The mRNA of Human Cytoplasmic Arginyl-tRNA Synthetase Recruits Prokaryotic Ribosomes Independently*

Received for publication, March 3, 2014, and in revised form, May 29, 2014. Published, JBC Papers in Press, June 4, 2014, DOI 10.1074/jbc.M114.562454

Fang Yang‡, Quan-Quan Ji‡, Liang-Liang Ruan‡, Qing Ye‡, and En-Duo Wang*1

From the 1Center of RNA Research, State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China and the 2School of Life Science and Technology, ShanghaiTech University, 319 Yue Yang Road, Shanghai 200031, China

Background: Two isoforms of human cytoplasmic arginyl-tRNA synthetase (hcArgRS) are produced from a single mRNA. There are two isoforms of cytoplasmic arginyl-tRNA synthetase (hcArgRS) in human cells. The long form is a component of the multiple aminoacyl-tRNA synthetase complex, and the other is an N-terminal truncated form (ΔNhArgRS), free in the cytoplasm. It has been shown that the two forms of ArgRS arise from alternative translational initiation in a single mRNA. The short form is produced from the initiation at a downstream, in-frame AUG start codon. Interestingly, our data suggest that the alternative translational initiation of hcArgRS mRNA also takes place in Escherichia coli transformants. When the gene encoding full-length hcArgRS was overexpressed in E. coli, two forms of hcArgRS were observed. The N-terminal sequencing experiment identified that the short form was identical to the ΔNhArgRS in human cytoplasm. By constructing a bicistronic system, our data support that the mRNA encoding the N-terminal extension of hcArgRS has the capacity of independently recruiting E. coli ribosomes. Furthermore, two critical elements for recruiting prokaryotic ribosomes were identified, the “AGGA” core of the Shine-Dalgarno sequence and the “A-rich” sequence located just proximal to the alternative in-frame initiation site. Although the mechanisms of prokaryotic and eukaryotic translational initiation are distinct, they share some common features. The ability of the hcArgRS mRNA to recruit the prokaryotic ribosome may provide clues for shedding light on the mechanism of alternative translational initiation of hcArgRS mRNA in eukaryotic cells.

Aminoacyl-tRNA synthetases (aaRS) catalyze the ligation of amino acids to the 3′ end of its cognate tRNA, generating aminoacyl-tRNAs that are transferred to the ribosome for translation, playing an essential role in the quality control system of protein biosynthesis (1). On the basis of structural characteristics, 20 aaRSs can be classified into two groups. Class I aaRSs have “HIGH” and “KMSKS” signature sequences located in the Rossmann fold active site, and class II aaRSs are characterized by three homologous motifs (2). In higher eukaryotes, a high molecular weight multi-aaRS complex (MSC) has been discovered, consisting of nine aaRSs and three auxiliary factors (p18, p38, and p43) (3). Currently, the MSC is considered to be a reservoir for these nine aaRSs and controls the flow of synthetases between their canonical translational functions and non-canonical functions (4, 5).

Arginyl-tRNA synthetase (ArgRS) is a member of the class I aaRSs (2) and requires the presence of the cognate tRNA for amino acid activation (6). Human cytoplasmic ArgRS (hcArgRS) is a particular synthetase in that it exists in two forms (7). The long form is a component of MSC, and its N-terminal 72-peptide extension is responsible for the interaction with other components of MSC (8, 9). The short form is an N-terminal extension truncated form (ΔNhArgRS) that is free in the cytoplasm (7). Although both forms of ArgRS exhibit similar catalytic characteristics in vitro (10), the exact biological function of the ΔNhArgRS still remains unclear. It has been hypothesized that the free enzyme provides Arg-tRNAArg as the substrate for arginyl-tRNA transferase (ATE1) (11). However, not enough experimental evidence has been found to support this hypothesis.

