Close Packing of Helices 3 and 12 of 16 S rRNA Is Required for the Normal Ribosome Function

Received for publication, August 14, 2006, and in revised form, October 20, 2006. Published, JBC Papers in Press, October 23, 2006, DOI 10.1074/jbc.M607725200

Matthieu G. Gagnon1,2, Alka Mukhopadhyay1, and Sergey V. Steinberg3

From the Département de Biochimie, Université de Montréal, Montréal, Quebec H3C 3J7, Canada

The along-groove packing motif is a quasi-reciprocal arrangement of two RNA double helices in which a backbone of each helix is closely packed within the minor groove of the other helix. At the center of the inter-helix contact, a GU base pair in one helix packs against a Watson-Crick base pair in the other helix. Here, based on in vivo selection from a combinatorial gene library of 16 S rRNA and on functional characterization of the selected clones, we demonstrate that the normal ribosome performance requires that helices 3 and 12 be closely packed. In some clones the Watson-Crick and GU base pairs exchange in their positions between the two helices, which affects neither the quality of the helix packing, nor the ribosome function. On the other hand, perturbations in the close packing usually lead to a substantial drop in the ribosome activity. The functionality of the clones containing such perturbations may depend on the presence of particular elements in the vicinity of the area of contact between helices 3 and 12. Such cases do not exist in natural 16 S rRNA, and their selection enriches our knowledge of the constraints imposed on the structure of ribosomal RNA in functional ribosomes.

Ribosomes are RNA-protein complexes that perform protein synthesis in all living organisms. In eubacteria, the ribosome consists of three RNA chains and of more than 50 different proteins. Although the major steps of protein synthesis have been known for a long time, the detailed picture of how the ribosome structure forms and how it functions is still to be drawn. The recent achievements in x-ray crystallography of individual ribosomal subunits (1–4) and of the whole ribosome (5, 6) have opened a new era in the study of the mechanisms of protein synthesis. Now, structure-function relationships can be analyzed for any region of the ribosome taken within its natural context. Of a special interest are those elements of the ribosome structure that have very similar, almost identical conformations and thus constitute recurrent motifs. Due to the similarity, recurrent motifs could play similar roles in the formation of the ribosome structure and, potentially, in the ribosome function.

Recently, we described a new, so-called along-groove packing motif (AGPM)4 found in a dozen places in the RNA of both ribosomal subunits (7). AGPM consists of two double helices closely packed via their minor grooves in the way that the sugar-phosphate backbone of one chain of each helix goes along the minor groove of the other helix (see Fig. 1). In each helix, the chain that is packed in the minor groove of the opposite helix is positioned closer to the center of the arrangement and is thus called internal, contrary to the other chain that is called external. The interaction between the two base pairs at the center of the inter-helix contact zone is responsible for about one-half of the inter-helix atom-atom contacts, including a framework of five H-bonds (see Fig. 2, A and B). Henceforth, we call these base pairs central. The whole arrangement is characterized by a quasi-reciprocity, which is limited by the fact that the close packing of the two helices requires that one of the central base pairs be GU, while the other one must be Watson-Crick (WC). The GU base pair should be oriented in the way that G and U belong to the external and internal strand, respectively (henceforth, the first nucleotide in each base pair corresponds to the external position). A violation of this requirement would result in a loss of the close packing of the helices and may ruin the whole arrangement (see Fig. 2C).

An essential feature of AGPM is that it is able to bring together elements distant from each other in the secondary structure. This ability and the fact that AGPM has been found in many parts of the ribosome structure make this motif an essential element of the ribosome architecture. Therefore, the elucidation of the general requirements for the formation of AGPM is important for understanding how the whole ribosome structure forms and how it functions.

In this report, we explored a possibility of alternative nucleotide arrangements in AGPM SU296 (see Fig. 3), which is formed between helices 3 and 12 of 16 S rRNA. Within this motif, the central base pairs occupy positions G27–C556 and U296–G27–C556. SU296 is positioned close to the root of the central pseudoknot at the region where three domains of the 30 S subunit the shoulder, the platform, and the neck come together. Each of two fragments 28−295 and 302−555 of 16 S rRNA starts

4 The abbreviations used are: AGPM, along-groove packing motif; WC, Watson-Crick; wt, wild-type; CAT, chloramphenicol acetyltransferase; GFP, green fluorescent protein.
Along-groove Packing Motif in 16 S rRNA

at SU296 and, after making a large loop, comes back to SU296 as a part of the other helix. Therefore, the formation of SU296 would strongly affect the set of possible conformations for each of the two fragments. Given that regions 28–295 and 302–555 form together the body and the shoulder of 30 S, SU296 is expected to play a major role in the formation of the tertiary structure of these two domains of the small ribosomal subunit. Thus, the closeness to several major domains and a probable involvement in the formation of their structures would make SU296 crucial for the integrity of the whole 30 S subunit.

