NOTES

Cloning and Expression of the Two Genes Coding for L-Serine Dehydratase from *Peptostreptococcus asaccharolyticus*: Relationship of the Iron-Sulfur Protein to Both L-Serine Dehydratases from *Escherichia coli*

ANTJE E. M. HOFMEISTER, † SUSANNE TEXTOR, AND WOLFGANG BUCKEL*

Laboratorium für Mikrobiologie, Fachbereich Biologie, Philipps-Universität Marburg, Marburg, Germany

Received 24 December 1996/Accepted 29 May 1997

The structural genes *sdhA* and *sdhB*, coding for the α - and β -subunits of the [4Fe-4S] cluster containing L-serine dehydratase from *Peptostreptococcus asaccharolyticus*, have been cloned and sequenced. Expression of modified *sdhB* together with *sdhA* in *Escherichia coli* led to overproduction of active His₆-tagged L-serine dehydratase. *E. coli* MEW22, deficient in the L-serine dehydratase L-SD1, was complemented by this *sdhBA* construct. The derived amino acid sequence of SdhBA shares similarities with both monomeric L-serine dehydratases, L-SD1 and L-SD2, from *E. coli* and with a putative L-serine dehydratase from *Haemophilus influenzae*, which suggests that these three enzymes are also iron-sulfur proteins.

L-Serine dehydratases and L-threonine dehydratases catalyze the irreversible overall deaminations of L-serine to pyruvate and L-threonine to 2-oxobutyrate. Most L-threonine dehydratases have been shown to contain pyridoxal-5'-phosphate as the prosthetic group. In contrast to L-threonine dehydratases, none of the bacterial L-serine dehydratases investigated to date has been conclusively proven to be dependent on pyridoxal-5'phosphate (6). An L-serine dehydratase unequivocally devoid of pyridoxal-5'-phosphate was discovered in the gram-positive anaerobe Peptostreptococcus asaccharolyticus (clostridial cluster XIII) (2) and characterized as an iron-sulfur protein (5). This enzyme was inactivated by exposure to air and could be specifically reactivated by incubation with Fe^{2+} under anaer-obic conditions. The dehydratase contains stoichiometric amounts of non-heme iron and acid-labile sulfur sufficient to form one $[4\text{Fe-4S}]^{2+}$ cluster per heterodimer (α , 30 kDa; β , 25 kDa). Electron paramagnetic resonance spectroscopic investigations supported the notion that the [4Fe-4S] cluster is involved in substrate binding and in catalysis by facilitating the elimination of the hydroxyl group of L-serine by a mechanism similar to that described for aconitase (8).

Cloning and sequencing of *sdhA* **and** *sdhB*. A preliminary report of part of this work has been published elsewhere (9). In an attempt to obtain a homologous probe for cloning, we derived four heterologous oligonucleotides from the N-terminal sequences of the α - and β -subunits of L-serine dehydratase, two in the forward (5' \rightarrow 3') direction and two in the backward direction. They were used to prime a PCR (17) from chromosomal DNA of *P. asaccharolyticus* ATCC 14963, which was isolated as described by Marmur (12). Only the oligonucleo-

tide combination β -forward and α -backward yielded a PCR product (0.75 kb), which indicated that the gene coding for the β -subunit (*sdhB*) preceded that coding for the α -subunit (*sdhA*) and that the amplified DNA was part of *sdhB*. This was confirmed by direct sequencing (1), since the N terminus of the derived amino acid sequence was identical to that of the β -subunit.

Southern blots (18) revealed strong hybridization of the digoxigenin-labeled *sdhB* probe with a single 3.6-kb *PstI* fragment and a single 4.2-kb *Hin*dIII fragment of *P. asaccharolyticus* chromosomal DNA. Both fragments were chosen to establish size-selected libraries in the vector pBluescript SK⁺ (Stratagene, Heidelberg, Germany), since they were likely to contain complete copies of *sdhA* and *sdhB*, which had been estimated at 1.65 kb on the basis of the molecular mass (55 kDa) of the combined heterodimeric L-serine dehydratase (5). For cloning procedures, *Escherichia coli* DH5 α (7) was used. Five recombinant plasmids were confirmed to contain a 3.6-kb *PstI* fragment which carried the whole *sdhB* gene and about 80% of *sdhA*. By the same method, one recombinant plasmid was demonstrated to contain a 4.2-kb *Hin*dIII fragment which carried complete copies of *sdhA* and *sdhB*.

A region on the *Hin*dIII fragment of almost 2.3 kb in length was sequenced (19) in both directions. The DNA sequence is shown in Fig. 1, in which *sdhB* and *sdhA* represent open reading frames of 669 and 879 bp encoding the β - and α -subunits of L-serine dehydratase, respectively. The DNA sequence revealed not only that *sdhB* precedes *sdhA* in the same orientation but that both genes occur in the same reading frame, directly adjacent to each other. Putative ribosome binding sites for translation initiation were found 6 bp upstream of *sdhB* and 5 bp upstream of *sdhA* in the coding region of *sdhB* (Fig. 1). The N-terminal sequences of the α - and β -subunits of L-serine dehydratase and their respective calculated molecular masses of 30,774 and 24,154 Da are in good agreement with the data reported previously for the enzyme after purification from *P. asaccharolyticus* (5). As expected, no known consensus se-

^{*} Corresponding author. Mailing address: Philipps-Universität Marburg, Fachbereich Biologie, Laboratorium für Mikrobiologie, D-35032 Marburg, Germany. Phone: 49 6421-281527. Fax: 49 6421-288979. Email: buckel@mailer.uni-marburg.de.

[†] Present address: Biological Laboratories, Harvard University, Cambridge, MA 02138-2097.

