Homology of aspartyl- and lysyl-tRNA synthetases

ALEXANDRA GAMPEL AND ALEXANDER TZAGOLOFF

Department of Biological Sciences, Columbia University, New York, NY 10027

Communicated by John Abelson, April 19, 1989

The yeast nuclear gene MSD1 coding for ABSTRACT mitochondrial aspartyl-tRNA synthetase has been cloned and sequenced. The identity of the gene is confirmed by the following evidence. (i) The primary structure of the protein derived from the gene sequence is similar to that of the yeast cytoplasmic aspartyl-tRNA synthetase. (ii) In situ disruption of MSD1 in a respiratory-competent haploid strain of yeast induces a pleiotropic phenotype consistent with a lesion in mitochondrial protein synthesis. (iii) Mitochondria from a mutant with a disrupted chromosomal copy of MSD1 are unable to acylate mitochondrial aspartyl-tRNA. The primary structures of the cytoplasmic and mitochondrial aspartyl-tRNA synthetases are similar to the yeast cytoplasmic lysyl-tRNA synthetase, suggesting that the two types of synthetases may have a common evolutionary origin. Searches of the current protein banks also have revealed a high degree of sequence similarity of the lysyl-tRNA synthetase to the product of the Escherichia coli herC gene and to the partial sequence of a protein encoded by an unidentified reading frame located adjacent to the E. coli frdA gene. Based on the sequence similarities and the map positions of the *herC* and *frdA* loci, we propose *herC* to be the structural gene of the constitutively expressed lysyl-tRNA synthetase of E. coli and the unidentified reading frame to be the structural gene of the heat-inducible lysyl-tRNA synthetase.

Most components of the mitochondrial protein synthetic machinery of yeast are encoded in nuclear genes and are imported into the matrix compartment where they engage in translating a small number of endogenous mRNAs. The exceptions are the tRNAs, the two ribosomal RNAs, and a single ribosomal protein, all of whose genes are located in mitochondrial DNA (1). To enlarge our current understanding of how the mitochondrial and other translational systems have diverged to accommodate the idiosyncracies of their respective genetic contexts, we have taken advantage of a collection of nuclear respiratory-defective nuclear gene mutants (*pet*) of *Saccharomyces cerevisiae* to clone and characterize the genes for mitochondrial aminoacyl-tRNA synthetases, ribosomal proteins, and various protein factors involved in chain initiation and elongation.

The aminoacyl-tRNA synthetase genes are of particular interest. Some of the mitochondrial enzymes are encoded by nuclear genes distinct from those coding for the corresponding cytoplasmic enzymes (2–5). There are exceptions, however. The histidyl- (6) and valyl-tRNA synthetase (7) genes code for both the mitochondrial and cytoplasmic enzymes. The factors determining whether a single synthetase can perform efficiently in two different subcellular compartments are not clear at present. Aminoacyl-tRNA synthetases are also of considerable interest from the standpoint of the evolution of functionally related enzymes. The present day diversity of the aminoacyl-tRNA synthetase family may have arisen as a result of duplication of a smaller number of genes. Although primary structure comparisons have revealed common sequence motifs among different bacterial and yeast synthetases, they are confined to short domains (8–10). Recently, however, more extensive sequence similarities have been noted among the leucyl- valyl-, isoleucyl- and methionyl-tRNA synthetases, suggesting that this group may share a common evolutionary history (11, 12).

In this communication we report the properties of mutants defective in the aspartyl-tRNA synthetase (EC 6.1.1.12) of yeast mitochondria. The wild-type MSD1 gene coding for the mitochondrial aspartyl-tRNA synthetase has been cloned by transformation of an *msd1* mutant. The amino acid sequence similarities of the yeast mitochondrial and cytoplasmic aspartyl-tRNA synthetases (13) and the recently described cytoplasmic lysyl-tRNA synthetase (EC 6.1.1.6; ref. 14) are sufficiently extensive to suggest a common evolutionary origin. These enzymes, which acylate their cognate tRNA substrates with polar amino acids, appear to comprise a subgroup of synthetases likely to have evolved through an early gene duplication event.

MATERIALS AND METHODS

Strains. E234 and E206 are noncomplementing respiratorydefective mutants assigned to the *pet* complementation group G94. Both mutants were obtained by mutagenesis of the respiratory-competent haploid strain *S. cerevisiae* D273-10B/A1 with ethyl methanesulfonate (15). The genotypes and origins of other wild-type and mutant strains are described in Table 1.

