Two Forms of Human Cytoplasmic Arginyl-tRNA Synthetase Produced from Two Translation Initiations by a Single mRNA†

Yong-Gang Zheng,[‡,§] Hui Wei,[§,||] Chen Ling,[‡] Min-Gang Xu,[‡] and En-Duo Wang*[,||]

State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, and Graduate School, Chinese Academy of Sciences, Shanghai 200031, China

Received August 23, 2005; Revised Manuscript Received October 18, 2005

ABSTRACT: Human cytoplasmic arginyl-tRNA synthetase (ArgRS) is a component of a macromolecular complex consisting of at least nine tRNA synthetases and three auxiliary proteins. In mammalian cells, ArgRS is present as a free protein as well as a component of the complex. Via an alignment of ArgRSs from different vertebrates, the genes encoding full-length human cytoplasmic ArgRS and an N-terminal 72-amino acid deletion mutant (hcArgRS and ΔNhcArgRS, respectively) were subcloned and expressed in *Escherichia coli*. The two ArgRS products were expressed as a soluble protein in *E. coli*. The level of production of ΔNhcArgRS in *E. coli* and its specific activity were higher than those for hcArgRS. By Western blot analysis, using an antibody against the purified ΔNhcArgRS, the two forms of ArgRS were detected in three human cell types. The 5′-end cDNA sequence, as confirmed by 5′RACE (5′-rapid amplification of cDNA ends), contained three start codons. Through mutation of the three codons, the two human cytoplasmic ArgRSs were found to be produced in different amounts, indicating that they resulted from two different translation initiation events. Here we show evidence that two forms of human cytoplasmic ArgRS were produced from two translational initiations by a single mRNA.

Aminoacyl-tRNA synthetases (aaRSs)[1] play a key role during genetic codon decoding by transferring amino acids to cognate tRNA (*1*). There are 20 aaRSs that can be divided into two classes of 10 members, on the basis of conserved sequences and characteristic structural motifs (*2*). In higher eukaryotes, at least nine different aaRSs together with three auxiliary proteins form a macromolecular complex (*3*−*5*). While the interaction and assembly of the macromolecular complex have been dissected (*6*−*8*), the function of the complex formation remains unclear. Genetic disruption of auxiliary factor p38 in mice, which is essential for complex formation, causes only neonatal lethality in lung deficiency (*9*, *10*). This result indicates that the formation of the complex is not essential for protein synthesis and cell proliferation.

Arginyl-tRNA synthetase (ArgRS, EC 6.1.1.19) is a member of class 1 aaRSs (*2*) and requires the presence of the cognate tRNA for amino acid activation (*11*). In rat liver cells, ArgRS exists both as a free protein and as a component of the complex (*12*, *13*). Compared to the free enzyme, ArgRS within the high-molecular weight complex has an extension at its N-terminus (*12*, *13*). The two forms of ArgRS exhibit similar catalytic characteristics (*11*, *14*, *15*). The N-terminal extension of ArgRS interacts with auxiliary factor p43, an interaction that is essential for the association of ArgRS with the complex (*14*, *15*). The association of p43 with ArgRS has been proposed to modulate the activity of ArgRS (*15*). However, more recent research demonstrates that p43 may not modulate tRNA aminoacylation properties (*14*). A model has been proposed to explain the existence of two forms of ArgRS, on the basis of the hypothesis of tRNA channeling. In this model, the complexed enzyme provides arginyl-tRNA for protein biosynthesis and the free enzyme supplies arginyl-tRNA as the substrate of arginyl-tRNA, protein arginyltransferase (*12*). In purified nuclei of both Chinese hamster ovary and rabbit kidney cells, only the long form of ArgRS is found, suggesting a more direct role in nuclear to cytoplasmic transport of tRNA (*16*).

It was thought that the free form of ArgRS could have arisen from a limited cleavage of the complex-bound ArgRS by an endogenous protease (*17*). However, N-terminal sequence analysis of the purified rat liver free ArgRS suggests that low-molecular weight ArgRS is probably a distinct translation product (*12*). Nevertheless, cloning of full-length ArgRS cDNA and Northern blot analysis of mRNA, in Chinese hamster ovary and human cells, suggest the existence of only one transcript corresponding to the high-molecular weight form of ArgRS (*18*, *19*). One possibility is that the free and complexed forms of ArgRS are translated from the same mRNA. In this regard, it has been reported that cytoplasmic and mitochondrial glycyl-tRNA synthetases (GlyRS) and alanyl-tRNA synthetase (AlaRS) of *Saccharomyces cerevisiae* are different translational products resulting from an alternative translation initiation (*20*, *21*).

† This work was funded by the Natural Science Foundation of China (Grant 30330180) and 973 (Grant 2005CB724600) and 863 (Grant 2004AA235091) projects of China.

* To whom correspondence should be addressed. Phone: 0086-21-54921241. Fax: 0086-21-54921011. E-mail: edwang@sibs.ac.cn.

‡ Graduate School.

§ These authors contributed equally to this work.

|| Shanghai Institutes for Biological Sciences.

[1] Abbreviations: ArgRS, arginyl-tRNA synthetase; hcArgRS, human cytoplasmic ArgRS; ΔNhcArgRS, N-terminal 72-amino acid deletion mutant of ArgRS; aaRS, aminoacyl-tRNA synthetase; 5′RACE, 5′-rapid amplification of cDNA ends; DTT, dithiothreitol.

