A Multifunctional Repeated Motif Is Present in Human Bifunctional tRNA Synthetase*

(Received for publication, December 2, 1997, and in revised form, February 2, 1998)

Seung Bae Rho‡, Jong Sang Lee‡, Eui-Jun Jeong§, Key-Sun Kim§, Yang Gyun Kim§, and Sunghoon Kim‡¶

From the ‡Department of Biology, Sung Kyun Kwan University, 300 Chunchundong, Jangangu, Suwon, Kyunggido 440-746, Korea, the §Structural Biology Center, Korea Institute of Science and Technology, Cheongryang Box 131, Seoul 139-751, Korea, and the ¶Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Tandem repeats located in the human bifunctional glutamyl-prolyl-tRNA synthetase (EPRS) have been found in many different eukaryotic tRNA synthetases and were previously shown to interact with another distinct repeated motifs in human isoleucyl-tRNA synthetase. Nuclear magnetic resonance and differential scanning calorimetry analyses of an isolated EPRS repeat showed that it consists of a helix-turn-helix with a melting temperature of 59 °C. Specific interaction of the EPRS repeats with those of isoleucyl-tRNA synthetase was confirmed by in vitro binding assays and shown to have a dissociation constant of approximately 2.9 μM. The EPRS repeats also showed the binding activity to the N-terminal motif of arginyl-tRNA synthetase as well as to various nucleic acids, including tRNA. Results of the present work suggest that the region comprising the repeated motifs of EPRS provides potential sites for interactions with various biological molecules and thus plays diverse roles in the cell.

Aminoacyl-tRNA synthetases are essential in protein synthesis, catalyzing the attachment of specific amino acids to cognate tRNAs. Despite the common catalytic role of these enzymes, cytoplasmic tRNA synthetases of higher eukaryotes differ from their lower eukaryotic or prokaryotic counterparts in forming a multienzyme complex (1–4). These multiprotein complexes are known to contain nine synthetases, which react with Glu, Pro, Ile, Leu, Met, Lys, Gln, Arg, and Asp (5). Although the presence of this complex in higher eukaryotic cells has been known for the last two decades, its functional significance and structural features remain elusive.

Among the tRNA synthetases of higher eukaryotes, glutamylyl- and prolyl-tRNA synthetase activities have been found linked in a single polypeptide (6, 7). In human glutamyl-prolyl-tRNA synthetase (EPRS) (1), two domains exhibiting each enzyme activity are connected by a linker that contains three tandemly repeated motifs of 57 amino acids (7). The human prolyl-tRNA synthetase lacking this linker peptide was still active, suggesting that it is not essential for catalytic activity (8). This leaves the possibility that the linker region may play a distinct role in the cell other than the catalytic function.

Peptide sequences homologous to these repeats have also been found in other tRNA synthetases, although they are present as a single copy. They are located in the N-terminal extensions of glycyll- (9, 10), tryptophanyl- (11–13), and histidyl-tRNA synthetases (14, 15), which have been found as free forms, and in the C-terminal extension of methionyl-tRNA synthetase (16), which has been found in the complex. Although it is not clear whether they all play the same role in the cell, the prevalence of these motifs among the eukaryotic tRNA synthetases implies their functional significance.

EPRS of Drosophila melanogaster contains six tandem repeats of these motifs between the two catalytic domains. Overexpression of these motifs in transgenic flies resulted in a decrease of fertility (17). This result further supports the physiological importance of these motifs. It has been proposed that they may serve as a template for assembly of multisynthetase complex or as an anchor to link the complex to other cellular components, such as protein synthesis machinery (6, 17).

The three repeated motifs of human EPRS were previously shown to interact with the two repeated motifs present in the C-terminal extension of human cytoplasmic isoleucyl-tRNA synthetase (IRS) (18). The tandem repeats in human IRS are distinct from those of EPRS, and their homologues have not been found in other tRNA synthetases (19, 20). In the present work, we further analyzed the molecular interactions of the EPRS repeats with those of IRS and with other cellular molecules. Our results suggest that the peptide region containing the repeated motifs of EPRS is multifunctional and allows for multiple interactions with nucleic acids as well as with at least two different tRNA synthetases. Functional meaning of these interactions will be discussed.

