

ORIGINAL PAPER

Bjørnar Synstad · Oddmund Emmerhoff
Reidun Sirevåg

Malate dehydrogenase from the green gliding bacterium *Chloroflexus aurantiacus* is phylogenetically related to lactic dehydrogenases

Received: 24 October 1995 / Accepted: 23 February 1996

Abstract The gene encoding malate dehydrogenase (MDH) from *Chloroflexus aurantiacus* was cloned, sequenced, and analyzed. The *mdh* gene corresponded to a polypeptide of 309 amino acids with a molecular mass of 32,717 Da. The primary structure and the coenzyme-binding domain showed a high degree of similarity to lactate dehydrogenase (LDH), whereas the conserved amino acids that participate in substrate binding were those typical of MDHs. Using PCR techniques, the *mdh* gene was cloned in the expression vector pET11a, and large amounts of active *C. aurantiacus* MDH were produced in *Escherichia coli* after induction with isopropyl β -D-thiogalactoside. The expressed enzyme thus obtained was purified and retained full activity at 55°C. High levels of expression of *mdh* were also observed when the gene and its flanking sequences were cloned into pUC18/19, indicating that the putative σ^{70} promoter sequences found upstream of the *C. aurantiacus mdh* functioned in *E. coli*. When these sequences were deleted, the expression in *E. coli* was reduced dramatically.

Key words *Chloroflexus aurantiacus* · Green bacteria · Phototrophic bacteria · Malate dehydrogenase · Lactate dehydrogenase · *mdh* · Gene expression · σ^{70} Promoter · Thermostable protein

Introduction

The phototrophic, green gliding bacterium *Chloroflexus aurantiacus* is a thermophile found in hot springs in nature. One of the characteristic features of the green gliding bacteria, which they share with the green sulfur bacteria, is the presence in their cells of organelle-like structures,

chlorosomes, where the bulk of the light-harvesting pigment, bacteriochlorophyll *c* or *d*, is organized. Results from 16S rRNA analyses indicate, however, that the green gliding bacteria and the green sulfur bacteria are phylogenetically only very distantly related to each other and to other groups within the Bacteria (Gibson et al. 1985; Woese 1987). The phylogenetic distance between the two groups is reflected in the composition of their reaction centers and in their physiology and metabolism. The green sulfur bacteria are obligately anaerobic and photoautotrophic, whereas the green gliding bacteria are metabolically versatile, capable of photoautotrophic, photoorganotrophic, and chemoorganotrophic growth.

Since the position of the green gliding bacteria the current phylogenetic tree suggests that this phylum diverged very early, more information about the molecular biology of this group might contribute to our understanding of its possible evolutionary origin, and to the origin of important biological processes such as photosynthesis and aerobicity. To obtain such information, we isolated and characterized malate dehydrogenase (MDH; L-malate:NAD oxidoreductase, EC 1.1.1.37) from various green bacteria (Rolstad et al. 1988; Charnock et al. 1992). Information about MDH from a thermophile like *C. aurantiacus* might also contribute to our understanding of thermophilic proteins in general.

MDH catalyzes the interconversion of malate to oxaloacetate and is a member of the NAD-dependent dehydrogenases, a large family of nucleotide-binding enzymes, including lactate dehydrogenase (LDH) and alcohol dehydrogenase (ADH). MDH and LDH from various prokaryotes and eukaryotes have been studied extensively and are multimeric enzymes organized as dimers or tetramers with subunit molecular mass of 30–38 kDa. Each subunit of MDH and LDH contains two structurally and functionally different domains. The dinucleotide binding domain is located in the N-terminal end and contains a specific, highly conserved sequence, the glycine motif. In MDH this motif is GXXGXXG, whereas in LDH it is GXGXXG (Wierenga and Hol 1983; Wierenga et al. 1986). Analysis of the amino acid sequence in the N-terminal end of MDH

B. Synstad · O. Emmerhoff · R. Sirevåg (✉)
Division of Molecular Cell Biology, Department of Biology,
University of Oslo, P.O. Box 1066, Blindern, N-0316 Oslo, Norway
Tel. +47-22854563; Fax +47-22854605
e-mail: reidun.sirevag@bio.uio.no

from *Chlorobium vibrioforme*, *Chlorobium tepidum*, *Hellobacterium gestii*, and *C. aurantiacus* has revealed that all contain the typical LDH glycine motif, GXGXXG, rather than that typical of MDH (Rolstad et al. 1988; Charnock et al. 1992). The domain containing the active site is located near the C-terminal end of the subunit and is conserved only to a slight degree in various MDHs, except for six specific, highly conserved amino acid residues (Arg-102, Arg-109, Asp-168, Arg-171, His-196, and Ala-246) that take part directly in substrate binding and catalysis. These residues are found also in LDH, except that in this enzyme the amino acid residues responsible for the substrate specificity, Arg-102 and Ala-246, are substituted by Gln and Thr, respectively (Wilks et al. 1988).

For further comparative and phylogenetic studies on MDH from green bacteria, it is necessary to have information on their complete amino acid sequences. Here we describe the isolation, cloning, and characterization of the *mdh* gene from *C. aurantiacus*, and its expression in *Escherichia coli*. The deduced amino acid sequence of *C. aurantiacus* MDH is compared to those of other MDHs and of LDHs from various sources.

Materials and methods

Bacterial strains, cloning vectors, and growth conditions

Chloroflexus aurantiacus J-10-fl was grown anaerobically at 55°C in Roux medium (Pierson and Castenholz 1974), with the modification that 10 mM Tris-HCl (pH 8.0) replaced glycyl-glycine buffer. Cultures were illuminated by two 60-W light bulbs at a distance of about 25 cm.

The lambda replacement vector EMBL3 (Frischauf et al. 1983) was used for construction of a genomic library. In vitro packaging extract was prepared from *E. coli* SM10 (Rosenberg et al. 1985), and *E. coli* NM539 (Frischauf et al. 1983) was used for recombinant lambda-phage propagation. DNA fragments were subcloned in pUC18/19 (Yanish-Perron et al. 1985) and transformed (Sambrook et al. 1989) into *E. coli* DH5 α (Hanahan 1983). For overexpression of *C. aurantiacus* MDH in *E. coli* BL21(DE3), the expression vector pET11a was used (Studier and Mofatt 1986).

E. coli BL 21(DE3) was grown in NZCYM medium (Sambrook et al. 1989). All other *E. coli* strains were grown in LB medium, supplemented with 10 mM MgSO₄ for recombinant lambda-phage propagation, and with 100 μ g ampicillin ml⁻¹ for *E. coli* DH5 α and *E. coli* BL21(DE3) transformants.

