Although previous studies have already shown that both cytoplasmic and mitochondrial activities of glycyl-tRNA synthetase are provided by a single gene, GRS1, in the yeast \textit{Saccharomyces cerevisiae}, the mechanism by which this occurs remains unclear. Evidence presented here indicates that this bifunctional property is actually a result of two distinct translational products alternatively generated from a single transcript of this gene. Except for an amino-terminal 23-amino acid extension, these two isoforms have the same polypeptide sequence and function exclusively in their respective compartments under normal conditions. Reporter gene assays further suggest that this leader peptide can function independently as a mitochondrial targeting signal and plays the major role in the subcellular localization of the isoforms. Additionally, whereas the short protein is translationally initiated from a traditional AUG triplet, the long isoform is generated from an upstream in-frame UUG codon. To our knowledge, GRS1 appears to be the first example in the yeast wherein a functional protein isoform is initiated from a naturally occurring non-AUG codon. The results suggest that non-AUG initiation might be a mechanism existing throughout all kingdoms.

In lower eukaryotes such as \textit{Saccharomyces cerevisiae}, AUG is the only codon recognized as the translational initiator, and the AUG codon closest to the 5'-end of mRNA preferentially serves as the start site for translation (1, 2). If the first AUG triplet is mutated, initiation can begin at the next available AUG on the message. The same rules apply to virtually all eukaryotes. However, there are many examples of cellular and viral mRNAs that can use “non-AUG” codons differing from AUG on the message. The same rules apply to virtually all eukaryotes, such as yeast, commonly have two genes that encode distinct sets of proteins for each aminoacylation activity; one functions in the cytoplasm and the other in the mitochondria. Each set aminoacylates the isoaccepting tRNAs within its respective cell compartment, and is sequestered from the isoacceptors in other compartments. As exceptions to this rule, two \textit{S. cerevisiae} genes, \textit{HTS1} (coding for histidyl-tRNA synthetase) (14) and \textit{VAS1} (coding for valyl-tRNA synthetase) (15), specify both the cytoplasmic and mitochondrial forms of their respective enzymes. mRNAs with distinct 5'-ends are alternatively transcribed from each of these genes. The mitochondrial form of the enzyme is translated from the first initiation AUG on the “long” message, whereas the cytosolic form is translated from the second in-frame AUG on the “short” message. These two isoforms cannot substitute for each other because of different localization (15). Similar observations have been made for the \textit{Arabidopsis thaliana} genes that encode alanyl-tRNA synthetase, threonyl-tRNA synthetase, and valyl-tRNA synthetase (16). It should be addressed that except for some algae, all aminoacyl-tRNA synthetases are encoded by nuclear genes and imported post-translationally into their respective compartments.

As with most yeast tRNA synthetases, two homologous nuclear genes encoding glycyl-tRNA synthetase (GlyRS), GRS1 and GRS2, have been identified in \textit{S. cerevisiae}. Paradoxically, GRS1 provides both cytoplasmic and mitochondrial functions, whereas GRS2 appears to be nonfunctional (17). However, unlike \textit{HTS1} and \textit{VAS1}, only one appropriate AUG triplet was identified in the 5' region of its open reading frame, leading to the assumption that the translational product initiated from this start site was a bifunctional enzyme (17). A similar example in which a single translation product is responsible for both mitochondrial and cytoplasmic functions is the yeast fumarase gene \textit{FUM1} (18, 19). In this instance, all fumarase molecules are targeted and processed in mitochondria before part of the

S ribosomal subunit (6). In contrast, there seems to be little sequence context effect in yeast (7), and the positive benefits of downstream secondary structures have not yet been clearly addressed. Perhaps for these reasons, yeast cannot efficiently use non-AUG codons as translation start sites (8). So far, examples of native non-AUG initiation have not been reported in the model yeast \textit{S. cerevisiae}.

In prokaryotes, there are typically 20 aminoacyl-tRNA synthetases, one for each amino acid (9–12). In eukaryotes, protein synthesis occurs not only in the cytoplasm, but also in organelles, such as mitochondria and chloroplasts (13). Compartmentalization of the protein synthesis machinery within the cytoplasm and organelles of eukaryotes leads to isoaccepting tRNA species that are distinguished by nucleotide sequence, subcellular location, and enzyme specificity. Thus, most eukaryotes, such as yeast, commonly have two genes that encode distinct sets of proteins for each aminoacylation activity; one functions in the cytoplasm and the other in the mitochondria. Each set aminoacylates the isoaccepting tRNAs within its respective cell compartment, and is sequestered from the isoacceptors in other compartments. As exceptions to this rule, two \textit{S. cerevisiae} genes, \textit{HTS1} (coding for histidyl-tRNA synthetase) (14) and \textit{VAS1} (coding for valyl-tRNA synthetase) (15), specify both the cytoplasmic and mitochondrial forms of their respective enzymes. mRNAs with distinct 5'-ends are alternatively transcribed from each of these genes. The mitochondrial form of the enzyme is translated from the first initiation AUG on the “long” message, whereas the cytosolic form is translated from the second in-frame AUG on the “short” message. These two isoforms cannot substitute for each other because of different localization (15). Similar observations have been made for the \textit{Arabidopsis thaliana} genes that encode alanyl-tRNA synthetase, threonyl-tRNA synthetase, and valyl-tRNA synthetase (16). It should be addressed that except for some algae, all aminoacyl-tRNA synthetases are encoded by nuclear genes and imported post-translationally into their respective compartments.