1	cgttta	agtcc	agta	ctaa	ataa	acct	gtee	atte	acata	acac	caac	teeta	agcai	tge	aaca	agta	ctgt	tcca	agtgg	80
81	aggcaa	atcct	gtaa	agtti	ttta	gttg	ctga	gtta	aaaat	tata	tcta	acaco	cttci	gca	ttta	aaag	cgat	actg	cgctc	160
161	ttggta	actte	ttcg	cctg	cttt	agcat	tcga	aata	agtta	atcg	gtgti	teegi	tacti	cegei	aaca	aaat	gaga	agca	attat	240
241	tacaac	cactg	ataa	aataa	acaa	atato	catt	getg	ggtgi	tggta	agee	tatto	gccta	atta	cttc	aaca	cett	tgaga	aaatc	320
321	cttta	cttct	tgtt	tgttg	gttt	gtaga	attto	catg	tteta	accto	cetei	tttai	tatti	aata	aata	aaati	tata	tatt	atatc	400
401	ctatca	tattc	atag	ettti	tatai	ttaaa	attt	gttgi	taaaa	atte	tatt	ttta	atagi	get	taaa	taat	cgat	tatti	taagt	480
481	tttta	taatc	tgtt	attaș	gttg	tttt	cataç	gttti	tataa	atta	tacaa	attt	tati	ttg	ctca	cttta	attt	ctata	acttt	560
567	agget	atrar		aatt	toot	cost	22t 2	ttoo	ta	acta	ə ə + +	++ c+	aata	+ +	~~~~	ataa	aat	a++ >	sdhB	→ ~~~
1	. accyc	acgage	igeac	aucc		ccat	aaca		aaca	geca	aacc		gera	Laat	gaaa	gegg	agge	grea	M	1
640	ACA GA	T TAT	AGT	GCA	TTT	GAA	GTC	ATG	GGG	CCT	ATT	ATG	GTC	GGT	CCA	TCA	AGT	TCA	CAC	699
2 700	T D	Y TGGA	S GCA	A TGT	F AAA	E ATT	GCA	MAAC	G GTT	PGCA	I ACG	MAGC	V ATT	G	P AGT	SAAC	S	S TAT	HAAC	21 759
22	T A	G	A	c	ĸ	I	A	N	v	A	т	s	I	v	S	N	N	Ŷ	N	41
760	CAA GI	T GAA	TTT	CAA	CTT	CAT	GGG	TCT	TTT	GCC	CAC	ACT	TTT	AAA	GGA	CAT	GGA	ACT	GAT	819
42 820	Q V CGA GC	Е Т ТТА	FGTT	GGT	GGA	н АТА	G	GGT	F TTT	A GAA	H CCA	GAT	GAT	GAT	G AGA	н АТА	GAAA	ACT	DTCT	61 879
62	R A	L	v	G	G	I	L	G	F	Е	Р	D	D	D	R	I	ĸ	т	s	81
880	TTT GA	A TTG	GCT	AAG	CAA	GCT	GGT	TTA	AAT	TAT	ATC	TTT	ACT	ACT	ACT	AAC	CTT	GGA	GAT	939
940	Р Е ААТ ТА	L T CAT	ACCC	K AAT	TCC	GTT	AAG	АТА	M GTT	TTC	TCA	TAC	CCA	AAT	GGT	GAA	GAA	GAA	U TAT	101
102	N Y	н	P	N	s	v	к	I	v	F	s	Y	P	N	G	Е	E	E	Y	121
1000	GTT AT	C GGC	TCG	TCA	ATT	GGT	GGT	GGA	GCT	ATG	AAG	ATT	GTA	AAC	ATA	AAT	GGT	ATT	GCC	1059
1060	ATT GA	A TTT	AGA	GGT	GAG	TAT	TCA	ACT	ATC	M	CTT	GAA	TAT	CCC	GAA	CAA	CGT	GGC	GTA	1119
142	I E	F	R	G	Е	Y	S	т	I	L	L	Е	Y	Ρ	Е	Q	R	G	v	161
1120	ATA TO	C TAT	GTA	TCT	TCT	CTA	TTA	ACA	GGA	AGT	GAA	TAC	AAT	ATT	GAG	TCC	CTT	AAT	ACA	1179
1180	AAG AA	G AAT	ÅÅÅ	TTG	ACC	AAT	ATT	GTA	ACC	TTG	ACC	GTG	GAG	ATA	GAT	AAG	CCA	CTC	ACT	1239
182	к к	N	к	L	т	N	I	v	т	L	т	v	Е	I	D	к	₽	L	т	201
1240	GAG AG	T TTA	AAA	TCT	GCA	ATA	CTT L	GGT	GTT	GAA	AGA	TTT	ACA	ACG	GCT	AAA V	TAT	GT <u>G</u>	GAG F	1299
202	1 0	-		0		-		<u> </u>	•	-	**				••		*	•	~	~~~
		sdhi	$a \rightarrow$																	
1300	<u>gt</u> a ta	<i>sdhi</i> G AT	$A \rightarrow G CT$	A AAT	r aci	A GCI	CG1	GAZ	A ATA	A ATA	A GAT	r GTC	G TGC	AA:	GAG	G AGA	A GGI	A ATT	ГААА	1359
1300 222	GTA TA V *	sdhi G AT M C GAC	$A \rightarrow G CT$	A AAT N	F ACA	A GCI A GAA	CG1 R GAA	GAZ E GAA	A ATA I ATT	A ATA I A ATA	A GAI D	r GTC V	G TGC	N ACC	GAG	G AGA	G GAN	A AT: I	ГААА К ата	1359 18 1419
1300 222 1360 19	GTA TA V * ATA TA I Y	sdhi G AT M C GAC D	$A \rightarrow G CT $ L TTA L	A AAT N GTA V	F ACZ T TTA L	A GCI A GAA E	r CG1 R GAA E	GAA E GAA E	A ATZ I ATT I	A ATA I AAA K	A GAT D AAC N	r GTO V TCT S	G TGO C CAC H	AAT N ACC T	r gao E ACA T	G AG R GAG E	A GGJ G GAA E	A AT I GAA E	I AAA K ATA I	1359 18 1419 38
1300 222 1360 19 1420	GTA TA V * ATA TA I Y AGA AA	sdhi G AT M C GAC D A AAA	$A \rightarrow$ G CTI L L L CTT	A AAT N GTA V GAT	TTA T TTA L GCG	A GCI A GAA E GTC	F CGI R GAA E ATT	GAA GAA E GAT	A ATA I ATT I GTT	A ATA I AAA K ATG	A GAT D AAC N CAT	TCT S GCC	G TGO C CAC H TCT	AAT N ACC T GCA	r gao E ACA T ACG	GAG R GAG E AAA	GAA G GAA E AAT	A AT I GAA E TTG	I AAA K ATA I ACT	1359 18 1419 38 1479
1300 222 1360 19 1420 39	GTA TA V * ATA TA I Y AGA AA R K	sdhi G AT M C GAC D A AAA K T GAT	$A \rightarrow$ G CTI TTA L CTT L GTT	A AAT N GTA V GAT D ACT	F AC T TTA L GCG A GAA	A GCI A GAA E GTC V TAC	GAA E ATT I AAA	GAA E GAA GAT D ATG	A ATA I ATT I GTT V ATT	A ATA I AAA K ATG M GAT	A GAT D AAC N CAT H GGA	r GTC V TCT S GCC A TTT	G TGC C CAC H TCT S GCA	ACC T GCA A	r gao E ACA T ACG T AGA	G AGA R GAG E AAA K ACT	A GGA G GAA E AAT N TAT	A AT I GAA E TTG L GAA	F AAA K ATA I ACT T TAT	1359 18 1419 38 1479 58 1539
1300 222 1360 19 1420 39 1480 59	GTA TA V * ATA TA I Y AGA AA R K CAA TC Q S	sdhi G AT M C GAC D A AAA K T GAT D	$A \rightarrow$ G CTI L CTT L GTT V	A AAT N GTA V GAT D ACT T	F ACZ T TTA L GCG A GAA E	A GCI GAA E GTC V TAC Y	F CGT R GAA E ATT I AAA K	F GAA E GAA GAT D ATG M	A ATA I ATT I GTT V ATT I	A ATA I AAA K ATG M GAT D	A GAT D AAC N CAT H GGA G	TCT S GCC A TTT F	G TGO CAC H TCT S GCA A	N ACC T GCA A AAG K	r gao E ACA T ACG T AGA R	GAG R GAG E AAA K ACT T	A GGA GAA E AAT N TAT Y	A AT GAA E TTG L GAA E	I AAA K ATA I ACT T TAT Y	1359 18 1419 38 1479 58 1539 78
1300 222 1360 19 1420 39 1480 59 1540	GTA TA V * ATA TA I Y AGA AA R K CAA TC Q S GCA AA	sdhi G AT M C GAC D A AAA K T GAT D T TCA	$A \rightarrow$ G CTI L CTT L GTT V GGC	A AAT N GTA V GAT D ACT T AAA	F ACZ T TTA GCG A GAA E TCC	A GCI GAA E GTC V TAC Y ATA	r CGT R GAA E ATT I AAA K GTT	F GAA E GAA D ATG M GGA	A ATZ I ATT GTT V ATT I GAT	A ATA I AAA K ATG M GAT D TTT	A GAT D AAC N CAT H GGA G CTT	GCC F GCC A TTT F GCC	G TGO CAC H TCT S GCA A AAA	N ACC T GCA A AAG K GCT	r gao E ACA T ACG T AGA R ATG	GAG R GAG E AAA K ACT T GCA	GGA GAA E AAT N TAT Y ATG	A AT GAA E TTG L GAA E GCT	I AAA K ATA I ACT T TAT Y TTT	1359 18 1419 38 1479 58 1539 78 1539
1300 222 1360 19 1420 39 1480 59 1540 79 1600	GTA TA V * ATA TA I Y AGA AA R K CAA TC Q S GCA AA A N TCA AC	sdhi G AT M C GAC D A AAA K T GAT D T TCA S T AGT	$A \rightarrow$ G CTI L TTA L CTT L GTT V GGC G GAA	A AAT N GTA V GAT D ACT T AAA K GTG	F AC T TTA GCG A GAA E TCC S AAT	A GCT A GAA E GTC V TAC Y ATA I GCT	r CGT R GAA E ATT I AAA K GTT V TCT	F GAA E GAA E GAT D ATG GGA G ATG	A ATA I ATT I GTT V ATT I GAT D GGT	A ATA I AAA K ATG GAT D TTT F AAA	A GAI D AAC N CAT H GGA G CTT L ATA	r GTC V TCT S GCC A TTT F GCC A GTC	G TGO CAC H TCT S GCA A AAA K GCA	ACC T GCA AAG K GCT A GCC	r GAG E ACA T ACG T AGA R AGA R ATG M CCT	GAG R GAG E AAA K ACT T GCA A ACT	A GGA GAA E AAT N TAT Y ATG M GCC	A AT GAA E TTG L GAA E GCT A GGT	F AAA K ATA I ACT T TAT Y TTT F TCA	1359 18 1419 38 1479 58 1539 78 1599 98 1659
1300 222 1360 19 1420 39 1480 59 1540 79 1600 99	GTA TA V * ATA TA I Y AGA AA R K CAA TC Q S GCA AA A N TCA AC S T	sdhi G AT M C GAC D A AAA K T GAT T CA S T AGT S	$A \rightarrow$ G CTI L CTT L GTT V GGC G GAA E	A AAT N GTA V GAT D ACT T AAA K GTG V	F ACJ T TTA GCG A GAA E TCC S AAT N	A GCT A GAA E GTC V TAC Y ATA I GCT A	F CGT R GAA E ATT I AAA K GTT V TCT S	F GAA E GAA D ATG M GGA G ATG M	A ATA I ATT I GTT V ATT I GAT D GGT G	A ATA I AAA K ATG M GAT D TTT F AAA K	A GAT D AAC N CAT H GGA G CTT L ATA I	r GTC V TCT S GCC A TTT F GCC A GTC V	G TGC CAC H TCT S GCA A AAA K GCA A	C AAT N ACC T GCA A AAG K GCT A GCC A	r gao E ACA T ACG T AGA R ATG M CCT P	G AGA R GAG E AAA K ACT T GCA A ACT T	A GGA GAA E AAT N TAT Y ATG M GCC A	A AT I GAA E TTG L GAA E GCT A GGT G	F AAA K ATA I ACT T TAT Y TTT F TCA S	1359 18 1419 38 1479 58 1539 78 1599 98 1659 118
1300 222 1360 19 1420 39 1480 59 1540 79 1600 99 1660	GTA TA V * ATA TA I Y AGA AA R K CAA TC Q S GCA AA A N TCA AC S T TCC GG	sdhi G AT M C GAC D A AAA K T GAT T GAT S T AGT S A ATC	$i \rightarrow$ f CTI L CTT L GTT V GGC G GAA E ATG M	A AAT N GTA V GAT D ACT T AAA K GTG V CCG	F ACJ T TTA GCG A GAA E TCC S AAT N GCA	A GCT A GAA E GTC V TAC Y ATA I GCT A TG M	F CGT R GAA E ATT I AAA K GTT V TCT S TTA	F GAA E GAA C GAT D ATG GGA G GGA GTA	A ATH I ATT I GTT V ATT I GAT GGT GCT	A ATA I AAA K ATG M GAT D TTT F AAA K GCC	A GAT D AAC N CAT H GGA G CTT L ATA I ACT	r GTC V TCT S GCC A TTT F GCC A GTC V GAA	G TGC CAC H TCT S GCA A AAA K GCA A AAA	N ACC T GCA A AAG K GCT A GCC A TAT	r GAG E ACA T ACG T AGA R ATG M CCT P AAT	GAG R GAG E AAA K ACT T GCA A CT T T T T T T	A GGJ GAA E AAT N TAT Y ATG GCC A GAC	A AT GAA E TTG L GAA E GCT A GGT G AGA	F AAA K ATA I ACT T TAT Y TTT F TCA S ACG	1359 18 1419 38 1479 58 1539 78 1599 98 1659 118 1719
1300 222 1360 19 1420 39 1480 59 1540 79 1600 99 1660 119 1720	GTA TA V * A ATA TA I Y AGA AA R K CAA TC Q S GCA AA A N TCA AC S T TCC GG S G C G A A CT AT	sdhi G AT M C GAC D A A AAA K T GAT T GAT S T AGT S A ATC I C CAA	$i \rightarrow$ f CTI L CTT L GTT V GGC G GAA E ATG M AAT	A AAT N GTA V GAT D ACT T AAA K GTG V CCG P GGT	F ACZ T TTA GCG A GAA E TCC S AAT N GCA A TTC	A GCT GAA E GTC V TAC Y ATA I GCT A ATG M TTG	r CGT R GAA E ATT I AAA K GTT V TCT S TTA L ACT	F GAA E GAT D ATG M GGA G ATG M GTA V TCA	A ATA I ATT I GTT V ATT I GAT GGT GCT A ATA	A ATA I AAA K ATG M GAT D TTT F AAA K GCC A GGA	A GAI D AAC N CAT H GGA G CTT L ATA I ACT T ATT	GTC V TCT S GCC A TTT F GCC A GTC V GAA E GGT	GCA CAC H TCT S GCA A AAA K GCA A AAA K CAA	C AAT N ACC T GCA A AAG K GCT A GCC A TAT Y GTC	F GAC E ACA T ACG T AGA R ATG M CCT P AAT N ATC	GAG GAG E AAA K ACT T GCA A ACT T T T T T T T T T T T T T T T T	A GGJ GAA E AAT N TAT Y ATG GCC A GAC D AAG	A AT I GAA E TTG L GAA E GCT A GGT G AGA R TAC	r AAA K ATA I ACT T TAT Y TTT F TCA S ACG T GCT	1359 18 1419 58 1539 78 1599 98 1659 118 1719 138
