Genetic dissection of protein–protein interactions in multi-tRNA synthetase complex

(Received for review November 24, 1998)

ABSTRACT Cytoplasmic aminoacyl-tRNA synthetases of higher eukaryotes acquired extra peptides in the course of their evolution. It has been thought that these appendices are related to the occurrence of the multiprotein complex consisting of at least eight different tRNA synthetase polypeptides. This complex is believed to be a signature feature of metazoans. In this study, we used multiple sequence alignments to infer the locations of the peptide appendices from human cytoplasmic tRNA synthetases found in the multisynthetase complex. The selected peptide appendices ranged from 22 aa of aspartyl-tRNA synthetase to 267 aa of methionyl-tRNA synthetase. We then made genetic constructions to investigate interactions between all 64 combinations of these peptides that were individually fused to nonsynthetase test proteins. The analyses identified 11 (10 heterologous and 1 homologous) interactions. The six peptide-dependent interactions paralleled what had been detected by crosslinking methods applied to the isolated multisynthetase complex. Thus, small peptide appendices seem to link together different synthetases into a complex. In addition, five interacting pairs that had not been detected previously were suggested from the observed peptide-dependent complexes.

Proteins are molecular fossils that may help to unravel the history and mechanism of evolution. Aminoacyl-tRNA synthetases (ARSs) are one of the ancient proteins evolved to decode genetic information to amino acids. Although all of the tRNA synthetases catalyze the same chemical reactions, these enzymes have accumulated a wide range of sequence and structural diversity throughout evolution (1). In particular, the tRNA synthetases of higher eukaryotes have acquired a few features that are not present in those of other organisms. Most intriguing is the presence of noncatalytic peptide appendices, the roles of which are yet to be elucidated. It has been thought that these peptides are involved in protein–protein interactions among the tRNA synthetases that form the multiprotein complex. This complex is another characteristic of the cytoplasmic tRNA synthetases of higher eukaryotes (2–4) and contains at least eight different tRNA synthetase polypeptides [Glu-Pro-tRNA synthetase (EPRS), IRS, LRS, MRS, QRS, RRS, KRS, and DRS; refs. 5–8], as well as three nonsynthetase proteins of 43 kDa, 38 kDa, and 18 kDa (5–11).

Although these complexes have been known for more than two decades, the structural organization and interactions between the components have not been well understood. Structural analysis with electron microscopy has shown that the complex forms an elongated U shape (8). The complex structure has been studied extensively by stepwise dissociation of the components by using nonionic detergent (12, 13), changing salt concentration (14), and chemical crosslinking (15, 16). The multi-ARS complex has been proposed to consist of three subdomains. The base subdomain consists of EPRS, IRS, and LRS, which are the enzymes of higher molecular weights. There are two arm subdomains on the top of the base subdomain. Subdomain I contains a dimer of DRS and monomers of MRS and QRS; subdomain II is made of dimers of KRS and RRS (16).

Most of the structural analyses on the multi-ARS complex have relied on biochemical approaches, as listed above. Because these methods address only the physical relationships of the protein components, other approaches are necessary to determine the molecular mechanisms and the peptide regions involved in the assembly or maintenance of the complex. We thought that a genetic approach would be suitable to investigate these questions. A yeast two-hybrid system has proven useful to analyze protein–protein interactions (17). This method was employed previously to analyze interactions of the repeated motifs of EPRS with the C-terminal repeated motifs of IRS (18), as well as with the N-terminal extension of IRS (19). The interaction between the repeats of EPRS and IRS was confirmed further by biochemical and biophysical methods (19). Interactions of p38 with the other ARS components recently have been reported by using two-hybrid analysis (11).

The role of the peptide appendices in the formation of the multi-ARS complex has been studied in a few cases such as the N-terminal sequences of DRS (20, 21), RRS (22, 23), and KRS (6). However, systematic studies on the structural features and molecular interactions of these peptides in eukaryotic ARSs are needed to understand their functional significance. In the present study, we focused on the peptide appendices deduced from the complex-forming ARSs. We investigated whether these regions actually are involved in protein–protein interactions among the complex-forming ARSs and, if so, how they are connected to each other. Finally, the interactions between the unique peptides of complex-forming tRNA synthetases determined by the genetic method in this work were compared with those previously obtained by biochemical methods.