As with most yeast tRNA synthetases, two homologous nuclear genes encoding glycyl-tRNA synthetase (GlyRS), GRS1 and GRS2, have been identified in \textit{S. cerevisiae}. Paradoxically, GRS1 provides both cytoplasmic and mitochondrial functions, whereas GRS2 appears to be nonfunctional (17). However, unlike \textit{HTS1} and \textit{VAS1}, only one appropriate AUG triplet was identified in the 5' region of its open reading frame, leading to the assumption that the translational product initiated from this start site was a bifunctional enzyme (17). A similar example in which a single translation product is responsible for both mitochondrial and cytoplasmic functions is the yeast fumarase gene \textit{FUM1} (18, 19). In this instance, all fumarase molecules are targeted and processed in mitochondria before part of the
products are sent by retrograde movement back to the cytosol (20). We wondered if GRS1 followed a similar pathway, or if a different mechanism was involved. In the work described here, we cloned the GRS1 gene and investigated its bifunctional control. To our surprise, two protein isoforms of different size and alternative translation initiation were alternatively translated from this gene. Even more remarkably, the first codon of one of the two protein isoforms is actually initiated from a UUG triplet, a codon that was believed inappropriate for initiation in yeast. The implications of this novel mechanism of translation initiation, particularly in yeast, are further discussed.

EXPERIMENTAL PROCEDURES

Construction of Plasmids—Cloning of GRS1 from S. cerevisiae followed standard protocols. The wild-type GRS1 (designated as wt GRS1) sequence was amplified by PCR as a 2.4-kb EagI-SalI fragment and cloned in a high-copy yeast shuttle vector (21). Various point mutations, such as M1T (ATG1 to GCG) in the preVAS1-GRS1 hairpin structure were inserted into position 300 to 12 of preGRS1-GRS1, a short DNA duplex coding for a putative RNA structure deduced from the duplex, where 5'-GCCCGCACGC-3', has a predicted stability of about 300 kCal/mol.

For construction of various derivatives, RJT3/II-1 was cotransformed with a test plasmid (carrying a LEU2 marker) and a second maintenance plasmid (carrying a TRP1 marker) that expresses a functional cytoplasmic GlyRS, but is defective in mitochondrial GlyRS activity. The second maintenance plasmid contained GRS1mRNA cloned in pRS14.) In the presence of 5-FOA, the first maintenance plasmid (containing a URA3 marker) was evicted from the cotransformants, whereas the second maintenance plasmid was retained. Thus, all cotransformants survived 5-FOA selections because of the presence of cytoplasmic GlyRS derived from the second maintenance plasmid. The cotransformants were further tested on YPG plates for their mitochondrial phenotypes at 30 °C, with results documented at day 3 following plating. Because a yeast cell cannot survive on glycerol without functional mitochondria, the cotransformants do not grow on YPG plates unless a functional mitochondrial GlyRS is present. Complementation assays for various VAS1 constructs were conducted essentially the same way as those for GRS1 constructs, except that a VAS1 knockout strain (designated as CW1) was used as the test strain.

RESULTS

Mapping the 5′-Ends of GRS1 Transcripts—It was previously shown that a single GRS1 gene encodes both cytoplasmic and mitochondrial glycl-TRNA synthetase activities in S. cerevisiae (17). However, unlike HTS1 and VAS1, no alternative in-frame AUG codon was found in the sequence upstream of the putative open reading frame (as diagrammed in Fig. 1A). Interestingly, a long stretch of uninterrupted coding sequence was found between the nearest upstream stop codon, UAA −57, and the putative translation initiation codon, AUG1. (The number −57 in UAA −57 refers to the −57 codon that precedes AUG1.) To gain insight into the mechanism of GRS1 expression, the in vivo transcription profiles of this gene were examined. The transcription start sites and 5′-terminal nucleotide sequences of GRS1 transcripts were determined using a technique known as 5′-RACE. Following reverse transcription and PCR-driven amplification, the double-stranded cDNA products were cloned and sequenced. Remarkably, only one cDNA product was identified on a 10% polyacrylamide gel (Fig. 1B), with its 5′-end mapped at position −88 relative to the A of AUG1 (Fig. 1A). In addition, sequence of the cDNA (∼400 bp determined) was identical to that of the corresponding segment in genomic DNA. Thus, the possibility of splicing occurring at the 5′-end of GRS1 mRNA could be ruled out.