Preparation and manipulation of DNA

DNA from *C. aurantiacus* was isolated according to the method of Sato and Miura (1963), with the modification that the DNA solution was extracted alternately with phenol and chloroform before RNase and proteinase K treatment. Recombinant lambda-phage DNA was prepared in small scale according to Davis et al. (1986). *E. coli* recombinant plasmids were routinely prepared by the alkaline lysis method (Sambrook et al. 1989). Plasmid DNA for sequencing was isolated and purified with the Magic and Wizard Minipreps DNA Purification System (Promega, Madison, Wisc., USA). DNA fragments from agarose gels were isolated and purified using the GeneClean system (Bio 101, La Jolla, Calif., USA) according to the manufacturer's instructions. Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (Beverly, Mass., USA) and were used as recommended by the supplier.

Construction of genomic library

For construction of a genomic library, DNA was partially digested with *Sau*3AI, and fractionated by NaCl-gradient centrifugation (Kaiser and Murray 1985). DNA fragments (12–20-kb) were pooled and ligated to lambda-EMBL3 vector cut with *Bam*HI and *Eco*RI. The ligated DNA was packaged using *E. coli* SMR10 packaging extract (Rosenberg et al. 1985), and the genomic library was amplified (Sambrook et al. 1989).

Preparation of oligonucleotides and labeling of probes

Based on the previously determined N-terminal amino acid sequence of MDH from *C. aurantiacus* (Rolstad et al. 1988), a 29mer oligonucleotide probe corresponding to Thr-16 to Glu-25 (excluding the third nucleotide in the codon for Glu-25) was synthesized using inosine in most of the wobble positions: 5'-ACIAC-IGCICA(C/T)TGGCTIGCIGCIAA(A/G)GA-3'. The oligonucleotide was radiolabeled with [γ -³²P] ATP (Amersham, Buckinghamshire, UK) using T4 polynucleotide kinase (New England Biolabs). Double-stranded DNA probes were labeled with [α -³²P] dCTP using the nick translation system of Promega.

Southern and plaque hybridization

DNA fragments separated by gel electrophoresis were transferred to Hybond N+ nylon membranes (Amersham) by capillary transfer as instructed by the manufacturer; plaques were lifted using Colony/Plaque Screen hybridization transfer membranes (Dupont, New England Nuclear, Boston, Mass., USA). Conditions for hybridization with the radiolabeled oligonucleotide probe were as described by Heinrich et al. (1993), except that the temperature was 55°C. Membranes were washed twice in 6 \times SSC containing 0.05% N-ethylmaleimide for 15 min at 60°C before being exposed to X-ray film at -80°C. Hybridization with double-stranded DNA probes was performed as recommended by the supplier (Amersham), and the membranes were washed twice in 2 \times SSPE, 0.1% SDS, and once in 1 \times SSPE, 0.1% SDS at 65°C before autoradiography.

DNA sequencing, sequence analysis, and polymerase chain reaction

DNA fragments subcloned in pUC18/19 were sequenced using the dideoxy method (Sanger et al. 1977), T7 DNA polymerase, fluorescein-15-³dATP, the Pharmacia PL Autoread kit (Voss et al. 1992), and an A.L.F. DNA sequencer (Pharmacia). Nucleotide and amino acid sequences were analyzed with the Sequence Analysis Software Package (version 7.0; Genetics Computer Group, Madison, Wisc., USA).

PCR reactions were performed on a Hybaid Thermal reactor (Hybaid Limited, Teddington, Middlesex, UK), using Taq polymerase (Promega) as recommended by the supplier. The amplification consisted of 25 cycles of denaturation at 94°C for 1 min, annealing at 60°C for 1 min, and extension at 74°C for 1 min.

Construction of plasmids

Digestion of *C. aurantiacus* DNA with *Kpn*I, *Acc*I, or *Fsp*I yielded three fragments that hybridized with the oligonucleotide of 8.5, 7, and 2 kb, respectively. Five lambda-clones of the genomic library gave strong hybridization signals with the same oligonucleotide probe.

The 2-kb *Fsp*I fragment was ligated into pUC18, yielding pMDH1, which contained a sequence that encoded an amino acid sequence identical to that previously determined for the N-terminal end of MDH, except that the first base in the N-formylmethionine start codon of *mdh* was lacking. The region upstream of *mdh* was analyzed by isolating the 8.5-kb *Kpn*I fragment from the same lambda-clone and digesting with *Hinc*II. The 540-bp *Fsp*I-*Kpn*I fragment of *mdh* (Fig. 1) was used as a probe to identify a partially overlap-

ping 1.1-kb *HincII* fragment, which was isolated and ligated into pUC18 to generate pMDH2. Plasmid pMDH3, containing the complete *mdh* and its upstream regions, was constructed by ligating the 1.1-kb *HincII* fragment of pMDH2 to the 1.0-kb *HincII*-*EcoRV* fragment of pMDH1 in pUC18.

pETMDH used for overexpression of *mdh* in *E. coli*, was generated by cloning *mdh* in pET11a. In a PCR reaction with pMDH3 as template, the sense primer, (p1, 5'-CCGACTT CATATGCGC-AAGAAGATTAGTATTATC-3') and the antisense primer, (p2, 5'-CCAGTCAATACCGCTAATGGTTGAAAAC-3'), (Fig. 1) were used to amplify the first 560 nucleotides of *mdh* and to produce within the start codon a novel *NdeI* site (underlined in the p1 sequence above). The resulting PCR product was digested with *NdeI* and *KpnI*, and the fragment obtained was, together with the 700-bp *KpnI*-*EcoRV* fragment (which includes the remainder of *mdh* and 350 bp downstream) ligated into pUC19 cut with *NdeI* and *SmaI*. The resulting plasmid was digested with *NdeI* and *BamHI* (*BamHI* site in the polylinker), and the 1,280-bp fragment containing the complete *mdh* gene and the downstream sequences were ligated into pET11a.

pMDH4 (Fig. 4) was constructed by digesting pMDH3 with *ScaI* and *BamHI* (polylinker site) and ligating into pUC18/19. In pMDH4, the *lacZ* promoter precedes the *mdh* gene, whereas in pMDH4r, the *mdh* gene is oriented in the opposite direction of the *lacZ* promoter.

pMDH5 and pMDH5r were made by deleting a region upstream of *mdh*. A sense primer (p3, 5'-GGC GAATTCAGGGA-TGAGTGAGTACGACACC-3') (Fig. 1) with an *EcoRI* tail (underlined), an antisense primer (p2), and pMDH4 as template were used. The resulting product including 65 bp upstream and the first 540 bp of the *mdh* gene, was cut with *EcoRI* and *KpnI*, ligated to the *KpnI*-*EcoRV* fragment (which includes the terminal part of *mdh*), and then ligated into pUC18/19, yielding pMDH5 and pMDH5r, respectively.