1300 222 1360 19 1420 39 1480 59 1540 79 1600 99 1660 119 1720 139	GTA TA V * ATA TA I Y AGA TA I Y AGA AA CAA TC QCA AA A N TCA ACC S G S G ACT AT T I	sdhi G AT M C GAC D A AAA K T GAT T CA S T AGT S A ATC C CAA Q	$A \rightarrow$ C CTI L TTA L CTT L CTT U GGC G GAA E ATG M AAT N	A AAT N GTA V GAT D ACT T AAA K GTG V CCG P GGT G	F ACI TTA GCG A GAA E TCC S AAT N GCA A TTC F	A GCT GAA E GTC V TAC Y ATA I GCT A TTG L	r CGT R GAA E ATT I AAA K GTT V TCT S TTA L ACT T	F GAA E GAA D ATG M GGA G ATG M GTA V TCA S	A ATA I GTT V ATT I GAT GGT G GCT A ATA I	AAA K AAA GAT D TTT F AAA K GCC A GGA G	A GAT D AAC N CAT H GGA G CTT L ATA I ACT T ATT I	r GTC V GCC A TTT F GCC A GTC V GAA E GGT G	G TGC CAC H TCT S GCA A AAA K GCA A AAA K CAA Q	N ACC T GCA A AAG K GCT A GCC A TAT Y GTC V	GAC E ACA T ACG T AGA R ATG M CCT P AAT N ATC I	GAG R GAG E AAA K ACT T GCA ACT T T T T T T T T T T T T	GAA GAA E AAT N TAT Y ATG GCC A GAC D AAG K	A AT I GAA E TTG GAA E GCT A GGT G AGA R TAC Y	F AAA K ATA I ACT T TAT Y TTT F CA S ACG T GCT A	1359 18 1419 58 1539 78 1599 98 1659 118 1719 138 1779 158
1300 222 1360 19 1420 59 1540 79 1660 99 1660 119 1720 139 1780	GTA TA V * ATA TA I Y AGA AA K CAA TC Q S GCA AA A N TCA AC S T TCC GG S G S G ACT AT T I ACC TT T	sdhi G AT M C GAC D A AAA K T GAT T GAT T A S T A A ATC C CAA Q C CAA	$I \rightarrow G$ G TTA L GTT U GGC G GAA E ATG M AAT N GGA	A AAT N GTA V GAT D ACT T AAA K GTG V CCG P GGT GCG Q CCG	F ACJ TTA GCG A GAA E TCC S AAT N GCA A TTC F GAG F	A GCT GAA E GTC V TAC Y ATA I GCT A ATG M TTG L GGT C	F CGT R GAA E ATT I AAA K GTT V TCT S TCA L ACT T GGT C	F GAA E GAA D ATG M GGA GGA GGA GTA V TCA S TCA S C	A ATA I GTT V ATT I GAT GGT G GCT A ATA I CAA	AAA K ATG GAT D TTT F AAA K GCC A GCA GCA GCT	A GAI D AAC N CAT H GGA G CTT L ATA I ACT T ATT I GAA P	GCC A TTT F GCC A TTT F GCC A GTC V GAA E GGT G GT C	G TGO CAC H TCT S GCA A AAA K GCA A AAA K CAA Q GGT C	ACC T GCA A AAG K GCT A GCC A TAT Y GTC V CC S	F GAC E ACA T ACG T AGA R ATG M CCT P AAT N ATC I GCT 2	GAG R GAG E AAA K ACT T GCA ACT T T T T T T T T T T C C	A GGJ GAA E AAT N TAT Y ATG GCC A GAC D AAG K CA	A AT I GAA E GAA E GCT A GGT G AGA R TAC Y ATG M	F AAA K ATA I ACT T TAT Y TTT F CA S ACG T GCT A GCA	1359 18 1419 38 1479 58 1539 78 1539 1539 1659 118 1719 138 1779 158 1339 270
1300 222 1360 19 1420 59 1540 79 1540 79 16600 99 16600 119 1720 139 1780 159 1840	GTA TA V * ATA TA I Y AGA AA R K CAA TC Q S GCA AA A N TCC GG S G ACT AT T I ACC TT T F GCA GC	sdhi G AT M C GAC D A AAA K T GAT T CAT S T AGT S T AGT I C CAA Q T GCC A T GCC	$I \rightarrow I$ G CTI L CTT L GTT V GGC G GAA E ATG M AAT N GGA GCA GCA GCA CTT	A AAT N GTA V GAT D ACT T AAA K GTG V CCG P GGT GCG A GTA	F ACF T TTA GCG A GAA E TCC S AAT N GCA A TCC F GAG GAA	A GCT A GAA E GTC V TAC Y ATA I GCT A TG L GGT G GT ATG	r CGT R GAA E ATT I AAA K GTT V TCT S TTA CGT GGT GGT TTA	F GAA E GAA E GAT D ATG M GGA G GA G GTA V TCA S TGT C GGT	A ATA I ATT I GTT V ATT I GAT GGT GGT A ATA I CAA QGA	I AAA K ATG GAT D TTT F AAA K GGA GGC G G G G G C C A ACT	A GAY D AACC N CAT H GGA GGA ATT I ACT T T ATT I GAA GTT	GTC V TCT S GCC A TTT F GCC A GTC V GAA E GGT G GAA	G TGC CAC H TCT S GCA A AAA K GCA A AAA K CAA Q GGT G CAA	N ACC T GCA A AAG K GCT A GCC A TAT Y Y GTC V TCC S GCT	r GAC E ACA T ACG T AGA R ATG M CCT P AAT I GCT I GCT A CTT	GAG R GAG E AAA K ACT T T GCA ACT T T T T T T T T C C AC	A GGA GAA E AAT N TAT Y ATG GCC A GAC D AAG K GCA A GCA	A AT I GAA E TTG GAA E GGT A GGT GGT A GGA TAC Y ATG M GCA	I AAA K ATA I ACT T TAT Y TCA S ACG T GCT A GCT A GCA AGC	1359 18 1419 38 1479 58 1539 78 1539 1539 1659 118 1779 158 1839 1839
1300 222 1360 19 1420 39 1480 59 1480 79 1600 99 1600 99 1660 119 1720 139 1720 139 1780 159 1840	GTA TA V * ATA TA I Y AGA AA R K CAA TC Q S GCA AA A N TCC GG S G ACT AT T I ACC TT T F GCA GC A CA A A	sdhi G AT M C GAC D A AAA K T GAT T TCA S T AGT T AGT C CAA Q T GCC A T GCG A	$\mathbf{q} \rightarrow \mathbf{G}$ \mathbf{G} \mathbf{CTT} \mathbf{L} \mathbf{GTT} \mathbf{V} \mathbf{GGC} \mathbf{G} \mathbf{G} \mathbf{G} \mathbf{G} \mathbf{A} \mathbf{T} \mathbf{M} \mathbf{G}	A AAT N GTA V GAT D ACT T AAA K GTG GGT G GGT G GCG A GTA V V	T AC/ T TTA GCG GCG A E TCC S AAT TCC S AAT TCC F GCA E GAA E CAA E	A GCT A GAA E GTC V V TAC Y ATA I GCT A TG GGT G GGT G ATG M	F CGT R GAA E AATT I AAA K GTT V V TCT S TTA L ACT T GGT GGT L	GAA E GAA C GAT D ATG G GA ATG G GTA V V TCA S TGT C GGT G G	A ATY I ATT I GTT V GTT U GGT G GCT A ATA I CCAA Q GGA G	I AAA K ATG M GAT D TTT F AAA G GCC A GCC G GCT A ACT T	A GAY D AAC N CAT H GGA G GCTT L ATA I ACT T ATT I GAA E GTT V	V GTCT S GCC A TTTT F GCCC A GTC V GAA E GGT G GAA E GGT C GAA E	G TGC CAC H TCT S GCA A AAA K CAA Q GGT G GGT G CAA Q	N ACC T GCA A AAG GCC A GCC A GCC V TAT Y GTC V C C S GCT A	r GAC E ACA T ACG T AGA R ATG M CCT P AAT N ATC I I GCT A CTT L	GAGA R GAG E AAAA K ACT T T T T T T T T T T T T T T T T T T	A GGJ G GAA E AAT N TAT Y ATG GAC D AAG GAC C A AAG GCA A A GCA A	I GAA E GAA E GGA GGT G GAA GGT A GGT TAC Y A A GCA A A GCA A	I AAA K ATA I ACT T TAT Y TTT F TCA S ACG T GCT A GCA A AGC S	1359 18 1419 38 1479 58 1539 78 1539 98 1659 118 1719 138 1779 158 1839 178 1839 198 198
1300 222 1360 19 1420 39 1480 59 1480 79 1600 99 1660 119 1720 139 1720 139 1780 159 1840 179	GTA TA V * ATA TA I Y * AGA AA R K CAA TC Q S GCA AA A N TCA AC S T TCC GG S G ACT AT T I T F GCA GC A C A A T C A C T T	sdhi G AT M C GAC D A AAA T GAT T GAT T AGT S A ATC S C CAA T GCC A T GCC A T GCC A T GCC	$\mathbf{A} \rightarrow \mathbf{G}$ CTJ L TTA L TTA L CTTT L GTT V GGC G GGAA E ATG M AATT N GGCA I L ATA L TTA L CTT L GTT V CTT L CTT C CTT C C C C C C C C C C C C C	A AAT N GTA V GAT D ACT T AAA K GGG GGG GGG GGG A GGTA V V AAT N	F ACJ T TTA GCG GCG A GCA F C C S AAT T C C F GCA E GAG E GAG E GTG V	A GCT A GAA E GTC V TAC V A TA A TG G GCT A TG G GGT G A TG G GT T T T G CTT L	r CG ^T R GAA E AATT I AAA K GTT V V TCT S TTA L ACT T GGT G GGA G GGA	F GAA E GAA E GAT D ATG G M GTA V V C GTA C GGT G CTT L	A ATA I ATT I GAT G GAT G GGT G GGT A ATA I CCAA Q GGA G G TA V	I AAA K ATG M ATG D TTT F AAA G GCT A GGC G GGA G GCT A ACT T T TGC C	A GAY D AAC N CAT H GGA G CTT L ATA I ACT T T AATT I GAA E GTT V GAT D	r GTC V TCT S GCC A TTT F GCC A GGT V GAA E GGT C GAA E C GAA E C CCA P	GCA CCC CAC H TCT S GCA A AAA AAA K CAA Q GGT G CAA Q CAA Q ATA T	N ACC T GCA A AG GCC A GCC A C C C C C S GCC S GCC A C C C S C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C	r GAC E ACA T ACG T AGA R ATG M CCT N AATG N AATC I GCT A CTT L CTT L GGA	GAGA R GAG E AAAA K ACT T T T T T T T T T T T T T T T T C ACA C A C	A GGJ G GAA E AAT N TAT Y ATG M GCC D GCC D C AAG GCA A A GCA A GCA A CT Y	I GAA E GAA E GGT G GGT G GGT TAC Y ATG M GCA A CAA	I AAA K ATA I ACT T T T T T T C A S C C T C C A S C C A C C A C C A C C X V	1359 18 1419 38 1479 58 1539 78 1539 98 1659 118 1719 138 1719 138 1779 138 1779 138 1839 178 1839 178 1839 258 1839 218 195 218 195 218 218 218 218 218 218 218 218
1300 222 1360 19 1420 59 1540 709 1540 99 1660 119 1720 139 1780 159 1840 179 1840 199 1960	GTA TA V * ATA TA I Y * AGA AA R K CAA CA Q S GCA AA A N TCA AC S T TCC GG S G ACT AT T F GCA GC A A ATC AC I T CCT TG	sdhi G AT M C GAC D A AAA T GAT T GAT T AGT C CAA T GCG A T GCG A T GCG A T ACT	$\mathbf{A} \rightarrow \mathbf{G}$ CTJ \mathbf{G} CTJ \mathbf{L} \mathbf{C} TTA \mathbf{L} CTT \mathbf{L} \mathbf{G} TT \mathbf{V} \mathbf{G} GTT \mathbf{V} \mathbf{G} GGC \mathbf{G} \mathbf{G} ATG \mathbf{M} \mathbf{A} ATG \mathbf{G} GGC \mathbf{C} TT \mathbf{L} \mathbf{L} \mathbf{G} TT \mathbf{V} \mathbf{G} GCC \mathbf{G} \mathbf{G} GCC \mathbf{G} \mathbf{G} GCCC \mathbf{G} \mathbf{G} GCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	A AAT N GTA V GAT D ACT T AAA K GTG GCG GCG GCG A GCCG A GCC A AAT N AGA	F ACJ TTA L GCG G GCA GCA E TCC S AAT TCC F GAG E GAG E GAG E GTG V AAT	A GCT A GAA E GTC V TAC V TAC V ATA I GCT A ATG G GCT G ATG G CTT L GCA	r CGT R GAA E ATT I AAA K GTT V TCT S TTA C TTA GGT T G GGT G GGA G GGA C C C C	F GAA E GAA D ATG GGA GGA GGA GGA TCA S TGT C GGT GGT L GGA	A ATA I ATT I GGTT V ATT GGTT G GGT G GGT A A TA A TA	A ATH I AAA K ATG GAT D TTT F AAA K GCC A A GGCA A C C ATA	A GAY D AACC N CCAT H GGA G CTT L ACT T ATA I ACT I GAA C GAT D AAT	r GTC V TCT S GCC A TTT F GCC A GTC V GAA E GGT G GGT G GAA E CCA P GCA	GCA CCC CAC H TCT S GCA A AAAA K CAA Q GGT G CAA Q Q CAA Q CAA I TTT	N ACC T GCA A AAG GCC A GCC A GCC S GCT Y C C S GCT A GCC A ATA	r GAC E ACA T ACG T AGA R ATG ATG M CCT P AAT N ATC CCT I CTT L GGA G TCT	GAGA R GAG E AAAA K T T T T T T T T T T T T T T T T	A GGJ G GAA E AAT N TAT Y ATG GCC A GCC A GCC A GCC A GCC A GCC A GCC A C GCC C C A C C C C	I GAA E TTGG GAA E GGT G AGA R CAA A ATG Q CAA Q TTG	I AAA K ATA I ACT T TTT F TCA S ACG T CCA S CCA A AGC S TAC S CCA X GCT	1359 18 1419 38 1479 58 1539 78 1599 98 1659 118 1719 138 1779 158 1839 178 1839 198 198 195 2019