Cloning of MSD1. The *msd1* mutant E234/L1 was grown in 10 ml of medium containing 2% galactose, 1% yeast extract, 2% peptone (YPGal medium) and transformed by the procedure of Beggs (19) with 5 μ g of a recombinant plasmid library consisting of partial Sau3A fragments (5–15 kb) of nuclear DNA of S. cerevisiae D273-10B/A1 ligated to the BamHI cloning site of the yeast/Escherichia coli shuttle vector YEp13 (20).

Aminoacylation Assays. The parental respiratory-competent haploid strain W303-1B and the mutant strain W303- ∇ MSD1 containing the disrupted allele MSD1::URA3 were grown to stationary phase in YPGa1 medium, and mitochondria were isolated by the procedure of Faye et al. (21) except that Zymolase 20,000 (Miles) instead of Glusulase was used for conversion of cells to spheroplasts. Mitochondria were lysed by addition of a 20% (wt/vol) solution of potassium cholate (pH 7.5) to a final concentration of 0.5% and centrifuged at 105,000 \times g_{av} for 30 min. The supernatant was dialyzed for several hours against 10 mM Tris chloride, pH 7.5/0.1 mM EDTA and used as a source of mitochondrial aminoacyl-tRNA synthetases. Total mitochondrial tRNAs were extracted and purified from wild-type yeast mitochondria and were acylated with a mixture of [³H]aspartic acid (200 mCi/mmol, Amersham; 1 Ci = 37 GBq) and [³H]serine (150 mCi/mmol, Amersham). The acylated products were analyzed by reverse-phase chromatography on RPC-5 (22) under previously described conditions (23).

Miscellaneous Procedures. Standard procedures were used for plasmid amplification in E. coli, digestion of DNA with

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. \$1734 solely to indicate this fact.

Table 1. Genotypes and sources of S. cerevisiae strains

Strain	Genotype	Source Ref. 16	
D273-10B/A1	$\alpha, \rho^+, met \delta$		
CB11	$a, \rho^+, adel$	Ref. 17	
W303-1B	α,ρ ⁺ ,ade2-1,his3-11,15,leu2-3, 112,ura3-1,trp1-1,can1-100	R. Rothstein*	
LL20	$\alpha, \rho^+, leu2-3, 112, his3-11, 15$	Ref. 18	
E234	α, ρ^+ met 6, ms d1-1	This study	
E206	α, ρ^+ met 6, msd 1	This study	
aE234	a,p ⁺ ade1,msd1-1	E234 × CB11	
E234/L1	$\alpha, \rho^+ leu 2-3, 112, msd 1-1$	$aE234 \times LL20$	
W303∇MSD1	$\alpha, \rho^{\circ}, ade2-1, his3-11, 15, leu2-3,$	This study	
	112,ura3-1,trp1-1,can1-100,		
	MSD1::URA3		

*Department of Human Genetics, Columbia University, New York.

restriction endonucleases, isolation and ligation of DNA fragments, 5'-end-labeling of DNA, and nick-translation (24). The procedures for Southern hybridizations and preparation of yeast nuclear DNA have been described (25). The chemical derivatization method of Maxam and Gilbert (26) was used for DNA sequencing. All of the sequences were obtained from single-stranded DNA restriction fragments labeled at the 5' end with $[\gamma^{-32}P]ATP$ (≈ 3000 Ci/mmol; ICN).

RESULTS

Phenotype of *msd1* **Mutants.** E234 and E206 are independent isolates in a large collection of *pet* mutants selected for lack of growth on the nonfermentable substrate glycerol. The two mutants do not complement one another but are complemented by a ρ^{0} tester (respiratory-deficient mutant of yeast lacking mitochondrial DNA), indicating that each has a recessive mutation in a nuclear gene necessary for oxidative metabolism. The mutant mitochondria are deficient in cytochromes *a*, *a*₃, and *b*. This phenotype is usually displayed by *pet* mutants impaired in mitochondrial protein synthesis (2–5, 11). A defect in translation was confirmed by the inability of E234 to incorporate [³⁵S]methionine into mitochondrial translation products in the presence of cycloheximide (data not shown).

Acylation assays suggested that the block in protein synthesis might be caused by a mutation in the mitochondrial aspartyl-tRNA synthetase (data not shown). However, the results were not clear. Extracts of E234 mitochondria were able to charge the aspartyl-tRNA, albeit not as efficiently as wild-type extracts. This is probably a consequence of the high concentrations of substrates used in the *in vitro* assays, which can compensate for mutations affecting changes in the K_m of the mutant enzyme for the amino acid, tRNA, or ATP (2, 4). To ascertain if G94 mutants have mutations in the mitochondrial aspartyl-tRNA synthetase gene, the gene was cloned and characterized as detailed in the following sections.