EXPERIMENTAL PROCEDURES

Preparation of the EPRS Motif—DNA encoding one of the EPRS repeats (EPRS-R1) was isolated by polymerase chain reaction and subcloned into pET28a (Novagen) using EcoRI and XhoI. The motif was expressed in Escherichia coli strain BL21 (DE3) grown in M9 minimal medium supplemented with 0.3% glucose and 0.1% 15NH4Cl. The His-tagged EPRS motif (from pHEPR1) (Table I) was purified using a Ni2+-column (Invitrogen), the His tag was cleaved off with thrombin, and the EPRS motif was further purified using C18 reverse phase high performance liquid chromatography (Waters 600). NMR samples containing approximately 2 mM isolated peptide in 90% H2O/H210% H2O were prepared by adjusting to pH 5.0 with NaOH. Samples for differential scanning calorimetry (DSC) were prepared by dialyzing the peptide overnight at 4 °C against 20 mM phosphate buffer (pH 5.0). DSC analysis was conducted at 1.0 mg/ml of the purified peptide. An extinction coefficient of 0.57 absorbance units/mg/ml was used to calculate the peptide concentration.
NMR Spectroscopy and DSC—The three-dimensional 1H-15N heteronuclear single quantum coherence-nuclear Overhauser enhancement spectroscopy and 1H-13C heteronuclear single quantum coherence-total correlation spectroscopy NMR spectra were recorded at 30 °C with a Varian UnityPlus 600-MHz spectrometer equipped with an actively shielded Nanodesc gradient. A gradient pumping system enabled the collection of NMR spectra with a gradient strength of 15 G/cm and a gradient duration of 6 ms. The EPRS repeats (Val173–Lys889) were generated (from pHEPR123+) (Table I). The expressed His-EPRS was purified by Ni²⁺ affinity and subsequently by Mono-S column chromatography (Amersham Pharmacia Biotech). Various DNA and RNA preparations labeled with 32P were used for the binding assays. Total yeast tRNA mixture was purchased from Boehringer-Mannheim. Minihelix (26) was obtained from the laboratory of Dr. Schimmel (Massachusetts Institute of Technology), and 50–60 base pair calf thymus DNA was prepared by sonication and gel electrophoresis. Nucleic acid preparations were labeled with 32P using polynucleotide kinase after removing 5′ phosphate with calf intestine phosphatase. The purified His-EPRS at various concentrations (0.004–3.6 μM) was mixed with 5% polyacrylamide solution. After preparation of the gel, the radioactively labeled nucleic acids were loaded into the vertically oriented well. After electrophoresis, the radioactive nucleic acids and their complexes were visualized by autoradiography.

Dissociation constants (Kᵦ) of the EPRS repeats and various nucleic acids were determined as described previously (27). The mobility shifts of the EPRS-nucleic acid complexes were recorded (23) and divided by the maximal possible shift to obtain the values of relative shift (R). The R values were subsequently divided by the respective concentrations of EPRS used for each lane. Scatchard plots of R versus [EPRS] give linear lines, the slopes of which are equivalent to ~1/Kᵦ. The dissociation constants for tRNA, minihelix, and carrier DNA were determined from values obtained three, two, and one independent experiments, respectively.

RESULTS

Structure and Conformational Stability of the EPRS Motif—Human EPRS contains three repeated units consisting of 57 amino acids (Fig. 1A). Secondary structure predictions of the motif suggested that it would form an α-helix (6, 17, 18). To determine the structure and stability in folded configuration, one of the repeated motifs (EPRS-R1) was subjected to NMR and DSC analyses. The chemical shift index of HCa in this motif indicates that the EPRS motif consists of two helices (residues 1–23 and 26–43) and a possible C-terminal β-sheet (residues 48–50) (Fig. 1B). The secondary structure profile of the NMR analysis is consistent with the computer prediction and suggests that it may form a helix-turn-helix or a dimer with intermolecular helix-helix interactions. However, the ratio of van’t Hoff to calorimetric enthalpy (28), which is close to 1, supports a helix-turn-helix fold rather than a dimeric fold. The DSC thermogram indicates that the EPRS motif unfolds at 59 °C and has an enthalpy of 51 kcal/mol, which further supports the compactness of the structure (Fig. 1C). The folding of this protein was completely reversible, with enthalpy recovery of higher than 95% for the second DSC scan.

