Isoleucyl-tRNA Synthetase from the Ciliated Protozoan *Tetrahymena thermophila*

DNA SEQUENCE, GENE REGULATION, AND LEUCINE ZIPPER MOTIFS

(Received for publication, November 15, 1991)

Csilla Csank† and Duane W. Martindale‡

From the Department of Microbiology, McGill University, Macdonald Campus, Ste Anne de Bellevue, Quebec, H9X 1C0 Canada

---

We have determined the nucleotide sequence of a protozoan aminoacyl-tRNA synthetase. The isoleucyl-tRNA synthetase (ileRS) gene (*ilsA*; formerly *cupC*) from *Tetrahymena thermophila* was sequenced and found to have eight introns, four transcription start sites, and a putative polypeptide of 1081 amino acids. A polypeptide 20 amino acids longer could be made if a transcribed in-frame ATG close to the start sites and with suboptimal sequence context is used. This gene was identified through hybridization and amino acid sequence similarity to the previously cloned and sequenced ileRS (cytoplasmic) gene from *Saccharomyces cerevisiae* [Englisch, U., Englisch, S., Markmeyer, P., Schischkoff, J., Sternbach, H., Kratzin, H., and Cramer, F. (1987) *Biol. Chem. Hoppe-Seyler* 368, 971–979; Martindale, D. W., Gu, Z. M., and Csank, C. (1989) *Curr. Genet.* 15, 99–106] with which it shares 47% of its amino acids. We also compared it to ileRS genes from *E. coli* and an archaebacterium. Two leucine zipper motifs were identified in the carboxyterminal domain of the polypeptide; one of these motifs is in the same area as the zinc finger motif found in the *E. coli* enzyme. The transcription pattern of the *ilsA* gene was monitored under various culture conditions and parallels changes in protein synthesis.

The aminocyl-tRNA synthetases (aaRSs) are the enzymes that unite nucleic acids with amino acids (reviewed in Ref. 23). Each of the 20 aaRSs catalyzes the formation of a high energy ester bond between the carboxyl group of a specific amino acid and the α-phosphate group of ATP creating an amino acid adenylate. The amino acid carboxyl group

---

*This work was supported by an Operating Grant (to D. M.) from the Natural Sciences and Engineering Research Council of Canada. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank†‡/EMBL Data Bank with accession number(s) M30942.

† Supported by a Postdoctoral Scholarship from the Natural Sciences and Engineering Research Council of Canada, a McConnell Fellowship from McGill University, and a graduate fellowship from the Desjardins Foundation (Bourse Girardin-Vaillancourt). Present address: National Research Council Canada, Biotechnology Research Inst., 6100 Ave. Royalmount, Montreal, Quebec, Canada H4P 2R2.

‡ To whom correspondence should be addressed. Tel: 514-398-7886; Fax: 514-398-7990.

1 The abbreviations used are: aaRS, aminocyl-tRNA synthetase; ileRS, isoleucyl-tRNA synthetase; kb, kilobase(s).

is then transferred to one of the hydroxyl groups of the 3′-terminal adenosine of an appropriate tRNA molecule. The fidelity of this reaction and thus of translation is maintained by accurate substrate recognition and discrimination and sometimes by an editing deacylase mechanism (20, 23). Despite common functions, the aaRSs vary in polypeptide size, amino acid sequence, and subunit composition (reviewed in Refs. 62 and 63). Comparisons of aaRS amino acid sequences and three-dimensional structures deduced by X-ray crystallography of representative aaRSs have revealed the presence of at least two classes of aaRSs, each with their own unique short sequence motifs (12, 17, 64). These two classes of enzymes, one that includes aaRSs for most hydrophobic amino acids, appear to have evolved from different ancestral enzymes (17).