Unlike most AGPMs, SU296 is not immediately involved in any interaction either with another part of rRNA or with a ribosomal protein (3, 6), which makes it a perfect candidate for studying the properties of AGPM within its natural structural context without a necessity of taking into consideration a potential role of neighboring regions. Using in vivo instant evolution approach applied to a combinatorial gene library of 16 S rRNA, we selected fourteen functional variants in which the identities of the nucleotides composing the two central base pairs of motif SU296 were different from those found in the wild-type (wt) Escherichia coli 16 S rRNA. Only in a few variants the efficiency was close to wt. Analysis of the nucleotide sequences of these variants helped us to identify the limits within which the wt sequence can be modified to preserve the ribosome function. In two variants, the GU and WC base pairs stayed as in wt, which made the close helix packing undisturbed. In two more variants, GU was replaced by WC, while WC was replaced by GU. Individually, each of two replacements GU → WC and WC → GU would result in a loss of the close packing between the helices. However, when both replacements occur simultaneously, the quality of the helix packing is not affected. In other clones, the nucleotide sequences do not allow the close packing of the two helices; correspondingly, the efficiency of such ribosomes is severely damaged. Based on these results, we can conclude that the close packing of the helices within SU296 is a major aspect of the ribosome functionality.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Media—For cloning and selection, we used the E. coli strain DH5α. Cultures were grown in the LB medium (8) or in the LB medium with 100 μg/ml ampicillin (LB-Amp-100) and with 50 μg/ml kanamycin, when appropriate (Sigma-Aldrich).

Plasmids—For cloning of combinatorial 16 S rRNA gene libraries and for selection of functional clones, plasmid pAMMG carrying the specialized ribosome system was constructed. This plasmid is analogous to the ones described elsewhere (9–17). Briefly, it contains a copy of the E. coli rnb operon (from pKK1192U) (18) under the transcriptional regulation of the lacIV5-inducible promoter. It also contains two genes with modified Shine-Dalgarno sequences, of a chloramphenicol acetyl-transferase (CAT) and of a green fluorescent protein (GFP), which were used, respectively, for selection and for quantifying the efficiency of the selected clones (supplemental “Methods”). To measure the level of miscoding of the selected clones, seven variants of plasmid pLuc were constructed (supplemental “Methods”).

Construction of the Combinatorial Gene Library and Selection of Functional Clones—To randomize the four nucleotides comprising the two central base pairs of the SU296 motif (Fig. 3C), we used an overlapping extension PCR procedure (19). In this way, the entire region comprising the SU296 motif (902 bp) was amplified by consecutive multistep PCR. The sequences of the primers used for different steps of PCR are given in supplemental “Methods.” Prior to selection, the transformants were grown for 1 h in the LB medium. The synthesis of the plasmid-encoded ribosomes was induced by addition of isopropyl-1-thio-β-D-galactopyranoside (BioShop Canada Inc.) to a final concentration of 1 mm. After incubation for 3.5 h, the library was plated on the selection plates containing 200 μg/ml chloramphenicol and 1 mm isopropyl-1-thio-β-D-galactopyranoside. The minimal inhibitory concentration of chloramphenicol for the cells expressing pAMMG-encoded rRNA was ~500 μg/ml. Out of a pool of 1.4 × 10^5 transformants, 50 survivors were obtained as chloramphenicol-resistant and taken for further analysis.

GFPuv3 Assay—For quantifying the efficiency of the selected clones, the GFPuv3 variant (20) of the GFPuv gene (BD Biosciences) was incorporated into pAMMG under the control of a synthetic lactococcal strong constitutive CP25 promoter (21). Compared with GFPuv, the GFPuv3 variant provided for a higher sensitivity of measuring the fluorescence. freshly transformed colonies of the selected clones were inoculated into 1 ml of the LB-Amp-100 medium containing 1 mm isopropyl-1-thio-β-D-galactopyranoside and incubated at 37 °C with shaking for 16 h. After the incubation, the cells were harvested, washed twice in 1 ml of the HN buffer (20 mm Heps, 0.85% NaCl, pH 7.4), and re-suspended in 1 ml of the HN buffer. The cell density (A_{600}) and the fluorescence (excitation = 497 nm, emission = 511 nm) of each clone was determined using a Packard Fusion α-FP plate reader. For each clone, the fluorescence was divided by A_{600} and presented as a percentage of wt.

Firefly Luciferase Assay—Each of the seven miscoding variants of plasmid pLuc (supplemental Tables S3 and S4) was co-transformed together with pAMMG-wt or with a mutant derivative of it into DH5α cells and plated on the LB agar containing 100 μg/ml ampicillin and 50 μg/ml kanamycin. To induce the specialized ribosome system, we proceeded as described above. To measure the level of miscoding, the luciferase assay was performed as described in the supplemental “Methods.”