10.1021/bi051675n CCC: $33.50
Published on Web 01/06/2006

The mechanism by which ribosomes most frequently select the initiation codon on an mRNA is the ribosomal scanning process (*22*, *23*). After recruitment, the 43S complex, composed of a 40S subunit of ribosome bound to eIF2-GTP/Met-tRNAi, eIF1A, and eIF3, scans the 5′-UTR (5′-untranslated region) of mRNA in a 5′ to 3′ direction to the initiation codon. A different translation initiation may result according to two models, the "leaky scanning model" and "re-initiation model". When upstream AUG triplets are not recognized by all of the scanning initiation complexes, some ribosomes will pass by leaky scanning and will initiate at the downstream AUG. The alternate mechanism would be by re-initiation. After translation at a small upstream open reading frame, the 40S ribosomal subunit remains connected to the mRNA and resumes scanning. There are two other mechanisms via which ribosomes may select the initiation codon. These are internal entry on an internal ribosome entry site and shunting of ribosomes. The selection of translation initiation is an important aspect of translational regulation (*24*, *25*). The question here is whether the two forms of ArgRS in mammalian cells could be produced from different initiation sites of the same mRNA.

Here we report that two forms of ArgRS are found in human cell cytoplasm and are the products of different translational initiations from a single transcript.

MATERIALS AND METHODS

Materials. Dithiothreitol (DTT), ATP, L-arginine, and anti-Flag M2 monoclonal antibody were purchased from Sigma. L-[^{3}H]Arginine (1 mCi/mL) was obtained from Amersham Biosciences. The plasmid (catalog number 97002RG) containing full-length cDNA of human liver cytoplasmic ArgRS was obtained from Invitrogen. T4 DNA ligase, restriction endonucleases, and the 5′RACE kit were obtained from Takara (Dalian, China). Ni−NTA Superflow was purchased from Qiagen Inc. *Escherichia coli* tRNA^{Arg} (1066 pmol/A_{260}) was purified from an overproducing strain constructed in our laboratory (*26*).

E. coli Strains, Human Cells, and Cell Culture. E. coli strain BL21-CodonPlus (DE3)-RIL [B F$^-$ *ompT hsdS* (r_B^- m_B^-) *dcm*$^+$ Tet^r *gal* λ (DE3) *endA* Hte (*argU ileY leuW* Cam^r)] (Statagene, Inc.) was used to express human cytoplasmic ArgRS. 293T human embryonic kidney fibroblast cells and HepG2 human hepatocellular carcinoma cells were gifts from Y. Wang (Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences) and grown in DMEM supplemented with 10% fetal bovine serum. 5HSY-5Y human neuroblastoma cells were a kind gift from Y. X. Ni (Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences).

Construction and Transformation of Plasmids. The pET-15b plasmid (Novagen), with a T7 promoter, was used for expression of ArgRS with an N-terminal His_6 tag in *E. coli* strain BL21-CodonPlus (DE3)-RIL. Plasmid pCDNA3 (Invitrogen, Inc.) was used to produce proteins in human embryonic kidney 293T cells. Using PCR, with the forward primer 5′-TTTTCATATGGACGTACTGGTGTCT-3′ and the reverse primer 5′-TTTTGGATCCTTACATCCTTTG-GACAGGTT-3′, the gene (*hc-argS*) encoding human cytoplasmic ArgRS was amplified from the plasmid (Invitrogen, catalog number 97002RG). The gene (Δ*Nhc-argS*) encoding the human cytoplasmic ArgRS with deletion of 72 amino acid residues at the N-terminus was amplified by using the forward primer 5′-AACATATGATTAACATTATTA-3′ and the reverse primer 5′-TTTTGGATCCTTACATCCTTTG-GACAGGTT-3′. The *Nde*I site in the primers is indicated by underlined letters and the *Bam*HI site the line above the sequence. The two amplified DNA fragments were digested with *Nde*I and *Bam*HI and inserted into the plasmid of pET-15b to obtain pET-*hc-argS* and pET-Δ*Nhc-argS*, respectively. The full-length cDNA of *hc-argS* was amplified by PCR. The forward primer was 5′-TTTCGGATCCGTC-CACTTGGCGAGTGAGACGCTG*ATG*GGAGG-3′, in which the line above the sequence marks the *Bam*HI site and the italics denote the ATG initiating codon A at the 5′-terminus of full-length cDNA of *hc-argS*. The reverse primer was 5′-AAAGAATTCTTA CTT GTC ATC GTC GTC CTT GTA GTC CATCCTTTGGACAGGTTTT, where the double underlining denotes the *Eco*RI site and the single underlining the portion encoding DYKDDDDK (Flag), which is recognized by M2. At the 5′-end sequence of the full-length cDNA of *hc-argS*, there are three ATG start codons (Figure 4). Start codons B and C are located within the same open reading frame with 213 bases between them. Start codon A is located upstream of codon B and in an open reading frame different from that of codon B or C. The mutations (A Kozak, A null, B null, and C null; see Figure 4) of full-length cDNA of *hc-argS* were generated by PCR using the forward primer containing a point mutation at or around start codons A−C and the reverse primer bringing in a Flag tag coding sequence described above (Figure 4). The amplified wild-type and mutant cDNA was digested by *Bam*HI and *Eco*RI and inserted into the same site in pCDNA3 to produce recombination plasmids, with which human embryonic kidney 293 cells were transfected.

The pCDNA-uORF-EGFP and pCDNA-uORF-A(null)-EGFP plasmids containing the gene of the enhanced green fluorescent protein (EGFP) and upstream open reading frame (uORF) of hcArgRS were used for assaying ORF translation as a mode of bypassing the start codon. With the forward primer TTTTGGATCCGCTGACCGTTTCCGCTTCCGT-CCACTTGGCGAGT (the *Bam*HI site underlined) and the reverse primer TTTTCTCGAGATTTAATCTCTTCTTC-CTGCTG (the *Xho*I site underlined), DNA fragments containing the uORF and uORF-A(null) were amplified by PCR from pCDNA-hcArgRS and pCDNA-hcArgRS-A(null), digested with *Bam*HI and *Xho*I, and inserted into the same sites in pCDNA3 to obtain plasmids pCDNA-uORF and pCDNA-uORF-A(null), respectively. Then with the forward primer TTTTCTCGAGGGGCCCGGGATCCACCGGCCG-GT and the reverse primer AAAACTCAGTTACTTATCGT-CGTCATCCTTGTAATCCTTGTACAGCTCGTCCATG-CCGAGAGTGA (the *Xho*I site underlined), the sequence encoding EGFP was amplified from pEGFP-N2, digested with *Xho*I, and inserted into plasmids pCDNA-uORF and pCDNA-uORF-A(null) to obtain plasmids pCDNA-uORF-EGFP and pCDNA-uORF-A(null)-EGFP, respectively. A Flag tag sequence was fused to the 3′-end of the EGFP ORF

so the products could be detected using the M2 anti-Flag antibody by Western blotting.