Overexpression of *C. aurantiacus mdh* in *E. coli* and purification of MDH

Expression of *mdh* in *E. coli* BL21(DE3)(pETMDH) was induced by isopropyl β -D-thiogalactoside (IPTG, 0.4 mM) when the OD₆₀₀ was 0.6–0.7. Samples were taken at various times and analyzed for enzyme activity and presence of MDH.

C. aurantiacus MDH produced in *E. coli* BL21(pETMDH) induced with IPTG for 180 min was purified. Cells were harvested by centrifugation, washed three times in 20 mM phosphate (pH 7.5) containing 1 mM DTT, and suspended in the same buffer (1 g wet weight of cells per ml). Cells were disrupted by ultrasonic treatment, and after centrifugation at 15,000 \times g for 10 min, the supernatant was centrifuged further at 40,000 \times g for 90 min at 4°C. This extract was incubated for 20 min at 55°C to denature *E. coli* proteins. After centrifugation at 15,000 \times g for 20 min at 4°C, the

clear supernatant was heated to 55°C and applied directly on a Matrix Gel Red A affinity column (Amicon, Lexington, Mass., USA) equilibrated with the same phosphate buffer at 40°C. Following washing with buffer until A₂₈₀ in the eluate was < 0.05, bound MDH was eluted with buffer containing 1 mM NADH and 1 mM oxaloacetate. The fractions containing the highest activity of MDH were analyzed by SDS-PAGE and used for further analysis.

Assay of enzyme activity and determination of protein concentration

E. coli crude extracts were prepared as follows. Samples (15 ml) were centrifuged; cells were washed three times in 90 mM K-phosphate (pH 7.5) and resuspended in 1 ml of the same buffer. After sonication, the cell debris was removed by centrifugation at 13,000 \times g for 20 min at 4°C. The supernatant was the crude extract. *C. aurantiacus* crude extracts were prepared and activity of MDH was measured as described by Rolstad et al. (1988). Protein concentration was determined by the method of Lowry et al. (1951), using bovine serum albumin as standard.

Thermostability (T_{1/2}) of MDH was determined by incubation at 55°C and 65°C. Aliquots of purified enzyme produced in *E. coli* were preincubated at 55°C for 5 min before being added to a five-fold volume of 20 mM phosphate (pH 7.5), preheated to the desired temperature. MDH activity was measured in samples taken at intervals. The possible effect of *C. aurantiacus* extracts on the thermostability of purified MDH was measured at 65°C. In these experiments, care was taken to ensure that the purified MDH accounted for over 90% of the total MDH activity in the mixture.

Gel electrophoresis and N-terminal amino acid sequencing

SDS-PAGE was performed in 10% gels by the method of Laemmli (1970), using low-MW standards (Bio-Rads, Richmond, Calif., USA). After electrophoresis, the gels were stained with Coomassie brilliant blue R-250. The N-terminal amino acid sequence of *C. aurantiacus* MDH produced in *E. coli* was determined as described by Charnock et al. (1992).

Results

Nucleotide sequence and amino acid analysis

Figure 1 shows the region of the *C. aurantiacus* genome containing *mdh* and its neighboring genes and the physical location of the various DNA fragments used to construct the necessary plasmids. The complete nucleotide

Fig. 1 The 3.0-kb region in the *Chloroflexus aurantiacus* genome containing *mdh* and part of *orfA* and *orfB*, showing the restriction sites and fragments used in the plasmid constructs. The boxes (open or filled) represent open reading frames and the arrows indicate direction of transcription. The black boxes represent the polylinker of the pUC18/19 vector. The region corresponding to the DNA probe used to identify the 1.1-kb *HincII* fragment is indicated by a bar, and the primers p1, p2 and p3 are marked

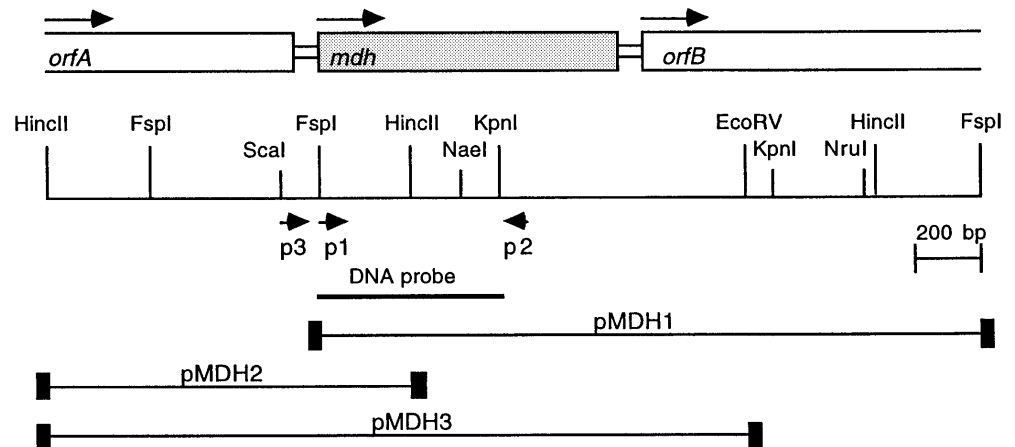


Fig. 2 Alignment of the amino acid sequences of MDH and LDH from various sources. The sequences were aligned using the program PILEUP (Devereux et al. 1984). The MDH sequences are from *Escherichia coli* (MDH Ec; McAlister-Henn et al. 1987), *Haloarcula marismortui* (MDH Hmar; Cendrin et al. 1993), *Chloroflexus aurantiacus* (MDH Caur). The LDH sequences are from *Thermotoga maritima* (LDH Tmar; Ostendorp et al. 1993), *Bacillus stearothermophilus* (LDH Bacst; Barstow et al. 1986); and dogfish (LDHm Squac; Taylor 1977). Conserved amino acids functioning in coenzyme binding, substrate binding, and catalysis are indicated by asterisks and numbered according to Eventoff et al. (1977)

