
```

FIG. 2. Nucleotide sequence of the *B. methanolicus* C1 *mdh* gene and upstream promoter sequence. The coding sequence has been given in triplets. >>.<<, inverted repeat; boldface type, A_n and T_n tracts; stippled sequence, possible −35 and −10 promoter regions; italic type, ribosome binding site.

Alignments. The N-terminal amino acid sequence of the *B. methanolicus* MDH protein had previously been shown to be similar to those of four other type III ADH proteins (27). In Fig. 3, it is shown that the similarity of MDH with these proteins extends throughout the entire sequence. In addition, it was found that the C-terminal part of the multifunctional *E. coli adhE* gene product (10) also exhibits significant similarity to this set of sequences and therefore belongs to this family of ADH. Exact overall similarity measures only 15%, although pairwise alignments (Table 1) indicate that similarity between two individual sequences amounts to at least 32% while maximal similarity measures 53%. The program CLUSTAL, which has been used in these studies, indicated the following possible phylogenetic relationship: {[(*Z. mobilis* ADH2—*S. cerevisiae* ADH4)—*B. methanolicus* C1 MDH]—*E. coli* POR}—(*E. coli* ADHE—*C. acetobutylicum* ADH1).

The N-terminal region of the studied sequences (starting at position 13 in the *B. methanolicus* C1 MDH sequence, Fig. 3) is of particular interest. Limited resemblance to a βαβ-dinucleotide-binding fold, which may form a category I nicotinamide coenzyme-binding region as known from other NAD(P)-dependent dehydrogenases (24), was observed. The calculated scores (varying from 6 to 7) for the ADP-binding βαβ-fold fingerprint, as have been proposed by Wierenga et al. (28), seem much too low to conclude beyond a doubt that this region will be involved in NAD(P) binding in this family of ADHs. Therefore, theoretical secondary structures of the peptide sequences involved were determined by using the programs based on the principles described by Novotny and Auffray (20). It was concluded that βαβ-like folds may be formed in the proteins from *S. cerevisiae* and *B. methanolicus* and in the ADHE protein from *E. coli*. In contrast, these theoretical predictions suggest that a leading β-peptide loop