A Second Isoform of GlyRS Generated from GRS1—A previous study suggested that an amino-terminal segment of GlyRS might direct its mitochondrial localization, because deletion of codons 2–12 from this gene selectively impaired its mitochondrial activity (17) (ADH/GRS1Δ2–12 in Fig. 2). However, these phenotypes were observed with protein expression under the control of a constitutive ADH promoter in a high-copy number vector, and might not faithfully reflect normal conditions in vivo. To mimic physiological conditions, various GRS1 se-
FIG. 1. Mapping the 5′-ends of GRS1 transcripts. A, the 5′-end DNA sequence of GRS1, including base pairs −150 to +120 or codons −60 to +40 relative to ATG1. The transcription start site of GRS1 is labeled on top of the sequence. ATG1 and the nearest upstream stop codon TAA are designated as −1. The amino acid residues deduced from the presences between TTT −56 and AGA are underlined. Diverging arrows highlight the potential formation of a hairpin structure upstream from the ATG start site. The numbers −57, −56, and −1 in TAA indicate the position of the missense mutation at this codon would impair the synthesis of the cytoplasmic and mitochondrial isofoms of GRS1. B, cDNA products of the 5′-terminal sequences of GRS1 transcripts. The relative migrating positions of the cDNA products (lane 1) and DNA markers (lane 2) on the 10% polyacrylamide gel are indicated on the left and right, respectively. The number −88 labeled on the left refers to nucleotide position −88 relative to ATG1.

Fig. 2. Cytoplasmic (Cyt) and mitochondrial (Mit) functions of GRS1 provided by two distinct isoforms of GlyRS. Complementation of the cytoplasmic and mitochondrial phenotypes of the grs1− strain was shown by the ability to grow on 5-FOA and YPG plates, respectively. A, a schematic summary of the various GRS1 constructs and their complementation activities. B, complementation assays on a 5-FOA plate. C, complementation assays on a YPG plate. ADH/GRS1−Δ2−12 and ADH/GRS1−12 each have a deletion in the sequence coding for residues 2–12 and are expressed from an ADH promoter. GRS1/GRS1 have a missense mutation (ATG1 to ACT) and an out-of-frame mutation (AAA−7 to TAAA), respectively. Mit denotes the native initiator methionine, whereas M* represents an artificial initiator methionine introduced at the 5′-side of codon 13. Deletion is shown as a dashed line. The open and solid boxes denote the ADH and GRS1 sequences, respectively.

To further prove our assumption, we inserted the presequence with the native promoter were cloned into a low-copy number yeast shuttle vector, pRS315, for further studies. (Normal conditions are referred to here as protein expression from a low-copy number vector under the control of its native promoter.) We supposed if AUG1 is the only initiator for GRS1, a missense mutation at this codon would impair the synthesis (and thus function) of both the cytoplasmic and mitochondrial enzymes, whereas a nonsense or an out-of-frame mutation placed upstream of AUG1 would have little effect on either function. As shown in Fig. 2, wt GRS1 allowed colony formation on both FOA (Fig. 2B) and YPG (Fig. 2C) plates, indicating the production of functional cytoplasmic and mitochondrial GlyRSs. When AUG1 was mutated to ACU, the resultant construct, GRS1/GRS1, lost the cytoplasmic function (Fig. 2B) but retained the mitochondrial function (Fig. 2C), suggesting that AUG1 serves as the translation initiator for the cytoplasmic form, but not the mitochondrial form. In contrast, when an out-of-frame mutation (AAA−7 to UAAA) was introduced 5′ to AUG1, the resultant construct, GRS1/GRS1, lost the mitochondrial activity (Fig. 2C) but retained the cytoplasmic function (Fig. 2B), suggesting that the mitochondrial precursor is translationally initiated at a site upstream of AAA−7. (Note that this out-of-frame mutation not only caused an out-of-frame reading, but also led to a stop codon at this site.) Thus, the mitochondrial function of GlyRS is provided by a distinct isoform initiated at a site 5′ to AUG1.