In Vitro Interaction and Binding Affinity between the EPRS and IRS Repeats—Interaction between the repeated motifs of EPRS and IRS was previously shown by genetic analysis (18). The interaction was further tested by an affinity co-purification experiment. The His-EPRS fusion protein was co-purified with GST-IRS, and the mixture was purified on a Ni²⁺-charged column matrix. GST-IRS was co-purified with His-EPRS but not with His-EMA1 II as a bound complex (Fig. 2). Similar results were obtained when the mixtures were purified by affinity of GST-IRS to glutathione column matrix (data not shown). These results confirmed the specific interaction between the repeats of EPRS and IRS.

We then determined the binding affinity between the EPRS and IRS repeated motifs by surface plasmon resonance using BIAcore system. Nonspecific binding of the isolated EPRS to the sensor chip was observed at 150 mM NaCl but not at 500 mM NaCl (data not shown). To avoid nonspecific binding of
EPRS onto the sensor chip and to better orient the peptide for the binding study, we fused the EPRS repeats to a biotin binding domain (pBEPR123) (Table I) and attached the biotinylated EPRS onto a streptavidin-coated chip at 500 mM NaCl. Subsequently, the purified GST-IRS (from pGIR12) (Table I) was passed over the EPRS-coupled streptavidin chip, and the binding was monitored (Fig. 3). Association of the two proteins was monitored by the increase in response unit. Nonspecific binding of GST-IRS fusion protein to the streptavidin chip was not observed under the experimental conditions employed (data not shown). The buffer without IRS was then added to the chip to monitor dissociation of the two proteins. Average association (k\(_{\text{on}}\)) and dissociation (k\(_{\text{off}}\)) rate constants measured from the response of sensogram were 1.25 ± 0.42 \times 10^3 M^{-1}s^{-1} and 3.08 ± 0.29 \times 10^{-3}s^{-1}, respectively. From these values, the dissociation constant between the two molecules was estimated to be 2.9 ± 1.2 μM.

Interaction of the EPRS Repeats with the N-terminal Motif in RRS—Human DRS and RRS contain unique N-terminal extensions. These peptides in mammalian enzymes have been shown

**Fig. 1.** Sequence, structure, and folding of the repeated motif of EPRS. A, three repeats of human EPRS (R1-R3) and spacer sequences are shown. B, the isolated EPRS-R1 was subjected to NMR analysis. The chemical shift index calculation and the estimation of secondary structure were generated according to the method of Wishart et al. (22). Briefly, chemical shifts of HCo were compared with random coil values. The values +1 and −1 indicate greater and less than random coil values, respectively. Grouping of −1 and +1 predicts α-helix and β-sheet, respectively. Boxes indicate the predicted α-helical regions. C, temperature dependence of the excessive heat capacity of EPRS-R1. Integration under the curve gives an enthalpy at transition temperature. The first scan was monitored at the rate of 1 °C/min up to 100 °C, and the determined T\(_{\text{m}}\) was 59 °C. The second scan was carried out after cooling down the solution containing the peptide to 25 °C.

**A. EPRS Repeats**

R1: 677-DSLVLNRYAVGQDVQRELCPEAEDVAAVQQLLSLKAQYKEKTGQZZYKPNPAPIGQNZISNSSIL
R2: 750-ESKSDYDEVAAQGEVVKKLKEKSPKAKTNAVECLLSLKAQYKEKTGKEPYIPQPPNLSSDDSNRSEENAPLRTF
R3: 827-EAKVLFDKVASQGEVVRKLKTEAKPKDQVDVIAVQELLQLKAQYKLGEVEKPYVSAT

**B. NMR Chemical Shift Index**

**C. DSC Thermogram**

![Graph](image)