Purification of hcArgRS from E. coli Transformants. E. coli BL21-CodonPlus (DE3)-RIL was transformed with the two plasmids, pET-*hc-argS* and pET-Δ*Nhc-argS*, to produce hcArgRS and ΔNhcArgRS, respectively. *E. coli* BL21-CodonPlus (DE3)-RIL transformants containing pET-*hc-argS* and pET-Δ*Nhc-argS* were grown at 25 °C to achieve an A_{600} of 0.6 in 1 L of LB with 100 μg/mL ampicillin and 60 μg/mL chloramphenicol and were induced by addition of isopropyl 1-thio-β-D-galactopyranoside (IPTG) to a final concentration of 200 μM, PMSF (phenylmethanesulfonyl fluoride) to a final concentration of 0.1 mM, and glucose to a final concentration of 0.04% (w/v). After a 4 h induction, the cells were harvested and washed with buffer A [10 mM imidazole, 300 mM NaCl, 10% glycerol, 20 mM β-mercaptoethanol, 0.5 mM PMSF, and 50 mM Na_2HPO_4 (pH 8.0)]. The cells, suspended in buffer A, were sonicated for 6 × 20 s at 15 W with a high-intensity ultrasonic processor (model 375W, Sonics and Materials Inc.). The crude extract was cleared of cellular debris by centrifugation at 150000*g* for 20 min. The supernatant was gently mixed with 1.5 mL of Ni−NTA Superflow resin for 30 min. The mixture was loaded on a minicolumn for gravity flow chromatography. The resin was washed with 15 mL of buffer B (components as for buffer A, except 20 mM imidazole) to remove nonspecific binding contaminants. Then the enzyme was eluted with 20 mL of buffer C (again components as for buffer A, except 250 mM imidazole). The eluted His_6-hcArgRS or His_6-ΔNhcArgRS fractions were pooled and dialyzed against 2 L of buffer D [20 mM potassium phosphate buffer (pH 7.5)], concentrated using a Centricon device (50−200 μL), gently mixed with an equal volume of glycerol, and stored at −20 °C. All steps of the purification were carried out at 4 °C. The production and purification of both His_6-ArgRS enzymes were followed by SDS−PAGE analysis. Protein concentrations were determined by the method of Bradford (*27*).

Enzyme Activity Analysis. The aminoacylation activity of the two purified ArgRS types was determined in a reaction mixture containing 50 mM Tris-HCl (pH 7.5), 12 mM $MgCl_2$, 4 mM ATP, 80 mM KCl, 0.5 mM DTT, 0.1 mM EDTA, 0.05 mg/mL BSA, 20 μM pure *E. coli* $tRNA^{Arg}$, and 0.1 mM [^{3}H]arginine (25 μCi/μmol). The reaction was initiated by adding 10−90 nM ArgRS. The kinetic parameters of enzymes were determined using various concentrations of the relevant substrates (*29*). One unit of aminoacylation activity is defined as the quantity of protein catalyzing the incorporation of 1 nmol of amino acid into tRNA under the conditions of the assay. The specific activity is defined in units per milligram of protein (*28*).

Analyzing ArgRS in Human Cells by Western Blots. To determine ArgRS levels in human cells, a crude extract from the cells was fractionated by SDS−PAGE. The primary IgY polyclonal antibody against ArgRS was obtained from the Sheng Longda Group Co. Purified ΔNhcArgRS was used as the antigen to immunize laying chickens. Anti-ArgRS immunoglobulin Y (IgY) was isolated from yolks of the eggs laid by these chickens using a two-step salt precipitation protocol. The HRP (horseradish peroxidase)-conjugated secondary antibody was also from Sheng Longda Group Co. The ArgRS bands were visualized by sequential treatment with anti-ΔNhcArgRS antibody, horseradish peroxidase-conjugated secondary antibody, and an enhanced chemiluminescence substrate kit (Amersham Biosciences).

293T cells were grown in DMEM supplemented with 10% fetal bovine serum at 37 °C in 35 mm dishes in a 5% CO_2 incubator, until they were approximately 80% confluent. The cells were transfected with the DNA variants, using the SuperFect Kit (Qiagen) according to the supplied protocol. After 24 h, cells were harvested by centrifugation at 600*g* for 5 min, washed twice with cold phosphate-buffered saline, and lysed with 20 μL of the chilled cell lysis buffer [20 mM potassium phosphate (pH 7.5), 10% glycerol, 2 mM dithiothreitol, 0.5 mM PMSF, and 1% Triton X-100]. The proteins produced in 293T cells contained a C-terminal Flag tag and could be detected by an anti-Flag antibody and a secondary antibody using Western blotting.

Mapping the 5′-End Sequence of the in Vivo Transcripts Encoding ArgRS. Identification of the 5′-end sequence of ArgRS mRNA was carried out by 5′-RACE using the 5′-Full RACE Core Set from Takara. The 5′-terminal sequence of ArgRS cDNA was obtained by RT from a poly(A) RNA pool, isolated from human 293T cells by an oligo(dT) cellulose column (Invitrogen). The 5′-terminal sequences of ArgRS mRNA were transcribed with reverse transcriptase into first-strand cDNAs using a 5′-end phosphorylated primer (5′-CACTAGCAGAGGAGG-3′) complementing a DNA sequence 72 bp downstream of start codon C in Figure 4. Then the first-strand cDNA products were circularized by RNA ligase, and the 5′-end of the cDNA was then amplified by two rounds of PCR from the circularized cDNA. For the first round of PCR, the forward primer was 5′-GGCTG-CATATCCAGATTTGGA-3′ and the reverse primer was 5′-CATGACCAAAGACCTCTTGTAG-3′. For the second round, the forward primer was 5′-AGATTTGGAAAATCCTC-CTCTG-3′ and the reverse primer was 5′-GACCTCTTG-TAGGCGGCTAA-3′. The amplified DNA fragments were inserted into a T vector (Sangon, Shanghai, China), as confirmed by DNA sequencing.