1300 222 1360 19 1420 59 1540 709 1540 99 1660 139 1780 139 1780 159 1840 179 1840 199 1900 219	GTA TA V * ATA TA I Y AGA AA R K CAA TC Q S GCA AA CAA N TCC GG S G ACT AT T T GCA GCA ACT AT T T F GCA GCA ACT AT T T F GCA GCA ACT AT C T T T C GC TT T C C GC TT C C C T C C C T C C C C C C C C C C	sdhi G AT M C GAC D A AAA K T GAT T CAT S T AGT S A ATC T GCC A A A ATT T ACT	$\mathbf{A} \rightarrow \mathbf{G}$ \mathbf{G} CTJ \mathbf{L} \mathbf{TTA} \mathbf{L} \mathbf{CTT} \mathbf{L} \mathbf{GTT} \mathbf{V} \mathbf{GGC} \mathbf{G} \mathbf{G} \mathbf{G} \mathbf{ATG} \mathbf{M} \mathbf{ATG} \mathbf{M} \mathbf{ATG} \mathbf{G} \mathbf{G} \mathbf{M} \mathbf{ATG} \mathbf{G} \mathbf{G} \mathbf{G} \mathbf{T} \mathbf{L} \mathbf{G} \mathbf	A AAT N GTA V GAT D AAT T AAA K GTG G G G G G G G G G C G A GTA V V AAT N AGA R R C C C C C C G C G A C T C C C C C C C C C C C C C C C C C	F ACL T TTA L GCG A GCA E TCC S AAT TCC S AAT N GCA E GAG GCA E GAG E GAA E S GTG V AAT N	A GCI A GAA GAA E GTC V V TAC V ATA A TTG G TTG A ATG ATG G GT C T L GCA A CTT	r CGT R GAA E ATT I AAAA K GTT V TCT S TTA CGTT C GGT G TTA L GGGA G GCA CC S	F GAA E GAA D ATG GGA ATG GGA GGA GGA CTT C GGT C CTT L L GGA G GGA CTT L	A ATA I ATT I GTT V ATT I GGT G GGT A ATA I CAA Q GGA GTA V V U U U	A ATH I AAA K ATG GAT D TTT F AAA K GCC A A GGCA A C C ATA I I C C ATA	A GAY D AACC N CAT H GGA G CTT L ATA I ACT T GAA ATT I GAA D D AAT N	V GTC V S GCC A TTT F GCC A GTC V GAA E GGT GGA C GAA E C C A C C GCA A A C C C C C C C C C C C	GCA CCC CAC H TCT S GCA A AAA K CAA Q GGT G CAA Q Q CAA Q CAA I TTT F F	N ACC T GCA A AAG GCT A GCC A TAT Y GTC V TCC S GCT A GCC A ATA I I	r GAC E ACA T ACG T AGA R ATG M CCT P AAT N ATC I GCT A CTT L GGA G G TCT S	GAGA R GAG E AAA K T T T T T T T T T T T T C ACA T T T T T	A GGJ GAA E AAT N TAT Y Y ATG GCC A GCC A GCC A GCC A GCC A GCC A C GCC C C C	A ATT I GAA E TTG GAA E GGT A GGT G AGA R R CAA Q CAA Q TTG L	I AAA K ATA I ACT T TAT Y TTT F TCA ACG T GCT A GCT A GCT A CGCT A CGCT A	1359 18 1419 38 1479 58 1539 78 1599 98 1659 118 1719 138 1779 158 1839 178 1839 1959 2019 2019 2019 2019 2019 2019
1300 222 1360 19 1420 39 1420 59 1480 99 1600 99 1600 139 1720 139 1720 159 1840 179 1900 199 1900 2020 239	GTA TA V * ATA TA I Y AGA Y AGA Y AGA K CAA TO Q GG AA N TCA AC S G ACC T T I ACC T GCA AC ACC T CCT T GCA AC AC T CCT T CCT T AC AC T T CCT T GCA GC CT T	sdhi G AT M C GAC D AAAA K T GAT T GAT T CAA S T GAT S T AGT C CAA A A ATC C CAA A T GCC A A A T T GCC A A A T GCC A A A T ACT A C A C A A A A A A A A A A A A A A A A	$\mathbf{A} \rightarrow \mathbf{G}$ \mathbf{G} CTI \mathbf{L} \mathbf{TTA} \mathbf{L} \mathbf{CTT} \mathbf{L} \mathbf{GTT} \mathbf{V} \mathbf{V} \mathbf{V} \mathbf{GGC} \mathbf{G} \mathbf	A AAT N GTA V GAT D ACT T AAA K GTG CCG G G G G G CCG A GTA V V AAT N AGA R R GAA E	F ACL T TTA L GCG A GAA F TCC S AAT N GCA A TTC F GAG GAA E GAA E GAA E GAA T TC S S AAT N AGT S	A GCI A GAA GAA E GTC V V TAC V ATA I GCT A ATG ATG ATG GGT G GCA A CTA L L	r CGT R GAA E ATT I AAAA K GTT V T CT S TTA L ACT T G GTTA G GTTA C S GTC C S GTG V	F GAA E GAA D ATG GGA GGA GGA GGA GGT GGT GGT L GGA GCA P	A ATY I GTT V ATT I GAT D GGT G GGT A ATA I CAA Q GGA G GTA V V TTT T F	A ATY I AAA K ATG M GAT D TTT F AAA K GGAT T C AA ACT T T C C ATA I GAT I D D	A GAY D AACC N CAT H GGA G CTT L ATA I ATT I GAA E GTT V V GAT D AAT N GAA E	V GTCT V S GCC A TTT F GCC A GTC V GCA A E GGA C C GAA E C C GCA A GTG C C GCA A C C GCC V V V V V V V V V V S GCC V V V V S GCC V V V S GCC V V V S GCC V V S GCC V V V S GCC V V V S GCC V V V S GCC V V V S GCC V V V V S GCC V V V S GCC V V V S GCC V V V S GCC V V V S GCC V V V S GCC V V V S GCC V V V S GCC V V V S GCC V V V S GCC V V V S GCC V V V S GCC V V V S GCC V V V S GCC V V V V V S GCC V V V V V S S C C V V V V V S C C C V V V S C C C V V V S S C C C V V V S C C C V V S C C C V V S C C C V V S S C C C V V S C C C V V S C C C V V S C C C V V S C C C V V V S C C C C	GCA CCACC H TCT S GCA A AAA A AAA K CAA Q GGT G CAA Q CAA Q TTT F GTA V	N ACC T T GCA A AAG GCT A GCC A TCC S GCT A C C TCC S GCT A ATA I I ATA I	r GAC E ACA T ACG T ACG T ACG T ACG M CCT P ACG M ATG M ATG GCT CTT L GCT CTT C GCT S GCA CTT S CA CTT S CA C CTT S CA C CCT C CTT	GAGA R GAG E AAA K ACT T T T T T T T T T T C C ACA T T T T	A GGJ G GAA E AAT N TAT Y ATG M GCC A A GCC A A GCC A GCA A GCA A GCA C C G GCC G GCC C G GCC C C C	A AT' I GAA E TTG GAA E GGT A GGT A GGA R TAC G AGA A AGA CAA Q TTG L GAA E	I AAA K ATA I ACT T T T T T T T T T T T T T T T T T T	1359 18 1419 58 1539 78 1599 98 1659 118 1719 138 1719 138 1779 138 1779 158 1839 1959 218 2019 238 2019 258
1300 222 1360 19 1420 39 1420 59 1540 99 1600 99 1600 99 1600 199 1720 139 1720 159 159 1900 159 1900 179 1900 2020 2050 2080	GTA TA V * ATA TA I Y AGA T Q S GCA AA CCA TCC GCA AA TCA ACC T TCA ACC T T F GCA AA ACC T T F GCA AC T T GCA AC T T GCA AC T T T T GCA AC ACC T CCT T CCT T CCT T CCT T CCT T GCA A A T CCT T GC T	sdhi G AT M C GAC D AAAA K T GAT T TCA S T AGT S A ATC I GCC A A T GCC A A ATT T AGT T AGT I GCC A A ATT I C C CAA C A C CAA C C CAA C C CAA C C CAA C C CAA C C CAA C C CAC C C CAA C C CAA C C CAC C C C C	$\mathbf{A} \rightarrow \mathbf{A}$ $\mathbf{G} \subset \mathbf{CTT}$ \mathbf{L} \mathbf{TTA} \mathbf{L} \mathbf{CTT} \mathbf{L} \mathbf{CTT} \mathbf{V} \mathbf{V} \mathbf{GGC} \mathbf{G} \mathbf{G} \mathbf{G} \mathbf{G} \mathbf{A} \mathbf{G} \mathbf{G} \mathbf{M} \mathbf{A} \mathbf{AT} \mathbf{I} \mathbf{CTT} \mathbf{L} \mathbf{ATG} \mathbf{GTT} \mathbf{V}	A AAT N GATA V GAT T AAA K GTG GGT GGT GGT GGTA V AAT N AAGA R GAA E AAT	F ACL' T TTA L GCG A A E TCC S AAT N GCA A A TTC F GAG E GTG GCA A TTC S GAA C A AT N S GAA	A GCI A GAA GAA E GTC V V TAC Y ATA I GCT A ATG ATG ATG GGT G GCA A L GCCA L CTA L GCC	r CGT R GAA E ATT I AAAA K GTT V TCT S TTA L ACT T G GGT G GGT G GTC S GTG V TTA	F GAA E GAA D ATG GGA GGA GGA GGA GGA GGA GGA GGA GGA	A ATY I GTT V ATT I GAT D GGT G GGT G GGT CAA Q GGA GTT V V TTT F F GAG	A ATY I AAA K ATG M GAT D TTT F AAA K GGCT A ACT T C C ATA I GAT D ACA	A GAY D AACC N CAT H GGA G CTT L L ATA I ATA I ATA I GGA E GTT V V GAT D AAT N GAA E GGA	V GTC V S GCC A TTT F GCC A GTC V GAA E GGT GGT GGA C C GAA E C C GCA A GTG V C TT	GCAC CAC H TCT S GCA A AAA A AAA K CAA Q GGT G CAA Q CAA Q CAA I TTT F GTA V GGC	N ACC T GCA A AAG GCT A GCC A TAT Y GTC V V C C S GCT A GCC A ATA I GGG GGG GGG	r GAC E ACA T ACG T ACG T ACG T ACG T A AAT N ATC I CCT CTT C CCT S CCT CTT S CCT	GAG R GAG E AAA K ACT T T T T T T T T T T C C ACA T T T T	A GGJ G GAA E AAT N TAT Y ATG M GCC A A GCC A A GCC A GCC A GCC A GCC C C C	A AT' I GAA E TTG GAA E GGT A GGT A GGA R TAC Y ATG GCA A AGA CAA Q TTG L GAA E GAA	r AAA K ATA I ACT T T T T T T T T T T T T T T T T T T	1359 18 1419 58 1539 78 1539 98 1659 118 1719 138 1779 138 1779 138 1779 158 1839 198 198 198 2019 218 2079 2139
1300 222 1360 39 1420 59 1540 79 1600 19 1720 139 1720 139 1720 159 1720 159 1840 159 1990 2599 2020 239 2020 239	GTA TA V * ATA TA I Y AGA AA K CAA TC Q S GCA AA A N TCA AC S T TCC GG S G ACT AT T F GCA GC ACT AT T F GCA GC ACT TG P C CCT TG P C CCT TG P C CCT GG A A ATC AC I T CCT TG P C CCT GG A A A ATC AC I T CCT TG C A A A C C T G C A A A A C C T C C C C C C C C C C C C C C C C C	schug G AT M C GAC D A AAA K T GAT D T TCC A ATC T AGT S T AGT S T AGT T C C CAA A ATC C CAC A ATC C CAC A CC A ATC C CAC A CC C CAC C C	A G CTU G CTU L CTTTA L CTTTA L CTTTA CTTTA GGTT V V GGC GGA ATG GGA ATG GGC TTCC F GTT V V ATG ATG ATG ATG ATG ATG ATG ATG	A AAT N GTA V GAT T AAA K GTG GGT GGT GGT GGTA V V AAT N AGA R GAA E AAT I	F ACL' T TTA L GCG A A E TCC S AAT N GCA A C F GAA E GTG GGA A AT N AGT S S GAA E S C A AT	A GCT A GAA E GTC V V TAC Y ATA G GCT A ATG G GCT A CTA L CTA A CTA A CTA A CTA	r CGT R GAA E ATT I AAA K GTT S TCT S TTA C GGT G GGT G GTG V TTA L C CTTA	F GAA E GAA E GAA E ATG G GAA G G G G G G C C C C C C C C C C C C C C	A ATH I ATT I GTT I GATA G GGT G GGT G GGA G GGA G GAG C TTT F F GAG G GAG E	A ATJ I AAA K ATG GGT D TTT F AAA K GGC A A C T T C C A A A C T T C C A T T C C A T A T	A GAX D AAC N CAT H GGA G GCT I ACT I ACT I ACT I GAA C GTT V GAT N GAA E GGA GGA G GGA G GC C C C C C C C C C C	r GTC V S GCC A TTT F GCC V GAA E GGT C V GAA E GGT C C GGA A C C A GTG C C C C C C C C C C C C C C C C C C	3 TGC CAC H TCT S GCA A AAA AAA K CAA Q GGT G CAA Q ATA I TTT F GTA V CGC C GGC C CAA Q CCAA C CAC C C C C C C C C C C	N ACC T GCA A AAG K GCT A GCC A TAT Y GTC S GCT A GCC A ATA I GGG G G C S	r GAA E ACA T ACA T ACA R ACA R ACA R ACG C T P AAT C CCT I GCT L GCA CTT C T S GCA C CTT L L	AGJ R GAG E AAA K A CT T T T T T T T T T T T T T T T T T	A GGG GAA E AAT Y ATG GGCC A GGCC A GGCC A AAG GCA C C C C C	A AT I GAA E TTG GAA E GGA GGA AGG AGA AGCA Q CAA Q GAA E CAA A CAA A CAA A CAA A CAA A CAA A C C C C C C C C C C C C C	r AAA K ATA I ACT T T T T T T T T T T T T T T T T T T	1359 18 1479 58 1539 78 1599 98 1659 118 1779 158 1839 178 1839 178 198 198 2019 238 2019 238 2079 258 2079 218 2079 218 2079 218 2079 218 2079 218 2079 218 2079 218 2079 218 2079 218 2079 218 2079 218 2079 258 2079 2078 2079 2078 2079 2078 2079 2078 2079 2078
