Analysis of selenocysteine (Sec) tRNA\(^{[Ser\text{Sec}]}\) genes in Chinese hamsters

Xue-Ming Xu \(^a\), Bradley A. Carlson \(^a\), Lark K. Kim \(^b\), Byeong J. Lee \(^b\), Dolph L. Hatfield \(^b\), Alan M. Diamond \(^c\) *

\(^a\) Section on the Molecular Biology of Selenium, Basic Research Laboratory, Division of Basic Sciences, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

\(^b\) Laboratory of Molecular Genetics, Institute for Molecular Biology and Genetics, Seoul National University, Seoul 151-742, South Korea

\(^c\) Department of Human and Nutrition and Dietetics, University of Illinois at Chicago, Chicago, IL 60612, USA

Received 1 June 1999; received in revised form 27 July 1999; accepted 18 August 1999; Received by A. Dugaiczyk

Abstract

Several recent observations have indicated that the primary structure of the Chinese hamster selenocysteine tRNA\(^{[Ser\text{Sec}]}\) is different than those of other mammalian species. These reports prompted us to investigate the gene sequence for this tRNA in Chinese hamsters. Southern blotting of Chinese hamster ovary (CHO) genomic DNA derived from cultured cells with a tRNA\(^{[Ser\text{Sec}]}\) probe indicated several hybridizing bands, and each of the corresponding genetic loci was isolated from a recombinant CHO library by molecular cloning. Sequence analysis of these regions indicated three likely pseudogenes and a single functional gene whose sequence differed from those of other mammals. Of these, only one pseudogene and the putative functional gene are actively transcribed following their microinjection into \(\textit{Xenopus}\) oocytes. The possibility that the functional CHO tRNA\(^{[Ser\text{Sec}]}\) evolved from an edited transcript is discussed. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Cloning; Pseudogenes; Selenium; Transcription

1. Introduction

The selenocysteine (Sec) tRNA\(^{[Ser\text{Sec}]}\) gene has been sequenced from several mammalian species, including human, mouse, rat, rabbit and cow [reviewed in Hatfield et al. (1999)]. With the exception of a pyrimidine transition at position 47c in the long extra arm of tRNA\(^{[Ser\text{Sec}]}\) [bases in tRNA\(^{[Ser\text{Sec}]}\) are numbered as in Sturchler et al. (1993)], all the sequences are identical. Direct sequence analysis of tRNA\(^{[Ser\text{Sec}]}\) from rat, cow, mouse and humans indicated primary structures that were colinear with the corresponding genes [reviewed in Hatfield, et al. (1999)]. An exception to this relationship was revealed by the direct sequence analysis of two bovine tRNA\(^{[Ser\text{Sec}]}\) species that differed from the corresponding gene, one by two pyrimidine transitions (‘CU’ to ‘UC’ at positions 12 and 13) of the tRNA (Diamond et al., 1981; Hatfield et al., 1982). tRNA editing has been speculated to be involved in the generation of this tRNA (Diamond et al., 1990), and further evidence has suggested that such putatively edited tRNA\(^{[Ser\text{Sec}]}\) species are likely to be a very rare component of the total tRNA\(^{[Ser\text{Sec}]}\) population (Diamond et al., 1981, 1993; Amberg et al., 1993). Amberg et al. failed to detect edited tRNA\(^{[Ser\text{Sec}]}\) species in the total bovine tRNA population, although chromatographically distinct tRNA\(^{[Ser\text{Sec}]}\) isoacceptors of relatively low abundance were detected, but not analyzed. Editing of tRNA\(^{[Ser\text{Sec}]}\) could therefore be similar to that reported for the likely editing of rat tRNA\(^{[Asp]}\), involving the conversion of a ‘CU’ to a ‘UC’ nucleotide pair at positions 32 and 33 (Béer et al., 1992).

Recently, two independent approaches have yielded data indicating that the tRNA\(^{[Ser\text{Sec}]}\) population of CHO cells may be unique. A novel approach involving the immunoprecipitation of CHO tRNAs containing the

Abbreviations: AE, activator element; CHO, Chinese hamster ovary; i\(^6\)A, N\(^\text{-}\)isopentenyladenosine; kb, kilobase(s); PCR, polymerase chain reaction; PSE, proximal sequence element; Sec, selenocysteine.