RESULTS

Purification of hcArgRS and ΔNhcArgRS. Two plasmids, pET-hc-*argS* and pET-ΔNhc-*argS*, were obtained. Both human cytoplasmic ArgRS genes could be expressed in the *E. coli* transformants. Most of ΔNhcArgRS and hcArgRS were produced from the *E. coli* transformants as soluble proteins. However, the amount of hcArgRS produced was low compared with the amount of ΔNhcArgRS. After Ni affinity chromatography, 10 mg of purified ΔNhcArgRS and 4.5 mg of hcArgRS were purified from 3.5 g of *E. coli* cell paste. hcArgRS was purified to 90% homogeneity and ΔNhcArgRS to 95% as single bands with molecular masses of 75 and 67 kDa, respectively, as determined by SDS−PAGE (Figure 1). The two human cytoplasmic ArgRS proteins could arginylate *E. coli* $tRNA^{Arg}$. The aminoacylation specific activities for hcArgRS and ΔNhcArgRS were 286 and 715 units/mg and the k_{cat} values 0.36 and 0.8 s^{-1}, respectively.

Two Forms of ArgRS Coexist in Extracts in Human Cells. We tested for the presence of ArgRS, using Western blots, in three human cell types: 5HSY-5Y neuroblastoma cells, 293T embryonic kidney fibroblast cells, and HepG2 hepa-

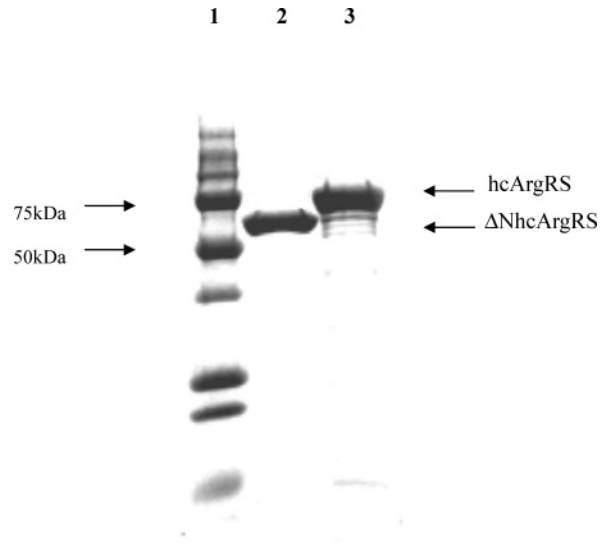

FIGURE 1: Analysis of purified hcArgRS and ΔNhcArgRS by SDS−PAGE. After one-step affinity chromatography on the Ni−NTA mini column, the purified ΔNhcArgRS (lane 2) and hcArgRS (lane 3) were compared with the protein markers, 250, 150, 100, 75, 50, 37, 25, 20, 15, and 10 kDa from top to bottom (Bio-Rad, lane 1). Two arrows indicate the two bands of hcArgRS and ΔNhcArgRS corresponding to molecular masses of 75 and 67 kDa, respectively.

tocellular carcinoma cells. Using an antibody against ΔNhcArgRS, two forms of human cytoplasmic ArgRS with different molecular masses were found to coexist in the three cell types (Figure 2). The larger band had the same molecular mass as hcArgRS, while the smaller one had the same molecular mass as ΔNhcArgRS, purified from *E. coli*. The hcArgRS level was the greatest in all cell types; in 5HSY-5Y, 293T, and HepG2 cells, the ratios of hcArgRS to ΔNhcArgRS were 2.96, 3.08, and 2.1, respectively. The subcellular localization experiment demonstrated that hcArgRS and ΔNhcArgRS were exclusively in the cytosol of 293T cells (*29*).

Mapping the 5′-End of Human Cytoplasmic ArgRS mRNA. Using 5′-RACE to obtain the 5′-end DNA fragment of human ArgRS cDNA, an only ∼300 bp DNA fragment could be amplified from 293T cells (Figure 3A). After TA subcloning, the DNA fragment was sequenced. The 5′-end of the cDNA was mapped to position −51 (the A of the ATG initiation codon of hcArgRS being position 1). The sequence of the 5′-end of the cDNA, encoding ArgRS (Figure 3B), was the same as that previously reported (*19*, *30*), which suggested that there is only one mRNA encoding ArgRS in human 293T cells.

Two Forms of ArgRS Produced from Two Translation Initiations. The combined plasmids, with the cDNA of hcArgRS and its four mutants, and with a Flag tag inserted into the plasmid pCDNA3, were separately transfected into human 293T cells. The ArgRS, produced by the recombinant plasmid in the 293T cells, was detected using an anti-Flag tag monoclonal antibody, as shown in Figure 4. Two forms of ArgRS were produced from the *hc-argS* cDNA, the amount of the smaller protein produced being less than the amount of the larger protein. When start codon B was mutated to ATT, the larger ArgRS product disappeared and only the smaller ArgRS was produced (corresponding to ΔNhcArgRS) with the amount being greatly increased. If start codon C, within the same reading frame as codon B, was similarly made "null", as compared with the wild-type cDNA, only the larger ArgRS (i.e., hcArgRS) was observed. There is an upstream initiation codon A, located in the 5′-UTR of *hc-argS* mRNA, from which a short upstream ORF (uORF) can be translated and which then leads to a bypass of codon B. When a mutant at start codon A (A null) was used to block translation initiation at this site, the amount of ΔNhcArgRS was decreased and the amount of hcArgRS increased. Conversely, when the sequence around start codon A was changed into an optimal Kozak context (*31*), the level of production of ΔNhcArgRS was increased while that of hcArgRS decreased. To provide more direct evidence that the scanning ribosomes initiate translation at the uORF, we fused this ORF to EGFP and transfected the reporter into HEK293T cells (Figure 5). Two proteins can be detected, and the small one has the same MW as the EGFP-Flag protein. When codon A was null, only small protein could be produced from the reporter. These results indicated that scanning ribosomes can initiate translation at codon A. Our results show that by using different start codons it was therefore possible to directly modulate the production of hcArgRS and ΔNhcArgRS.