1300 222 1360 19 1420 39 1480 59 1540 79 1600 99 1660 199 1720 179 1780 179 1780 179 1900 2199 2020 2399 2020 2399 2020 259 21400	GTA TA V * ATA TA ATA TA K CAA AA C Q S G C A A A A C A C A C A C A C A C A C A C	sdb.y G AT M C GAC M C GAC D A AAA K T TGAT D T TGAT S T TGAT S T AGT C T AGT GC T GCA A T GCA A T GCA T A GGT T G GGT T G GC S A C S	A G CTU G CTU L CTTTA L CTTTA L GTT V V GGC G GGA ATG GGC T L ATG CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CTT V V CGC C G C C C T V V C CTT V V C CTT V V C C C C C C C C C C C C C	A AAT N GTA V GAT D ACT T AAA K GTG GCG G GCG A A GTA N AGA R AAT I I ATA I	F ACJ T TTA L GCG A GAA E TCC S S AAT TCC S S AAT TTC F GAA GCA A C GAA E GAA E GAA T TC S S AAT A N AGT A A A A A A A A A A A A A A A A A A	A GCT A GAA E GTC V V TAC Y ATA G GCT A A GGT C TA G GCT A CTA A CTA A CTA A CTA A CTA A CTA A CTA A CTA A CTA A CTA A CTA A CTA A CTA C C C C	r CGT R GAA E ATT I AAA K GTT S TTA ACT T T GGT G GGT G GTG Q TTA L GAT L D	F GAA E GAA E GAA TG ATG M GGA G G G G TG TCA S TGT C G G GGT G CCA F AGA R TTC C CCA F	A ATJ I ATT I GTT I GATA G GGT G GGT G GGT G GGA G GGA G GAG G GAG E TTA L	A ATJ I AAA K ATG GGAT D TTT F AAAA K GGCA C AAA ACT T C C C ATA I GAT D ACA T AAA K	A GAX D AAC N CAT H GGA G GCT T AAT T AAT T AAT T AAT T AAT T AAT T AAT T AAT T GAA E GGA GGA GGA G GGA G GGA G GGA G C GGA C T	r GTC V S GCC A TTTT F GCC V GAA E GGT C C GGA A GTG C C C C C C C C C C C C C C C C C C	GCA CCAC H TCT S GCA A AAA AAA K CAA Q GGT G GGC CAA Q ATA I F F GTA V GGCC G GGC C A D	N ACC T GCA A AGG K GCT A A GCC A TAT T T T C C C A A TAT I GGG G GGG GGG T C A X X X X X X X X X X X X X X X X X X	r GAA E ACA T ACG T AAG AGA R AAG R AAG R AAG R AAG CCTT L GCT S GCA A CTT L L ttat	GAGJ R GAG E AAA K T T T T T T T T T T T T T T T T T	A GGG GAA E AAT Y ATG GAC A GGCC A AAG GCA C C A AAG GCA A GCA A GCA C GGAC GGAC GGAC GGGA GGAA GGAA GGAA GGAA GGAA A A GGAA A AAG GAA A A A GAA A A A A GAA A A A A GAA A A A GAA A A GAA A A GAC A A A GAC A A A GAC A A A GAC A A A GAC A A A GAC A A A GAC A A A GAC A A A GAC A A A GAC A A GAC A A GAC A A GAC A A GAC A A GAC A A GAC A A GAC A A GAC A A GAC C A A GAC GAC	A AT I GAA E TTG GAA E GGA A GGA AGA A CAA Q GAA CAA Q GAA E GAA A CAA A CAA A CAA A CAA A CAA A CAA A C C C C C C C C C C C C C	I AAA K ATA I ACT T TAT T TAT F F TCA S ACG T TCA S ACG C T A ACG S TAC S CT Q CT A AGC S CT V V AAA K Kaaaaaa	1359 18 1419 58 1539 78 1599 98 1659 118 1719 138 1779 158 1839 178 1839 198 2019 238 2019 238 2019 238 2019 238 2204 238
1300 222 1360 19 1420 39 1480 59 1540 79 1600 39 1720 139 1720 139 1720 139 1900 219 1900 219 2020 229 2020 259 2140 259 2245	$\begin{array}{c} {\rm GTA} & {\rm TA} \\ {\rm V} & \ast \\ {\rm ATA} & {\rm TA} \\ {\rm ATA} & {\rm TA} \\ {\rm I} & {\rm Y} \\ {\rm AGA} & {\rm AA} \\ {\rm R} & {\rm K} \\ {\rm CAA} & {\rm TC} \\ {\rm Q} & {\rm S} \\ {\rm GCA} & {\rm AA} \\ {\rm A} & {\rm N} \\ {\rm TCA} & {\rm AC} \\ {\rm S} & {\rm TC} \\ {\rm TCC} & {\rm GG} \\ {\rm ACT} & {\rm ATC} \\ {\rm ACC} & {\rm TT} \\ {\rm T} & {\rm T} \\ {\rm F} \\ {\rm GCA} & {\rm GCA} \\ {\rm ACC} \\ {\rm CTT} & {\rm GC} \\ {\rm CTT} & {\rm GC} \\ {\rm CTT} & {\rm GC} \\ {\rm T} & {\rm G} \\ {\rm T} \\ {\rm C} \\ {\rm CTT} \\ {\rm CT} \\ {\rm CT} \\ {\rm CT} \\ {\rm CT} \\ {\rm CTT} \\ {$	sdhydd ar	A G CTI G CTTI L TTA L CTTT L GTT V CGG G GAA E ATG GGA G CTT L ATA GGA CTT V V CTT V V V V V CATA CTT CTT CTT CTT CTT CTT CT	A AAT N GTA V GAT D ACT T AAA K GTG GCG A AAT N AAA V AAT N A AAT I ATA I I	F ACL T TTA GCG GA GCG A GCA E C C C S AAT TTCC S AAT TTCC S GAA E GCA E GAG E AGT S C AAT N AGT S C AAT N A C C G C G C G C G C G C G C G C G C	A GCT A GGAA E GTC V TAC Y ATA I GCT A TTG GGT G GGT G GCT A CTA L CTA L CTA A CTA A CTA A CTA A CTA CTA CTA CTA	r CGT R GAA E ATT I AAT K S GGT GGT C GGT GGT C GGG GGG GGG GGG GGG	GAA E GAA E GAT D ATG GGA GGA GGA GGA GGA CTT L C GGT CTT L C GGA CTT C C CTT E GGA TGT C C TGT C C TGT C C TGT C C TGT C C TGT C C C TGT C C C C	A ATH I ATT I GTT V ATT I G GT D GGT A ATA ATA ATA ATA ATA CAA Q GGT V V TTT F GGG GGT TTA L Latag	A ATJ I AAA K ATG GAT D TTT F AAA GGA GGCT A A GGCC A A GGCT A T TGC C C A T T GGAT D ACA T A AA K K	A GAN D AACC N CAT H GGA G CTT L ATA T T ATT T ATT T GAA E GAA GAA GGA GGA GGA GGA GGA GAG GAG	r GTC V TCT S GCC A TTT F GCC A GCC A GCC A GCC C C C C C C C C C	GGCA CCAC H TCT S GCA A A A A A A A A A A A A A A A A A A	N ACC T GCA A AGG K GCT A A GCC A TAT T A T A T A T A T A T A T A T A	r GAX E ACA T ACA T ACG T ACG T ACG T ACG T ACG T ACG CCT I CCTT L CCTT L CCTT L CCTT L CCTT L CCTT C	G AGJ R GAG GAG E AAAA K ACT T T GCA A ACT T T T F ACA A T C A C A C A C A T C T C A C A C	A GGJ G GAA E AAT N TAT Y ATG M GCC A A GGC D AAG GAC D GGC A GCA A GCA A GCA A GCA A GCA A GCA C A C	A ATT I GAA E GGA GGT G A GGT A GGT A A GGT A A A CAA Q Q CAA Q CAA CAA A CAA A CAA A CAA A CAA A CAA A CAA A CAA A CAA A C CAA A C CAA A C CAA A C CAA A C CAA A C CAA A C CAA A C CAA A C CAA A C CAA A C CAA A C CAA A C CAA A C	T AAA K ATA I ACT T T TTT F CA S CT T CA S CT T CA S CCA A A CC T T CA S CCA C CCA C CCA C CCA C CCA C CCA C CCA C C CCA C	1359 18 1419 58 1539 98 1659 118 1719 138 1779 158 1839 178 1839 178 1839 218 2019 238 2028 228 228 228 228 228 228
1300 222 1360 19 14200 59 1540 99 1540 99 16600 99 17200 139 1720 139 1780 139 1780 139 1780 139 1780 139 2020 209 209 209 209 229 2240 2245	GTA TA V * ATA TA I Y AGA N T Y AGA AR K CAA CO S GCA AN TCA AC TCC GG ACC T T T ACC T GCA AC GCA T GCA AC GCA AC GGA AA GGA ACC GC T T T T T T T T T GC T	sdb.y G AT MC GAC D M T GAT D T T TAGT T TCAT T TCAT T TCAT T TCC A ATC C CAC T GCC A ATC T GCC A ATT T ACT G T GCC S A CAA Q CAA Q CAA Q CAA Q CAA Q CAA Q CAA <td>A G CTI G CTTI L TTA L CTTT L GTT V V GGC GAA E ATG GGA AAT N GGA CTT L ATG GTT V ATG M AAT CTT L CTT C C C C C AAT C CTT L CTT C C C C C C C C C C C C C</td> <td>A AAT N GTA V GAT D ACT T AAA K GTG V V CCG GGT G GCG A GTA V V N AGA R R AAT I ATA I ATA</td> <td>F ACJ T TTA GCG GCG A GCA E TCC S AAT TCC F GAA E GAA E GAA C GTG GTG GAA R E AGA R C Laas</td> <td>A GCT A GAA E GTC V V TAC GT A A ATA A TG GCT A A A GGT CTT L CTT L CTT A A AGA A R A CCTA A A AGA A CTA CTA A CTA A A A A A A A A A A A</td> <td>r CGT R GAA E ATT I AAT K S GGT C GGT GGT TTA L GGA GGG GGG GGG GGG GGG GGG GGG C CC S GTG V TTA L D A AT T T D A A A A A A A A A A A A A</td> <td>GAA E GAA E GAT D ATG GGA G GGA CTA C C CTT L C GGT C CTT L C GGA CTT C C CTT C C CTT F C CCA C CTT C C C C C C T C C C C C C C</td> <td>A ATH I ATT I GTT V ATT I GGT GGT GGT GGT GGT GGA GGA GTA V TTT F GAG E TTA L L catag</td> <td>A ATJ I AAA K ATG M GAT D TTT F AAA GGA GGCT A A A A C C C A A A C C C A A A C T T T GCC C C A A A A C GCT T T T T T T T T T T T T T T T T T</td> <td>A GAN D AACC N CAT H GGA CTT L ATA T T ATT T ATT T ATT T GAA E GAA G GAA GGAA G</td> <td>r GTC V TCT S GCC A TTT F GCC A GCC A GCC GAA E C GGT GGT C CCA CCA CCA CCA CCA CCT L GGA GCC CCA CCT C CCA CCT C CCA CCC CCC CCC C</td> <td>GCAC CAC H TCT S GCA A AAA K CAA Q GGT CCAA Q CCAA Q CCAA Q CCAA Q CCAA Q CCAA Q GGT TTT T F GGCA G GGC GGC C CCAC C CAC C CAC C H H TCT S S C C A A AAA A A A A A A A A A A A A</td> <td>ACC N ACC T GCA A AAG GCC A A GCC A GCC V V V TCC S GCT TAT TCC S GCC A A TAT I GGG G GGC G C C A A S C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C C A A C C C C A A C C C C A A C C C C C C A C C C C A C</td> <td>r GAX E ACA T ACA T ACA T AGA R R ATG M CCT I CCT I CCT C CCT L CCT L L CCT L L CCT L L CCT L CCT L CCT L CCT C CCT L CCT L</td> <td>GAG R GAG E AAAA K T T T T T T T T T T T T T T T T</td> <td>A GGJ G GAA E AAT N TAT Y Y ATG M GCC A A GCC A A GCC A A GCC C D GGC C C G GCC C C GCC C C C C C</td> <td>A ATT I GAA E GGT GGA GGT G GGT G GGT G GGT A GGT A C A GGT TAC Y A TG GGA A C CAA Q Q Q C CAA S CAA A C CAA C A C CAA C GGT A C C C C C C C C C C C C C C C C C C</td> <td>r AAA K ATA A ATA I ACT T TTAT Y Y TTT F F TCA A ACG T T GCT A AGC S CTT A AGC S CTT Y AAA K X AAAA K X AAAA K X</td> <td>1359 18 1419 38 1479 58 1539 98 1599 98 1599 118 1719 138 1779 178 1839 198 1839 198 1859 198 2019 218 2019 218 2019 228 2204 2294 2296</td>	A G CTI G CTTI L TTA L CTTT L GTT V V GGC GAA E ATG GGA AAT N GGA CTT L ATG GTT V ATG M AAT CTT L CTT C C C C C AAT C CTT L CTT C C C C C C C C C C C C C	A AAT N GTA V GAT D ACT T AAA K GTG V V CCG GGT G GCG A GTA V V N AGA R R AAT I ATA I ATA	F ACJ T TTA GCG GCG A GCA E TCC S AAT TCC F GAA E GAA E GAA C GTG GTG GAA R E AGA R C Laas	A GCT A GAA E GTC V V TAC GT A A ATA A TG GCT A A A GGT CTT L CTT L CTT A A AGA A R A CCTA A A AGA A CTA CTA A CTA A A A A A A A A A A A	r CGT R GAA E ATT I AAT K S GGT C GGT GGT TTA L GGA GGG GGG GGG GGG GGG GGG GGG C CC S GTG V TTA L D A AT T T D A A A A A A A A A A A A A	GAA E GAA E GAT D ATG GGA G GGA CTA C C CTT L C GGT C CTT L C GGA CTT C C CTT C C CTT F C CCA C CTT C C C C C C T C C C C C C C	A ATH I ATT I GTT V ATT I GGT GGT GGT GGT GGT GGA GGA GTA V TTT F GAG E TTA L L catag	A ATJ I AAA K ATG M GAT D TTT F AAA GGA GGCT A A A A C C C A A A C C C A A A C T T T GCC C C A A A A C GCT T T T T T T T T T T T T T T T T T	A GAN D AACC N CAT H GGA CTT L ATA T T ATT T ATT T ATT T GAA E GAA G GAA GGAA G	r GTC V TCT S GCC A TTT F GCC A GCC A GCC GAA E C GGT GGT C CCA CCA CCA CCA CCA CCT L GGA GCC CCA CCT C CCA CCT C CCA CCC CCC CCC C	GCAC CAC H TCT S GCA A AAA K CAA Q GGT CCAA Q CCAA Q CCAA Q CCAA Q CCAA Q CCAA Q GGT TTT T F GGCA G GGC GGC C CCAC C CAC C CAC C H H TCT S S C C A A AAA A A A A A A A A A A A A	ACC N ACC T GCA A AAG GCC A A GCC A GCC V V V TCC S GCT TAT TCC S GCC A A TAT I GGG G GGC G C C A A S C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C A A C C C C A A C C C C A A C C C C A A C C C C C C A C C C C A C	r GAX E ACA T ACA T ACA T AGA R R ATG M CCT I CCT I CCT C CCT L CCT L L CCT L L CCT L L CCT L CCT L CCT L CCT C CCT L CCT L	GAG R GAG E AAAA K T T T T T T T T T T T T T T T T	A GGJ G GAA E AAT N TAT Y Y ATG M GCC A A GCC A A GCC A A GCC C D GGC C C G GCC C C GCC C C C C C	A ATT I GAA E GGT GGA GGT G GGT G GGT G GGT A GGT A C A GGT TAC Y A TG GGA A C CAA Q Q Q C CAA S CAA A C CAA C A C CAA C GGT A C C C C C C C C C C C C C C C C C C	r AAA K ATA A ATA I ACT T TTAT Y Y TTT F F TCA A ACG T T GCT A AGC S CTT A AGC S CTT Y AAA K X AAAA K X AAAA K X	1359 18 1419 38 1479 58 1539 98 1599 98 1599 118 1719 138 1779 178 1839 198 1839 198 1859 198 2019 218 2019 218 2019 228 2204 2294 2296