				Glycine motif **** **		Asp 53 *
MDH Ec				MKVAVLGAAG	GIGQALALLL	KTQLPSGSEL SLYDIAPV..
MDH Hmar				M TKVSVVGAAG	TVGAAAGYNI	ALR.DIADEV VFVDIPDKED
MDH Caur				MR KKISIIIG.AG	FVGSTTAHWL	AAK.ELG.DI VLLDIVEGVV
LDH Tmar					MKIGIVG.LG	RVGSSTAFAL LMK.GFAREM VLIDVDKK..
LDH Bacst				MKNNGG	ARVVVIG.AG	FVGASYVFAL MNQ.GIADEI VLIDANES..
LDHm Squac	ATLTKDKLIGH	LATSQEPRSY	NKITVVG.VG	AVGMACAISI	LMK.DLADEV	ALVDVMED..
						Arg 102 *
						Arg 109 *
MDH Ec	.TPGVAVDLS	HIPTAVKIKG	FSGEDATP.A	LEGADVLLIS	AGVRRKPGMD	RSDLFNVNAG
MDH Hmar	DTVGQAADTN	HGIAYDSNT.	RVRQ.GGYED	TAGSDVVVIT	AGIPRQPQGT	RIDLADGNAP
MDH Caur	Q..GKALDLY	EASPIEGFDV	RVTGTNNYAD	TANSDIVVVT	SGAPRKPGMS	REDLIKVNAD
LDH Tmar	RAEGDALDLI	HGTPFTRRA.	NIYA.GDYAD	LKGSDDVIVA	AGVPQKPGET	RLQLLGRNAR
LDH Bacst	KAIGDAMDFN	HGKVFAPKPV	DIWH.GDYDD	CRDADLVVIC	AGANQKPGET	RLDLVDKNIA
LDHm Squac	KLKGEEMDLQ	HGSLFLHTAK	IVSG.KDYSV	SAGSKLVVIT	AGARQQEGES	RLNLVQRNVN
						Asp 168 *
						Arg 171 *
MDH Ec	IVKNLVQQVA	KTCPKACIGI	ITNPVNTTVA	IAAEVLKKG	VYDKNKLLGV	TT.LDIIRSN
MDH Hmar	IMEDIQSSLD	EHNDDYISLT	TSNPVD....	LLNRHLYEAG	DRSREQVIGF	GGRLDSARFR
MDH Caur	ITRACISQAA	PLSPNAVIIIM	VNNPLD....	AMTYLAAEVS	GFPKERVIGQ	AGVLDAAARYR
LDH Tmar	VMKEIARNVS	KYAPDSIVIV	VTNPVD....	VLTYFFLKES	GMDPRKVFSG	GTVLDTARLR
LDH Bacst	IFRSIVESVM	ASGFQGLFLV	ATNPVD....	ILTYATWKF	GLPHERVIGS	GTILDTARFR
LDHm Squac	IFKFIIPNIV	KHSPDCIILV	VSNPVD....	VLTYVAVKLS	GLPMHRIIGS	GCNLD SARFR
				His 195 *		
MDH Ec	TFVAELKQKQ	PGEVEVPVIG	GHSQVTTILPL	LSQ.VPGVSF	TEQEV.....ADLTK
MDH Hmar	YVLSEEFDAP	VQNVEGTILG	EHGDAQVPVF	SKVRVDGT..	...DDPEFSG	D..EKEQLLG
MDH Caur	TFIAMEAGVS	VEDVQAMLMG	GHGDEMVPPL	RFSTISGIPV	SEFI.....	APDRLAQIVE
LDH Tmar	TLIAQHCGFS	PRSVHVYVIG	EHGDSEVPVW	SGAMIGGIPL	QNMCIQKQK	DSKILENFAE
LDH Bacst	FLLGEYFVA	PQNVHAYIIG	EHGDETELPVW	SQAYIGVMPI	RKLVESKGE	AQKDLERIFV
LDHm Squac	YLMGERLGVH	SCSCHGWVIG	EHGDSVPSVW	SGMNVASIKL	HPBELGTNK..	DKQDWKLLHK
				Ala 246 *		
MDH Ec	RIQNAGTEVV	EAKAGGGSAT	LSMQAAARF	GLSLVRALQG	EQGVVECAVY	EGD.GQ.YAR
MDH Hmar	DLQESAMDVI	E...RKGATE	WGPARGVAHM	VEAILH..DT	GEVLPASVKL	EGEFGHED.T
MDH Caur	RTRKGGGEIV	NL.LKTGSAY	YAPAAATAQM	VEAVLK..DK	KRVMPVAAYL	TGQYGLND.I
LDH Tmar	KTKRAAYEII	E...RKGATH	YAIALAVADI	VESIFF..DE	KAVLTLVSVL	EDYLGVKD.L
LDH Bacst	NVRDAAAYQII	E...KKGATY	YGIAMGLARV	TRAILH..NE	NRLLTVSAYL	DGLYGERD.V
LDHm Squac	DVVD SAYEVI	K...LKGYTS	WAIGLSVADL	AETIMK..NL	CRVHPVSTMV	KDFYGIKDNV
MDH Ec	FFSQPLLLGK	NGVEERKSIG	TLSAFEQNAL	EGMLDTLKKD	IALGQEFVNK	
MDH Hmar	AFGVVPRLGS	NGVEEIVEWD	.LDDYEQDLM	ADAAEKLSAQ	YDKIS	
MDH Caur	YFGVPVILGA	GGVEKILELP	.LNNEEMALL	NASAKAVRAT	LDTLKSL	
LDH Tmar	CISVPVTLGK	HGVERILELN	.LNNEELEAF	RKSASILKNA	INEITAEENK	HQNTSG
LDH Bacst	YIGVPAVINR	NGIREVIEIE	.LNDDEKNRF	HHSAAATLKS	LARAFTR	
LDHm Squac	FLSLPCVLND	HGISNIVKMK	.LKPDEEQQL	QKSATTLWDI	QKDLKF	

sequence of the *mdh* gene and its surrounding regions is deposited in the EMBL data library under accession number X89038.