DISCUSSION

The aaRSs complex is found in cells from fruitflies to humans (*3*−*5*). However, two forms of ArgRS are found in only vertebrates, suggesting that the two forms of ArgRS were developed in a later stage of evolution (*12*, *13*, *32*). The two forms of ArgRS have been reported in studies of rat liver, nearly more than three decades ago (*31*). Only one form of ArgRS cDNA has been cloned from CHO, mouse,

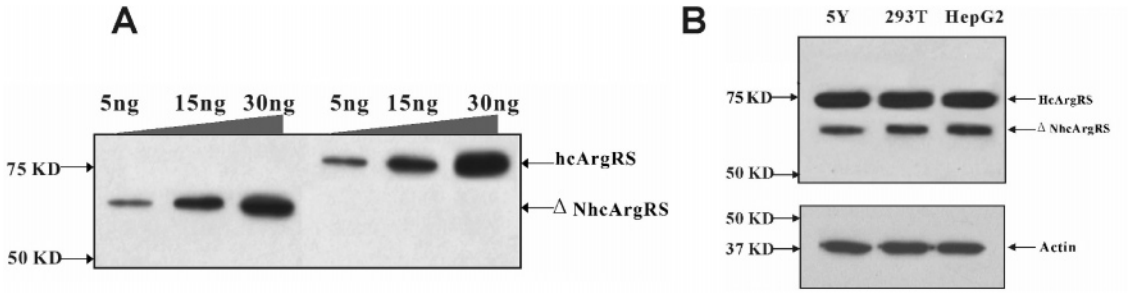

FIGURE 2: Two forms of ArgRS in three kinds of human cells. (A) The sensitivity of the polyclonal antiserum for hcArgRS and ΔNhcArgRS was analyzed by comparative Western blot analysis. The results indicate that the antibody has the same sensitivity for both purified recombinant proteins. (B) Two forms of human cytoplasmic ArgRS were detected by Western blotting using the polyclonal antibody against ΔNhcArgRS on 20 μg of extracts of 5HSY-5Y human neuroblastoma cells, 293T human embryonic kidney fibroblast cells, and HepG2 human hepatocellular carcinoma cells. In 5HSY-5Y, 293T, and HepG2 cells, the hcArgRS:ΔNhcArgRS ratios were 2.9, 3.08, and 2.1, respectively.

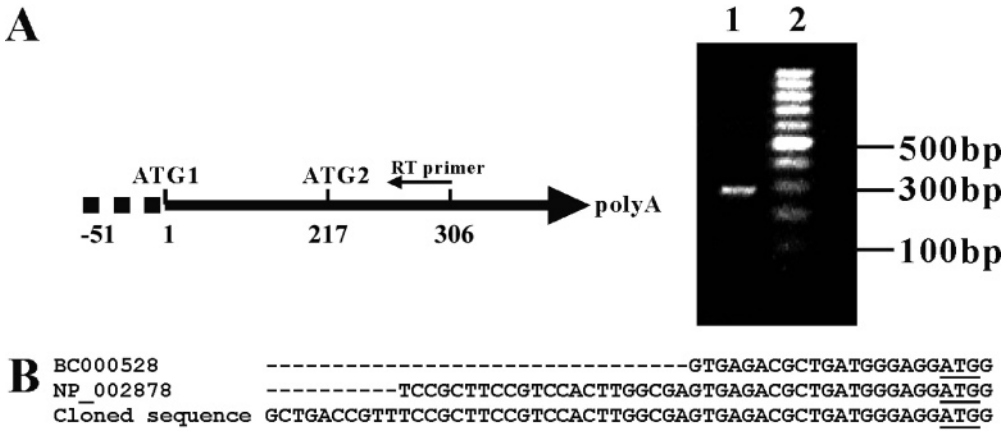

FIGURE 3: Mapping the 5′-end of the hcArgRS mRNA. (A) The scheme represents the mRNA of hcArgRS and the relative positions of the two ATG codons and RT primer (left). The cDNA fragment at the 5′-end of the transcripts (lane 1) and the DNA marker (lane 2) are indicated at the right. (B) The DNA sequence of the 5′-end of the cDNA is compared with the two reported sequences (the accession numbers of the two sequences are given at the left). The initiation codon ATG corresponding to hcArgRS is denoted with underlined letters.