Sequences have been determined of genes representing all 20 aaRSs (18, 33). These genes have been obtained mainly from the eubacterium *Escherichia coli* and from the yeast *Saccharomyces cerevisiae* (see Refs. 63, 33). Examples of sequenced aaRS genes from other organisms are limited and include examples from an insect (8), mammals (22; 35; 69), and recently, from an archaebacterium (36). To improve our understanding of aaRSs, their conserved features and their idiosyncrasies, it will be instructive to examine these enzymes and their genes from a number of species.

In this report we describe the identification and full sequence of an isoleucyl-tRNA synthetase (ileRS) gene (*ilsA*) from the ciliated protozoan *Tetrahymena thermophila*. The ciliated protozoa are organisms that diverged relatively early from the main stream of eukaryotic evolution (67). They are very complex unicellular organisms that harbor two functionally distinct nuclei (reviewed in Ref. 38). This is the first reported sequence of an aaRS gene from a protozoan and one of the few reported sequences of an enzyme-encoding gene from a ciliate. We analyze its deduced polypeptide sequence and compare it to the available ileRS sequences from *E. coli* (78), *S. cerevisiae* (49), and the archaebacterium, *Methanobacterium thermoaurotrophicum* (36). In addition, we examine its expression under different physiological conditions.

**MATERIALS AND METHODS**

**RESULTS**

Isolation—cDNA (pC8) representing the *T. thermophila ilsA* gene (formerly *cupC*; 47) was originally isolated during a study of genes active during the sexual stage (conjugation) of

2 Portions of this paper (including "Materials and Methods," Figs. 1, 2, 4, 6, and 7, and Table 1) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.
the T. thermophila life cycle (45). The pC8 cDNA was found to contain sequences of an RNA transcript present and peaking in abundance during early conjugation but also present at lower levels relative to total RNA during normal cell growth. Almost no transcript was present in total RNA from starved cells prior to conjugation.

The insert in pC8 harbors one EcoRI site and hybridized to two genomic EcoRI fragments that represent a single chromosomal location for this gene (Fig. 1 in miniprint). The ilsA gene is located on micronuclear chromosome 2 (47) and Southern analyses of micronuclear and macronuclear DNA digests indicated that no rearrangement or major sequence elimination occurs in or near this gene during differentiation of the somatic macronucleus (47). The ilsA gene (pC8) was found to hybridize under stringent conditions to genomic DNA from S. cerevisiae indicating that the sequence is evolutionarily conserved (Fig. 1).

Cloning and Sequencing—The structural organization of the ilsA gene, restriction maps, and clones used for sequencing are illustrated in Fig. 2 (see miniprint). Both genomic DNA and cDNA representing the ilsA gene were sequenced. Both strands of the coding region (98 and 92% of each, respectively) were sequenced. Regions not confirmed by the opposite strand were confirmed using the sequence of an independent clone. The complete nucleotide sequence and deduced amino acid sequence of the ilsA gene are shown in Fig. 3. The ilsA gene spans approximately 4500 nucleotides, from transcription start sites to the polyadenylation site, and is composed of nine exons and eight introns. The ilsA gene's introns have an average G+C content of 15%. This is much lower than the average G+C of its coding regions (38%).

Transcription Start Sites—Transcription start sites were deduced by primer extension analysis of polyadenylated mRNA (Fig. 4 in miniprint) using a primer complementary to nucleotides +8 to +25 of the ilsA gene's sense strand (see Fig. 3). Three major start sites (positions −60, −66, and −70; underlined in Fig. 3) and a minor start site (position −72; also underlined in Fig. 3) were found (Fig. 4). The same start sites were used when the gene was transcribed during normal logarithmic growth and during conjugation (Fig. 4). The most complete cDNA clone went to position −31 (Fig. 3) and no 3′ intron acceptor dinucleotide (AG) is present between position −31 and the −72 start site.

Open Reading Frame Analysis—The ilsA gene has a long open reading frame of 3243 nucleotides (Fig. 3). The open reading frame ends at a TGA and, as in several other ciliates, TAA and TAG code for glutamine instead of termination (26, 32).