MATERIALS AND METHODS

Determination of Unique Peptide Appendices. Sequences of the complex-forming human tRNA synthetases were aligned...
with those of the corresponding enzymes from various organisms by using a MACAW program (24) to determine the unique peptide appendices present in each of the complex-forming human tRNA synthetases. Secondary structures of the determined peptide appendices were predicted by using the GOR, DPM, GIBRAT, and HOMOLOGUE methods installed in the ANTHEPROT program (25) and the PAIRCOIL programs (26). The previous reports on the peptide appendices also have been consulted to determine the peptide appendices, especially, in the cases of DRS (20, 21), RRS (22, 23), QRS (27), IRS (28), and EPRS (29). These peptide regions were predicted to be predominantly α-helices and were used for the interaction mapping.

**Subcloning of Human ARSs for Two-Hybrid Assay.** The gene coding for the N-terminal 22 aa (T5–E26) of human DRS was synthesized chemically. The other cDNAs encoding the unique ARS peptides (Fig. 1) were isolated by PCR by using the specific primers and template cDNAs. The templates used were pM116 for KRS (30), pM182 for RRS (this study), pM191 for QRS (this study), pM184 for MRS (H.M. and K.S., unpublished work), and pM208 for LRS (H.M. and K.S., unpublished work). The pM182 that contains the N-terminal part of RRS was constructed by cloning of the 5′ rapid amplification of cDNA ends/PCR product with human fetal brain (Clontech). The pM191 carrying the full-length cDNA of human QRS was constructed by cloning the PCR product of the published sequence (27). The forward and backward primers contained EcoRI and SalI sites, respectively, so that the PCR products could be cloned into pLexA and pJG4–5 vectors (19, 31). For the deletion mapping of the interaction motifs in RRS and DRS, the corresponding primers were synthesized and used to isolate cDNA fragments. These fragments were subcloned into LexA or B42 fusion vectors as described above. The B42 plasmid containing the full DRS was cleaved with EcoRI and NcoI to delete the N-terminal part of the DRS protein. The staggered ends of the digested plasmid were filled with Klenow fragment and self-ligated. The resulting plasmid thus expressed the C-terminal 192 aa from M309 to P500 of DRS and was used for the interaction mapping. The deletion constructs of IRS and EPRS have been described (18). The sequences of the constructs were confirmed, and expression of the LexA and B42 hybrid proteins was tested by immunoblotting with monoclonal antibodies specific to LexA and hemagglutinin (which was tagged between B42 and the fused polypeptides), respectively (data not shown). The alkaline phosphatase-conjugated secondary antibody was used to detect the hybrid proteins by using disodium 3-(4-methoxyspiro[1,2-dioxetane-3,2′-(5′-chloro)tricyclo-[3.3.1.1.3,7]decan]-4-yl)phenylphosphate according to the manufacturer’s instructions (Boehringer Mannheim).

**Determination of Interaction.** Interactions between the unique peptide regions of the complex-forming ARSs were determined by the peptide-dependent induction of the reporter genes, LEU2 and lacZ. The LexA- and B42-hybrid proteins were coexpressed in yeast strain EGY48 (31), and positive interactions between the two-hybrid proteins were detected by cell growth on the leucine-depleted synthetic yeast medium containing 2% (wt/vol) galactose and by the formation of blue colony on the same synthetic medium containing 5 mM X-Gal (5-bromo-4-chloro-3-indolyl β-D-galactoside), 2% (wt/vol) galactose, and 2% (wt/vol) raffinose (18, 31).