The *mdh* gene contains 930 bp including the stop codon, and encodes the polypeptide of 309 amino acids (Fig. 2). The molecular mass of the polypeptide is 32.7

kDa, which is slightly lower than that of 35 kDa estimated by gel electrophoresis (Rolstad et al. 1988). Between the nucleotides at positions -69 to -114, upstream of the start codon of *mdh*, two putative -10 and -35 promoter sequences (1 and 2 in Fig. 5) were found. The exact start site of transcription remains to be determined. A putative

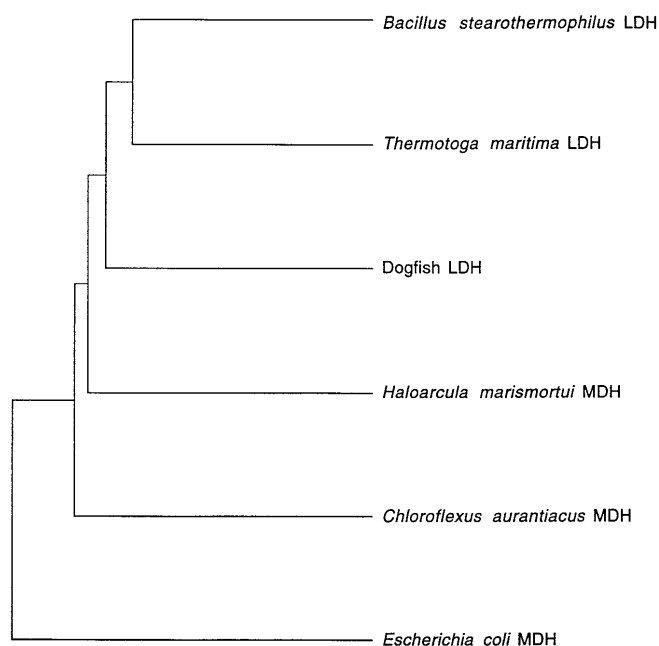


Fig. 3 A dendrogram showing the relationship between various MDHs and LDHs. The dendrogram was constructed on basis of binary alignments of the amino acid sequences shown in Fig. 2., using the program PILEUP (Devereux et al. 1984)

Shine-Dalgarno sequence, AGGAGGAG, ends 7 bp upstream of the start codon (see Fig. 5). An inverted repeat with the potential of forming an 11 bp stem with a 14-bp loop structure starts 6-bp downstream of the *mdh* stop codon (see Fig. 5) and might function in transcript processing or termination.

The open reading frame *orfA* located upstream of *mdh* probably extends beyond the *HincII* site on the chromosome and appears to be transcribed in the same direction as the *mdh* gene. *orfA* ends with a stop codon 87 bp upstream of *mdh* and encodes a putative polypeptide of at least 270 amino acids. *orfA* shows sequence similarity to an ORF of unknown function, *orfW*, in *E. coli* and *Pseudomonas aeruginosa* positioned upstream of the *clpB* gene (Squires et al. 1991; Hobbs et al. 1993). *orfA* from *C. aurantiacus* also shows sequence similarity to another ORF in *E. coli*, *orfX*, located upstream of *ams*, which encodes RNaseE (Claverie-Martin et al. 1991; Mudd et al. 1990).

Another ORF, *orfB*, starts 61 nucleotides downstream of *mdh* and ends outside the 3-kb region. It encodes a putative polypeptide of at least 321 amino acids, but no significant similarity to any other known sequence was found. No putative ribosome binding site was found upstream of the putative start codon of *orfB*.

Results from searches in protein databases (Swiss-Prot and NBRL) using the deduced amino acid sequence of *mdh*, indicated that this sequence was similar to sequences of MDH as well as LDH from various prokaryotes and eukaryotes. An alignment of the MDH sequence of *C. aurantiacus* to MDH from *E. coli* and the Archaeon *Halobacterium marismortui*, and to LDH from *Thermo-*

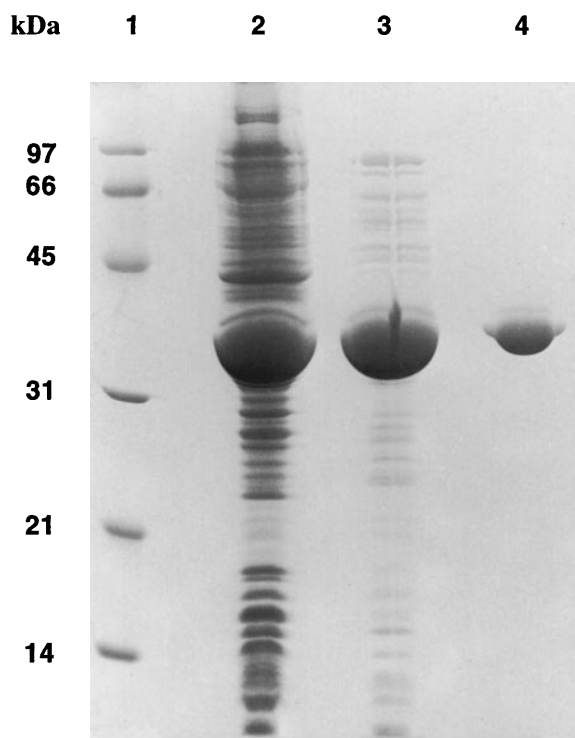


Fig. 4 Overexpression and purification of *Chloroflexus aurantiacus* MDH produced in *Escherichia coli*. Lanes: 1 Size markers, 2 crude extract, 3 crude extract after heat treatment, 4 eluate from a Matrix Gel RedA affinity column

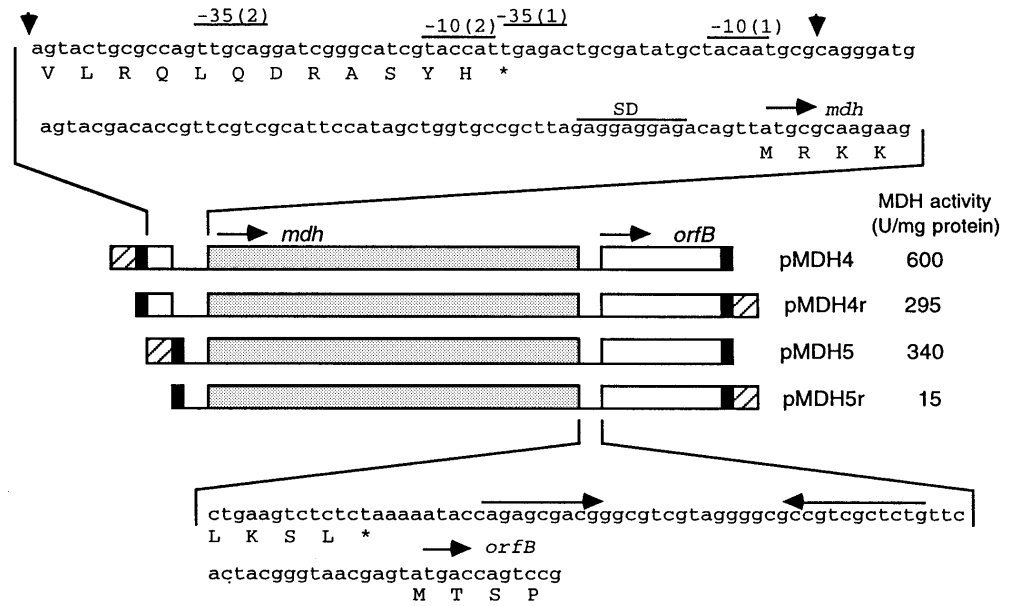
toga maritima, *Bacillus stearothermophilus*, and dogfish is shown in Fig. 2. A general survey of the alignment shows that the primary structure of MDH from *C. aurantiacus* is similar to MDH as well as to LDH. However, the dendrogram presented in Fig. 3 and constructed on the basis of the same amino acid sequences, clearly places MDH from *C. aurantiacus* apart from that of *E. coli*, but together with MDH from *H. marismortui* and with the LDHs.