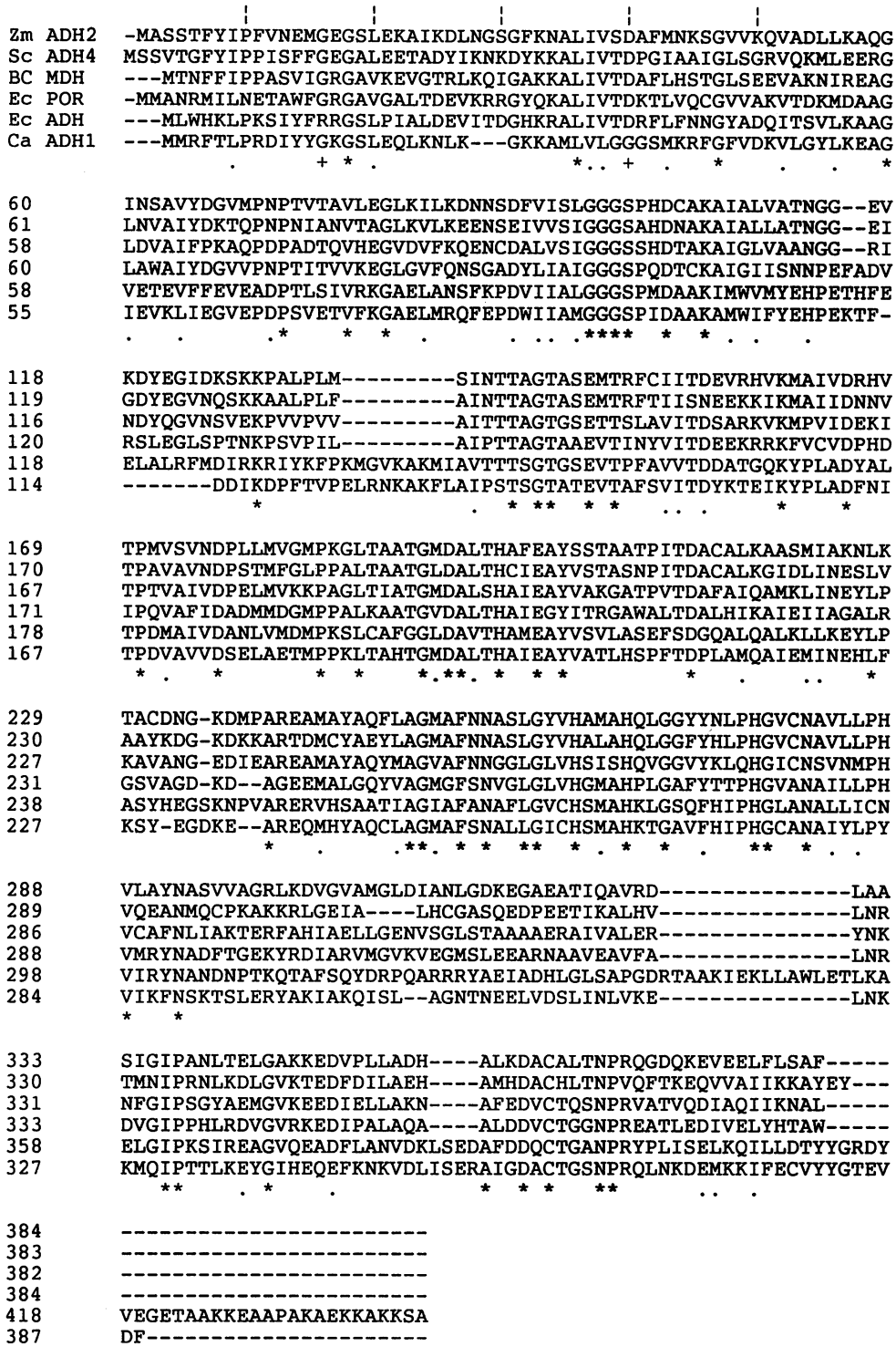

FIG. 3. Alignments of *Z. mobilis ADH2* (8), *S. cerevisiae* ADH4 (29), *B. methanolicus* C1 MDH (this study), *E. coli* POR (7), *E. coli* ADH (C-terminal part of ADHE, starting at residue 449) (10), and *C. acetobutylicum* ADH1 (31). Similarities are shown as determined by the program CLUSTAL (13). Conditions for the calculation were as follows: gap penalty, 4; cutoff, 10; variable gap penalty, F; window, 10; and filtering, 2.5. Symbols: *, exact match across all sequences; ·, conservative matches across all sequences; +, the N-terminal sequence, delimited by this symbol, resembles a βαβ-dinucleotide-binding fold. Amino acid residue numbering is shown at the start of each line.

TABLE 1. Individual similarities of type III ADHs[a]

Type III ADH	% Similarity					
	Z. mobilis ADH2	*S. cerevisiae* ADH4	*B. methanolicus* C1 MDH	*E. coli* POR	*E. coli* ADH[b]	*C. acetobutylicum* ADH1
Z. mobilis ADH2		69	63	61	51	54
S. cerevisiae ADH4	53		61	58	51	57
B. methanolicus C1 MDH	44	46		58	54	53
E. coli POR	42	39	40		54	56
E. coli ADH	33	32	35	34		61
C. acetobutylicum ADH1	37	37	36	36	46	

[a] The similarity of each pair of sequences was calculated from aligned sequences generated by the program CLUSTAL. In the lower left half of the table, absolute values are given as a percentage of the shortest sequence in each pair. In the upper right half of the table, conserved matches were included in these percentages.
[b] C-terminal part of the bifunctional enzyme ADHE.

may not be formed in the *E. coli* POR and the *Z. mobilis* ADH2 proteins. The *C. acetobutylicum* ADH1 sequence seems quite atypical in the distal portion of the hypothetical fold.