**RESULTS**

**Determination of Peptide Appendices in Complex-Forming ARSs.** Sequences of eight human ARS polypeptides found in the multi-ARS complex were compared with those of other species (lower eukaryotes, eubacteria, and archaebacteria) to identify the peptide appendices that are unique to higher eukaryotes (data not shown). In addition to these sequence comparisons, we also took account of the continuity of the predicted secondary structures and inferred the higher eukaryote-specific domains of these enzymes whose locations and lengths are summarized in Fig. 1.

ARSs of smaller molecular weights, including DRS (500 aa; ref. 20), KRS (597 aa; ref. 30), RRS (659 aa; ref. 22, 32), QRS (792 aa; ref. 27), and MRS (900 aa; ref. 33), commonly contain the unique N-terminal peptides, although their sequences and lengths are different from each other. In contrast, ARSs of larger molecular weights, including LRS (1,176 aa; H.M. and K.S., unpublished work) and IRS (1,266 aa; ref. 28), contain the unique peptides in their C-terminal extensions. In particular, the C-terminal extension of IRS contains two tandem repeats of about 90 aa. The bifunctional tRNA synthetase (EPRS; 1,440 aa) is the largest polypeptide containing three unique repeats of 57 aa in the noncatalytic linker region between the two catalytic domains (29).

The motifs homologous to the repeats of EPRS have been found as a single copy in the C-terminal region of MRS (Fig. 1, hatched box), as well as in the N-terminal extensions of GRS, IRS, and WRS that have not been found in the multi-ARS complex. We previously found that the motif in the N-terminal extension of GRS also interacted with the C-terminal repeats of IRS as the repeated motif in the EPRS (18), although its interaction strength was weaker than that of the three repeats of EPRS. Thus, we did not include the C-terminal motif of MRS in the interaction test, because it would give the same result as the single repeated unit of EPRS. The cDNA fragments encoding the unique peptide appendices were isolated by PCR and used for the yeast two-hybrid analyses. The ARS polypeptides expressed as fusion proteins with LexA and B42 to test their interactions in a reciprocal manner (31). Sequences and expression of the hybrid proteins were confirmed by DNA sequencing and immunoblotting, respectively (data not shown).

**Interactions of Peptide Appendices.** Interactions between pairs of the human ARS peptide appendices were investigated for 64 combinations in the two-hybrid assay by monitoring the simultaneous induction of lacZ and LEU2 reporter genes. Expression of the LEU2 and lacZ genes was detected by cell growth on the leucine-depleted medium and the formation of blue colonies on the X-Gal containing medium, respectively. The results determined by the two reporter systems were consistent with each other as shown in Fig. 2.

**Fig. 1.** Peptide appendices in human complex-forming ARSs. A group of eight complex-forming human ARS polypeptides are listed according to their amino acid lengths. The unique peptide appendices of ARSs are indicated by black boxes, and their locations used for the interaction studies are described. The C-terminal extension of IRS contains two repeats of about 90 aa, and the internal linker region of EPRS consists of three repeats of 57 aa. IRS-R2 and EPRS-R3 indicate the second and third units of their repeats, respectively. Although a single copy of the motif homologous to the repeats of EPRS is also present in Q841–K897 of MRS (indicated as a hatched box), it was not used for the experiment, because the same result as the motif of EPRS was expected (18).
Of 64 combinations (28 heterologous pairs in two directions and 8 homologous pairs), 8 heterologous (I–E, E–P, I–L, I–O, R–L, R–K, R–Q, R–M, and O–M) and 1 homologous (M–M) pairs had positive interactions. Most of the peptides (except for DRS-N and KRS-N) showed the ability to interact with more than a single peptide. For instance, the N-terminal 267-aa extension of human MRS interacted not only with two heterologous peptides, RRS-N and QRS-N, but also with its homologous partner (Fig. 2). Similarly, RRS-N interacted with LRS-C, KRS-N, QRS-N, and MRS-N. The C-terminal appendix of IRS interacted with EPRS-R, LRS-C, and QRS-N. LRS-C interacted with IRS-C and RRS-N. The interaction between IRS and EPRS, but not the interaction of IRS-C with LRS-C and QRS-N, was reported previously (18).