To further prove our assumption, we inserted the presequence back into construct ADH/GRS1−Δ2−12 and tested for its effect. Consistent to the data mentioned above, the resultant construct, ADH/GRS1−Δ2−12, restored the mitochondrial function of ADH/GRS1−Δ2−12. However, when an intact open reading frame of GRS1 without the prescence was expressed from the ADH promoter, the overexpressed GlyRS rescued both the cytoplasmic and mitochondrial defects of the grs1− strain (Fig. 2), suggesting that codons 2–12 are important for the mitochondrial phenotype of the “overexpressed” GlyRS molecules (compare ADH/GRS1 and ADH/GRS1−Δ2−12 in Fig. 2, B and C). Thus, residues 2–12 appear to constitute a weak, cryptic, transit signal that normally does not play a role in mitochondria import, but can be recruited to do so when overexpressed.

Initiation of the Mitochondrial Isoform from a Native Non-AUG Codon of GRS1—Because there is no other in-frame AUG
triplet in the sequence 5’ to AUG1 (Fig. 1), the possibility that a non-AUG codon differing from AUG by one nucleotide serves as the translation start site for the mitochondrial form was investigated. There are 9 such triplets at GAA−25 and AUG1, including UUG−23, AUC−10, AUU−10, AUU−33, UUG−23, AUC−10, AUU−10, AAG−3, and AUU−33 (Fig. 1). To determine whether any of these serves as the initiator, a nonsense mutation was introduced into this region to map its approximate location. As shown in Fig. 3, a nonsense mutation at GAA−25 (resulting in GRS1−25(stop)) did not compromise the mitochondrial function of this gene (Fig. 3C), suggesting that the alternative initiator was located downstream of GAA−25. Together with the results shown in Fig. 2, the alternative initiation site must be between GAA−25 and AAA−7. There are three candidates, UUG−23, AUC−18, and AUU−10, located within this region. Among these three codons, AUU−10 is the least likely candidate, because translation initiated at this site would lead to a protein with a very short leader peptide. For this reason, we focused on AUC−18 and UUG−23. Mutation at AUC−18 (resulting in GRS1−18(L)) had little effect on the mitochondrial activity (Fig. 3C), whereas mutation at UUG−23 (GRS1−23(L)) abolished the mitochondrial activity. To provide further evidence that UUG−23 is the initiator, a stop codon, UAA, was inserted between codons −23 and −22, and the resultant construct (GRS1−23(stop)−22) was tested for its mitochondrial activity. Consistent to our hypothesis, insertion of a stop codon immediately downstream of UUG−23 specifically knocked out its mitochondrial activity. In contrast, insertion of a stop codon, UAA, between codons −24 and −23 (resulting in GRS1−24(stop)−23) had no obvious effect on the mitochondrial function. Together, our results suggest that UUG−23 is the translation initiator for the mitochondrial isoform of GlyRS.

Note that none of these mutations in the presequence compromised the cytoplasmic function of GRS1 (Fig. 3B).

To rule out the possibility that mutation at UUG−23 (such as UUG−23 to UUA) caused an unexpected effect in the secondary structure of GRS1 mRNAs, UUG−23 was mutated to a classical AUG translation initiation codon and then assayed for its effect on complementation activity. As expected, the mutation did not compromise the mitochondrial activity of GRS1 (GRS1−23/23M in Fig. 3C), suggesting that it is the initiator activity, but not the RNA secondary structure, that was damaged by the mutation of UUG−23 to UUA. In addition, the relative mRNA levels produced from each mutant were assessed by RT-PCR and found to be similar, suggesting that the mutations did not alter mRNA stability.

Partitioning of GlyRS Isoforms Determined by the Signal Peptide—to provide further insight into the mechanism that governs the subcellular localization of the GlyRS isoforms, we asked whether the leader peptide of the precursor form functions independently as a mitochondrial targeting signal and if the short form of GlyRS could be converted to a mitochondrial enzyme by fusion of a heterologous signal peptide without overexpression. As previously shown, VAS1c, which has its mitochondrial targeting peptide (residues 1–46) deleted, encodes only the cytoplasmic form of valyl-tRNA synthetase and thus, cannot rescue the mitochondrial defect of the VAS1 knockout strain CW1 (22). Substitution of the native VAS1 promoter in VAS1c with the GRS1 presequence enabled the fusion to provide both the cytoplasmic and mitochondrial activities (preGRS1−VAS1c in Fig. 4), whereas mutation of the non-AUG initiator in the fusion construct (preGRS1−23/23c−VAS1c in Fig. 4) selectively impaired its mitochondrial activity. Thus, the leader peptide is in itself a fully functional mitochondrial targeting signal.