FIG. 1. Sequences of the two genes, *sdhA* and *sdhB*, coding for both subunits of L-serine dehydratase from *P. asaccharolyticus*. The determined DNA sequences and the deduced amino acid sequences of the two subunits are shown. The two genes occur in the same reading frame, directly adjacent to each other. Putative ribosome binding sites for translation initiation are underlined.

quence from pyridoxal-5'-phosphate-dependent enzymes (15, 16) was detectable.

Similarity of the amino acid sequences deduced from sdhAand sdhB with both L-serine dehydratases from *E. coli*. A comparison of the deduced amino acid sequences of sdhA and sdhB with all the sequences available in the nonredundant protein database by BLAST network service revealed identities with the two L-serine dehydratases present in *E. coli* (20, 22) and a putative L-serine dehydratase from *Haemophilus influenzae* (Fig. 2). This putative L-serine dehydratase was identified by 68% identity of the deduced amino acid sequence with those of the two L-serine dehydratases from *E. coli*, L-SD1 and L-SD2 (3). Each of these three proteins consists of only one subunit, which is nearly as large as the two subunits of L-serine dehydratase from *P. asaccharolyticus* together. They all showed 24% overall identity to SdhBA of *P. asaccharolyticus*. There is, however, an extended central region (amino acids N102 to V220 of the β -subunit and M1 to A99 of the α -subunit, i.e., M221 to A319 in Fig. 2) in the *P. asaccharolyticus* enzyme which has no counterpart in these three other enzymes. In the C-terminal part of SdhA, the α -subunit of the *P. asaccharolyticus* enzyme, three conserved cysteines were found (Fig. 2), which should be sufficient for the coordination of the [4Fe-4S] cluster as postulated for the mechanism of the enzyme (8).