There is an argument related to how the short form of ArgRS is produced in mammalian cells. A previous study indicated that the free form in rat liver extracts was derived from the complex form by a limited endogenous proteolysis (10). To resolve this issue, the endogenous low molecular weight ArgRS was separated and purified from rat liver tissue, and an N-terminal sequence analysis suggested that it is probably a distinct translation product (11). A Northern blot analysis revealed that there is only one single transcript of approximate 2.2 kb in human cells, corresponding to the high molecular weight form of ArgRS (13). Furthermore, 5′ RACE (Rapid Amplification of cDNA Ends) was conducted to map the 5′ end of hcArgRS mRNA, which confirmed that there is only one mRNA encoding ArgRS in human 293T cells (14). Moreover, it has been
HcArgRS mRNA Recruits E. coli Ribosomes

demonstrated that the two forms of hcArgRS were produced from two translational initiations by a single mRNA. The free form was a product initiated at a downstream, in-frame AUG start codon (14).

Here our data showed that the in-frame alternative initiation of hcArgRS could take place not only in mammalian cells but also in prokaryotic cells. Two forms of hcArgRS were also observed when the gene encoding the full-length hcArgRS was overexpressed in Escherichia coli cells, and the short form corresponds to ΔNhcArgRS in human cells. Moreover, our data showed that the mRNA encoding the N-terminal extensional region of hcArgRS has the capacity of independently recruiting the E. coli ribosome.

EXPERIMENTAL PROCEDURES

Materials—Pyrobest DNA polymerase and the dNTP mixture were purchased from TaKaRa. The DNA fragment rapid purification kit and the plasmid extraction kit were purchased from Tiangen Biotech. T4 ligase and restriction endonucleases were obtained from MBF Fermentas. Oligonucleotide primers were synthesized by Invitrogen. The KOD Plus mutagenesis kit was purchased from TOYOBO. Nickel-nitrilotriacetic acid (Ni-NTA) Superflow was purchased from Qiagen Inc. The His6 tag monoclonal antibody and luciferase antibody were from Sigma. The GST monoclonal antibody was purchased from Epitomics. The GST monoclonal antibody and luciferase antibody were from Sigma.

Cloning and Mutagenesis—The plasmids used for expressing N-terminal His6-tagged hcArgRS and ΔNhcArgRS in E. coli were pMT7H6-hcArgRS and pMT7H6-ΔNhcArgRS, which have been constructed previously in our laboratory (15). The vector pGEX-4T-1 was used to construct GST-fused hcArgRS. The amplified gene encoding full-length hcArgRS with Pyrobest DNA polymerase was inserted into the vector between the BamHI and EcoRI restriction sites. The single-point mutation, insertion, and deletion of hcArgRS from different plasmids were performed according to the protocol of the KOD Plus mutagenesis kit. DNA sequencing was determined by Invitrogen.

Overexpression of the Gene Encoding hcArgRS in E. coli—BL21(DE3) cells were transformed with a plasmid carrying the gene encoding His6-tagged or GST-fused hcArgRS, respectively. The transformants containing each plasmid were grown at 37 °C to reach an A600 of 0.6 in 50 ml of Luria broth medium containing 100 μg/ml ampicillin and were induced by adding isopropyl 1-thio-β-D-galactopyranoside (IPTG) to a final concentration of 200 μM. After a 5-h induction at 22 °C, the cells were harvested and suspended in PBS buffer with 0.5 mM PMSF. The whole cell extracts (WCE) of E. coli were obtained after supersonic treatment. The His6-tagged hcArgRS could be purified by Ni-NTA affinity chromatography. Similarly, the N-terminal His6-tagged ΔNhcArgRS was obtained by the same procedure.