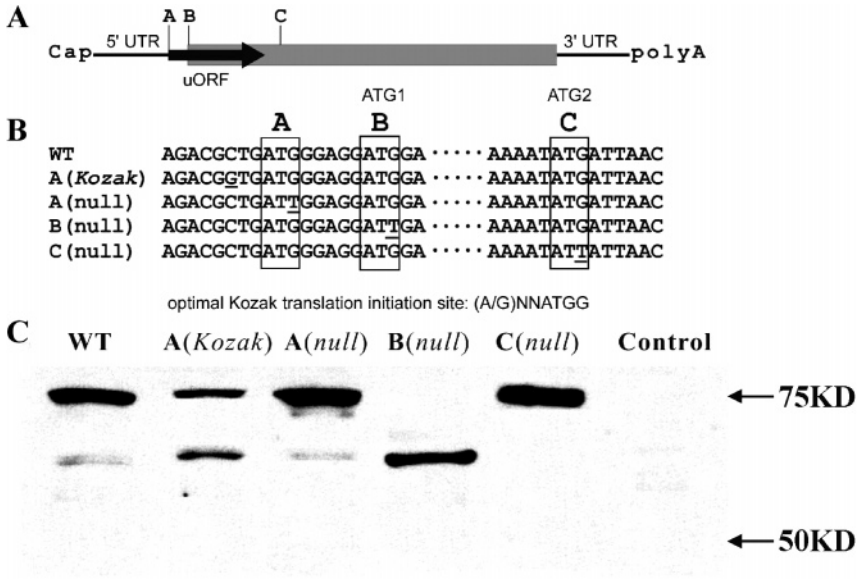

FIGURE 4: Two forms of ArgRS produced from two translational initiations. (A) There is an upstream ATG codon (codon A) in the 5′-UTR of the mRNA. ATG codons B and C are in the same ORF and correspond to the start codons of the ORF encoding hcArgRS and ΔNhcArgRS, respectively. We hypothesized that the scanning ribosomes could initiate translation at codons B and C, respectively, to produce two proteins. Ribosomes also can initiate translation at codon A and then bypass codon B. After translation of the uORF, a part of the ribosomes will scan down to codon C to re-initiate translation. (B) Three start codons and their mutations. (C) Effect of the mutations of the start codons on the production of hcArgRS and ΔNhcArgRS. The wild-type cDNA of hcArgRS yielded two ArgRSs (WT lane). Null of the codon B [lane B(*null*)] or C [lane C(*null*)] resulted in the disappearance of hcArgRS or ΔNhcArgRS. Null of the codon A [lane A(*null*)] induced an increase for hcArgRS and a decrease for ΔNhcArgRS. The change three nucleotides upstream of codon A from C to G toward an optimal Kozak context [lane A(*Kozak*)] increased the level of production of ΔNhcArgRS. The crude extraction of 293T cells transfected with vector PCDNA3 was used as a control.

and human cells (*18*, *19*, *30*). Also, a single human cytoplasmic ArgRS mRNA of 2.2 kb has been detected by Northern blot analysis (*19*). However, two forms of cytoplasmic ArgRS are found in some mammalian cells (*12*, *13*, *32*). Amino acid sequencing revealed that the long form of the enzyme had a 72-amino acid extension at the N-terminus, compared with the short form (*12*). This N-terminal extension of ArgRS is essential for this enzyme to be assembled into the complex (*12*−*15*). Amino acid sequence homology analysis demonstrated that the 72-amino acid extension, followed by a Met (73 amino acids in *Cricetulus longicaudatus*), is highly conserved from *Danio rerio* (zebrafish) to *Homo sapiens* (Figure 6), although it differs from those of *Caenorhabditis elegans* and *Drosophila* (data not shown). This suggests that two forms of ArgRS might exist in vertebrates. Alignment of the primary structure of ArgRS from different vertebrate cells suggests there is also likely to be two forms of cytoplasmic ArgRS in human cells (Figure 6).

It was reported that the expression of cytoplasmic ArgRS from Chinese hamster ovary cells in *E. coli* produces inclusion bodies and that the fraction of soluble enzyme has a reduced (100-fold) specific activity compared with that of the enzyme produced in yeast (*11*, *14*). In this work, we subcloned human cytoplasmic cDNA encoding ArgRS in two forms in *E. coli*. *E. coli* strain BL21-CodonPlus (DE3)-RIL was used, which contains extra copies of the rare *E. coli argU*, *ileY*, and *leuW* tRNA genes that can correct codon

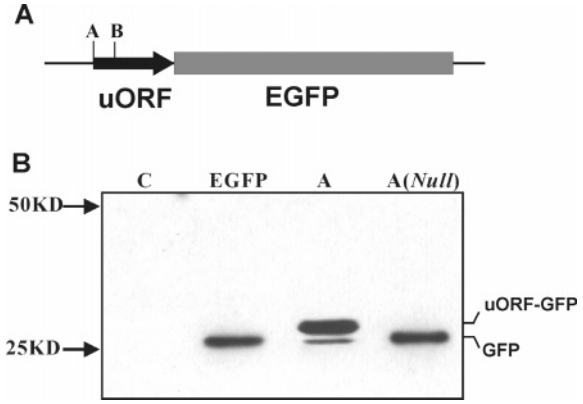

FIGURE 5: Translation initiated at the uORF. (A) Schematic structure of the mRNA encoding the fused proteins with EGFP from the reporter plasmid. (B) Detection of proteins produced from the reporter by Western blotting. In lane A, two proteins can be detected using the M2 antibody, and the small one has the same molecular mass as the EGFP. In lane A(*Null*), only the small protein could be detected when codon A was mutated.

bias in *E. coli*. This allowed expression of the genes encoding human cytoplasmic ArgRS. In addition, the transformants containing the genes encoding human cytoplasmic ArgRS were cultured in LB medium at 25 °C. Under the conditions described above, most of the human ArgRS, either full-length hcArgRS or ΔNhcArgRS, was produced in a soluble form. ΔNhcArgRS, without the N-terminal extension, was produced at a greater level and had a higher turnover number than hcArgRS. By using the purified ΔNhcArgRS as the antigen, we obtained an antibody which could be used to test ArgRS in different human cells using Western blots. It was found that the two forms of ArgRS exist in all three screened human cell types. The hcArgRS:ΔNhcArgRS ratio for the three kinds of human cells was 2:1 to 3:1. This suggested that the two forms of ArgRS are ubiquitous in vertebrates, when taken together with other reported results (*11*−*14*).

RNA analysis and cDNA cloning support the contention that only one form of ArgRS mRNA exists in human cells (*19*, *30*). N-Terminal amino acid sequence analysis of the purified free ArgRS from rat liver suggests that low-molecular weight ArgRS is probably a distinct translational form, rather than a proteolytically derived product (*12*). Many reports indicate that more than one product could be produced from a single transcript by a different translational initiation (*24*, *25*). Recently, evidence has shown that yeast cytoplasmic and mitochondrial forms of GlyRS and AlaRS are also two distinct translational products from a single transcript (*20*, *21*).