Cloning and Sequence* of the *MSD1* Gene. E234/L1, an *msd1* mutant with a *leu2* marker, was transformed with a plasmid library of wild-type yeast DNA cloned in the *LEU2*bearing shuttle vector YEp13 (20). Two respiratory-competent clones were obtained, one of which was also complemented for the leucine auxotrophy (E234/L1/T1). The respiratory-competent phenotype of E234/L1/T1 was ascertained by segregation tests to be a function of an autonomously replicating plasmid. The complementing plasmid (pG94/T1) was isolated from the yeast transformant and amplified in *E. coli*. The restriction map of pG94/T1 indicated the presence of a 10.7-kilobase (kb) nuclear DNA insert (Fig. 1).



FIG. 1. Restriction maps of pG94/T1 and of derivative plasmids. The locations of the Sph I (S), EcoRI (E), and HindIII (H) sites in the 10.7-kb insert of pG94/T1 are indicated. The reading frame corresponding to MSDI is depicted by the solid bar. The different regions cloned in the shuttle vector YEp351 are indicated by the lines above the map. Complementation of the respiratory deficiency of E234/L1 is indicated by the plus sign, and lack of complementation, by the minus signs.

To localize the gene, different regions of the pG94/T1 insert were transferred to the shuttle vector YEp351 (27), and the derivative plasmids were tested for their ability to restore growth of E234/L1 on glycerol. The results of these transformations are summarized in Fig. 1. The smallest plasmid capable of complementing E234/L1 (pG94/ST4) had a 2.4-kb DNA fragment defined by the *Sph* I site and the end of the pG94/T1 insert. Further dissection of this region resulted in loss of complementation. The inability of either pG94/ST7 or pG94/ST8 to complement the *pet* mutant suggested that the gene spans the two closely spaced *Hind*III sites of pG94/ST4.

The DNA sequence of the insert in pG94/ST4 and the 5' flanking sequence present in pG94/ST3 were determined by the method of Maxam and Gilbert (26). All of the restriction sites used for 5'-end-labeling were crossed from adjacent sites, and most of the sequence was verified from the complementary strands. Analysis of the sequence disclosed a single long open reading frame initiated by an ATG codon 41 nucleotides downstream of the *Sph* I site. The reading frame is 1974 nucleotides long and ends with an opal terminator. The predicted protein consists of 658 amino acids and has a calculated molecular weight of 75,472. The aminoterminal 17 residues are rich in basic and hydroxylated amino acids suggestive of a mitochondrial import signal.

The alignment of the amino acid sequence deduced from MSD1 with the sequence of the yeast cytoplasmic aspartyltRNA synthetase shown in Fig. 2 was obtained by the MFALGO program (30). The two proteins share 21% identical amino acid residues. Regions with substantially greater sequence conservation are located in the middle (residues 150-286 of the mitochondrial enzymes and residues 233-365 of the cytoplasmic enzymes) and at the carboxyl-terminal (residues 545-625 in the mitochondrial enzymes and residues 471-548 in the cytoplasmic enzymes) parts of the protein where there are 29% and 30% identical residues, respectively. The sequence similarity indicates that the two proteins are homologous and suggests that MSD1 is the structural gene for the yeast mitochondrial aspartyl-tRNA synthetase. In spite of their homology, the two yeast aspartyl-tRNA synthetases have diverged significantly as evidenced by the numerous gaps needed to align the two sequences.

^{*}The sequence reported in this paper has been deposited in the GenBank data base (accession no. M26020).