The transcription start site analysis reveals that the first possible translation initiation site (an in-frame ATG at −60) is located very close to the transcription start sites. Indeed, the most abundant transcription product of the ilsA gene begins at the A residue of this putative start codon. It is only six bases and ten bases downstream of the two other abundant transcription start sites, is 12 bases downstream of the minor mRNA start site, and probably does not support efficient translation (see "Discussion").

The open reading frame originating from the ATG at +1 encodes a deduced polypeptide of 1081 amino acids (Fig. 3) with a putative molecular mass of 124 kDa and a pI of 6.3. The frequency of preferred codons (Fpr) within the ilsA gene was computed (48). The Fpr of this gene (0.85) is slightly lower than that of the highly expressed ciliate histone and actin genes (Fpr = 0.90; 48). If translation occurs from the ATG at −60, the putative polypeptide is 20 amino acids longer. The region separating the first and second start codons does not have a codon bias typical of T. thermophila genes (Fpr of 0.47; 48). The 1100 nucleotides of DNA upstream of the ilsA start sites (Fig. 3) is devoid of open reading frames with appropriate codon usage, and has a high A+T content (75%) typical of T. thermophila intergenic regions (31).

Gene Expression—When T. thermophila cells that have been growing exponentially in a rich medium are washed and resuspended in starvation buffer (10 mM Tris, pH 7.4), the rates of RNA and protein synthesis rapidly decline, and by 2 h are reduced to 6% or less of their original rates (46). Polysomes also dissociate (24), and most cells complete their cell division cycles and become arrested at macronuclear G1 phase (51). As RNA and protein synthesis decline, the level of the ilsA mRNA product decreases (45).

When different mating types of starved T. thermophila cells are mixed they undergo synchronous mating (conjugation). Soon after cells pair, micronuclei begin to undergo meiosis. A burst in RNA synthesis occurs almost immediately upon mixing cells, and this is followed by a rise in the rate of protein synthesis that peaks during meiotic prophase (46). Polyome formation is also induced (46). After the period corresponding to meiotic prophase, protein synthesis rates are again reduced. Several conjugation-specific genes are transcribed maximally during meiotic prophase, as is the conjugation-induced ilsA gene (45). Primer extension analysis confirms previous findings (45) that ilsA transcripts make up a larger fraction of the polyadenylated RNA during early conjugation (3–5 h) than during normal growth.3

We wanted to examine whether ilsA gene expression is coordinately increased with protein synthesis in a situation other than meiotic prophase. We chose to examine ilsA gene expression during nutritional shift-up. When starved T. thermophila cells are refed, most mRNA is loaded onto polysomes within 2 h (15, 24) and a rapid induction of protein synthesis occurs (7). About 40% of the protein synthesis occurring in the first hour after refeeding is that of ribosomal proteins (15, 24). This is paralleled by an increase in ribosomal protein mRNA abundance that is much higher in newly refeled cells than in exponentially growing cells (1, 25, 58). Fig. 5 shows a Northern blot analysis of total RNA extracted from T. thermophila cells that were growing exponentially, from cells starved for 24 h, or from starved cells that were refeled. Approximately equal amounts of RNA (10 μg) were size-separated on an agarose gel, transferred to nylon membrane, and hybridized with the cDNA probe, pC8 (Fig. 5). An mRNA transcript size of 3.5 kb was determined for the ilsA gene (Fig. 5).

This transcript was present during normal growth and was greatly reduced in starved cells (Fig. 5; Refs. 45, 68). Upon refeeding, the ilsA transcript rapidly reappeared and 0.5–1 h after refeeding made up a larger fraction of total RNA than it did during logarithmic growth. When the refeled cells started dividing, the relative abundance of the transcript declined to levels somewhat less than that found in growing cells prior to starvation. This may be because of the slower growth rate reached by refeled cells. The results described above suggest that transcription of the ilsA gene parallels changes in protein synthesis rates. The transcription pattern of the ilsA gene also resembles that of ribosomal protein genes (1, 25, 58).