Subunit Association Assay—Cells containing pAMMG or a mutant derivative of it were induced and grown in 500 ml of the LB-Amp-100 media. The ribosomal particles were isolated from the cellular lysates prepared by freezing and thawing method (22). The lysates were applied onto a 15–40% (w/v) linear sucrose gradient in 20 mm Tris-HCl (pH 7.8), 5 mm MgCl_2, 100 mm NH_4Cl, and 6 mm β-mercaptoethanol, followed by centrifugation at 20,000 rpm for 20 h in a Beckman rotor SW28 at 4 °C. Fractions containing the 30 S subunits and the 70 S ribosomes were collected using an ISCO gradient fractionator. RNA was extracted from the appropriate fractions with phenol. The proportion of the plasmid-encoded and host-encoded 16 S rRNA was determined by the primer extension method (12, 23) (supplemental “Methods”).
Sequencing—Sequencing of the selected clones was performed on the LI-COR DNA sequencing system (Département de Biochimie, Université de Montréal) using primer 5′-AATTATTTATGACAATCTGTGGGGAATCG-3′ for reading positions 27, 296, and 301 and primer 5′-GATGGTGGGTTAAGCCCTACG-3′ for reading position 556. These primers were labeled at the 5′-end with IRDye-800 (LI-COR Biosciences).

Computer Modeling—The models were based on the crystal structure of the E. coli 30 S subunit (6) (PDB code 2AVY). Fragments 26–30 and 553–558 of helix 3 and 293–304 of helix 12 were taken from the x-ray structure and were used as a starting conformation into which the nucleotide replacements corresponding to a particular selected clone were introduced. Modeling was done interactively, using the InsightII/Discover package (version 2000, Molecular Simulations Inc., San Diego, CA). Each model was submitted to a restrained energy minimization using the AMBER force field (amber.ch.ic.ac.uk/amber). During the minimization, bp 30–553 and 26–558 of helix 3, bp 293–304 of helix 12, as well as nucleotides 298 and 299 of the tetraloop capping helix 12 were immobilized. Visualization was done on a Silicon Graphics Fuel computer.

RESULTS

Experimental System and Library Design—To study structure-function characteristics of SU296 within the 30 S subunit, we used the instant evolution approach that consisted in selection of functionally active ribosomes from a combinatorial library in which several nucleotides of 16 S rRNA were randomized. The selection was done in vivo with use of a specialized ribosome system pAMMG analogous to the ones previously described (9–17). The system contained genes of rRNA and of two reporter proteins, the chloramphenicol acetyltransferase (CAT) and the green fluorescent protein (GFPuv3) (20). The rRNA gene was needed for incorporation and expression of the designed combinatorial library, while CAT and GFPuv3 were used, respectively, for selection of active clones in the presence of chloramphenicol and for measurement of the ribosome activity. In the 16 S RNA gene, the anti-Shine-Dalgarno sequence was modified as suggested (14, 24), and the corresponding modification of the Shine-Dalgarno sequence was introduced into the genes of both reporter proteins. In the cell, this system co-exists without interference with the normal cellular protein-synthesis machinery (14). In the combinatorial library, all four positions composing the two central base pairs (Figs. 1 and 2A) were fully randomized (Fig. 3), which provided for 256 possible variants.

Cloning and Selection of Functional Clones—After cloning the library, a pool of ~1.4 × 10^5 transformants was obtained. A random sequencing of 12 transformants showed that all four positions of the central base pairs in the SU296 motif were reasonably randomized, except that there was some bias toward cytosines and against adenines, accounting for 38 and 7% of the identities of randomized nucleotides, respectively (data not shown). Out of 50 clones that were selected as survivors in the presence of 200 μg/ml chloramphenicol, 25 had the wt sequence, most probably, due to an incomplete digestion of plasmid pAMMG during the cloning of the library. The other 25 clones contained 14 unique sequences (Table 1), which demonstrated a mild level of redundancy. In no case did mutations touch non-randomized nucleotides.

Characterization of the Selected Clones—For the selected clones, the GFPuv3 activity varied between 2 and 100% of that of wt (Table 1, column GFPuv3). The selected clones were further characterized as to their capacity for the normal protein synthesis. First, for each clone, we determined the proportion of the plasmid-encoded 16 S rRNA in the population of 30 S subunits (column 30 S). This proportion mostly depends on the
relation between the rates of formation and degradation of the specialized 30 S subunit and indirectly characterizes the stability of its structure. Second, we determined the proportion of the specialized ribosomes within the total 70 S ribosome population (column 70 S). This proportion depends on several factors, including some not related to the mutated ribosomes, like the availability of mRNAs carrying the specialized Shine-Dalgarno sequence. However, if the general conditions are maintained and if the number in column 30 S is comparable to that of wt, a low number in column 70 S will suggest a damaged initiation of translation.