At the 5′-end of cDNA of human cytoplasmic ArgRS, there are three ATG start codons; start codons B and C are located in one open reading frame. Upstream of start codon B is an ATG start codon A, from which a short upstream open reading frame (uORF, 24 codons) could be translated (Figure 4). From start codon B or C, hcArgRS and ΔNhcArgRS are translated separately. The hcArgRS:ΔNhcArgRS ratio, as from wild-type mRNA, was 10:1. When start codon B was mutated to ATT, all the recruited ribosomes scan to codon C to initiate translation, the amount of ΔNhcArgRS increasing, while that of hcArgRS could not be detected. However, when start codon C was mutated to ATT, the ribosomes used only start codon B and only hcArgRS can be detected, while ΔNhcArgRS disappeared. If translation was carried out from start codon A, ribosomes could overpass codon B, and the translation was stopped upstream of start codon C. A part of ΔNhcArgRS was translated by re-initiation of ribosomes at start codon C, which is why when codon A was mutated into the optimal initiating Kozak sequence, a small amount ArgRS was found, the hcArgRS:ΔNhcArgRS ratio then being 1:1. The mutation of start codon A to null slightly decreased the level of production of ΔNhcArgRS initiating at start codon C. The reporter constructed by fusion of the uORF to EGFP confirmed the initiation of translation at the uORF (Figure 5). The data showed that the two forms of human cytoplasmic ArgRS were produced from different translational initiations within a single transcript.

ΔNhcArgRS could still be detected when the A site was mutated. There may be other mechanisms involved, such as leaky scanning, which lead to the short form of ArgRS. To date, no experimental results that shed light on the biological roles for the two forms of ArgRS are available. The catalytic rate of ΔNhcArgRS was somewhat greater than that of hcArgRS, but overall, they are considered to be at a similar level (*11*, *14*, *15*). Regulation of translational initiation is required for the fine-tuning of protein levels during cell proliferation and differentiation, and for spatial and temporal regulation of protein expression during embryogenesis (*24*, *25*, *33*). Many external stimuli, such as the fluctuation in nutrient levels and availability, growth factors, nitrogen, and inducers of differentiation, can all affect this process (*25*). Considering the fact that ArgRS provides a substrate for

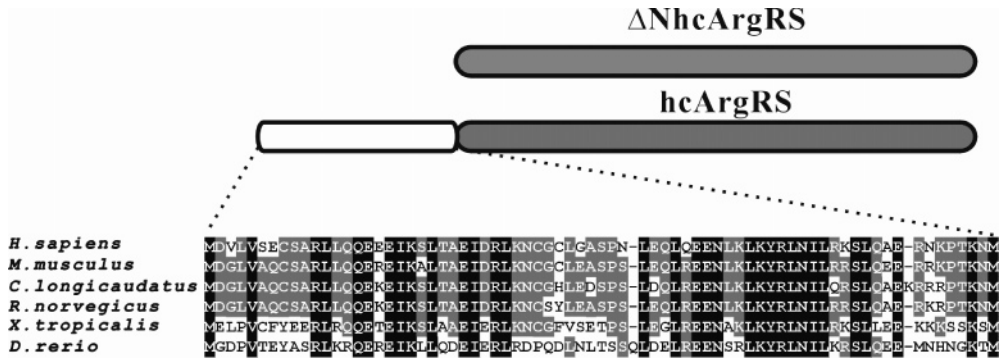

FIGURE 6: Conservation of the N-terminal extension from *Xenopus tropicalis* to *H. sapiens*. The 72-amino acid extension was found from *X. tropicalis* to *H. sapiens* (73 amino acids in *C. longicaudatus*) and to interact with p43 and then to be assembled in the aaRS complex (*11*−*15*). Just after the extension is a Met, and translation initiated from this Met can produce the short form of ArgRS (*12*).

arginyl-tRNA protein transferase and considering the model of Sivaram et al. (*12*), there may be a regulatory mechanism requiring two different ArgRS proteins. We tested whether the expression of the two ArgRSs could be modulated by amino acid starvation. However, limitation of Arg in the medium had no effects on the ratio of the two proteins in 293T cells (data not shown).