**Multiple Interactions of N-terminal Extension in RRS.** The 72-aa N-terminal extension of RRS fused to LexA showed interactions with three different peptides of LRS, KRS, and MRS fused to B42 (Fig. 2). RRS-N was predicted to form helices from M1 to K30 and from L41 to N67 (data not shown). To determine the peptide region of RRS involved in the interactions with the heterologous peptides, this extension was divided further into three fragments, M1 to C32, Q15 to Y53, and S38 to N72. All three peptide fragments were stably expressed in yeast as fusion proteins with LexA (data not shown). The peptide spanning from Q15 to Y53 interacted with all of the three tested peptides (Fig. 3), suggesting that the interactions would be localized in a small peptide motif. We currently do not know whether these three peptides interact with RRS-N simultaneously or competitively.

**Interaction of the N-terminal Extension of DRS with Repeated Sequences of IRS.** It was surprising that human DRS-N did not interact with any of the tested peptides in the two-hybrid assay (Fig. 2). The homologous N-terminal extension of rat DRS was shown to be necessary to associate the enzyme to the multi-ARS complex by comparing the behaviors of the full-size and N-terminal truncated DRS (34, 35). We thus expected that the human DRS-N would play a similar role by interacting with other peptide appendages.

It reminded us that RRS-N interacted with the three repeats of EPRS but not with its single unit (19). Among the eight polypeptides of tRNA synthetases found in the multi-ARS complex, two tRNA synthetases, EPRS and IRS, contain repeated sequences (Fig. 1). If DRS-N interacts with one of these repeats, this interaction would be missed, because only single repeated units of the two enzymes were used to obtain the data shown in Fig. 2. We thus tested whether the repeated sequences of EPRS and IRS create a different pattern of interaction with other ARS peptides, including DRS-N. Although the EPRS repeats interacted with RRS as shown previously (19), the IRS repeats interacted with DRS-N (Fig. 4A). The interaction of DRS-N with IRS was refined further by using the C-terminal peptides of IRS containing different numbers of repeats (Fig. 4B). Interaction was shown only with the C-terminal peptide of IRS containing the two repeats and flanking sequence, further supporting the idea that the sequence repetition and flanking sequences are important to determine the interaction.

**Heterologous Interactions of the N- and C-terminal Regions of DRS.** Although our results showed that DRS-N interacts with IRS (Fig. 4), recent two-hybrid analysis using the full-length DRS suggested its interaction with EPRS (11). This result suggested that the N-terminal extension and core domain of DRS may have a different interaction partner. We tested whether the N- and C-terminal regions of DRS would interact respectively with IRS and EPRS. As expected, DRS-N interacted with IRS-C, whereas the C-terminal 192-aa peptide of DRS interacted only with the EPRS peptide containing the repeats and flanking sequence (Fig. 5). No interaction of DRS with other ARSs was observed (11). These results suggest that the interactions of the core bodies of ARSs may not necessarily be the same as those of the peptide appendages.
methods. The R–K and EP–I pairs also were suggested by the two experiments could be complementary. Of the 11 interacting pairs identified in this work, 6 (R–K, Q–R, EP–R, I–Q, I–L, and EP–I) were detected by genetic and crosslinking methods. The R–K and EP–I pairs also were suggested by other methods (4, 13). These results suggest that many interactions between ARSs would be mediated by their appendices, although additional interactions may take place via their core domains. The remaining five pairs (I–D, M–R, L–R, M–Q, and M–M) were found only by the two-hybrid analyses, and the homologous (R–R, K–K, and D–D) and heterologous (K–D, Q–D, M–D, EP–D, Q–K, L–K, EP–K, and I–R) pairs were found only by the crosslinking method. The fact that some pairs were not detected by the crosslinking methods may have been a result of crosslinkers that did not reach the interacting regions or peptides that did not have chemical groups reactive to the crosslinkers. On the contrary, the pairs that were observed only by the crosslinking methods were likely in close proximity, but these pairs do not make specific interactions via their respective appendices.