* Corresponding author. Tel.: 1-312-996-0462; fax: 1-312-413-0319.
E-mail address: adiamond@uic.edu (A.M. Diamond)
modified residue N⁶-isopentenyladenosine (i⁶A) and subsequent generation of cDNA from these immunoprecipitated tRNAs was reported (Warner et al., 1998). Several cDNAs isolated in this manner were sequenced, and among these were cDNA{i⁶A}{Ser}. This analysis indicated that the corresponding tRNAs contained two pyrimidine transitions at positions 11 and 12 that differ from all other tRNA{i⁶A}{Ser} genes sequenced from higher vertebrates (Hatfield et al., 1999). Confirmation of this observation was obtained by primer extension analysis of total CHO tRNA (Moustafa et al., 1998). The similarity between the sequences reported for the putatively edited human tRNA{i⁶A}{Ser} gene as a hybridization probe, the CHO tRNA (pyrimidine transitions at positions 11 and 12) prompted the analysis of the tRNA{i⁶A}{Ser} genetic loci reported in this manuscript.

Results and discussion

3.1. Detection and cloning of all CHO tRNA{i⁶A}{Ser} loci

Southern blot analysis was performed on CHO genomic DNA using a PCR-amplified product representing the human tRNA{i⁶A}{Ser} gene as a hybridization probe. Multiple hybridizing bands were detected in PstI-digested DNA (Fig. 1, lane 5). Given the complexity of the hybridization pattern, each locus representing the bands obtained in the Southern blot was isolated for analysis by molecular cloning. A recombinant CHO genomic library was screened with the human tRNA{i⁶A}{Ser} probe, resulting in the recovery of 19 positively hybridizing clones. All 19 clones were analyzed by restriction-enzyme digestion, and the results indicated that each could be assigned to one of four loci based on shared digestion products. Representative clones from each class were digested with PstI and compared to PstI-digested CHO genomic DNA by Southern blotting (Fig. 1). Of the four hybridizing bands detected in CHO genomic DNA, three corresponded to the PstI-digested fragments detected in the representative clones. One band in ... designated with an arrow in Fig. 1, most certainly represents the PstI fragment from a fourth class of recombinant clone containing a truncated 5'-flank (see below). These data indicate that every genetic loci

2. Materials and methods

All molecular cloning procedures, including plaque lifting, Southern blotting, subcloning, restriction enzyme digestions and gel electrophoresis, followed standard techniques or the manufacturer’s protocol unless otherwise stated. Nitrocellulose membranes were from Schleicher & Schuell, Hybond-N nylon membranes from Amersham, 2⁵P-dCTP (Spec. Act. 6000 Ci/mmol) from NEN, 3²P-α-UTP (Spec. Act. 3000 Ci mmol) from Amersham, Prime-it Random Primer Labeling Kit and QuickHyb Hybridization Solution from Stratagene, ABI dye terminase kit from Perkin Elmer, restriction endonucleases from New England BioLabs and DNA purification columns from Qiagen. A lFIX II CHO cell line library (Stratagene) was screened with a 3²P-dCTP labeled 193 bp fragment encoding the human tRNA{i⁶A}{Ser} gene (Hatfield, et al., 1999). Following hybridization for 2 h, membranes were washed twice for 10 min each in 1× SSC, 0.5% SDS at room temperature and exposed to a Phosphor screen for 0.5–2 h or to X-ray films for 2–10 h. Bacteriophage yielding positive hybridization signals were plaque-purified and amplified, and phage DNA was isolated using Qiagen columns. Fragments encoding the Sec tRNA{i⁶A}{Ser} gene obtained by restriction endonuclease digestion were subcloned into pUC 19 vector for sequencing. Sequence alignment and analysis were performed using GCG-Lite + Clustalw in NIH Network. Microinjection of each subcloned fragment of Chinese hamster DNA, the Xenopus tRNA{i⁶A}{Ser} gene, the SS RNA gene and 3²P-α-UTP into Xenopus laevis oocytes and extraction and gel electrophoresis of the resulting transcripts have been described in detail elsewhere (Lee et al., 1987; Park et al., 1995).

Fig. 1. Hybridization of CHO genomic and recombinant phage DNAs. Lanes 2–4 contain PstI-digested phage DNA from clones designated CSIR2-4, respectively, and lane 5 contains PstI-digested CHO genomic DNA. Lane 1 contains DNA markers consisting of λ, digested with HindIII. The arrow Designates the PstI fragment in genomic DNA that was not found in recombinant phage DNA and most certainly represents the functional tRNA{i⁶A}{Ser} gene designated CSIR1 (see text).
detectable by Southern blotting was recovered during the
library screening.