We next tested whether the short form of GlyRS can be converted into a mitochondrial enzyme by a well characterized mitochondrial targeting signal (22). Fig. 5 showed that the leader peptide of valyl-tRNA synthetase could direct the otherwise cytoplasmic glycine enzyme into mitochondria. Further mutagenesis assays confirmed that the mitochondrial and the cytoplasmic functions are provided by proteins initiated from the AUG initiators in the VAS1 and GRS1 portions, respectively (compare preVAS1c−VAS1c and preVAS1c−GRS1c in Fig. 5). Thus, the subcellular locations of the GlyRS isoforms depend solely on the presence or absence of an appropriate
investigated the initiating capacity of UUG translation. To confirm our results obtained by functional tests, we —lation indeed translated from the same transcript. These data suggested that the GlyRS isoforms are available solely for translation of the cytoplasmic isoform, and additional evidence that there are no alternative transcripts signal peptide under normal conditions.

Translation of the Isoforms from a Single mRNA —To provide additional evidence that there are no alternative transcripts available solely for translation of the cytoplasmic isoform, and that the two isoforms are indeed generated from the same transcript identified by 5’-RACE, a sequence coding for a stable RNA secondary structure was inserted at the start of the transcription initiation site and tested for its effect on protein synthesis. As shown in Fig. 6, cytoplasmic and mitochondrial functions were simultaneously affected by the putative hairpin structure, suggesting that no alternative mRNAs are available solely for the translation of the cytoplasmic isoform. (A negligible number of small colonies were observed in H-GRS1 on 5-FOA and YPG, probably because of incomplete inhibition of translation.) These data suggested that the GlyRS isoforms are indeed translated from the same transcript.

Western Blot Analyses of Non-AUG-initiated Protein Translation —To confirm our results obtained by functional tests, we investigated the initiating capacity of UUG translation with a different approach, Western blot. Various GRS1 presequences were fused in-frame to lexA and the expression profiles of the fusion constructs were tested. (These fusions were expressed under the control of a constitutive ADH1 promoter.) As shown in Fig. 7, the LexA protein produced from the lexA construct was specifically recognized by anti-LexA antibody (compare lanes 1 and 2 in Fig. 7B). Unexpectedly, when the wild-type GRS1 presequence was fused to lexA (resulting in preGRS1-lexA), we observed only one protein band with a molecular mass similar to that of the LexA protein but could not detect the larger fusion protein (lane 3 in Fig. 7B). A similar profile was seen with the initiator mutant preGRS1L-lexA (lane 4 in Fig. 7B). We wondered whether the non-AUG-initiated protein was actually produced under these conditions or if the LexA fusion generated has been processed by matrix processing peptidase in the mitochondria. To distinguish between the two possibilities, the non-AUG initiator in the GRS1 portion (UUG) was substituted for the sequence 5’- to ATG1 of GRS1. preGRS1M-lexA has a mutation (ATG1 to GCG) in the GRS1 portion, whereas preGRS1-GRS1M1T has a mutation (ATG1 to ACT) in the GRS1 portion. The solid and striped boxes indicate the GRS1 and VAS1 sequences, respectively.
that of LexA was produced from preGRS1-lexAM1T in which the AUG1 in the lexA portion was mutated (AUG to AGU) (lane 3 in Fig. 7C), suggesting that the GRS1-LexA fusion protein has been efficiently processed. The protein band became much stronger when the UUG was changed to AUG (preGRS1L−23M-lexAM1T, lane 5 in Fig. 7C). In contrast, when UUG23 was mutated to a codon unsuitable for initiation (preGRS1L−23M-lexA1T, lane 4 in Fig. 7C), the protein band reduced dramatically. We surmised that the insignificant protein band that appeared in lane 4 probably came from one of the non-AUG codons (that differ from AUG by one nucleotide) between UUG23 and AUG1. However, these non-AUG initiators are probably too inefficient to serve as functional initiators in complementation assays. To obtain a more accurate estimation of the relative initiation efficiencies, various amounts of total proteins were loaded and compared. As shown in Fig. 7D, the frequency of initiation at UUG23 is around 5% relative to those at AUG1 and AUG23 (compare lanes 1, 5, and 9). Thus, AUG23 appears to be as efficient as AUG1 in initiation (compare lanes 5 and 9 in Fig. 7D). In addition, the relative amounts of mRNA expressed from preGRS1-lexAM1T, preGRS1L−23M, lexA1T, and preGRS1L−23M-lexA were similar (Fig. 7E), suggesting that it is not the stability of the mRNAs that was affected by the mutations.