APS KRS1 herC	M S O D E M I V KALVEE SA BE PAQVILGE DG KPLSK KALKK LOKKE O E KORKE IE IR M S O O D N V KAAAEG VA NILHEDE ATGEN VSKSELKK RIKOROVEAKKAAK M SEOHAQGADAV V D L NNELKT RREKLANLR	50 48 30
MSD1	R L ÁD F ÞE A NA I I K K K F L F R L D T S T I K Q L K G L S	45
APS	A L ÓL E <u>Á E I</u> R E I A R E K K A A A E D T A I K D N Y G K L P L	80
KRS1	K Á A A Ó P K PIA S K K K I T D L F A D L D P S Q Y F E T R S R Ó I Ó E L R K T H E P N P Y P H K F	97
herC	E ÓG I Í ÁF PIND F R R D H T S D Ó L H ÁB E F D	54
MSD1	SGQKI [V L NGWIEQKPK RVGK NL IF	69
APS	IQSRDSDRTGQKRVKFVDL DE AKDSDKEVLFRARVHN TRQQGATLAF	127
KRS1	HVSISNPEFLAKYAHLKKGETLPEEK	142
herC	GKENEELEALNIE VAVAGRMMTRRIMGKASF	85
MSD1	GLLRDSNGDITO LVDNKSLLKGFTLEDVVDAVGILSLK	107
APS	LTLRDQASLIOGLV <u>KANKEG</u> TISKNMVKWAGSLN <u>LESIVL</u> VRGIV	172
KRS1	VVLHGDGVEVQLM SOLODVCDPDSVEKDDHDLLKRGDIVGVJEGVUGR	188
herC	VTLQDV GGRTQLV VARDDLPE GVVNEGFKKWDLGDJLGAKGKLFK	130
MSD1	RKLSNEDADEYEVQLEDITVLNASNK	133
APS	KKVDEPIKSATVQNLEIHITKIYTISETPEALPIILLED ASRSEAE	217
KRS1	TQP <u>KK</u> GGEGEVSVFVV <u>SR</u> VQLLTPCL HM	215
herC	T KTGELSIHC TELRLLTKAL RP	152
MSD1	KPA QMQDFKLISALIYPPEFRYLLQL RNPKYDDFLKKRSISIISKEI RNSFINNF	182
APS	AEAAGLPVVNLDTRL DYRLYIDL RTVTNDAIFRIGAGYCELFREYLAITK	265
KRS1	LPADHFGFKDQETRYRKRYLDLIMNNKDARNRFI TRSEIIRYIRRFLDQR	264
herC	LPDKFHGLQDQEARYR\RQRYLDLISNDEISRNRFI KVRSIQIILSGIROFFMYNR	201
gX	MSETASWQPSALSIPNLLK RAAIMAEIRRFFADR	33
MSD1 APS KRS1 herC gX	DFTEVETPMLFKATPEGAREFLVPTRITKRSDGKPSFYA LDOSPOOVKOL KFTEVENTPKLLGADPSEGGSSVFEVITYFKI GKAY LAOSPOOFNKOO KFTEVETPMM NVIAGGATA KPFTTHHNDLDNDMY MRIAPELFLKO GEMEVETPMM OVIPGGAAA RPFTTHHNDLDNDMY LRIAPELFLKO GEMEVETPMM OVIPGGAAA RPFTTHHNDLDLDMY LRIAPELYLKR GVLEVETPCMSQATVTDIHLVPFETRFVGPGHSQGMNEWLMTSPEYHMKR	231 308 309 246 83
MSD1	LMAASGYN KYYQQMA (RICFR) DED LRAD ROPEFTO YD MEMAFAN SE DYMKII	279
APS	LIIVADFER VYEIG (PYFRALENSN THRHMT) EFT (GLDMEMALFEEHY) HEVLDTL	358
KRS1	LVYG GLDN VYEIG ROVFRALEDD MT HNPEFTT (GETYQAYADIVY) (DLMDMT)	357
herC	LYYG GFER YFEI (MRINFRN 5GI) SYR HNPEFTM HELLYMAYADIYK DLLIELT	294
gX	LLYYAG CGPVF[0] LCRSFRNEJEMGRY HNPEFTM LEWYRPHYDMY RLMN EV	131
MSD1 APS KRS1 herC gX	EK T YSGY) S ELFYFI ELMFSEM V KEITGS Y LIKYH P D P A D P A K E LELN KELDKRFAH E IE LV R ELMFSEM V KEITGS Y LIKYH P D P A D P A K E LELN F <u>SRPH KR</u> INMIE <u>E L</u> D LLLQQYLDC P A A E SESYQQA D DLLQQYLDC P A A E SESYQQA	301 381 404 332 152
MSD1	KIG T LVDPAKKIEN G T VSI FRIMTYEQAM T SYGIDK POLRAPDLKIIIN LGEFIN	350
APS	KIQYPVEEFKLPKDGKMVRLT YKEGIEMLRAAGKEI GDFE	420
KRS1	DEKVFNVKFPSGDQLH TAETGEFLKKIILVDN KLEC	438
herC	I KKYRPETDMA DLDNFDSAKAIAESI GIHV	362
MSD1	AFSHLMK <mark>KFPYFE</mark> VIILRSAFSNMEEYKERWSFLTNNSNYNYRVPIVLP	399
APS	DL <u>IS</u> TEME <u>IKFLG</u> KLV	434
KRS1	PPPLTMARMLDDKLV	452
herC	EKSWG LGRTVTEIF	376
MSD1	I E N D E Q A N S N W F E N F H A I A T F E N P H L I T K F L K L K K G D I V C G C T R E P N H S	448
MSD1	I F E N P T P L G R L R Q L V L Q S E H G K N I Y H A V N K D V A S W I V D F P L F S P V I I E D K	498
MSD1	SGKK <u>IE</u> KL AYPEYJEKD RLCSTHHPFITMVKLKDYEKLEKTPEKCLGR	543
APS	RDKJY DTDFYJILD KFPLEIRPFYTMPDPANPK YSN	468
KRS1	GIELE DTCIINPTFIFGHPIQMMSPLAKYJSRDQPGLCE RFE	490
herC	EEVAEAHLIQPFFITEYPAEV <u>SPLA</u> RRNDVNPEITD RFE	415
MSD1	HY DL V VINGVELGGGSTRIHDPRLQDYIFEDILKIDNAYELFGHLLN	589
APS	SY <u>DFFMRGE</u> ELGGG <u>AQRIHDHAL</u> LQERMKAHG <u>LSPE</u> DPGL <u>KDYCD</u>	513
KRS1	VFVATKETICNAYTELNDPFDQRARFEEQQABQGKDQGDDEAQLVD	538
herC	FFIGGREIGNGFSELNDAEDQAQRFLDQVAAKDAGDDEAMFYD EDYVT	463
MSD1	AFDMGTPPHAGFAJGFDRMCAMICETESJRDVIAFPKSITGADLVVKSPS	639
APS	GFJSYGCCPPHAGGGIGLERVVMFYLDLKNIRRASLFPRDPKRLRP	557
KRS1	ALEYGLPPTGGWGCGIDRLAMFLTDSNTTREVLLFPTLKPDVLREEVKKE	588
herC	ALEHGLPPTAGLGIGIDRWVMLFTNSHTIRDVILFPAMRPVK	505
MSD1	VIPESILEPYNIKYSNSKK	658
KRS1	EEN	591