Expression of the *C. aurantiacus mdh* gene in *E. coli* and purification of MDH

Expression of *mdh* in *E. coli* (pETMDH) after induction with 0.4 mM IPTG led to the overexpressed protein comprising about 60% of the total protein in the extract (Fig. 4, lane 2). MDH had a molecular mass of approximately 35 kDa. The specific activity of MDH in the crude extract was 620 U (mg protein)⁻¹.

C. aurantiacus MDH was more thermostable than the *E. coli* proteins, allowing the use of a simple heat treatment at 55°C as an initial step in the purification of this enzyme from *E. coli*. The specific activity of MDH in the extract after such treatment was 825 U (mg protein)⁻¹ and increased to 1,000 U (mg protein)⁻¹ after passage of the heat-treated extract through the Matrix gel Red A affinity column, with a recovery of 25%. Analysis by SDS-PAGE of the different stages in the purification of MDH are

Fig. 5 Activity of MDH in *Escherichia coli* DH5 α transformed with pUC18/19 plasmids containing the complete *mdh* gene from *Chloroflexus aurantiacus*. The nucleotide sequences upstream and downstream of *mdh* are shown. The sequence deleted to produce pMDH5 and pMDH5r is marked, and the putative -10 and -35 and Shine-Dalgarno (SD) sequences are indicated. A potential stem-loop structure downstream of *mdh* is indicated by converging arrows. The black and hatched boxes represent the polylinker and *lacZ* promoter, respectively



shown in Fig. 4. After affinity chromatography, the sample (lane 4) was almost pure except for a faint band just above the major protein band. Determination of the N-terminal amino acid sequence of the purified *C. aurantiacus* MDH produced in *E. coli* confirmed that it was identical to that reported by Rolstad et al. (1988).

Incubation of the purified MDH at 55°C for 60 min had no effect on the activity of the enzyme, whereas at 65°C, the activity was reduced by 50% after only 9 min. However, addition of crude extract from *C. aurantiacus* to the purified enzyme before incubation at 65°C increased the stability of the enzyme. Under these conditions, the time required to reduce the activity 50% was 24 min.

Analysis of the putative promoter regions of *mdh*

The importance of the regions upstream of *C. aurantiacus mdh* for its expression in *E. coli* is shown in Fig. 5. High specific activity of *C. aurantiacus* MDH was observed in cells of *E. coli* transformed with pMDH4, in which the *lacZ* promoter precedes the *mdh* gene and its upstream region. In pMDH4r, the *mdh* gene and its upstream region are oriented in the opposite direction relative to the *lacZ* promoter; and expression in cells containing this plasmid was approximately 50% lower than in cells containing pMDH4. Since a significant amount of thermostable MDH was produced without the *lacZ* promoter, a *C. aurantiacus* promoter upstream of the *mdh* gene is probably active in *E. coli*. Both pMDH5 and pMDH5r have a deletion that removed the two putative $-10/-35$ regions upstream of *mdh*. In cells containing pMDH5, in which the *lacZ* promoter is located in front of the *mdh* gene, the specific activity of MDH was approximately the same as for those containing pMDH4r. However, when the *mdh* gene with the upstream deletion was oriented in the opposite direction relative to the *lacZ* promoter as in pMDH5r, the specific activity was very low (Fig. 5).

Discussion

Here we report the cloning, sequencing, and characterization of the *mdh* gene from *C. aurantiacus* and its expression in *E. coli*. The transcriptional orientation of the flanking genes, *orfA* and *orfB*, was the same as that of *mdh*. Two potential $-10/-35$ promoter regions upstream of *mdh* were located within the terminal region of *orfA*. The spacing between the -10 and -35 elements in these potential promoters was only 11 and 12 bp, respectively and thus differed significantly from the $-35/N17/-10$ consensus typical of most *E. coli* σ^{70} promoters (Harley and Reynolds 1987). As can be seen from Fig. 5, the region containing these potential promoter elements was present in pMDH4 and pMDH4r, but deleted in pMDH5 and pMDH5r. The high enzyme activity obtained in cells harboring pMDH4 indicates that transcription appears to be initiated from the *lacZ* promoter and from a putative *C. aurantiacus* promoter. The relatively high levels of enzyme activity obtained also when the *mdh* gene and its upstream region were oriented in the opposite direction relative to the *lacZ* promoter (pMDH4r), indicate that the *C. aurantiacus mdh* promoter functions in *E. coli*. In cells harboring pMDH5r, a low specific thermostable MDH activity was still detected in spite of the deletion of the potential promoters, possibly because the deleted DNA region contains enhancer sequences for an unidentified promoter closer to the *mdh* gene.

Analysis of the MDH sequence from *C. aurantiacus* (Fig. 2) identified, in addition to the glycine motif and Asp-53 involved in coenzyme binding, also the six amino acids important in substrate binding and catalysis. In the active site of the enzyme, His-195 and Asp-168 form a proton-relay system and play an important role in the protonating/deprotonating of the substrate by the imidazole ring of His-195 during catalysis (Birktoft and Banaszak 1983). Arg-109 polarizes the carbonyl bond of the substrate

(Clarke et al. 1986), and Arg-171 binds and orients the substrate via the carboxylate group of the substrate (Hart et al. 1987). This overall mechanism in catalysis is similar in MDH and LDH, and these amino acids (asterisks in Fig. 2) are present in both enzymes. The amino acid residues in positions 102 and 246 are responsible for the substrate specificity, and accordingly, they are different in MDH and LDH. LDH has glutamine in position 102 and threonine in position 246, whereas in MDH these positions are occupied by arginine and alanine, respectively. In MDH, the positively charged Arg-102 forms a counterion for the second carboxylate group in malate/oxaloacetate. When this amino acid is replaced by the neutral glutamine in *E. coli* MDH (Nicholls et al. 1992) and in *Haloarcula marismortui* MDH (Cendrin et al. 1993), the substrate specificity of the mutant enzymes is shifted toward pyruvate. Alanine instead of threonine in position 246 in MDH is necessary since a smaller residue is needed to accommodate oxaloacetate, which is larger than pyruvate, the substrate of LDH (Wilks et al. 1988). Despite the relatively high sequence similarity of *C. aurantiacus* MDH to LDHs in general, the presence of Arg-102 and Ala-246 shows that the *mdh* gene reported here encodes an enzyme with the amino acid residues necessary to bind oxaloacetate/malate.