Identification of the ilsA Gene as an Isoleucyl-tRNA Synthetase Gene—We have previously reported the cloning and sequencing of the cytoplasmic ilRS gene from S. cerevisiae (49). This gene was obtained from an S. cerevisiae genomic DNA library using the T. thermophila ilsA probe pC8. It is

3 C. Csank and D. W. Martindale, unpublished results.
**T. thermophila Isoleucyl-tRNA Synthetase**

**FIG. 3. DNA (sense strand) and amino acid sequence of the ilsA gene.** The putative open reading frame starts at position 1 and amino acids are placed below the first codon position. Lower case letters represent intron sequences. Arrows indicate the boundaries of sequenced cDNA. Transcription start sites and a polyadenylation site are underlined with double lines. The HIGH and KMSKS peptides are underlined with single lines.
identical except for minor strain differences to the ilsA gene isolated independently by transformation of a yeast strain and hence the starvation. The ilsA gene from S. cerevisiae is shown in matrix form in Fig. 6 (miniprint). Over the region where the S. cerevisiae sequence overlaps with the pC8 probe the identity is 63%. The high similarity between the coding regions of the two genes may be expected because S. cerevisiae and T. thermophila have similar codon biases (48).

The ilsA polypeptide is aligned with its S. cerevisiae, E. coli, and M. thermoaerotrophicum counterparts in Fig. 7 (miniprint). The deduced amino acid sequences of the T. thermophila ilsA gene and the ilsA gene from S. cerevisiae share 47% of their amino acids and their percentage similarity, which includes functionally conserved residues, is 61%. The amino acid sequences are most similar (56% identical amino acids) over their amino-terminal two-thirds and least similar (28% identical amino acids) in the remaining third closest to the carboxyl ends. This trend is also noticed when the DNA sequences are compared (Fig. 6). We did pairwise comparisons for each of the ileRSs and determined amino acid identities and similarities (Table 1, see miniprint). The two eukaryotic ileRSs are more similar to the archaeabacterial ileRS than to the homologous enzyme from E. coli (see also Ref. 36). Differences in size (see Fig. 7B) between these homologous enzymes appear to be the result mainly of differences at the carboxyl termini of the amino acid sequences. Computer analysis of the T. thermophila ileRS identified, within the carboxyl region, the presence of two leucine zipper motifs (42) consisting of 4 or 5 leucines spaced at 7-residue intervals with adjacent basic regions (Fig. 8; see "Discussion"). The aligned S. cerevisiae amino acid sequence reveals some conservation of these motifs (Fig. 8). α-Helical wheel diagrams (66) of the T. thermophila leucine zippers are shown in Fig. 9.

We also compared the T. thermophila ileRS polypeptide to aaRS sequences specific for the other branch-chained aliphatic amino acids (valine, leucine, and methionine). Similarities between the aaRSs for these amino acids as well as highly conserved regions have been previously noted (29, 49). We aligned the T. thermophila sequences to the other aaRS sequences (4, 13, 14, 19, 29, 37, 49, 70, 73) using the GCG programs BestFit and GAP, and calculated the percentage identity by counting the number of shared amino acids and dividing this number by the length of the T. thermophila polypeptide. As observed for the S. cerevisiae enzyme (49), the T. thermophila polypeptide is as similar to the valRSs from S. cerevisiae (21%), E. coli (18%), and Bacillus subtilis (22%) as it is to the ileRS from E. coli (22%). It shares 16% of its amino acids with leuRSs and 14% of its amino acids with metRSs from S. cerevisiae and E. coli. Conserved regions shared between these enzymes, including the T. thermophila ileRS, have been described in a previous publication (49).