**Fig. 2.** Co-purification of EPRS and IRS. The His-EPRS (30 kDa) and His-EMAP II (41 kDa) fusion proteins were mixed with GST-IRS (62 kDa). The mixtures were subjected to Ni\(^{2+}\) affinity purification. Co-purification of GST-IRS with His-EPRS, but not with His-EMAP II, is shown by the arrow.
to be responsible for the formation of multisynthetase complex (29–31). The N-terminal motifs of human DRS (32, 33) and RRS (34, 35) are predicted to form $\alpha$-helices (Fig. 4A). Because EPRS is a component of the multisynthetase complex, it is possible that the repeated motifs of EPRS are involved in multivalent protein-protein interactions with other synthetases. We tested the possibility as to whether the N-terminal extensions of DRS and RRS interact with the EPRS repeats using the yeast two hybrid method.

The motifs present in the four tRNA synthetases were fused to the DNA binding domain, LexA, as well as to the transcriptional activator, B42 (Fig. 4A; Table I). Interactions between all of the sixteen combinations between the four unique motifs of EPRS-R3, IRS-R2, DRS-N, and RRS-N were tested. Among the tested combinations, only the pair between the EPRS and IRS motifs showed positive interaction (Fig. 4B). Subsequently, interactions of the DRS and RRS motifs with the EPRS protein were further studied using the peptides containing different numbers of the repeated motifs (pLEPR series) (Table I). The N-terminal motif of RRS showed an interaction with the EPRS containing three repeats but did not bind to those containing fewer than three (Fig. 4C). In contrast, the N-terminal motif of DRS did not show an interaction with any of the EPRS motifs tested. These results suggest that the three repeats of EPRS make additional interactions with the N-terminal motif of RRS.

Nucleic Acid Binding of the EPRS Repeats—Most eukaryotic tRNA synthetases show an affinity to polyanions, such as nucleic acids (36, 37) and heparin (38). The repeated motifs of EPRS contain a high proportion of positively charged amino acids (about 20%), especially in the central region of the repeated motifs. Therefore, we expected that the EPRS motifs may be able to bind to nucleic acids. Interaction of the EPRS motifs with various nucleic acids was investigated by affinity co-electrophoresis. $^{32}$P-labeled yeast total tRNA, minihelix Ala, and calf thymus DNA of 50–60 base pairs were used for the binding assays. His-EPRS (from pHEPR123+) (Table I) was mixed with the radioactively labeled nucleic acids and subjected to electrophoresis. His-EPRS formed complexes with all of the tested nucleic acids (Fig. 5). The observed apparent dissociation constants were approximately 0.016 μM for yeast tRNA, 0.031 μM for minihelix, and 0.027 μM for calf thymus DNA.
DNA when estimations were done following the previously described procedure (27). These results suggest that the EPRS repeats poorly distinguish different nucleic acids, although they showed a slight preference for tRNA over calf thymus DNA and minihelix. However, the EPRS repeats did not bind to single stranded poly(dA)30 (data not shown), suggesting that they may recognize helical backbone of double stranded nucleic acids.

**DISCUSSION**

The physiological significance of the association between the different eukaryotic tRNA synthetases has been a long-stand-
ing question. The structure and function of the multisynthetase complex are thought to be related to the newly evolved unique peptides present in the tRNA synthetases of higher eukaryotes (3, 4). To get an insight into the functions of these peptides, we focused on the repeated motifs of EPRS located between the two catalytic domains and investigated their structure and interactions. The results of NMR and DSC suggested that these repeats form a stably folded helix-turn-helix (Fig. 1, B and C). The well defined structure of these repeats suggests that they may not be a simple linker of the two catalytic domains. The marginal and spacer regions of the repeated motifs in EPRS contain a high proportion of proline and small amino acids, such as glycine, alanine, and serine. The peptides rich in these residues may form flexible turns between the repeated motifs (Fig. 1A). Thus, the repeated motifs of EPRS may form a higher order conformation rather than a linear structure. The complex three-dimensional structure of these repeats may be responsible for the multiple interactions with various molecules, as shown in this work.