FIG. 2. Alignment of aspartyl- and lysyl-tRNA synthetase sequences in single-letter code. The proteins encoded by MSD1 (yeast mitochondrial aspartyl-tRNA synthetase), APS [yeast cytoplasmic aspartyl-tRNA synthetase (13)], KRS1 [yeast cytoplasmic lysyl-tRNA synthetase (14)], herC (28), and gX (29) were aligned by the MFALGO program (30) and further adjusted by eye for the maximal number of identical or conserved residues among the five sequences. Identical amino acids between two or more sequences are boxed.

Disruption of MSD1. The reading frame coding for the mitochondrial aspartyl-tRNA synthetase was disrupted with the yeast URA3 gene to create the MSD1::URA3 allele (Fig. 3). The one-step gene disruption procedure (31) was used to substitute the MSD1::URA3 mutant gene for the wild-type gene in the respiratory-competent strain W303-1B. Respiratory-deficient, uracil-independent transformants were verified by back-crosses to aE234 to have acquired the mutant allele. The respiratory-deficient transformants with mutations allelic to msd1 were not complemented by a ρ° tester, indicating that they had sustained mutations (deletions) in mitochondrial DNA. The loss of wild-type mitochondrial DNA has been shown (25) to be a secondary effect of tight

mutations in genes coding for components of the mitochondrial protein synthetic system.

More direct evidence for the presence of the MSD1::URA3allele in one of the respiratory-deficient transformants (W303 ∇ MSD1) was obtained by analysis of its genomic DNA by the Southern blot-hybridization technique (Fig. 3). A probe containing the entire MSD1 gene detected a 6.2-kb fragment in the EcoRI digest of wild-type DNA. In the mutant, the size of this band, 7.4 kb, is consistent with an insertion of the 1.2-kb fragment with the URA3 gene at the HindIII sites of MSD1. When wild-type DNA was digested with Sph I and EcoRI, the probe detected a single band of 1.7 kb. This band consisted of two different fragments which



FIG. 3. Southern analysis of W303-1B and W3037MSD1 genomic DNA. (Lower) Restriction maps of the wild-type strain W303-1B and of the mutant W303VMSD1 transformed with the mutant allele MSD1::URA3. Only the locations of the Sph I (S), EcoRI (E), and HindIII (H) sites are indicated. The MSDI gene is depicted by the solid bar, and the yeast URA3 gene is depicted by the open bar. (Upper) Total nuclear DNAs purified from W303-1B and W303- ∇ MSD1 were digested with either *Eco*RI or a combination of *Sph* I and *Eco*RI and separated by electrophoresis on a 1% agarose gel. After transfer to nitrocellulose, the blot was hybridized with a nick-translated fragment (probe) containing the entire sequence of MSD1. Lanes: 1, EcoRI digest of DNA from W303-1B; 2, EcoRI digest of DNA from W303VMSD1; 3, Sph I/EcoRI digest of DNA from W303-1B; 4, Sph I/EcoRI digest of DNA from W303VMSD1. The migrations of DNA standards are indicated in kilobases in the right-hand margin.