The complete amino acid sequence of MDH has been determined for a large number of prokaryotes and eukaryotes, and the *mdh* gene from four different bacteria has been completely sequenced. The MDH sequences from *E. coli* (Vogel et al. 1987; McAlister-Henn 1987) and *Salmonella typhimurium* (Lu and Abdelul 1993) show relatively high sequence identity with the mitochondrial isozymes of mammalian MDH, whereas the amino acid sequence of MDH from *Thermus flavus* (Nishiyama et al. 1986) and *Thermus aquaticus* B (Nicholls et al. 1990) are more similar to the mammalian cytoplasmic isozyme. From these sequences, it appears that the various bacterial MDHs diverge into two main phylogenetic groups. The search in protein databases using the *C. aurantiacus* MDH sequence indicated that the primary sequence of this enzyme is more similar to LDH from eubacteria and eukaryotes than to other MDHs. When MDH of *C. aurantiacus* was compared to the various LDH sequences available, it showed the highest sequence identity (36%) to *Thermotoga maritima*, a member of the phylum Thermotogales, the earliest branch of the 16S rRNA phylogenetic tree of the Bacteria. This finding is noteworthy since it supports and strengthens the proposal that the green gliding bacteria branched off very early in evolution. Of the MDH sequences, *C. aurantiacus* MDH showed the highest sequence identity to that of the Archaeon *Haloarcula marismortui*. This MDH sequence has previously been shown to be more similar to various LDHs than to other MDHs (Cendrin et al. 1993), and in the phylogenetic tree of MDH and LDH presented by Goward and Nicholls (1994), MDH from *H. marismortui* is grouped together with the LDH sequences. These findings together with results from the dendrogram presented in Fig. 3, strengthen the hypothesis of a link between the evolution of MDH

and LDH and also support the 16S rRNA phylogenetic tree in which the green gliding bacteria branch off early, just after Thermotogales. Recently, the *mdh* genes from two species of green sulfur bacteria, *Chlorobium vibrioforme* and *C. tepidum*, were isolated and sequenced (this lab, unpublished results). The primary structure of these two MDHs showed more than 50% identity with the MDH sequence of *C. aurantiacus*. To make this evolutionary picture more complete, further MDH sequences from different bacterial species representing different phyla need to be obtained and analyzed.

The cloning of the *mdh* gene into the expression vector pET11a and subsequent transformation of *E. coli* BL21(DE3) led to high expression of the gene, as visualized on SDS-PAGE. This system was used to produce large amounts of *C. aurantiacus* MDH in *E. coli*, and the specific activity of this purified MDH was approximately the same as that obtained with MDH purified from *C. aurantiacus* (Rolstad et al. 1988). The purified, plasmid-encoded MDH was stable at 55°C, which is in agreement with the optimum temperature for growth of the organism and for enzyme activity. At 65°C, the thermostability of the enzyme was significantly lower, but addition of cell-free extract from *C. aurantiacus* increased the stability almost threefold, possibly because of factors in the extract that confer thermoprotection on MDH and other proteins. Such thermoprotection is probably more pronounced in vivo and may partly explain why *C. aurantiacus* is able to grow at temperatures up to 70°C. It is well established that many organisms induce the formation of specific proteins, chaperones, in response to environmental stress such as high temperature. The hyperthermophile *Pyrodictium occultum* and the thermophile *Thermoanaerobacter brockii*, for example, produce large amounts of chaperones when grown at the upper temperature border (Phipps et al. 1991; Truscott et al. 1994).

Acknowledgements We would like to thank the following persons at the Biotechnology Center in Oslo: Dr. Tom Kristensen and Dr. Marianne Wright for performing the automated DNA sequencing, Dr. Eshrat Babaie for making the oligonucleotides and DNA primers, and Dr. Knut Sletten for determining the N-terminal amino acid sequence. This research was supported by a grant from the Norwegian Research Council.

References

- Barstow DA, Clarke AR, Chia WN, Wigley D, Sharman AF, Holbrook JJ, Atkinson T, Minton NP (1986) Cloning expression and complete nucleotide sequence of the *Bacillus stearothermophilus* L-lactate dehydrogenase gene. *Gene* 46: 47–56
- Birktoft JJ, Banaszak LJ (1983) The presence of a histidine-aspartic acid pair in active sites of 2-hydroxyacid dehydrogenases. *J Biol Chem* 258: 472–482
- Cendrin F, Chroboczek J, Zaccai G, Eisenberg H, Mevarech M (1993) Cloning, sequencing, and expression in *Escherichia coli* of the gene coding for malate dehydrogenase of the extreme halophilic archaeobacterium *Haloarcula marismortui*. *Biochemistry* 32: 4308–4313