**FIG. 5.** Northern blot analysis of total RNA harvested from T. thermophila that were logarithmically growing (L), starved for 24 h (S), or refed for various times after 24 h of starvation. The probe pC8 was used (see Fig. 2). Approximately 10 μg of RNA was loaded in each lane. RNA from 0.5 h was slightly underloaded and hence the RNA transcript is under-represented.
FIG. 8. Leucine zipper motifs and adjacent basic regions. Aligned sequences are as in Fig. 7. Thin lines underline repeated leucines; the thin dotted lines underline hydrophobic amino acids at the same position at 7-residue intervals in the S. cerevisiae, M. thermooautotrophicum, or E. coli ileRS. Thick dotted lines underline basic residues. The E. coli zinc finger-like region is underlined by thick solid lines.

served short amino acid motifs (HIGH and KMSKS) that interact with ATP lie within this structure (21, 33, 34, 49, 59, 64).

The T. thermophila ileRS has two leucine zipper motifs (42, 71) in the carboxyl third of the protein (Figs. 7 and 8) that are somewhat conserved in the other ileRSs (Fig. 8). These are typical leucine zippers in that they form amphipathic helices that can associate as dimers (42, 56, 66). The aaRSs, including ileRSs, often occur in multi-enzyme complexes (2, 10, 27). Since leucine zippers can form heterodimers (42, 56), they may be involved in the formation of such complexes (35, 72). Leucine zippers have also been shown to form coiled coils (56); we calculated the probability of coiled coils being formed by the ileRS’s leucine zippers using a recently described algorithm (43). Both of T. thermophila’s leucine zippers, as well as the yeast equivalents, were found to have amino acid sequence distributions typical of coiled coils. Intriguingly, the amino-terminal helices of several other aaRSs have been found to form either a coiled coil (E. coli serRS; 12) or an amphipathic helix (human aspRS; 35) (yeast valRS; 9).

Leucine zippers found in DNA binding proteins have a basic region on their amino-terminal side that contacts DNA (71). In the T. thermophila ileRS, leucine zipper-1 has a short basic stretch following the last leucine of the leucine repeats, while leucine zipper-2 is preceded by a long basic stretch of amino acids (Fig. 8) present also in the S. cerevisiae ileRS and the M. thermooautotrophicum ileRS. The carboxyl terminus of the E. coli ileRS has a zinc finger-like motif (see Ref. 55) that aligns with the basic region of T. thermophila ileRS’s leucine zipper-2. Since zinc finger motifs are also found in DNA binding proteins, this suggests the presence of structurally different nucleic acid binding regions that may have been acquired independently by the eubacterial enzyme and the eukaryotic ancestral enzymes.

Unlike the other sequenced ileRS genes, the T. thermophila gene has eight introns. As aaRS sequences from eukaryotes emerge it will be useful to compare the position of their introns with those of the T. thermophila ileRS to see whether they interrupt sequences at conserved positions and whether they are positioned between structural or functional domains of the aaRSs, as is implicated by the positions of some 1isA introns.
Numerous studies have indicated that the levels and synthesis of aaRSs in cells are positively correlated with protein synthesis rates (reviewed in Refs. 54, 57). This suggests that the production of aaRSs and other components of the translational machinery may be co-regulated at the transcriptional level. Indeed, in T. thermophila the transcription pattern of the ilsaA gene parallels changes in protein synthesis. When cells are starved, protein synthesis is greatly reduced (46). When cells are refed, protein synthesis resumes and so does ilsaA transcription. The ilsaA transcript peaks in abundance within 1 h after a nutritional shift-up and then declines to transcript levels observed in growing cells (Fig. 5). The ilsaA transcription pattern during refeeding resembles that of the ribosomal protein genes (1, 25, 58), suggesting concerted activation of these genes during refeeding. When cells mate and undergo conjugation, a transient increase in the protein synthesis rate occurs in early meiotic prophase (46). In our studies (45), we find a similar transient increase in the transcription of the ilsaA gene as deduced from mRNA abundance. Nuclear run-on experiments have shown that the increase in abundance of the ilsaA RNA transcript during conjugation is the result of new transcription rather than of the stabilization of the mRNA (68). It is probable that conjugation in T. thermophila induces the transcription of several components of the translational apparatus that are needed to provide the cell with the required components for protein synthesis. The ilsaA transcription start sites used were identical in conjugating and growing cells indicating that changes in transcript levels of the ilsaA gene are not due to transcription from different start sites. It is also of interest to note that, unlike many other mRNAs, the ilsaA transcript stays on poly-somes and is not degraded during a heat shock (41). This is presumably because the ileRS gene is needed for the synthesis of heat shock proteins. It is likely that the ileRS gene will be controlled in a similar manner to other genes involved in translation and will share, for example, promoter elements, and transcription factors with these genes.