N-terminal Sequencing—Edman degradation sequencing was used to determine the N-terminal sequence of the short form of hcArgRS produced in E. coli. The His6 tag was inserted at the C-terminal of the GST-fused hcArgRS by KOD Plus mutagenesis. The constructed plasmid (pGEX-hcArgRS-His6) was transformed into BL21(DE3) cells, and the overexpression of the transformants was performed as described previously (15). The purified protein had over 95% purity by SDS-PAGE assay. The sequence of the N terminus was determined using a PROCISETM4929cLC protein sequencer (ABI).

Western Blot Analysis—Whole cell extracts of E. coli were separated by SDS-PAGE, and the protein bands on the gel were transferred onto PVDF membranes. After blocking with 5% (w/v) nonfat dried milk, the membrane was incubated with the appropriate antibody overnight at 4 °C. The membrane was washed three times with phosphate-buffered saline plus 0.05% Tween 20 (PBST), and the secondary antibody conjugated to horseradish peroxidase was added and incubated at room temperature for 2 h. Super Signal chemiluminescent substrate (Thermo Scientific) was sprayed onto the washed PVDF membranes, and images were acquired using a Fujifilm LAS-4000 system configured for chemiluminescence.

RESULTS

Alternative Translational Initiation of hcArgRS mRNA Occurs in E. coli—We overexpressed the gene encoding full-length hcArgRS in E. coli cells with the pMT7H6-hcArgRS plasmid constructed previously (15). After IPTG induction, the target band corresponding to the full-length hcArgRS (~72 kDa) on the gel after SDS-PAGE was observed and could be purified by Ni-NTA affinity chromatography (15). The obtained enzyme was N-terminal His6-tagged. However, we also observed another band on the gel that appeared between 55 and 72 kDa after IPTG induction (Fig. 1A) that was similar to the molecular weight of ΔNhcArgRS found in human cytoplasmic cells. Furthermore, after induction of the overexpressed gene encoding hcArgRS fused with GST at its N-terminal, we again observed the band between 55 and 72 kDa on the gel (Fig. 1A), in addition to the band corresponding to the fused protein GST-hcArgRS. To test whether the smaller band was from hcArgRS, we examined the bands by Western blot analysis using the monoclonal antibody of hcArgRS and observed two bands in each E. coli WCE containing the gene of the full-length hcArgRS (Fig. 1B). The monoclonal antibody of hcArgRS recognized the band between 55 and 72 kDa, suggesting that another form of hcArgRS was produced when wild-type hcArgRS (His-tagged or GST-tagged) was produced in E. coli cells.

To further reveal the existence of another form of hcArgRS in the E. coli transformants containing the full-length ArgRS gene, N-terminal sequencing was performed. Firstly, a His6 tag was added to the C-terminal of GST-hcArgRS, and the recombined plasmid was named pGEX-hcArgRS-His6. The E. coli BL21(DE3) cells were transformed with pGEX-hcArgRS-His6 to produce ArgRS (Fig. 2A). The fused protein GST-hcArgRS-His6 could be over-produced in BL21 transformants, but it was found in pellet (Fig. 2B, lane 2) and could not be purified. While the short form of hcArgRS with His6-tag at the C terminus was partially soluble (Fig. 2B, lane 3) and was successfully purified by Ni2+ -NTA affinity chromatography (Fig.
**HcArgRS mRNA Recruits E. coli Ribosomes**

2B, lane 5). Surprisingly, the N terminus of the purified protein with a molecular mass of 55 kDa was MINIIISRLQE, which exactly matched the N terminus of ΔNhArgRS in human cells (Fig. 2C). Previous data from our laboratory suggested that ΔNhArgRS was produced from alternative translational initiation of the same mRNA in human cells (14). Our current results suggest that the alternative translational initiation of hcArgRS mRNA may also occur in *E. coli* transformants. However, the short form of hcArgRS produced in *E. coli* transformants may be produced from the degradation of the full-length protein. Therefore, we mutated the potential translational reinitiation codon of Met73 to Ala (M73A) in hcArgRS in pGEX-hcArgRS-His6. Interestingly, the short form of hcArgRS disappeared completely (Fig. 3A). Our data show that alternative translational initiation indeed takes place from the AUG codon of Met73 and that the short form of ArgRS produced in *E. coli* is not a product of degradation. In Fig. 3B, we changed the expressing vector pMF77H6 to pET28a, and we also observed the production of the short form corresponding to ΔNhArgRS, suggesting that the sequences of different vectors did not affect the occurrence of translational reinitiation from the Met73 codon. Therefore, the internal sequence of hcArgRS mRNA might be responsible for the alternative translational initiation in *E. coli* cells.