It is remarkable that, although possibly similar NAD-binding folds may be present in the N-terminal region of (most of) these proteins, only a few residues have been fully conserved in this region. This is in contrast to five other regions in these proteins (positions 94 to 103, 137 to 145, 182 to 201, 248 to 279, and 355 to 365 in the *B. methanolicus* C1 MDH sequence, Fig. 3), as observed when aligned by the program CLUSTAL. Each of these regions was used as a probe with the programs PATSCAN and ISEARCH in the SWPROT Release 16 protein data bank to investigate whether similar regions of unrelated proteins would generate some information on the function of these primary sequences. None of the utilized probes resulted in the recognition of entries with significantly positive scores. In effect, relatively simple sequence probes could be designed which would select only the six ADH sequences from the entire protein data bank. Therefore, each of these probes (Table 2) could serve as an identification marker for these type III ADH proteins.

Properties of cloned *B. methanolicus* MDH purified from *E. coli*. The purified cMDH preparation (Table 3) was judged essentially homogeneous on the basis of the results of SDS-PAGE analysis (data not shown). With this preparation, the first 15 residues in the N-terminal amino acid sequence of cMDH were determined to be fully identical to those in the known sequence of purified *Bacillus* MDH (27) and showed that the N-terminal methionine residue was also missing from the cMDH protein. The elution characteristics of cMDH on the phenyl-Sepharose and Q-Sepharose columns, as well as the retention time on the Superose-12 column, were similar to those of the *Bacillus* MDH protein. In addition, the two enzyme preparations comigrated on a denaturing polyacrylamide gel (data not shown). These results indicated that cMDH and *Bacillus* MDH possessed similar native and subunit molecular weights (M_r of 430,000 and 43,000, respectively [27]).

TABLE 2. Unique protein sequences for type III ADHs[a]

Region (positions)	Sequence
94–103	GGGS..D..K
190–201	G.DA..H..E.Y
254–268	N...G..H...H..G
355–365	A..D.C...NP

[a] Sequence probes were used with PATSCAN in the SWPROT Release 16 data bank. Dots indicate nondiscriminating positions in the conserved regions as indicated for the *B. methanolicus* C1 MDH protein sequence.

TABLE 3. Purification of cloned MDH from *E. coli*[a]

Sample	Protein (mg)	Total activity (U)	Sp act (U mg^{-1})	Recovery (%)	Purification (fold)
Crude extract	106	130	1.2	100	1.0
Phenyl-Sepharose	70	122	1.7	94	1.4
Q-Sepharose	16	59	3.6	45	3.0
Superose-12	8	28	3.5	22	2.9

[a] Formaldehyde reductase assay. One unit (U) of activity was defined as the amount of enzyme catalyzing the oxidation of 1 μmol of NADH per min.

The substrate specificities of cMDH and *Bacillus* MDH were also similar, although cMDH displayed a slightly lower relative activity with *n*-butanol as a substrate (data not shown). The specific formaldehyde reductase activity of purified cMDH (3.5 μmol min^{-1} mg of protein^{-1}; Table 3) was considerably lower than the value reported for purified *Bacillus* MDH (19.6 μmol min^{-1} mg of protein^{-1} [5]). Partially purified cMDH (after the phenyl-Sepharose column chromatography step) was stimulated only 1.5- to 2-fold in the presence of an excess amount of purified activator protein isolated from *B. methanolicus* C1. Purified cMDH enzyme preparations could no longer be activated. This is in contrast to results with (partially) purified *B. methanolicus* MDH fractions, which were stimulated approximately eightfold in the presence of a saturating amount of activator protein (4).

Metal analysis of cMDH and *Bacillus* MDH (data not shown) revealed that cMDH contained only 0.3 mol of zinc and 0.3 mol of magnesium per mol of subunits. In contrast, *Bacillus* MDH contained 0.8 mol of zinc and 0.8 mol of magnesium per mol of subunits. Spectrophotometric analysis of purified cMDH revealed an absorption peak in the 300- to 350-nm region. This indicated that cMDH, like *Bacillus* MDH, contained tightly associated NADH. The absorption level of cMDH at 340 nm, however, was considerably lower than the level previously observed for *Bacillus* MDH. This suggested that part of the cMDH enzyme preparation either existed in the oxidized state (containing bound NAD) or lacked a bound coenzyme.