REFERENCES

1. Ibba, M., and Söll, D. (2000) Aminoacyl-tRNA synthesis, *Annu. Rev. Biochem. 69*, 617−650.
2. Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990) Partition of tRNA synthetase into two classes based on mutually exclusive sets of sequence motif, *Nature 26*, 223−228.
3. Mirande, M. (1991) Aminoacyl-tRNA synthetase family from prokaryotes and eukaryotes: Structural domains and their implications, *Prog. Nucleic Acid Res. Mol. Biol. 40*, 95−142.
4. Kisselev, L. L., and Wolfson, A. D. (1994) Aminoacyl-tRNA synthetases from higher eukaryotes, *Prog. Nucleic Acid Res. Mol. Biol. 48*, 83−142.
5. Yang, D. C. (1996) Mammalian aminoacyl-tRNA synthetases, *Curr. Top. Cell Regul. 34*, 101−136.
6. Quevillon, S., Robinson, J. C., Berthonneau, E., Siatecka, M., and Mirande, M. (1999) Macromolecular assemblage of aminoacyl-tRNA synthetases: Identification of protein−protein interactions and characterization of a core protein, *J. Mol. Biol. 285*, 183−195.
7. Robinson, J. C., Kerjan, P., and Mirande, M. (2000) Macromolecular assemblage of aminoacyl-tRNA synthetases: Quantitative analysis of protein−protein interactions and mechanism of complex assembly, *J. Mol. Biol. 304*, 983−994.
8. Rho, S. B., Kim, M. J., Lee, J. S., Seol, W., Motegi, H., Kim, S., and Shiba, K. (1999) Genetic dissection of protein−protein interactions in multi-tRNA synthetase complex, *Proc. Natl. Acad. Sci. U.S.A. 96*, 4488−4493.
9. Kim, J. Y., Kang, Y. S., Lee, J. W., Kim, H. J., Ahn, Y. H., Park, H., Ko, Y. G., and Kim, S. (2002) p38 is essential for the assembly and stability of macromolecular tRNA synthetase complex: Implications for its physiological significance, *Proc. Natl. Acad. Sci. U.S.A. 99*, 7912−7916.
10. Kim, M. J., Park, B. J., Kang, Y. S., Kim, H. J., Park, J. H., Kang, J. W., Lee, S. W., Han, J. M., Lee, H. W., and Kim, S. (2003) Downregulation of FUSE-binding protein and c-myc by tRNA synthetase cofactor p38 is required for lung cell differentiation, *Nat. Genet. 34*, 330−336.
11. Lazard, M., Agou, F., Kerjan, P., and Mirande, M. (2000) The tRNA-dependent activation of arginine by arginyl-tRNA synthetase requires inter-domain communication, *J. Mol. Biol. 302*, 991−1004.
12. Sivaram, P., and Deutscher, M. P. (1990) Existence of two forms of rat liver arginyl-tRNA synthetase suggests channeling of aminoacyl-tRNA for protein synthesis, *Proc. Natl. Acad. Sci. U.S.A. 87*, 3665−3669.
13. Vellekamp, G., and Deutscher, M. P. (1987) A basic NH2-terminal extension of rat liver arginyl-tRNA synthetase required for its association with high molecular weight complexes, *J. Biol. Chem. 262*, 9927−9930.
14. Guigou, L., Shalak, V., and Mirande, M. (2004) The tRNA-interacting factor p43 associates with mammalian arginyl-tRNA synthetase but does not modify its tRNA aminoacylation properties, *Biochemistry 43*, 4592−4600.
15. Park, S. G., Jung, K. H., Lee, J. S., Jo, Y. J., Motegi, H., Kim, S., and Shiba, K. (1999) Precursor of pro-apoptotic cytokine modulates aminoacylation activity of tRNA synthetase, *J. Biol. Chem. 274*, 16673−16676.
16. Nathanson, L., and Deutscher, M. P. (2000) Active aminoacyl-tRNA synthetases are present in nuclei as a high molecular weight multienzyme complex, *J. Biol. Chem. 275*, 31559−31562.
17. Vellekamp, G., Sihag, R. K., and Deutscher, M. P. (1985) Comparison of the complexed and free forms of rat liver arginyl-tRNA synthetase and origin of the free form, *J. Biol. Chem. 260*, 9843−9847.
18. Lazard, M., and Mirande, M. (1993) Cloning and analysis of a cDNA encoding mammalian arginyl-tRNA synthetase, a component of the multisynthetase complex with a hydrophobic N-terminal extension, *Gene 132*, 237−245.
19. Girjes, A. A., Hobson, K., Chen, P., and Lavin, M. F. (1995) Cloning and characterization of cDNA encoding a human arginyl-tRNA synthetase, *Gene 164*, 347−350.
20. Chang, K. J., and Wang, C. C. (2004) Translation initiation from a naturally occurring non-AUG codon in *Saccharomyces cerevisiae*, *J. Biol. Chem. 279*, 13778−13785.
21. Tang, H. L., Yeh, L. S., Chen, N. K., Ripmaster, T., Schimmel, P., and Wang, C. C. (2004) Translation of a yeast mitochondrial tRNA synthetase initiated at redundant non-AUG codons, *J. Biol. Chem. 279*, 49656−49663.
22. Pestova, T. V., Kolupaeva, V. G., Lomakin, I. B., Pilipenko, E. V., Shatsky, I. N., Agol, V. I., and Hellen, C. U. (2001) Molecular mechanisms of translation initiation in eukaryotes, *Proc. Natl. Acad. Sci. U.S.A. 98*, 7029−7036.
23. Hellen, C. U., and Sarnow, P. (2001) Internal ribosome entry sites in eukaryotic mRNA molecules, *Genes Dev. 15*, 1593−1612.
24. Meijer, H. A., and Thomas, A. A. (2002) Control of eukaryotic protein synthesis by upstream open reading frames in the 5′-untranslated region of an mRNA, *Biochem. J. 367*, 1−11.
25. Dever, T. E. (2002) Gene-specific regulation by general translation factors, *Cell 108*, 545−556.
26. Wu, J. F., Wang, E. D., Wang, Y. L., Gilbert, E., and Gangloff, J. (1999) Gene Cloning, Overproduction and Purification of *Escherichia coli* $tRNA^{Arg}2$, *Acta Biochim. Biophys. Sin. 31*, 226−232.
27. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem. 72*, 248−254.
28. Liu, M. F., Huang, Y. W., Wu, J. F., Wang, E. D., and Wang, Y. L. (1999) Effect of the cysteine residues on the activity of arginyl-tRNA synthetase from *Escherichia coli*, *Biochemistry 38*, 11006−11011.
29. Ling, C., Yao, Y. N., Zheng, Y. G., Wei, H., Wang, L., Wu, X. F., and Wang, E. D. (2005) The C-terminal appended domain of human cytosolic leucyl-tRNA synthetase is indispensable in its interaction with arginyl-tRNA synthetase in the multi-tRNA synthetase complex, *J. Biol. Chem. 280*, 34755−34763.
30. Strausberg, R. L., et al. (2002) Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences, *Proc. Natl. Acad. Sci. U.S.A. 99*, 16899−16903.
31. Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes, *Cell 44*, 283−292.
32. Som, K., and Hardesty, B. (1975) Isolation and partial characterization of an aminoacyl-tRNA synthetase complex from rabbit reticulocytes, *Arch. Biochem. Biophys. 166*, 507−517.
33. Kuersten, S., and Goodwin, E. B. (2003) The power of the 3′ UTR: Translational control and development, *Nat. Rev. Genet. 4*, 626−637.

BI051675N