Consistent to our complementation assays (Fig. 6), when the duplex described in the previous section was inserted at the 5'-end of the preGRS1-lexA fusion, synthesis of both the LexA fusion protein from various preGRS1-lexA fusions. A, a schematic summary of various preGRS1-lexA fusions. B, Western blot analysis of non-ATG-initiated protein translation. The initiating activity of the non-ATG initiator of GRS1 was shown by its ability to start the synthesis of a LexA fusion protein from various preGRS1-lexA fusions. C, Western blot analysis by chemiluminescence-based method of the expression profiles of various preGRS1-lexA fusions. D, comparison of initiation efficiencies by chemiluminescence-based Western blot analysis. E, RT-PCR. The relative amounts of mRNA generated from each construct were measured by RT-PCR using a set of primers complementary to −30 to −10 bp of GRS1 and +370 to +390 bp of lexA, respectively. F, RT-PCR. The relative amounts of mRNA generated from each construct were measured by RT-PCR using a set of primers complementary to −30 to −10 bp of GRS1 and +370 to +390 bp of lexA, respectively. G, RT-PCR. The relative amounts of mRNA generated from each construct were measured by RT-PCR using a set of primers complementary to −30 to −10 bp of GRS1 and +370 to +390 bp of lexA, respectively.
to UUG−23 was mutated (A to U or C), the mitochondrial function of the resultant GRS1 constructs appeared to be normal (data not shown). We did not change the nucleotide at position +4 relative to UUG−23, because it is already a less favorable nucleotide, T.

**DISCUSSION**

In the present work, we have discovered a non-AUG translation initiator occurring naturally in the yeast *S. cerevisiae*. Many cellular and viral messages have been shown to use non-AUG codons, such as ACG, CUG, and AUU, as translation start sites (23). In some cases, non-AUG codons serve as exclusive translation initiators, such as in the mRNAs for the *Arabidopsis AGAMOUS* gene (24) and the latently expressed kaposin locus of Kaposi's sarcoma-associated herpesvirus (25). However, in most cases, non-AUG codons act as alternative translation initiators. Recognition of these initiators by mammalian ribosomes is often facilitated by particular sequences around the initiators, in particular nucleotides in position −3 and +4 (5). In contrast, sequence context has little effect on translation initiation in yeast. As a result, yeast genes were not previously expected to be capable of using non-AUG triplets as translation start sites (8). In this regard, our discovery of a non-AUG initiator occurring naturally in the yeast GRS1 gene is particularly striking (Fig. 3). The question arose as to why UUG−23 can serve as the alternative initiator if sequence context is irrelevant. Analysis of the nucleotide sequence 5′ to AUG reveals a weak secondary structure having a predicted stability of ΔG around −2 kcal/mol 15 nucleotides downstream of UUG−23 (Fig. 1), an optimal position to stimulate recognition of the weak upstream start site (6). However, mutations that destabilize the already weak structure did not affect the initiation ability of UUG−23 (Fig. 7C). Perhaps the sequence downstream of the UUG initiator does not assume a structure in vivo as predicted. More detailed experiments are necessary for mapping the sequences or structural elements required for efficient non-AUG initiation.

In mammalian cells, the 40 S ribosomal subunit often skips a weak upstream initiator, frequently a non-AUG initiator or an AUG codon with suboptimal sequence context, and continues scanning toward the 3′-end until it encounters an AUG triplet within a favorable sequence context. As a result of this “leaky” scanning, two protein isoforms can be alternatively translated from a single transcript (26–29). Although leaky scanning has been observed frequently in mammals, there are so far only two known examples in yeast: MOD5 (coding for isopentenyl pyrophosphate: tRNA isopentenyl transferase) (30) and CCA1 (coding for ATP (CTP): tRNA nucleotidyltransferase) (31). In these two instances, leaky scanning occurs because the first AUG codon in the two genes is positioned too close to the 5′-end of the mRNA, which makes it inaccessible to the initiating ribosome. Because the first initiator (UUG−23) in GRS1 is relatively inefficient, it is possible that the downstream AUG initiator is also recognized through leaky scanning. It is interesting to note that when UUG−23 was mutated to AUG, the resultant construct rescued both the cytoplasmic and mitochondrial defects of the GRS1 knockout strain (Fig. 3). We surmised that the cytoplasmic function of this construct probably results from the single translation product initiated from AUG−23, which has been processed in the mitochondria and then leaks back to the cytoplasm. Upon insertion of a stable hairpin at the beginning of the transcript, the production of both isoforms was simultaneously inhibited (Figs. 6 and 7), lending further support to the hypothesis that recognition of the alternative initiators follows the traditional mechanism of cap-dependent ribosomal scanning and ruling out the possibility of internal ribosomal entry, a mechanism previously discovered in picornavirus (32), and recently shown in *S. cerevisiae* (33).