were indistinguishable in this gel system: a 1.7-kb Sph I-EcoRI fragment with the 5' end of the gene and a second 1.7-kb EcoRI fragment with the rest of the gene and 3' flanking sequence. As predicted, insertion of the 1.2-kb fragment with the URA3 gene at the HindIII site increased the size of the Sph I-EcoRI fragment from 1.7 kb in wild type to 2.9 kb in W303VMSD1.

Acylation Activities of Wild-Type and Mutant Mitochondria. Additional evidence that MSD1 codes for mitochondrial aspartyl-tRNA synthetase was obtained by testing the ability of wild-type and W303VMSD1 mitochondria to acylate aspartyl-tRNA. Aminoacyl-tRNA synthetases, extracted from mitochondria of W303VMSD1 and from the parental respiratory-competent strain W303-1B, were tested for their ability to charge the yeast mitochondrial aspartyl-tRNA and, as a control, servl-tRNA. The wild-type extract charged the mitochondrial aspartyl-tRNA and a small amount of contaminating cytoplasmic aspartyl-tRNA (Fig. 4). Three other radioactive peaks corresponding to the cytoplasmic servitRNA and to two mitochondrial servl-tRNA isoacceptors were also detected. The identical tRNA species, except for the mitochondrial aspartyl-tRNA, were acylated by the extract obtained from the mutant mitochondria (Fig. 4).

Homology of Aspartyl- and Lysyl-tRNA Synthetases. A three way comparison of the two aspartyl-tRNA synthetases and the lysyl-tRNA synthetase showed the three enzymes to have similar primary structures. Based on the alignments shown in Fig. 2, 23% of the residues are identical in the lysyl-tRNA synthetase and the cytoplasmic aspartyl-tRNA synthetase, whereas only 14% of the residues are identical in



FIG. 4. Acylation of aspartyl- and seryl-tRNAs by extracts of wild-type and mutant mitochondria. Mitochondrial tRNAs were acylated with a mixture of ³H]aspartic acid and [³H]serine in the presence of wild-type (W303-1B) and mutant (W303- ∇ MSD1) mitochondrial extracts, and the acylated products were analyzed by chromatography on RPC-5. The identity of the radioactive peaks corresponding to the cytoplasmic aspartyl- (cyt. asp.), mitochondrial aspartyl-(mit. asp.), cytoplasmic seryl-(cyt. ser.), and mitochondrial seryl-tRNA synthetases (mit. ser.) were established in earlier studies (G. Macino and A.T., unpublished data; compare ref. 23).

the lysyl-tRNA synthetase and the mitochondrial aspartyltRNA synthetase. The largest number of positional identities occur in the middle and carboxyl-terminal regions of the proteins. In the two regions combined there are 26% identical residues between the cytoplasmic aspartyl- and lysyl-tRNA synthetases and 24% identical residues between the mitochondrial aspartyl-tRNA synthetase and the lysyl-tRNA synthetase (Table 2). In addition, 28% of the amino-terminal 50 residues of the cytoplasmic aspartyl- and lysyl-tRNA synthetases are identical. The sequence of the amino-terminal regions of the lysyl-tRNA synthetase could not be aligned with the mitochondrial aspartyl-tRNA synthetase sequence.

Sequence Similarity of the Yeast Lysyl-tRNA Synthetase and Two Putative E. coli Gene Products. Searches of the current protein sequence data bases revealed that the yeast aspartyland lysyl-tRNA synthetases are homologous to two putative E. coli proteins: the herC gene product and a protein encoded

Table 2. Amino acid identities in the yeast aspartyl- and lysyl-tRNA synthetases and the proteins encoded by the *E. coli herC* and gX genes

	Mit Asp	Cyt Asp	Cyt Lys	HerC	GX	
Mit Asp	658	63 (30%)	51 (24%)	51 (29%)	31 (22%)	
Cyt Asp	118 (21%)	557	53 (26%)	51 (26%)	27 (20%)	
Cyt Lys	75 (14%)	128 (23%)	591	130 (60%)	38 (29%)	
HerC	80 (16%)	84 (17%)	201 (39%)	505	43 (33%)	
GX	33 (22%)	27 (15%)	37 (24%)	45 (30%)	152	