**DISCUSSION**

Proteins are adapted to different environments by amino acid replacements and recruitment of additional domains to their core bodies. Modular arrangement is prominent in the primary and tertiary structures of ARSs (36–39). Their core bodies essentially consist of two modules responsible for catalysis and tRNA recognition. Some ARSs like *Escherichia coli* methionyl-tRNA synthetase (40) and alanyl-tRNA synthetase (41) have dispensable domains that are required for oligomerization. In this paper, we focused on such extra domains of ARS, especially higher euakaryote-specific domains that must have been acquired in the course of evolution (42). Recent advances in genome sequencing enable us to perform systematic comparison of ARS sequences from various organisms and to infer these extra domains from eight ARSs (Fig. 1). Using these defined peptides, we systematically examined the interaction abilities of these extra domains.

It has been thought that the acquired domains are related to the occurrence of the multi-ARS complex. Indeed, there is some experimental evidence supporting this relationship (21, 22, 43). The work presented here was designed to address this possibility in all eight human ARSs that form the complex. Interactions of the peptide appendices were determined by using a yeast two-hybrid system. With 8 different ARS peptides, 28 heterologous and 8 homologous combinations are possible. Because we fused each peptide with both LexA and B42 in the two-hybrid system, the 28 heterologous pairs were analyzed in two directions, and thus a total of 64 combinations were searched exhaustively (Fig. 2). In some cases, the interaction was sensitive to the polarity of the fused proteins. For example, LRS-C and KRS-N fused to LexA and QRS-N fused to B42 did not show any interaction, whereas the same peptides fused to the counterpart did (Fig. 2). The polarity effect of the fused proteins has been observed frequently in the two-hybrid assay. The negative interaction of these hybrid proteins may have resulted from a conformational change or steric hindrance that interfered with the interaction, and thus we took only the positive results. Because we used single-repeated motifs in IRS and EPRS in the exhaustive search (Figs. 1 and 2), we separately tested the interaction between the complete repeats of IRS and EPRS as well as another six domains (Fig. 4). From this combinatorial analysis, we observed one additional interaction (I–D) and confirmed the interaction previously reported (EP–R; ref. 19). Together, a total of 10 heterologous and 1 homologous interaction pairs were found.

ARSs seem to have developed idiosyncratic ways to interact with other molecules. RRS-N showed the ability to interact with several different peptides. EPRS and IRS increased the strength and versatility of their interactions by sequence repetition (Figs. 2 and 4). MRS contains the N-terminal extension that can make a homologous interaction (Fig. 2) and an additional C-terminal motif homologous to the EPRS repeats (Fig. 1) that may be involved in extra protein–protein interactions.
interaction (18). The interactions occurring at both extremes of MRS may explain its flexible stoichiometry in the multi-ARS complex (44). The multivalency of the interactions or the multiplicity of the sequence units may help to stabilize the multi-ARS complex. Alternatively, different peptides may compete for binding to a particular ARS in vivo, and this competition may facilitate the dissociation of the component from the complex.

The N-terminal truncated DRS did not associate to the multi-ARS complex, suggesting the role of the N-terminal region in the complex-formation. However, its N-terminal extension did not take the connected heterologous domain to the complex (34). These results imply that an additional interaction in the core domain is also important in the complex. Analyses of the interactions with the N- and C-terminal peptides of DRS showed that the two regions interact with IRS and EPRS, respectively (Fig. 5). Our conclusions concerning different interactions between the appendices and core domains of the ARS components are supported further by the fact that these extra signal also has been proposed for the N-terminal extension of MRS (49). Thus, the appended domains have other functions. Perhaps they contribute to its assembly. Although our results showed that many peptide appendices are bound to the Arc1p that facilitate the assembly of the multi-ARS complex (11), it is worthwhile to note a recent report showing that p38 plays an important role in the assembly of the multi-ARS complex (18). The interactions occurring at both extremes of MRS may explain its flexible stoichiometry in the multi-ARS complex.