Contrary to a previous report, which suggested that a single translation product of GRS1 provides both cytoplasmic and mitochondrial GlyRS functions (17), we presented strong evidence that this translational product (initiated from AUG3) is solely a cytoplasmic enzyme when expressed at normal levels. Localization of the GlyRS molecule into mitochondria requires the presence of an additional leader sequence that serves as a transit signal (Fig. 2). Like many other mitochondrial targeting signals (34), this signal peptide is rich in positively charged (−30%) and hydroxylated (−17%) residues, and devoid of acidic residues. The reporter gene assays confirmed that the leader peptide alone carries all the relevant information required for mitochondrial targeting and thus can successfully convert a cytoplasmic passenger into a mitochondrial player (Fig. 4). On the other hand, the cytoplasmic form, because of lack of the leader peptide, functions strictly in the cytoplasm when expressed from a low-copy number vector under the control of its native promoter (Fig. 2) or an analogous promoter (Fig. 4). Despite this, the leaderless enzyme can be forced into mitochondria when overexpressed (Fig. 2). These observations suggested that the GlyRS preprotein contains two adjacent mitochondrial transit signals, one in the leader peptide and the other in the amino-terminal part of its catalytic body. The former appears to play a predominant role in mitochondrial localization of the preprotein, whereas the latter is a weak cryptic signal, which normally does not take part in protein import (Fig. 2), but can be recruited to do so when overexpressed (Fig. 2). A similar scenario has been observed in COXV2, the yeast gene coding for the precursor form of the mitochondrial cytochrome c oxidase subunit Va. Although the leader peptide is required for mitochondrial import of this enzyme under normal conditions, overexpression of a leaderless form enabled it to overcome this requisite (35). In contrast to these observations, although VAS1 also encodes two distinct protein isoforms through alternative transcription and translation, the cytoplasmic form cannot complement its mitochondrial function even when it is overexpressed (22).

Because yeast ribosomes have now been found to recognize a naturally occurring non-AUG start site, it is conceivable that more examples of non-AUG-mediated protein translation will soon be identified in the yeast. Complementary to our viewpoint, an interesting report published very recently argued that a AUG codon may serve as the exclusive translation initiator for the gene encoding acidic ribosomal P2A protein in the yeast *Candida albicans* (36). In addition, preliminary screening of genomic sequences for GlyRS genes in other low eukaryotes revealed the presence of only one such gene in both *C. albicans* and *Schizosaccharomyces pombe*, and even more incredibly, only one suitable AUG triplet in each of the two genes. Protein sequence alignment suggested that the primary translation products initiated from their putative AUG initiators share high homology with the cytoplasmic GlyRS enzyme of *S. cerevisiae*, in particular at their amino termini (data not shown), implying that they only serve a cytoplasmic function. It is thus likely that, as with *S. cerevisiae*, the production of a second, mitochondrial, isoform might involve the utilization of an upstream in-frame non-AUG initiator. Prediction of their upstream sequences with PSORTII (37) showed high mitochondrial localization potential, lending further support to our assumptions. These observations suggest that non-AUG initiation might be a mechanism more widespread than previously thought. In addition, recent investigations have discovered that tRNA synthetases are capable of functions other than protein synthesis, including roles in cellular fidelity, tRNA
processing, RNA splicing, RNA trafficking, apoptosis, and transcriptional and translational regulation (38). These non-canonical functions might require additional protein interaction domains and may often take place in locations other than the cytoplasm, such as mitochondria or the nucleus (39–41). In this sense, alternative translation at in-frame non-AUG codons might contribute to the multifunctional phenotype of a gene by leading to the production of distinct protein isoforms with such signal peptides or additional protein interaction domains. Cumulatively, our discovery has opened a new avenue to the identification of new open reading frames that start exclusively or alternatively with non-AUG initiators, and novel gene functions that might have been previously overlooked in the yeast S. cerevisiae.

Acknowledgments—We thank Paul Schimmel, Robert Turner, Brian Nordin (The Scripps Research Institute), and Grace Lin (National Central University) for helpful discussion and critical reading of the manuscript, and Huey-Nan Wu (Academia Sinica, Taiwan) for technical support.