The entries on the diagonal refer to the number of residues in each protein. Entries below the diagonal refer to the number of identical residues over the entire lengths of the proteins based on the alignments in Fig. 2. Entries above the diagonal refer to the number of identical residues in the two regions defined by residues 150 through 286 and 545 through 625 of the mitochondrial aspartyl-tRNA synthetase (Mit Asp). In the case of the gX product (GX), only residues 150 through 286 have been analyzed because the sequence of the carboxyl-terminal region is not available. The percentages have been calculated on the basis of either total (below the diagonal) or the number of residues in the two regions analyzed (above the diagonal). Cyt Asp and Cyt Lys, cytoplasmic aspartyl- and lysyl-tRNA synthetases.

by an unidentified open reading frame designated gX. The herC has been shown (28) to be part of an operon that includes the gene for the translational termination factor RF2 (32). The yeast lysyl-tRNA synthetase and the protein encoded in herC have 39% identical residues (Fig. 2 and Table 2). Although the herC gene product and the yeast aspartyl-tRNA synthetases also have similar sequences, the number of positional identities is lower (16-17%).

E. coli has been reported to have two different genes for lysyl-tRNA synthetase-lysS and lysU. The lysU gene located at 92-93 min of the map codes for a heat-inducible enzyme (33). The lysS gene at 62.1 min codes for the constitutively expressed lysyl-tRNA synthetase (34). It is significant that herC has been located at 62 min (28, 32). Based on the primary sequence similarity of the herC product and the yeast lysyl-tRNA synthetase and the nearly identical map locations of herC and lysS, we propose that herC is the structural gene for the constitutive lysyl-tRNA synthetase of E. coli.

The other protein sequence retrieved from the data base is encoded by an open reading frame located at 94 min of the E. coli map next to frdA (29). Only the amino-terminal coding region of the reading frame has been sequenced. Of the 152 amino acid residues reported for this putative protein, 45 (30%) are identical to herC, 37 (24%) to the yeast lysyl-tRNA synthetase, 27 (15%) to the yeast cytoplasmic aspartyl-tRNA synthetase, and 33 (22%) to the yeast mitochondrial aspartyltRNA synthetase (Table 2). This gene could encode either the E. coli aspartyl-tRNA synthetase or the heat-inducible lysyltRNA synthetase. The structural gene of the E. coli aspartvltRNA synthetase has not been mapped. It may be significant that the location of the unidentified gene (94 min) is not too dissimilar from the location of the lysU gene (92–93 min).

DISCUSSION

Mutations in MSD1 induce a respiratory-deficient phenotype but are not lethal to yeast, indicating that the aspartyl-tRNA synthetase encoded by this gene functions exclusively in mitochondrial translation. Both the yeast mitochondrial and cytoplasmic aspartyl-tRNA synthetases are homologous to the recently described yeast cytoplasmic lysyl-tRNA synthetase (14). The primary structure similarity of the two different types of synthetases suggests that their respective genes may have evolved from a common ancestral gene. According to this evolutionary scenario of gene duplication and sequence divergence, the precursor gene could have coded for another synthetase capable of acylating both the lysyl- and aspartyltRNAs. Alternatively, the ancestral synthetase may have been capable of acylating only the lysyl-tRNA, the aspartyltRNA, or a tRNA with acceptor specificity for some other amino acid. Assuming the ancestral enzyme to have been selective in its choice of tRNA substrate, a second inference of a gene duplication event leading to the appearance of the two or more types of synthetases is that proteins were originally built from fewer amino acids. In the present example, the prediction would be that either lysine or aspartic acid was not part of the make up of earlier proteins.

Two E. coli proteins were found to be homologous to the aspartyl- and lysyl-tRNA synthetases. It seems fairly certain that the product of herC is the constitutively expressed lysyl-tRNA synthetase of E. coli. The yeast lysyl-tRNA synthetase and herC are 39% identical in overall sequence. Furthermore, lysS, one of the two genes encoding lysyltRNA synthetases of E. coli has the same map location as herC (28).

The partial sequence of the protein encoded by the unidentified reading frame adjacent to frdA also shares a large number of amino acid identities with the lysyl-tRNA synthetase. It is therefore likely that the product of this gene is the heat-inducible lysyl-tRNA synthetase of E. coli (33). Independent of its tRNA specificity, this protein is some 150-230 residues shorter at the amino-terminal end than the other four aminoacyl-tRNA synthetases, suggesting that the core structure (35) involved in the acylation activity of both types of synthetases resides in the carboxyl-terminal two-thirds of their polypeptide chains. This conclusion is also supported by the observation that the amino-terminal region is the least well conserved in the five different sequences examined.