The mRNA Encoding the N-terminal Extension Region of hcArgRS Has the Capacity to Independently Recruit the *E. coli* Ribosome—To clarify the mechanism of translational reinitiation of hcArgRS mRNA in *E. coli* transformants, we inserted a TAA stop codon at the 3′ end of the GST ORF in pGEX-hcArgRS-His6 to construct two ORFs encoding GST and hcArgRS-His6, respectively (Fig. 4A). When the recombinant plasmid with two ORFs was overexpressed in *E. coli*, we could observe the production of both proteins, GST and a short form of hcArgRS, ΔNhArgRS (Fig. 4A). This result suggests that the reinitiation of mRNA encoding hcArgRS from the Met73 codon still occurs. Similar as above, no ΔNhArgRS was observed when Met73 was mutated to Ala (data not shown). These data show that mRNA encoding the N-terminal sequence of hcArgRS can recruit the *E. coli* ribosome independently. However, the production of ΔNhArgRS in *E. coli* cells may be attributable to cryptic promoter activity of mRNA. Therefore, we deleted the tac promoter of pGEX-hcArgRS-His6 with a GST stop codon insertion and observed that both GST and ΔNhArgRS diminished (Fig. 4A). These data indicate that a cryptic promoter does not exist in the mRNA encoding N-terminal sequence of hcArgRS and that ΔNhArgRS is the product of translational reinitiation from the alternative initiation site. Furthermore, we inserted a DNA sequence encoding an N-terminal 72-peptide of hcArgRS at the 5′ end of a luciferase gene to construct a chimeric sequence called “N72-luciferase.” The chimeric sequence replaced the hcArgRS-His6 gene in the plasmid pGEX-hcArgRS-His6, including a GST stop codon insertion (Fig. 4B). When the construct was overexpressed in *E. coli*, we observed that the luciferase protein and the GST protein were produced (Fig. 4B). Concomitantly, when the initiation codon of the luciferase gene was mutated, the luciferase protein was no longer detected (Fig. 4B). These data suggest that the alternative translational initiation is determined by the mRNA encoding the N-terminal 72-peptide of hcArgRS and has no relationship with the mRNA encoding the internal sequence of ΔNhArgRS, which further confirms the phenomenon that the mRNA encoding the N-terminal extension region of hcArgRS can independently recruit the *E. coli* ribosome.