The structure and function of the unique peptides in the eukaryotic tRNA synthetases have been extensively studied in the case of the N-terminal extension of eukaryotic DRS, which is one of the components of the complex. The N-terminal truncated DRS did not associate to the multisynthetase complex in the cell (29), suggesting a role of the N-terminal extension as a complex association domain. However, this domain, when fused to a heterologous protein, did not drive the protein to the complex. It is thus implied that additional interactions via the connected catalytic domain are required to determine the complex formation of DRS. The N-terminal extension of DRS was also shown to bind to tRNA and elongation factor, thereby facilitating the transfer of the bound aspartyl-tRNA for protein synthesis (39, 40). These results suggest that the N-terminal extension of DRS plays versatile roles in the cell as a domain responsible for association to tRNA synthetases and other molecules and also as a channel for tRNA delivery.

The known functions of the DRS extension can be extrapolated to the case of the EPRS repeats. We demonstrated in the present work that the EPRS repeats are involved in specific interaction with IRS (Figs. 2 and 4). The determined dissociation constant between the EPRS and IRS repeats (Fig. 3) is about 100-fold higher than that between DRS and the multisynthetase complex (30). Assuming that all of the components are bound to the complex with a similar binding stability, the interaction between the repeats of EPRS and IRS may not be strong enough to keep the two proteins within the complex. Thus, assembly of the two enzymes to the native complex should be aided by additional interactions. The association of these proteins to the complex can be further contributed to by the interactions with other molecules attracted to the repeated motifs as well as by the interactions of the catalytic domains of EPRS and IRS. In this work, we showed that the EPRS repeats also bind to the N-terminal extension of RRS (Fig. 4C). The interaction between the repeats of EPRS with IRS may be stabilized by their additional interaction with RRS.

The nucleic acid binding activity of the EPRS repeats (Fig. 5) can function as a channel of tRNA for catalysis and protein synthesis, stabilizer of protein-protein interaction, and cellular localization signal of the synthetase complex. Yeast Arc1p protein bound to cytoplasmic methionyl-tRNA synthetase and glutamyl-tRNA synthetase showed the binding activity to tRNA (41). The proposed role of this protein was to facilitate the binding and delivery of tRNA to the catalytic site of the bound tRNA synthetases. Perhaps the EPRS repeats were a functional homologue of yeast Arc1p and later genetically fused to a catalytic domain of various tRNA synthetases during evolution.

Although the EPRS repeats were not essential for the activity of the connected tRNA synthetase (8), overexpression of these repeats in Drosophila resulted in sterility (17). Similarly, a mutation of the gene for yeast Arc1p induced slow growth and cold sensitivity of the cell, although it was not essential for the activities of the bound tRNA synthetases (41). These results together imply that the in vivo function of the EPRS repeats could be more significant and pleiotropic than what they appear to be from the experimental results of isolated in vitro systems.

The results of the present study indicate that there is much similarity between the EPRS repeats and the N-terminal extension of DRS in their structure and interactions. Many other eukaryotic tRNA synthetases also contain unique peptides of similar structural and functional features. All of these peptides are thus expected to play versatile roles in the cell in a similar fashion. The functional versatility would result from the potential of these peptides to accommodate multiple interactions with diverse cellular molecules.

Most of the conserved residues among the EPRS repeats are hydrophobic and positively charged (6). Because interactions between eukaryotic tRNA synthetases have been reported to be driven by hydrophobic interactions (2, 29), these protein-protein interactions are expected to be mediated by hydrophobic residues of the repeated motifs. In contrast, interactions with nucleic acids are likely to involve charged residues. Thus, the repeated motifs of EPRS may be simultaneously involved in protein-protein and protein-nucleic acid interactions using different residues in the motifs. In summary, all of the proposed roles of the unique eukaryotic peptides, including the EPRS repeats, may not be mutually exclusive and can be sequentially or simultaneously achieved in the cell via their versatility of interaction. The complex formation of the tRNA synthetases could provide a means for these peptides to work together in a concerted manner.

Acknowledgments—We thank Drs. K. Manoharan, K. Shiba, S. Martinis, and K. Musier-Forsyth for critical reading and comments and C. Lee and S. G. Park for materials.

REFERENCES
Repeated Motifs in Human Bifunctional tRNA Synthetase


NMR 4, 845–858