REFERENCES
1. Andraesen, P. H., Dreisig, H., and Kristiansen, K. (1984) Eur. J. Biochem. 140, 485-492
2. Bec, G., Kerjan, P., Zha, X. D., and Wailer, J.-P. (1989) J. Biol. Chem. 264, 21131-21137
3. Boorstein, W. R., and Craig, E. A. (1989) Methods Enzymol. 180, 347-369
4. Borgford, T. J., Brand, N. J., Gray, T. E., and Fersht, A. R. (1989) J. Biol. Chem. 264, 21131-21137
5. Bruns, P. J., and Bruns, P. J. (1989) J. Biol. Chem. 264, 21131-21137
6. Bruns, P. J., and Bruns, P. J. (1989) J. Biol. Chem. 264, 21131-21137
7. Chatton, B., Walter, P., Chatton, B., Walter, P., and Leisinger, T. (1991) J. Biol. Chem. 266, 10570-10577
8. Jordan, X., Chatton, B., Buhler, M., Buhler, M., Chapon, C., and Waller, J.-P. (1987) J. Biol. Chem. 262, 7189-7194
9. Karrer, K. M. (1986) in The Molecular Biology of the Ciliated Protozoa (Gell, T. G., ed) Academic Press Inc, New York
10. Kozak, M. (1981) Nucleic Acids Res. 9, 5233-5252
11. Kozak, M. (1989) J. Cell. Biol. 108, 229-241
12. Kraus, K. W., Good, P. J., and Hallberg, R. L. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 383-387
13. Landscul, W. H., Johnson, P. F., and McKnight, S. L. (1988) Science 240, 1759-1764
14. Lupas, A., Van Dyke, M., and Stock, J. (1991) Science 252, 1162-1164
15. Martindale, D. W., Allis, C. D., and Bruns, P. J. (1982) Exp. Cell Res. 140, 227-236
16. Martindale, D. W., Allis, C. D., and Bruns, P. J. (1983) Mol. Biol. Cell. 3, 1857-1865
17. Martindale, D. W., Allis, C. D., and Bruns, P. J. (1985) J. Protozool. 32, 644-649
18. Martindale, D. W., Martindale, H. M., and Bruns, P. J. (1986) Nuclear Acids Res. 14, 1341-1354
19. Martindale, D. W. (1989) J. Protozool. 36, 29-34
20. Martindale, D. W., Gu, Z. M., and Csaki, C. (1989) Curr. Genet. 15, 99-106
21. Meusdoeffer, F., and Fink, G. R. (1983) J. Biol. Chem. 258, 2629-2629
22. Mowat, D., Pearlman, R. E., and Engberg, J. (1974) Exp. Cell Res. 84, 282-286
23. Najarian, D., Dianman, E. M., and Hopper, A. K. (1987) Mol. Cell. Biol. 7, 185-191
24. Natsoulis, G., Hilger, F., and Fink, G. R. (1986) Cell 46, 235-243
25. O’Shea, E. K., Rukwadi, R., and Kim, P. S. (1989) Science 243, 538-542
26. Parker, J., Flashner, M., McKeever, W. G., and Neidhardt, F. C. (1974) J. Biol. Chem. 249, 1044-1053
T. thermophila Isoleucyl-tRNA Synthetase