DISCUSSION

The *B. methanolicus* C1 *mdh* gene was cloned by using immunological identification methods and was effectively expressed in *E. coli*. The properties of the enzyme isolated from the *E. coli* clone MC1061 pDV72 (cMDH) were found to be different from those of native *Bacillus* MDH. Its metal content was lower, and the enzyme was partly oxidized or lacked the bound NAD(H) cofactor. Concomitantly, the in vitro stimulatory effect of purified activator protein was lost. The first 15 amino acid residues in the N-terminal sequences

of MDH and cMDH were identical and agreed with the deduced sequence from the single ORF that was found on the 2,500-bp cloned insert. Interestingly, both native MDH and cMDH were found to lack the starting methionine residue.

MDH is highly expressed in *B. methanolicus* as well as in *E. coli*; for both organisms, enzyme levels exceeding 20% of total cellular protein have been observed. In *E. coli*, this is partly due to a gene dosage effect, because a high-copy-number cloning vector was used. The A_n/T_n tracts in the promoter region, the highly conserved ribosome binding site, and a possible 5′ secondary mRNA structure may contribute as well. MDH and activator protein in *B. methanolicus* are coordinately and constitutively expressed (2) and inversely controlled by the growth rate as observed in chemostat cultures. MDH expression is virtually absent from cells grown in the presence of complex medium components such as tryptone or yeast extract. These observations could be reproduced in defined media by the addition of certain single amino acids, notably glycine, threonine, and isoleucine; activator protein was still present under these conditions (data not shown). MDH expression in *B. methanolicus* may therefore be under negative control, exerted by specific amino acids or dissimilatory products. The current understanding of this regulation is too limited to suggest whether regulatory proteins or upstream *mdh* promoter elements could be involved in this process. Some regions with possible secondary structures have been made visible in Fig. 2, and deletion analysis should be performed to demonstrate their potential involvement.

B. methanolicus MDH belongs to a novel family of ADHs including *Z. mobilis* ADH2, *S. cerevisiae* ADH4, *E. coli* POR, and the *C. acetobutylicum* ADH1. This group of enzymes forms a third and separate family of ADHs (type III), distinct from the horse liver-type ADH (type I) and the *Drosophila*-type ADH (type II). Alignment studies indicated that the C-terminal part of the bifunctional *E. coli* fermentative ADH (ADHE) also belongs to the type III enzymes, bringing the current number of members to six. Five regions of remarkable amino acid sequence conservation can be distinguished (Fig. 3). In low-stringency searches in the SWPROT Release 16 protein data bank, none of these regions resulted in the finding of significantly similar regions from nonrelated proteins. Instead, four relatively simple search patterns could be composed (Table 2), each of which could serve as a probe to uniquely find type III ADH entries from the SWPROT data bank.

All six type III ADH proteins require the presence of coenzyme NAD(P) for activity. A primary sequence with resemblance to the proposed fingerprint of a βαβ-dinucleotide-binding fold (28) could be identified in the N-terminal part of the aligned sequences. However, the highly conserved Gly-X-Gly-X-X-Gly sequence, typical for the structure, has been retained fully only in the *E. coli* POR sequence. The second glycine, which seems to form the most important residue, has been implicated in the binding of the dinucleotide without obstruction from an amino acid side chain at this position. This residue indeed has been conserved in all sequences. The first glycine residue, reportedly involved in the forming of a tight first βα turn (24), is present in all sequences, except that of the *E. coli* ADHE. This protein is different from the other type III proteins, since it has been fused at its N terminus with a coenzyme A-dependent acetaldehyde dehydrogenase protein (10). Therefore, the hypothetical βαβ fold is actually not N terminally located in this protein. The third glycine residue, which allows a close interaction between the β strands and the α helix, is present only in the *E. coli* POR. The negatively charged aspartic acid residue (position 37 in *B. methanolicus* C1 MDH) at the C terminus of the second β strand, which may form a hydrogen bond with the 2′-hydroxyl group of the adenine ribose of NAD^+, has been conserved in all studied sequences, except that of the *C. acetobutylicum* ADH1. This enzyme seems quite aberrant in the distal portion of the possible βαβ-fold region, which cannot be explained by its preference for NADPH (24).