REFERENCES
1. Sherman, F., Stewart, J. W., and Schweingruber, A. M. (1980) Cell 20, 215–222
2. Donahue, T. F., and Cigan, A. M. (1988) Mol. Cell. Biol. 8, 2955–2963
3. Konzak, M. (1989) Mol. Cell. Biol. 9, 5073–5080
4. Peabody, D. S. (1987) J. Biol. Chem. 262, 11847–11851
5. Kozak, M. (1991) J. Biol. Chem. 266, 19867–19870
6. Kozak, M. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 8301–8305
7. Kozak, M. (1999) Gene (Amst.) 234, 187–208
8. Clements, J. M., Laz, T. M., and Sherman, F. (1988) Mol. Cell. Biol. 8, 4533–4536
9. Carter, C. W., Jr. (1993) Annu. Rev. Biochem. 62, 715–748
10. Martinis, S. A., and Schimmel, P. (1996) in Escherichia coli and Salmonella Cellular and Molecular Biology (Neidhardt, F. C., ed) 2nd Ed., pp. 887–901, American Society of Microbiology, Washington, D. C.
11. Giege, R., Sissler, M., and Florentz, C. (1998) Nucleic Acids Res. 26, 5017–5035
12. Pelchat, M., and Lapointe, J. (1999) Biochem. Cell Biol. 77, 343–347
13. Marechal-Drouard, L., Weil, J. H., and Dietrich, A. (1992) Annu. Rev. Cell Biol. 8, 115–131
14. Natsoulis, G., Hilger, F., and Fink, G. R. (1986) Cell 46, 235–243
15. Chatton, B., Walter, F., Ebel, J. P., Lacroute, F., and Fasiolo, F. (1988) J. Biol. Chem. 263, 52–57
16. Souciet, G., Menand, B., Oresna, J., Cosset, A., Dietrich, A., and Wintz, H. (1999) Eur. J. Biochem. 266, 848–854
17. Turner, E. J., Lovato, M., and Schimmel, P. (2000) J. Biol. Chem. 275, 27681–27688
18. Wu, M., and Trangolo, A. (1987) J. Biol. Chem. 262, 12275–12282
19. Stein, L., Peleg, Y., Even-Ram, S., and Pines, O. (1994) Mol. Cell. Biol. 14, 4770–4778
20. Sass, L., Blachinsky, E., Karnieli, S., and Pines, O. (2001) J. Biol. Chem. 276, 46111–46117
21. Sikorski, R. S., and Haeter, P. (1989) Genetics 132, 19–27
22. Wang, C. C., Chang, K. J., Tang, H. L., Hsieh, C. J., and Schimmel, P. (2003) Biochemistry 42, 1646–1651
23. Kozak, M. (1989) Gene (Amst.) 83, 156–166
24. Riechmann, J. L., Lain, S., and Garcia, J. A. (1990) Virology 177, 170–176
25. Sadler, R., Wu, L., Forghani, B., Renne, R., Zhong, W., Herndier, B., and Gannem, D. (1999) J. Biol. Chem. 274, 5722–5730
26. Acland, P., Dixon, M., Peters, G., and Dickson, C. (1990) Nature 343, 662–665
27. Saris, C. J., Domen, J., and Berra, A. (1991) EMBO J. 10, 655–664
28. Hans, S. R., Sloan-Brown, K., and Spotts, G. D. (1992) Genes Dev. 6, 1229–1240
29. Packham, G., Brimmell, M., and Cleveland, J. L. (1997) Biochem. J. 328, 807–813
30. Slusher, L. B., Gillman, E. C., Martin, N. C., and Hopper, A. K. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 9789–9793
31. Wolfe, C. L., Lou, Y. C., Hopper, A. K., and Martin, N. C. (1994) J. Biol. Chem. 269, 13561–13566
32. Pelletier, J., and Sonenberg, N. (1988) Nature 334, 320–325
33. Zhou, W., Edelman, M. G., and Mauro, P. V. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 1511–1516
34. Neupert, W. (1997) Annu. Rev. Biochem. 66, 863–917
35. Dircks, L. K., and Poyton, R. O. (1990) Mol. Cell. Biol. 10, 4984–4986
36. Abramczyk, D., Tchorzewski, M., and Granowski, N. (2003) Yeast 20, 1045–1052
37. Nakai, K., and Horton, P. (1999) Trends Biochem. Sci. 24, 34–36
38. Martinis, S. A., Plateau, P., Cavarelli, J., and Florentz, C. (1999) Biochimie (Paris) 81, 683–700
39. Loidl, E., and Dahlberg, J. E. (1998) Science 282, 2082–2085
40. Sarkar, S., Azad, A. K., and Hopper, A. K. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 14366–14371
41. Schimmel, P., and Wang, C. C. (1999) Trends Biochem. Sci. 24, 127–128
Translation Initiation from a Naturally Occurring Non-AUG Codon in
Saccharomyces cerevisiae
Kuang-Jung Chang and Chien-Chia Wang

J. Biol. Chem. 2004, 279:13778-13785.
doi: 10.1074/jbc.M311269200 originally published online January 19, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M311269200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 40 references, 22 of which can be accessed free at
http://www.jbc.org/content/279/14/13778.full.html#ref-list-1