Identification of Crucial Sequences Responsible for Alternative Translational Initiation in *E. coli*—To identify the crucial elements responsible for the alternative translational initiation in *E. coli*, we constructed a series of DNA fragments that encoded truncation mutants of the N-terminal extension of hcArgRS in the plasmid pGEX-hcArgRS-His6, with a GST stop codon insertion, as shown in Fig. 5A. Western blot analysis was used to detect alternative translational initiation. When the DNA sequence encoding the N-terminal 24, 48, and 60 amino acid residues was deleted, respectively, each construct could be expressed to produce ΔNhArgRS, similar to the wild-type plasmid (Fig. 5B). However, upon deletion of the DNA sequence encoding the 72-peptide at the N-terminal, the ΔNhArgRS protein band was not observed. However, the GST protein was still detected. Therefore, these results imply that the mRNA sequence encoding the 60th to 72nd amino acid residues proximal to the translational reinitiation site (Met73) may be responsible for recruiting ribosomes in *E. coli*. To test this possibility, we deleted the 36-nucleotide sequence encoding positions 61–72 of hcArgRS (Δ61–72), which is proximal to the reinitiation site, and, as expected, ΔNhArgRS was not observed (Fig. 5B), indicating that the mRNA fragment immediately in front of the reinitiation site plays a crucial role for recruiting the prokaryotic ribosome.
We further analyzed the 36 nucleotide sequence proximal to the Met73 codon in the N-terminal extensional region to gain more information and found that the fragment contains an "AGGA" sequence that is the core of the Shine-Dalgarno (SD) sequence (Fig. 6A). When the AGGA sequence was mutated to TCCT, the expression of \( \Delta \text{Nh} \text{cArgRS} \) declined dramatically compared with the wild-type gene (Fig. 6B, lanes 1 and 4), suggesting that the SD-like sequence is important for recruiting the prokaryotic ribosome. In addition, there is an "A-rich" AAAAAAT sequence just proximal to Met73 (Fig. 6A). When the A-rich sequence was mutated to GCGACG, we could not observe the appearance of \( \Delta \text{Nh} \text{cArgRS} \) (Fig. 6B, lane 2), which suggests that the sequence plays a more crucial role for effective translational reinitiation. The A-rich sequence may provide a more open and flexible mRNA structure that is easier for the \( E. \text{coli} \) ribosome to initiate and load on the alternate in-frame initiation site.

**DISCUSSION**

The mechanism for alternative translation initiation is an important layer of gene expression control. The protein products of alternative translation initiation are usually required for stress response (16). Our previous investigation has revealed that there are two forms of ArgRS in human cells and that the truncated form is the product of alternative translation initiation (14). Although the exact function of \( \Delta \text{Nh} \text{cArgRS} \) in the cytoplasm is still unknown, clarifying the mechanism of alternative translation initiation may be helpful to understand the role of this truncated protein. We have shown that a short uORF (upstream ORF) is located in the 5' UTR of the ORF encoding full-length hcArgRS and that the existence of the
uORF facilitates the production of the ΔNhArgRS protein (14). However, when we mutated the uORF, ΔNhArgRS was still detected, which implies that there may be other mechanisms involved in the process of translation reinitiation. Furthermore, when the N-terminal extension region of hArgRS without the uORF was inserted into the 5′ end of the luciferase gene and the plasmid was transfected into 293T cells, two forms of luciferase protein were also produced. One form was the N-terminal 72-peptide extension of hcArgRS-fused luciferase, and the other was the luciferase protein alone, which suggests that some elements on the mRNA encoding N-terminal extension might be responsible for translation reinitiation.3

In this research, we show that the alternative translational initiation of human ArgRS also occurs in E. coli transformants and that the translational reinitiation site is the in-frame AUG codon encoding Met73, which is the same as that in human cells. In addition, the mRNA encoding the N-terminal extensional region of hArgRS can independently recruit E. coli ribosomes. Furthermore, two crucial elements in this region that are responsible for reinitiation in E. coli were identified. One element is the SD-like sequence located in front of the reinitiation AUG codon. It is well known that the prokaryotic translational initiation is triggered by the binding of the 3′ end of the 16S rRNA to the complementary sequence termed Shine-Dalgarno, recruiting the prokaryotic ribosome to the 5′ UTR of target mRNA (17). In E. coli, some genes are frequently expressed from polycistronic messengers. Ribosomes have the potential to initiate at internal cistrons, and the presence of an SD-like