58. Rosendahl, G., Andraesen, P. H., and Kristiansen, K. (1991) Gene (Amst.) 88, 161–167
59. Rould, M. A., Perona, J. J., Söll, D., and Steitz, T. A. (1989) Science 246, 1135–1142
60. Sambrook, J., Fritsch, E. F., and Maniatis, T. (ed) (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
61. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
62. Schimmel, P., R. and Söll, D. (1979) Annu. Rev. Biochem. 48, 605–630
63. Schimmel, P. (1987) Annu. Rev. Biochem. 56, 125–158
64. Schimmel, P. (1991) Trends Biochem. Sci. 16, 1–3
65. Sedman, S. A., Gelsembiuk, G. W., and Mertz, J. E. (1990) J. Virol. 64, 453–457
66. Segrest, J. F., DeLoof, H., Dohman, J. G., Brouillette, C. G., and Anantharamaiah, G. M. (1990) Proteins Struct. Funct. Genet. 8, 103–117
67. Sogin, M. L., Elwood, H. J., and Gunderson, J. H. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 1383–1387
68. Stargell, L. A., Karer, K. M., and Gorovsky, M. A. (1990) Nucleic Acids Res. 18, 6837–6839
69. Tsui, F. W. L., and Siminovitch, L. (1987) Nucleic Acids Res. 15, 3349–3367
70. Tzagoloff, A., Akai, A., Kurkulos, M., and Repetto, B. (1988) J. Biol. Chem. 263, 850–866
71. Vinson, C. R., Sigler, P. B., and McKnight, S. L. (1989) Science 246, 911–916
72. Walter, P., Weygand-Durasevic, I., Sanni, A., Ebel, J.-P., and Fasiolo, F. (1989) J. Biol. Chem. 264, 17126–17130
73. Webster, T., Tsai, H., Kula, M., Mackey, G. A., and Schimmel, P. (1984) Science 226, 1515–1517
Figure 2. Restriction map of llsA genomic DNA (A) and cDNA (B). Inserts of clones and subclones used for sequencing are shown. Arrows indicate the direction of sequencing from the MluI-universal primer site, except for pCC101 where the reverse primer was used. Fully sequenced nested deletions used for sequence determination are shown below clones. The 4.6 kb region from the transcription start sites to the polyadenylation site is boxed in (A) with blank boxes indicating introns and hatched boxes exons. In (B) the open reading frame is shown as a thick line.

Figure 4. Transcription start sites. The results of primer extension using conjugation mRNA (see text) and a primer complementing nucleotides -6-19 (see Figure 3) are shown in lane 1. The numbers represent nucleotide position from Figure 3. The results were in lane 1 the mRNA from growing cells was used. Lane 2 is the T7 RNA from growing cells was used. Lane 3 is the y RNA from growing cells was used. The sequence of genomic DNA sequenced by longer sequencing with P6-dNTP using the same primer as in lane 1, but the primer was not end-labeled. Note: All end-labeled primer extension products (compilments of the RNA) migrate half a residue below a T-residue of the genomic DNA standard. The genomic DNA standard sequencing reactions done with end-labeled primer and unlabeled dNTP also run half a nucleotide faster.

Figure 6. DNA similarity matrix of the T. thermophila llsA gene (cDNA) and the llsA gene from S. cerevisiae. The Pustell DNA Matrix (BD) (Bash value=6; window size=30) was used. Note: similarities do not extend beyond regions shown.

Figure 7. Amino acid sequence alignment of the llsA5s of T. thermophila (Tt), S. cerevisiae (Sc), P. thermoleophilia (Pt) and E. coli (Ec). Identical amino acids are boxed. RLH and KMBK regions are boxed with dashed lines, the two leucine zipper motifs (Leu zip-1 and Leu zip-2) of T. thermophila are indicated, with the repeating leucines indicated with arrows. The cysteines of the E. coli zinc finger motif are indicated with small "hats." A schematic diagram of the four llsA's. Regions shared by all four llsA's are shown as black boxes. Leu zip-1 (Leu zip-2), zf (E. coli zinc finger) inverted triangles point to T. thermophila intron positions.