Expression of *sdhA* and *sdhB* in *E. coli*. Cloned copies of *sdhA* and *sdhB* were placed under the control of inducible T7 promoter on vector pCRII (Invitrogen, Leek, The Netherlands) to yield pAH2 and transformed into *E. coli* BL21 (DE3) (21) by electroporation. After aerobic growth on M9ZB medium (21) in the presence of ampicillin (0.2 mg/ml) at 37°C, the



FIG. 2. Alignment of the amino acid sequences of four L-serine dehydratases by Clustal's method. The sequences of the two subunits of L-serine dehydratase from *P. asaccharolyticus* in the order $\beta \rightarrow \alpha$ (SdhPa), the two one-subunit L-serine dehydratases from *E. coli* (L-SDIEc and L-SD2Ec), and a putative L-serine dehydratase from *H. influenzae* (L-SDHi) (3) are shown. The C-terminal amino acid of SdhB is V220 and the N-terminal amino acid of SdhA is M221, as indicated by a vertical line. The four conserved cysteines are marked by arrows.

cells were incubated in M9 medium supplemented with all amino acids except methionine. Cells were labeled with 1.2 MBq of L-[³⁵S]methionine for 5 min after induction of T7 RNA polymerase with isopropylthiogalactoside (0.8 mM) and inhibition of bacterial RNA polymerase with rifampin (200 µg/ml). Crude extracts of labeled cells separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (10) contained two radiolabeled proteins with molecular masses of 25 and 30 kDa, as expected for the β - and α -subunits of L-serine dehydratase, respectively (Fig. 3). Scanning of the two bands by a phosphorimager showed the α -subunit to be labeled four times more brightly than the β -subunit. This result can be partially explained by the fact that the α -subunit contains three times the amount of methionines (12 methionines) as the β subunit (4 methionines). Therefore, in E. coli, sdhA and sdhB are expressed in a proportion close to 1:1.

In order to facilitate purification of the enzyme, the L-serine dehydratase was overproduced in E. coli XL1-Blue (18) as a protein in which a His₆ tag was attached to the N terminus of the β -subunit by cloning the L-serine dehydratase genes into vector pQE30 (Qiagen, Hilden, Germany), yielding pST1. E. coli cells carrying pST1 were grown aerobically at 37°C in Luria-Bertani medium (18) in the presence of ampicillin (100 μ g/ml) and tetracycline (15 μ g/ml) to an optical density at 578 nm of 1.0, induced by 1 mM isopropylthiogalactoside for at least 1 h before harvesting. Since the enzyme is inactivated by air, all following steps were performed under an atmosphere of 95% N₂-5% H₂. Cells were sonicated in 50 mM potassium phosphate buffer, pH 8.0, containing 300 mM NaCl and centrifuged at 90,000 \times g for 30 min. As measured by standard assay, the supernatant had a specific activity of 233 nkat/mg, which was 16-fold higher than that determined in extracts of P. asaccharolyticus cells (5). The His₆ affinity tag allowed the purification of the modified enzyme by affinity chromatography on Ni-NTA (nickel-nitrilotriacetate) resin according the protocol of the QIA express (Qiagen) system (elution of protein from the resin by a step gradient from 50 to 500 mM imidazole in 50 mM potassium phosphate, pH 7.0, containing 5 mM glycine). Unfortunately, copurification of the α -subunit, which contained no His tag, did not occur in the required 1:1 proportion. Nevertheless, an active enzyme with a specific activity comparable to that of the wild-type enzyme (His₆-SDH, 4.8 μ kat/mg; wild type, 6.0 μ kat/mg) (5) with a purification factor of 20 and a yield of 5% was obtained.