This research was supported by Grant PCM-8116680 from the National Science Foundation.

- 1. Dujon, B. (1982) in Molecular Biology of the Yeast Saccharomyces cerevisiae: Life Cycle and Inheritance, eds. Strathern, J. N., Jones, E. W. & Broach, J. R. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 505-653.
- Koerner, T. J., Myers, A. M., Lee, S. & Tzagoloff, A. (1987) J. Biol. Chem. 262, 3690-3696.
- Pape, L. K., Koerner, T. J. & Tzagoloff, A. (1985) J. Biol. Chem. 3. 260, 15362-15370.
- 4. Myers, A. M. & Tzagoloff, A. (1985) J. Biol. Chem. 260, 15371-15377
- 5. Tzagoloff, A., Vambutas, A. & Akai, A. (1989) Eur. J. Biochem. 179, 365-371.
- Natsoulis, G., Hilger, F. & Fink, G. R. (1986) Cell 46, 145-154.
- Chatton, B., Walker, P., Ebel, J.-P., Lacroute, F. & Fasiolo, F. (1988) J. Biol. Chem. 263, 48-57
- Webster, T., Tsai, H., Kula, M., Mackie, G. A. & Schimmel, P. 8. (1984) Science 226, 1315–1317.
- Barker, D. G. & Winter, G. (1982) FEBS Lett. 145, 191-193. Q
- 10. Houtondji, C., Dessen, P. & Blanquet, S. (1986) Biochimie 68, 1071-1078
- Tzagoloff, A., Akai, A., Kurkulos, M. & Repetto, B. (1988) J. Biol. 11. Chem. 263, 850-856.
- Heck, J. D. & Hatfield, G. W. (1988) J. Biol. Chem. 263, 868-877. 12.
- Sellami, M., Fasiolo, F., Dirheimer, G., Ebel, J.-P. & Gangloff, J. 13. (1986) Nucleic Acids Res. 14, 1657-1666.
- Mirande, M. & Waller, J.-P. (1988) J. Biol. Chem. 263, 18443-18451. 14.
- Tzagoloff, A., Akai, A. & Needleman, R. B. (1975) J. Bacteriol. 15. 122, 826-831.
- Tzagoloff, A., Akai, A. & Foury, F. (1976) FEBS Lett. 65, 391-396. 16.
- ten Berge, A. M. A., Zoutewelle, G. & Needleman, R. B. (1974) Mol. Gen. Genet. 131, 113-121. 17.
- Orr-Weaver, T. L., Szostak, J. W. & Rothstein, R. J. (1981) Proc. 18. Natl. Acad. Sci. USA 78, 6354-6358.
- Beggs, J. D. (1978) Nature (London) 275, 104-109. 19.
- Broach, J. R., Strathern, J. N. & Hicks, J. B. (1979) Gene 8, 20. 121-133.
- Faye, G., Kujawa, C. & Fukuhara, H. (1974) J. Mol. Biol. 88, 21. 185-203.
- Pearson, R. L., Weiss, J. F. & Kelmers, A. D. (1971) Biochim. 22. Biophys. Acta 228, 770-774.
- Macino, G. & Tzagoloff, A. (1979) Mol. Gen. Genet. 169, 183-188. 23.
- 24. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
- Myers, A. M., Pape, L. K. & Tzagoloff, A. (1985) EMBO J. 4, 25. 2087-2092.
- Maxam, A. M. & Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 26. 560-564.
- 27. Hill, J. E., Myers, A. M., Koerner, T. J. & Tzagoloff, A. (1986) Yeast 2, 163-167.
- Kawakami, K., Jonsson, Y. H., Bjork, G. R., Ikeda, H. & Naka-28. mura, Y. (1988) Proc. Natl. Acad. Sci. USA 85, 5620-5624.
- Cole, S. T. (1982) Eur. J. Biochem. 122, 479-484. 29
- Woodbury, N. W. & Doolittle, R. F. (1980) Mol. Evol. 15, 129-148. 30.
- 31.
- Rothstein, R. J. (1983) *Methods Enzymol.* 101, 202–211. Craigen, W. J., Cook, R. G., Tate, W. P. & Caskey, C. T. (1985) 32. Proc. Natl. Acad. Sci. USA 82, 3616-3620.
- 33. Van Bogelen, R. A., Vaughn, V. & Neidhardt, F. C. (1983) J. Bacteriol. 153, 1066-1068.
- Emmerich, R. V. & Hirshfield, I. N. (1987) J. Bacteriol. 169, 34. 5311-5313
- 35. Schimmel, P. (1987) Annu. Rev. Biochem. 56, 125-158.