References:
3 F. Yang, Q.-Q. Ji, L.-L. Ruan, Q. Ye, and E.-D. Wang, unpublished data.
structure in front of the downstream cistron is believed to enhance the translational reinitiation efficiency (18). In our case, the N-terminal extensional region of hcArgRS behaves similar to that of the intercistron space, which favors the initiation of the downstream cistron. When the SD-like sequence is mutated, the efficiency of reinitiation declined dramatically. However, a negligible amount of ΔNhcArgRS was still detected (Fig. 6B). This result is in agreement with a previous study showing that, in the absence of an SD sequence, translation reinitiation in E. coli still occurred but that its efficiency was about 10% of that found in the SD-carrying counterpart (18). The other determinant in the N-terminal region of hcArgRS responsible for reinitiation in E. coli is the A-rich sequence located just proximal to the reinitiation triplet. Translation efficiency is strictly correlated with the fraction of mRNA molecules in which the ribosome binding site is unfolded, indicating that initiation is completely dependent on spontaneous unfolding of the entire initiation region (19). In this study, the A-rich region that is in close proximity to the reinitiation triplet will probably avoid the formation of a more complex duplex in the ribosome binding site, and the exposure of this region makes it available for E. coli ribosome binding. Additionally, it has been reported that AU-rich sequences within the 5′ UTR enhance translation and stabilize mRNA in E. coli (20). Our investigation is consistent with the idea that the reinitiation efficiency is influenced by the primary sequences present in the neighborhood of the reinitiation triplet in E. coli (21).

In this study, a mouse monoclonal antibody against hcArgRS was raised to detect the different isoforms of hcArgRS in E. coli cell lysate when hcArgRS was overexpressed in E. coli. Although the E. coli ArgRS shares sequence homology with ΔNhcArgRS, the monoclonal antibody failed to recognize the E. coli enzyme (data not shown). Therefore, we exclude the possibility that the observed small band in Fig. 1B is the E. coli enzyme.

It has rarely been reported that translational reinitiation occurs in both eukaryotic and prokaryotic cells simultaneously. Therefore, our findings may help to deepen the understanding related to the phenomenon of alternative translational initiation. The majority of eukaryotic cellular mRNAs initiates translation by the cap-dependent scanning mode of translation initiation, which depends on the recognition of the 5′ cap structure (22). However, even though the mechanisms of AUG selection differ fundamentally between bacteria and eukaryotes, there are important structural and mechanistic features of initiation conserved between these two kingdoms (22). Similar to the base-pairing between the SD sequence upstream of the initiation codon and the complementary nucleotides of 16 S rRNA in prokaryotes, the mRNA–tRNA base-pairing mechanism for translation initiation in eukaryotes has also been observed (23–26). It has been demonstrated that a 9-nucleotide element found in the mouse Gtx homeodomain mRNA facilitates translation initiation by base-pairing to 18 S rRNA (24). In addition, the mRNA/18 S rRNA hybridization for translation reinitiation in a eukaryotic system has also been described (25). The internal ribosome entry site of human IGFlR mRNA may recruit 40 S ribosome by a eukaryotic equivalent of the SD-like interaction (26). In eukaryotes, the A-rich sequence at the 5′ UTR also plays an important role for translation initiation. It has been reported that the poly(A) tract in front of the AUG codon in Saccharomyces cerevisiae enhances translation initiation (27). An unstructured A-rich element mediates internal initiation via recruitment of the poly(A) binding protein to the 5′ UTR of invasive growth messages (12). As mentioned previously, although the uORF favors the internal initiation of ArgRS mRNA, there is still another mechanism responsible for the production of ΔNhcArgRS in human cells. Whether there are some determinants in the N-terminal extensional region for the internal initiation of hcArgRS mRNA in human cells will be investigated in the future.

In addition, from a practical point of view, the ability of the hcArgRS mRNA to recruit E. coli ribosomes may be used to further develop an expression vector in which the N-terminal extension could ensure the expression of a gene of interest by coupling its translation to that of the upstream cistron. This type of vector could be applied in the field of overexpression optimization of multiple-subunit proteins in E. coli cells.

Acknowledgments—We thank Mei-Qin Xue and Xian Xia for help with the experiments.

REFERENCES

characterization of cDNA encoding a human arginyl-tRNA synthetase. Gene 164, 347–350