The N-terminal βαβ-dinucleotide-binding fold of the type III ADH proteins thus forms an interesting protein domain. At the moment, it is not known whether the tightly associated (noncovalent) NAD(H) cofactor in each MDH subunit of the *B. methanolicus* MDH (27) is located in the discussed binding fold. This certainly would put constraints on the tertiary structure of the fold, with consequences for the primary sequence. Also, the interaction between the *B. methanolicus* activator protein and the MDH enzyme, resulting in reoxidation of cofactor NADH with concomitant reduction of external free coenzyme NAD (1), requires further investigation.

Type I ADH proteins may contain two atoms of zinc per subunit. The catalytic zinc ligands to one histidine and two cysteine residues, and the structural zinc atom, probably only present in enzymes from eukaryotic origin, is bound by four closely spaced cysteine residues (14, 21). These seven residues, which are conserved in most members of type I ADH proteins, could not be found in type III ADH sequences. Many members of a group of DNA-binding proteins have been assumed to contain zinc, on the basis of the presence of a defined number of conserved cysteine and histidine residues in consistently spaced patterns (zinc fingers, zinc clusters, and zinc twists) (26). However, no such structures could be found in type III ADH proteins. The lack of similarity with other known metal-binding proteins precluded the determination of a possible zinc-binding site in the *B. methanolicus* MDH. The histidine, cysteine, glutamic, or aspartic acid residues are, nevertheless, the most likely candidates to be involved in interactions with zinc. X-ray crystallography, nuclear magnetic resonance structure determinations, and site-specific mutagenesis should be used to determine how and which metal atoms are incorporated in type III ADHs.

ACKNOWLEDGMENTS

We thank H. Enequist and B. Kazemier for technical assistance during the sequencing of the *mdh* gene, J. Van Beeumen for the determination of the N-terminal sequence of cMDH, and M. van Geel for the studies on the amino acid regulation of MDH synthesis. The valuable discussions with W. Harder, P. Haima, W. Vrijbloed, and R. van den Bergh are gratefully acknowledged.

This investigation was supported by the Biotechnology Action Programme of the Commission of the European Communities (contract no. BAP-0267-NL) and the Programme Committee for Industrial Biotechnology (PCIB; The Netherlands).

REFERENCES

1. **Arfman, N.** 1991. Ph.D. thesis. University of Groningen, Haren, The Netherlands.
2. **Arfman, N., K. J. de Vries, H. R. Moezelaar, M. M. Attwood, G. K. Robinson, M. van Geel, and L. Dijkhuizen.** 1992. Environmental regulation of alcohol metabolism in thermotolerant methylotrophic *Bacillus* strains. Arch. Microbiol. **157**:272–278.
3. **Arfman, N., L. Dijkhuizen, G. Kirchhof, W. Ludwig, K.-H. Schleifer, E. S. Bulygina, K. M. Chumakov, N. I. Govorukhina, Y. A. Trotsenko, D. White, and R. J. Sharp.** 1992. *Bacillus*

methanolicus sp. nov., a new species of thermotolerant, methanol-utilizing, endospore-forming bacteria. Int. J. Syst. Bacteriol. **42:**439–445.
4. **Arfman, N., J. Van Beeumen, G. E. de Vries, W. Harder, and L. Dijkhuizen.** 1991. Purification and characterization of an activator protein for methanol dehydrogenase from thermotolerant *Bacillus* spp. J. Biol. Chem. **266:**3955–3960.
5. **Arfman, N., E. M. Watling, W. Clement, R. J. van Oosterwijk, G. E. de Vries, W. Harder, M. M. Attwood, and L. Dijkhuizen.** 1989. Methanol metabolism in thermotolerant methylotrophic *Bacillus* strains involving a novel catabolic NAD-dependent methanol dehydrogenase as a key enzyme. Arch. Microbiol. **152:**280–288.
6. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
7. **Conway, T., and L. O. Ingram.** 1989. Similarity of *Escherichia coli* propandiol oxidoreductase (fucO product) and an unusual alcohol dehydrogenase from *Zymomonas mobilis* and *Saccharomyces cerevisiae*. J. Bacteriol. **171:**3754–3759.
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. **Dijkhuizen, L., N. Arfman, M. M. Attwood, A. G. Brooke, W. Harder, and E. M. Watling.** 1988. Isolation and characterization of thermotolerant *Bacillus* strains. FEMS Microbiol. Lett. **52:** 209–214.
10. **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 *E. coli*. Gene **85:**209–214.
11. **Harley, C. B., and R. P. Reynolds.** 1987. Analysis of *E. coli* promoter sequences. Nucleic Acids Res. **15:**2343–2361.
12. **Hawkes, R., E. Niday, and J. Gordon.** 1982. A dot-immunobinding assay for monoclonal and other antibodies. Anal. Biochem. **119:**142–147.
13. **Higgins, D. G., and P. M. Sharp.** 1988. CLUSTALL: a package for performing multiple sequence alignments on a microcomputer. Gene **73:**237–244.
14. **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.
15. **Koo, J. S., H. M. Wu, and D. M. Crothers.** 1986. DNA bending at adenine · thymine tracts. Nature (London) **320:**501–506.
16. **Laemmli, U. K., and K. Favre.** 1973. Maturation of the head of bacteriophage T4. I. DNA packaging events. J. Mol. Biol. **80:**575–599.
17. **Lozinski, T., W. T. Markiewics, T. K. Wyrzykiewicz, and K. L. Wierzchowski.** 1989. Effect of the sequence-dependent structure of the 17bp AT spacer on the strength of consensus-like *E. coli* promoters *in vivo*. Nucleic Acids Res. **17:**3855–3863.
18. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **McAllister, C. F., and E. C. Achtberger.** 1989. Rotational orientation of upstream curved DNA affects promoter function in *Bacillus subtilis*. J. Biol. Chem. **264:**10451–10456.
20. **Novotny, J., and C. Auffray.** 1984. A program for prediction of protein secondary structure from nucleotide sequence data: application to histocompatibility antigens. Nucleic Acids Res. **12:**243–255.
21. **Peretz, M., and Y. Burstein.** 1989. Amino acid sequence of alcohol dehydrogenase from the thermophilic bacterium *Thermoanaerobium brockii*. Biochemistry **28:**6549–6555.
22. **Plaskon, R. R., and R. M. Wartell.** 1987. Sequence distributions associated with DNA curvature are found upstream of strong *E. coli* promoters. Nucleic Acids Res. **15:**785–795.
23. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
24. **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.
25. **Sridhara, S., and T. T. Wu.** 1969. Purification and properties of lactaldehyde dehydrogenase in *E. coli*. J. Biol. Chem. **244:**5233–5238.
26. **Vallee, B. L., J. E. Coleman, and D. S. Auld.** 1991. Zinc fingers, zinc clusters and zinc twists in DNA-binding protein domains. Proc. Natl. Acad. Sci. USA **88:**999–1003.
26a. **Van Beeumen, J.** Personal communication.
27. **Vonck, J., N. Arfman, G. E. de Vries, J. Van Beeumen, E. F. J. van Bruggen, and L. Dijkhuizen.** 1991. Electron microscopic analysis and biochemical characterization of a novel methanol dehydrogenase from the thermotolerant *Bacillus* sp. C1. J. Biol. Chem. **266:**3949–3954.
28. **Wierenga, R. K., P. Terpstra, and W. G. J. Hol.** 1986. Prediction of the occurrence of the ADP-binding βαβ-fold in proteins using an amino acid sequence fingerprint. J. Mol. Biol. **187:**101–107.
29. **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.
30. **Yoon, K. H., and M. Y. Pack.** 1990. Nucleotide sequence of the *Zymomonas mobilis* alcohol dehydrogenase II gene. Nucleic Acids Res. **18:**187.
31. **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.
32. **Zagursky, R. J., K. Baumeister, N. Lomax, and M. L. Berman.** 1985. Rapid and easy sequencing of large linear double stranded and supercoiled DNA. Genet. Anal. Technol. **2:**89–94.