3.2. Nucleotide sequence analysis of rRNA<sup>3</sup> loci in
CHO DNA

Sequence analysis of representative clones from each
class was performed, and the results are presented in
Fig. 2. An additional clone, designated CSr1, lacked
the entire 5' flanking region extending to position 8
within the coding sequence. This explained why the
rRNA<sup>3</sup> hybridizing region within this clone did not
correspond to the 3.2 kb band detected in genomic DNA.
In order to obtain the 5' flanking region, a 416 bp
fragment of this locus was isolated from CHO genomic
DNA using a primer complementary to the 3' flank by
ligation-mediated PCR (Siebert et al., 1995). Sequencing
of the resulting fragment indicated that it contained
247 bp of 5' flanking region, including that missing in
3.2. Nucleotide sequence analysis of tRNA<sup>[Ser]Sec</sup>
loci in the original CSr1 clone. The sequence alignments of
CHO DNA all four loci (CSr1–4) are presented in Fig. 2 along
with the previously sequenced mouse tRNA<sup>[Ser]Sec</sup> gene
(Ohama et al., 1994). Examination of the sequences presented in Fig. 2
indicates that the gene designated CSr1 is the source
of the rRNA previously detected by reverse transcription
(Warner et al., 1998) and primer extension (Moustafa
et al., 1998). Previous primer extension analysis of CHO
tRNA indicated that the tRNA<sup>[Ser]Sec</sup> from this species
had a 'TC' sequence at positions 11 and 12 (Moustafa
et al., 1998), and the sequences reported from Warner

<table>
<thead>
<tr>
<th>CSr1</th>
<th>CSr2</th>
<th>CSr3</th>
<th>CSr4</th>
</tr>
</thead>
<tbody>
<tr>
<td>3---</td>
<td>---5</td>
<td>---7</td>
<td>---10</td>
</tr>
<tr>
<td>1---</td>
<td>---5</td>
<td>---7</td>
<td>---10</td>
</tr>
<tr>
<td>+100</td>
<td>+50</td>
<td>+100</td>
<td>+100</td>
</tr>
<tr>
<td>+100</td>
<td>+100</td>
<td>+100</td>
<td>+100</td>
</tr>
<tr>
<td>+200</td>
<td>+200</td>
<td>+200</td>
<td>+200</td>
</tr>
<tr>
<td>+300</td>
<td>+300</td>
<td>+300</td>
<td>+300</td>
</tr>
</tbody>
</table>

Fig. 2. Alignment of Chinese hamster recombinant DNA sequences and the mouse tRNA<sup>[Ser]Sec</sup> gene. CSr1–CSr4 sequences are aligned with the mouse tRNA<sup>[Ser]Sec</sup> (MSR). The coding region of the gene (+1 to +87) is shaded, and the TCA anticodon (+36 to +38) is shaded slightly darker. Upstream regulatory regions involved in expression of Sec rRNA<sup>[Ser]Sec</sup> genes near −29 (TATA box), at approximately −64 (PSE) and at approximately −220 (AE) are underlined in CSr1 as is the Pol III termination sequence, +103 to +106. Asterisks show the highly conserved region in the 5' flank. To maximize the homology between sequences, dots were used to indicate gaps between aligned sequences and to indicate unidentifiable bases at the 5' and 3' ends. The homologies of the rRNA coding regions of CSr2, CSr3 and CSr4 to CSr1 are 70, 74 and 91%, respectively. Positions 13 and 14 are indicated by + and correspond to positions 11 and 12 in the rRNA<sup>[Ser]Sec</sup> gene as discussed in Section 1. The GenBank Accession Nos for CSr1, CSr2, CSr3 and CSr4 are AF133045, AF133046, AF133047 and AF133048, respectively.
et al. (1998) are also only consistent with the CSIR1 sequence. Other differences between the sequences reported by Warner et al. and that of CSIR1 are likely due to artifacts introduced by the methods employed in their study. These include a 5'-terminal base, which is a C in the cDNA and a G in CSIR1, and a T in the tRNA sequence at position 58, which corresponds to an A in CSIR1. The conclusion that the differences observed in position 1 and position 58 between the present study and that of Warner et al. are due to artifacts in the techniques employed in the latter study is based on evidence that all known tRNA[^Ser^Sec] genes sequenced to date contain a G in the first position of the gene and an A (which is 1-methyladenosine in tRNA[^Ser^Sec]) at position 58 (Hatfield et al., 1999). Further evidence that CSIR1 is the functional gene is the presence of three regulatory regions, a TATA box near −29, a proximal sequence element (PSE) at approximately −64, an activator element (AE) at approximately −220, and a Pol III termination signal at +103, which are characteristic of all mammalian tRNA[^Ser^Sec] genes sequenced to date (Hatfield et al., 1999 and see Fig. 2 and figure legend).

The other two loci that are consistent with the primer extension data, CSIR2 and CSIR3, contain multiple single base differences/deletions from CSIR1 and are therefore likely to represent pseudogenes. The fourth locus, CSIR4, is also likely to be a pseudogene as it contains a total of seven positional differences from the functional CSIR1 gene and, more importantly, was undetectable by primer extension (Moustafa et al., 1998).