Fusion of *sdhB* with *sdhA* and expression in *E. coli*. Since *sdhB* and *sdhA* occur in the same reading frame, directly adjacent to each other (Fig. 1), and since the derived amino acid sequence of both genes together shows significant similarity to L-SD1 and L-SD2 from *E. coli*, each of which consists of only one subunit, we wanted to investigate the effect of a fusion of



FIG. 3. Autoradiogram after expression of *sdhBA* in *E. coli* in the presence of [³⁵S]methionine. Lane 1, *E. coli* harboring the "empty" vector (negative control); lane 2, *E. coli* harboring the vector with the two structural genes of the L-serine dehydratase (both subunits [$\alpha = 30$ kDa, $\beta = 25$ kDa] can be detected as labeled protein bands); lane 3, marker proteins with sizes in kDa.

sdhB to sdhA. The stop codon (TAG) was thus changed to a glycine codon (GGT) by site-directed mutagenesis following the protocol of the Sculptor in vitro mutagenesis system (Amersham, Braunschweig, Germany). The resulting fusion protein was overproduced in E. coli XL1-Blue (18) carrying six histidine residues at the N terminus. Under denaturing purification conditions, according to the protocol of the QIA express system, the fusion protein was detected by SDS-PAGE. Anaerobic extracts of aerobically grown E. coli cells harboring the sdhBA-containing plasmid with the mutation (pST2) had a specific L-serine dehydratase activity of 233 nkat/mg. The purification of the His₆ fusion protein under nondenaturing conditions was not as successful as that of the His₆-L-serine dehydratase, probably because of weak binding to the Ni-NTA resin. It was purified up to a specific activity of 1.8 µkat/mg, but clear detection of the protein by SDS-PAGE was not possible. Like the β -His₆-L-serine dehydratase, the His₆ fusion protein was inactivated by air. The half-life observed by air inactivation of the fusion protein (2 to 3 h) was somewhat shorter than that of the β -His₆-protein and of the wild type enzyme (6 h) (5). Both modified L-serine dehydratases were stable under anaerobic conditions for at least 10 h.

Functional complementation of an E. coli mutant deficient in L-SD1. In order to examine whether the [4Fe-4S] cluster containing L-serine dehydratase from P. asaccharolyticus can functionally replace one of the corresponding enzymes from E. coli under physiological conditions, an expression vector carrying sdhBA (pST1) was introduced into E. coli MEW22 (22). This strain is deficient in L-SD1 and can be distinguished from the parent strain MEW1 (14) by its inability to grow on a minimal medium with L-serine, glycine, and L-leucine as the major carbon sources (SGL medium) (22). Synthesis of β-His₆-L-serine dehydratase from P. asaccharolyticus in MEW22 restored the mutant's ability to form visible colonies on solid SGL medium within 4 days, whereas MEW22 cells harboring the plasmid that lacked *sdhBA* did not grow in the same period. Nevertheless, the growth rate of the complemented MEW22 was lower than that of its parent strain, MEW1, which formed visible colonies on SGL medium within 1 day. To confirm that complementation was not due to MEW22 mutations, the plasmid carrying sdhBA was isolated and retransformed in MEW22. Growth studies with these retransformants showed the same results.

Discussion. We demonstrated that the [4Fe-4S] cluster containing L-serine dehydratase from *P. asaccharolyticus* is structurally and functionally related to the two L-serine dehydratases, L-SD1 and L-SD2, present in *E. coli*. Furthermore, Lserine dehydratase from *P. asaccharolyticus* was shown to complement an *E. coli* mutant deficient for L-SD1 in its ability to grow on a minimal medium containing L-serine, glycine, and L-leucine as the only major carbon sources.

The α -subunit of L-serine dehydratase from *P. asaccharolyticus* contains four conserved cysteine residues, three of which may be involved in coordinating three irons of the [4Fe-4S] cluster. Interestingly, the conserved cysteine residues followed spacings found neither in the ferredoxins nor in aconitase (23). This finding was unexpected, since previous electron paramagnetic resonance spectroscopic investigations of L-serine dehydratase from *P. asaccharolyticus* had suggested a mechanism of catalysis similar to the one elucidated for aconitase (8).

The structural genes for the α - and β -subunits of L-serine dehydratase from *P. asaccharolyticus* occur in the same reading frame, directly adjacent to each other. Together with the sequence homology with L-SD1 and L-SD2 from *E. coli*, both of which consist of only one subunit, this finding suggested that the single subunit of the enzymes from *E. coli* might have

resulted from a gene fusion during the course of evolution. In support of this view, an active recombinant fusion protein of L-serine dehydratase from *P. asaccharolyticus* was created.

The iron-sulfur-dependent mechanism of deamination of L-serine has long escaped recognition due to the unexpected oxygen lability of the involved enzymes. The described sequence similarities between the enterobacterial and clostridial enzymes support our earlier speculation (6) that in Bacteria, L-serine is dehydrated in general by the aid of an iron-sulfur cluster, whereas L-threonine requires pyridoxal-5'-phosphate for its deamination to 2-oxobutyrate in an otherwise completely analogous reaction. The question of whether the two different mechanisms evolved just by chance or as a result of a chemical necessity arises (13). We favor the latter explanation since, due to the hyperconjugative effect of the methyl group, the hydroxyl group is much more easily eliminated from threonine than from serine. Hence, in order to get the most efficient enzymes in serine, the hydroxyl group has to be activated, whereas lowering the pK of the α -hydrogen in threonine by forming a Schiff's base with pyridoxal-5'-phosphate may optimize the catalytic power of L-threonine dehydratases. In mammals, in which L-serine also is deaminated by pyridoxal-5'phosphate-dependent enzymes, several iron-sulfur-dependent dehydratases (e.g., the two extant iron-sulfur-containing fumarases found in E. coli) (4) have apparently been abandoned during evolution, due probably to their oxygen sensitivity. An apparent exception to these considerations is represented by the pyridoxal-5'-phosphate-dependent D-serine dehydratase from E. coli. The function of this enzyme, however, has been attributed to detoxification rather than to a rate-limiting step in energy metabolism (11).

Nucleotide sequence accession number. The nucleotide sequence for sdhA and sdhB has been deposited with GenBank under accession no. U76260.

This work was supported by grants from the Deutsche Forschungsgemeinschaft and the Fond der Chemischen Industrie.

We are indebted to E. B. Newman (Concordia University, Montreal, Canada) for the gift of the two *E. coli* strains, MEW1 and MEW22.

REFERENCES

- Casanova, J.-L., C. Pannetier, C. Jaulin, and P. Kourilski. 1990. Optimal conditions for directly sequencing double-stranded PCR products with Sequenase. Nucleic Acids Res. 18:4028.
- Collins, M. D., P. A. Lawson, A. Willems, J. J. Cordoba, J. Fernandez-Garayzabal, P. Garcia, J. Cai, H. Hippe, and J. A. E. Farrow. 1994. The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. Int. J. Syst. Bacteriol. 44:812–826.
- Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. Tomb, B. A. Dougherty, J. M. Merrick, K. McKennedy, G. Sutton, W. FitzHugh, C. Fields, J. D. Gocayne, J. Scott, R. Shirley, L. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Philips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nguyen, T. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrmann, N. S. M. Geoghagen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Fraser, H. O. Smith, and J. C. Venter. 1995. Whole genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269:496– 512.
- Flint, D. H., M. H. Emptage, and J. R. Guest. 1992. Fumarase A from Escherichia coli: purification and characterization as an iron-sulfur cluster containing enzyme. Biochemistry 31:10331–10337.
- Grabowski, R., and W. Buckel. 1991. Purification and properties of an ironsulfur-containing and pyridoxal-phosphate-independent L-serine dehydratase from *Peptostreptococcus asaccharolyticus*. Eur. J. Biochem. 199:89– 94
- Grabowski, R., A. E. M. Hofmeister, and W. Buckel. 1993. Bacterial L-serine dehydratases: a new family of enzymes containing iron-sulfur clusters. Trends Biochem. Sci. 18:297–300.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166:557.
- 8. Hofmeister, A. E. M., S. P. J. Albracht, and W. Buckel. 1994. Iron-sulfur

cluster-containing L-serine dehydratase from *Peptostreptococcus asaccharolyticus*: correlation of the cluster type with enzymatic activity. FEBS Lett. **351**:416–418.

- Hofmeister, A. E. M., R. Grabowski, and W. Buckel. 1994. Iron-sulfur clusters as alternatives to pyridoxal-5'-phosphate in bacterial L-serine dehydratases: cloning of the gene encoding the β-subunit of the enzyme from *Peptostreptococcus asaccharolyticus*, p. 217–221. *In* G. Marino, G. Sannia, and F. Bossa (ed.), Biochemistry of vitamin B₆ and PQQ. Birkhäuser Verlag, Basel, Switzerland.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685.
- Marceau, M., S. D. Lewis, and J. A. Schafer. 1988. The glycine-rich region of Escherichia coli D-serine dehydratase; altered interactions with pyridoxal 5'-phosphate produced by substitution of aspartic acid for glycine. J. Biol. Chem. 263:16934–16941.
- Marmur, J. 1961. A procedure for isolation of deoxyribonucleic acid from microorganisms. J. Mol. Biol. 3:208–218.
- 13. Monod, J. 1972. Zufall und Notwendigkeit, 4th ed. Piper, Munich, Federal Republic of Germany.
- Newman, E. B., D. Dumont, and C. Walker. 1985. In vitro and in vivo activation of L-serine deaminase in *Escherichia coli* K-12. J. Bacteriol. 162: 1270–1275.
- 15. Ogawa, H., K. Konishi, and M. Fujioka. 1989. The peptide sequences near the bound pyridoxal phosphate are conserved in serine dehydratases from rat

liver, and threonine dehydratases from yeast and *Escherichia coli*. Biochim. Biophys. Acta **139**:139–141.

- Ogawa, H., T. Gomi, K. Konishi, T. Date, H. Nakashima, K. Nose, Y. Matsuda, C. Peraino, H. C. Pitot, and M. Fujioka. 1989. Human liver serine dehydratase, cDNA cloning and sequence homology with hydroxyamino acid dehydratase from other sources. J. Biol. Chem. 264:15818–15823.
- Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491.
- 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.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467.
- Shao, Z., and E. B. Newman. 1993. Sequencing and characterization of the sdaB gene from Escherichia coli K-12. Eur. J. Biochem. 212:777–784.
- Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorf. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. 185:60–89.
- Su, H., B. F. Lang, and E. B. Newman. 1989. L-Serine degradation in *Escherichia coli* K-12: cloning and sequencing of the *sdaA* gene. J. Bacteriol. 171:5095–5102.
- Zeng, L., P. C. Andrew, M. A. Hermodson, J. E. Dixon, and H. Zalkin. 1990. Cloning and structural characterization of porcine heart aconitase. J. Biol. Chem. 265:2814–2821.