- Charnock B, Refseth UH, Sirevåg R (1992) Malate dehydrogenase from *Chlorobium vibrioforme*, *Chlorobium tepidum*, and *Hellobacterium gestii*: purification, characterization, and investigation of dinucleotide binding by dehydrogenases by use of empirical methods of protein sequence analysis. *J Bacteriol* 174: 1307–1313
- Clarke AR, Wigley DB, Chia WN, Barstow D, Atkinson T, Holbrook JJ (1986) Site-directed mutagenesis reveals the role of mobile arginine residue in lactate dehydrogenase catalysis. *Nature* 324: 699–702
- Claverie-Martin F, Diaz-Torres M, Yancey SD, Kushner SR (1991) Analysis of the altered mRNA stability (*ams*) gene from *E. coli*. *J Biol Chem* 266: 2843–2851
- Davis LG, Dibner MD, Batky JF (1986) Basic methods in molecular biology. Elsevier, Amsterdam New York
- Devereux J, Haerberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 12: 387–395
- Eventoff W, Rossman MG, Taylor SS, Torff HJ, Meyer H, Keil W, Kiltz HH (1977) Structural adaptations of lactate dehydrogenase isozymes. *Proc Natl Acad Sci USA* 74: 2677–2681
- Frischauf A, Lehrach H, Poustka A, Murray N (1983) Lambda replacement vectors carrying polylinker sequences. *J Mol Biol* 170: 827–842
- Gibson J, Ludwig W, Stackebrandt E, Woese CR (1985) The phylogeny of the green photosynthetic bacteria: absence of a close relationship between *Chlorobium* and *Chloroflexus*. *Syst Appl Microbiol* 6: 152–156
- Goward CR, Nicholls DJ (1994) Malate dehydrogenase: a model for structure, evolution, and catalysis. *Protein Sci* 3: 1883–1888
- Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* 166: 557–580
- Harley CB, Reynolds RP (1987) Analysis of *E. coli* promoter sequences. *Nucleic Acids Res* 15: 2343–2361
- Hart KW, Clarke AR, Wigley DB, Chia WN, Barstow DA, Atkinson T, Holbrook JJ (1987) The importance of arginine-171 in substrate binding by *Bacillus stearothermophilus* lactate dehydrogenase. *Biochim Biophys Res Commun* 146: 346–353
- Heinrich JH, Beecher DJ, MacMillen JD, Zilinskas BA (1993) Molecular cloning and characterization of the *hblA* gene encoding the B component of hemolysin BL from *Bacillus cereus*. *J Bacteriol* 175: 6760–6766
- Hobbs M, Collie ESR, Free PD, Livingston SP, Mattick JS (1993) PilS and PilR, a two-component transcriptional regulatory system controlling expression of type 4 fimbriae in *Pseudomonas aeruginosa*. *Mol Microbiol* 7: 669–682
- Kaiser K, Murray N (1985) The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries. In: Glover DM (ed) DNA cloning techniques. A practical approach, vol 1. IRL Press, Oxford, pp 1–44
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of bacteriophage T4. *Nature* 227: 680–685
- Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) Protein measurement with the Folin phenol reagent. *J Biol Chem* 193: 265–275
- Lu CD, Abdelal AT (1993) Complete sequence of the *Salmonella typhimurium* gene encoding malate dehydrogenase. *Gene* 123: 143–144
- McAlister-Henn L, Blaber M, Bradshaw RA, Nisco SJ (1987) Complete nucleotide sequence of the *Escherichia coli* gene encoding malate dehydrogenase. *Nucleic Acids Res* 15: 4993
- Mudd EA, Krisch HM, Higgins CF (1990) RNase E, an endonuclease, has a general role in the chemical decay of *Escherichia coli* mRNA: evidence that *rne* and *ams* are the same genetic locus. *Mol Microbiol* 4: 2127–2135
- Nicholls DJ, Sundaram TK, Atkinson T, Minton NP (1990) Cloning and nucleotide sequence of the *mdh* and *sucD* genes from *Thermus aquaticus* B. *FEMS Microbiol Lett* 70: 7–14
- Nicholls DJ, Miller J, Scawen MD, Clarke AR, Holbrook JJ, Atkinson T, Goward CR (1992) The importance of arginine 102 for the substrate specificity of *Escherichia coli* malate dehydrogenase. *Biochem Biophys Res Commun* 189: 1057–1062
- Nishiyama M, Matsubara N, Yamamoto K, Iijima S, Uozumi T, Beppu T (1986) Nucleotide sequence of the malate dehydrogenase gene of *Thermus flavus* and its mutation directing an increase in enzyme activity. *J Biol Chem* 261: 14178–14183
- Ostendorp R, Liebl W, Schurig H, Jaenicke R (1993) The L-lactate dehydrogenase gene of the hyperthermophilic bacterium *Thermotoga maritima* cloned by complementation in *Escherichia coli*. *Eur J Biochem* 216: 709–715
- Phipps BM, Hoffmann A, Stetter KO, Baumeister W (1991) A novel ATPase complex selectively accumulated upon heat shock is a major cellular component of thermophilic archaeobacteria. *EMBO J* 10: 1711–1722
- Pierson BK, Castenholz RW (1974) A phototrophic gliding filamentous bacterium of hot springs, *Chloroflexus aurantiacus* gen. and sp. nov. *Arch Microbiol* 100: 5–24
- Rolstad AK, Howland E, Sirevåg R (1988) Malate dehydrogenase from the thermophilic green bacterium *Chloroflexus aurantiacus*: purification, molecular weight, amino acid composition, and partial amino acid sequence. *J Bacteriol* 170: 2947–2953
- Rosenberg SM, Stahl MM, Kobayashi I, Stahl FW (1985) Improved in vitro packaging of coliphage lambda DNA. A one-strain system free from endogenous phage. *Gene* 38: 165–175
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
- Sanger F, Nicklen S, Coulson RA (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463–5467
- Sato H, Miura K (1963) Preparation of transforming deoxyribonucleic acid by phenol treatment. *Biochim Biophys Acta* 72: 619–629
- Squires CL, Pedersen S, Ross BM, Squires C (1991) ClpB is the *Escherichia coli* heat shock protein F84.1. *J Bacteriol* 173: 4254–4262
- Studier FW, Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* 189: 113–130
- Taylor SS (1977) Amino acid sequence of dogfish muscle lactate dehydrogenase. *J Biol Chem* 252: 1799–1806
- Truscott KN, Hoj PB, Scopes RK (1994) Purification and characterization of chaperonin 60 and chaperonin 10 from the anaerobic thermophile *Thermoanaerobacter brockii*. *Eur J Biochem* 222: 277–284
- Vogel RF, Entian K-D, Mecke R (1987) Cloning and sequence of the *mdh* structural gene of *Escherichia coli* coding for malate dehydrogenase. *Arch Microbiol* 149: 36–42
- Voss H, Wiemann S, Wirkner U, Schwager C, Zimmermann J, Stegemann J, Erfle H, Hewitt NA, Rupp T, Ansoerge W (1992) Automated DNA sequencing system resolving 1000 bases with fluorescein-15-*dATP as internal label. *Methods Mol Cell Biol* 3: 153–156
- Wierenga RK, Hol GJ (1983) Predicted nucleotide binding properties of p 21 protein and its cancer-associated variant. *Nature* 302: 842
- Wierenga RK, Tepstra P, Hol GJ (1986) Prediction of the occurrence of the ADP-binding fold in proteins using an amino acid sequence fingerprint. *J Mol Biol* 187: 101–107
- Wilks HM, Hart KW, Feeney R, Dunn CR, Muirhead H, Chia WN, Barstow DA, Atkinson T, Clarke AR, Holbrook JJ (1988) A specific, highly active malate dehydrogenase by redesign of a lactate dehydrogenase framework. *Science* 242: 541–544
- Woese CR (1987) Bacterial evolution. *Microbiol Rev* 51: 221–274
- Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33: 103–119