3.3. The CHO tRNA[^Ser^Sec] gene and one pseudogene are actively transcribed in Xenopus oocytes

As can be seen in Fig. 2, the sequences 5' of the tRNA coding region of CSIR2 included TATA boxes at positions −35 and −45. In order to evaluate whether each of the cloned loci (CSIR1–4) could support transcription, each was individually microinjected into Xenopus oocytes along with a 5S RNA gene and ^32P-α-UTP, and nascent RNA was extracted and analyzed by gel electrophoresis (Fig. 3). Each sample was co-injected with a 5S RNA gene in order to demonstrate that the injected oocytes were transcriptionally active. Lane 1 in the figure shows the results of such an analysis following injection of the Xenopus tRNA[^Ser^Sec] gene, presented for comparison. The results included in Fig. 3 indicate that CSIR1 and CSIR2 were efficiently transcribed in this system (lanes 2 and 3), while CSIR3 and CSIR4 were not (lanes 4 and 5). Active transcription of CSIR1 was anticipated, in accordance with the above discussion, suggesting that this locus was likely to be the functional CHO tRNA[^Ser^Sec] gene, and the mobility of the resulting transcript in the gel is similar to that seen for the corresponding Xenopus tRNA. The CSIR2 transcript detected following transcription in Xenopus oocytes is very likely to be that encoded by the tRNA[^Ser^Sec] pseudogene as it migrated in the gel with a reduced mobility (lane 3) compared to the CSIR1 and Xenopus tRNAs, consistent with the use of a Pol III transcription termination site approximately 200 nucleotides 3' of the pseudogene. However, we have been unable to detect a tRNA[^Ser^Sec] hybridizing species of this apparent molecular weight, approximately 300 nucleotides, by northern blot analysis of total CHO tRNA (unpublished observation). The in vitro transcription of precursor tRNA from pseudogenes, which are undetectable in vivo, has been reported previously (Pirtle et al., 1993, for example).

3.4. Evolution from an edited tRNA[^Ser^Sec] transcript?

Further inspection of the sequences aligned in Fig. 2 reveals a highly similar region of approximately 20 nucleotides immediately 3' and adjacent to the tRNA sequences within three of four loci characterized from the CHO genome (CSIR1, CSIR3 and CSIR4) and in the mouse genome. Collectively, these data are sugges-
tive of the generation of two of the pseudogenes from a functional tRNA\[Ser\]Sec by reverse transcription and subsequent reinsertion into the genome. Transcription of tRNA\[Ser\]Sec is unusual in that it begins at the first position of the tRNA coding sequence (Lee et al., 1987), and it is noted that there is limited homology among the 5′-flanking regions of CSR1-4. The homologous 3′ flanking regions are likely to represent RNA sequences that are typically removed by processing to generate an end compatible with CCA addition and were apparently present during reverse transcription.

It is interesting to speculate that CSR4 may have been the primordial tRNA\[Ser\]Sec gene in hamsters, from which CSR1 and 3 later evolved. CSR4 contains the ‘CT’ sequence present in all other mammalian tRNA\[Ser\]Sec genes at positions 13 and 14 (designated positions 11 and 12 in the tRNA\[Ser\]Sec gene product as described in Section 1). These data raise the possibility that editing of a once active CSR4 gene yielded a tRNA with ‘TC’ at positions 11 and 12, which was then reintroduced to the CHO genome prior to 3′-end processing. It has been noted that editing of plant mitochondria tRNA, also involving a C-to-U transition, must precede 3′-end maturation (Marchfelder et al., 1996; Marechal-Drouard et al., 1996; Kunzmann et al., 1998). This hypothesis would provide a feasible explanation of the sequence differences between the CHO tRNA\[Ser\]Sec functional gene and that of all other mammalian species.

3.5. Conclusions

1. Four genetic loci that hybridize to the tRNA\[Ser\]Sec gene represent one functional gene and three pseudogenes.
2. The previously detected CHO tRNA\[Ser\]Sec, which differs in primary sequence when compared to all other tRNA\[Ser\]Sec genes described form higher vertebrates, is not an edited transcript.
3. Both the functional tRNA\[Ser\]Sec gene and one pseudogene are actively transcribed when injected into Xenopus oocytes.
4. The functional gene for tRNA\[Ser\]Sec in hamsters may have evolved from an edited transcript.

Acknowledgements

The authors would like to acknowledge the technical assistance of Judith Murray. This work was supported by Grant # RO1 CA51153 to A.M.D. and the Molecular Medicine Research Group Program, MOST, Korea (98-MM-02-A-03) to B.J.L.

References


Beter, H., Lee, M.C., Sekiya, T., Kuchino, Y., Nishimura, S., 1992. Two nucleotides next to the anticodon of cytoplasmic rat tRNA(Asp) are likely generated by RNA editing. Nucleic Acids Res. 20, 2679–2683.