Next, for each clone, the accuracy control was performed. Using previously reported techniques (12, 25, 26), we measured the efficiency of seven different miscoding events, of the opal, ochre, and amber stop codon read-through, of the −1 and +1 frameshift, as well as of two types of amino acid misincorporation. Each time, the particular miscoding event was required for the expression of the reporting protein firefly luciferase from the specially constructed derivative of its gene, and the activity of this protein was taken as a measure of the efficiency of the miscoding. For each clone, the average and the highest levels of the seven measured miscoding events are given in Table 1 (three last columns). The complete set of data on translational fidelity of the mutated ribosomes is presented in the supplemental Table S1.

Potentially Negative Role of Uridine 1192—When this work was in progress, Rodriguez-Correa and Dahlberg (27) demonstrated that mutation C1192U, which is also present in the 16 S rRNA encoded in pAMMG, can show synthetic lethality with some other mutations in 16 S rRNA. To test a potentially damaging role of U1192 in the selected clones, we created six additional mutants through introduction of the reverse mutation U1192C into six clones with varying activities. The activities of all new mutants are presented in the supplemental Table S5. They were within the same ranges of activities as of their U1192-containing analogs, being, on average, 7% higher. These results demonstrate that the observed phenotypes are essentially independent of the C1192U mutation. These results also show that mutation C1192U provides a mild negative effect on the ribosome function, which is independent of the packing between helices 3 and 12.

Analysis of the Selected Clones—Although the activity of all selected clones was sufficient for survival in the presence of chloramphenicol, the efficiency of the protein synthesis varied substantially among the clones. Based on the characteristics studied, we divided all clones in three groups seen in Table 1.
Group I consists of clone A1 with the wt sequence and of three more clones whose functional characteristics are comparable to those of wt. Group II consists of only one clone, A2, which in all aspects is as good as the Group I clones, except for a slightly lowered miscoding level. All other clones form Group III; in each clone of this group, some functional aspects are essentially damaged compared with those in Group I and II clones.

Analysis of the selected clones demonstrates that the ribosome efficiency strongly depends on the identity of the central base pairs of SU296. Thus, we have noticed that in all Group I and Group II clones, both central base pairs are either GU or WC. On the other hand, all clones in which at least one central base pair is different from GU and WC belong to Group III. One can say that the presence of any kind of base pair different from GU and WC results in a serious damage to the ribosome function. Also, in most Group III clones, the fraction of the special-ized 30S subunits is substantially lower than in wt, which indicates that the formation of the 30 S subunit is compromised. This observation is in agreement with the discussed earlier suggestion that the integrity of SU296 is critical for the integrity of the structure of the whole 30S subunit.

Close Packing in Group I Clones—From Table 1, one can also learn that in all Group I clones and only in them, one central base pair is GU, whereas the other one is WC. Therefore, the co-existence of a GU and WC central base pair in SU296 should be considered as a major factor of the normal ribosome functioning. Two clones A1 and A8 differ only by the identity of bp 27–556, which in these clones is GC and CG, respectively. As one can see in Fig. 2 (A and B), replacement GC → CG preserves both the close helix packing and the inter-helix H-bonds. Interestingly, in the other two Group I clones, A5 and A7, GU has moved to the other helix, being replaced by either GC (A5) or CG (A7). Such an interchange of the central base pairs is not expected, however, to affect the quality of the helix packing. Indeed, due to the quasi-reciprocity of the arrangement, none of the two helices has a preference of harboring GU or WC. Therefore, if the close packing of the two helices exists in clones A1 and A8, it should be preserved in clones A5 and A7 as well.

To elucidate additional details related to the GU ↔ WC interchange, we superposed the juxtaposition of the central base pairs in clone A1 with the same juxtaposition rotated for 180° around the center of pseudo-symmetry of the arrangement (Fig. 4). This superposition would simulate the GU ↔ GC interchange that occurs in clone A5 compared with A1. The superposition shows that, after such an interchange, none of the four nucleotides would remain in its position. The displacement of each nucleotide can be well approximated by a rotation around a particular point. The external purines rotate for ~12° around their C1’ atoms, whereas the internal pyrimidines rotate for ~17° around their N1 atoms. The direction of the rotation is the same for nucleotides of the same base pair and opposite for nucleotides of different base pairs. In each case, the direction of the rotation is such that it would allow the bases to rearrange into the newly formed base pair GC instead of GU and GC instead of GC.

Although the rotation of each nucleotide is driven by the changes of the hydrogen bond scheme within each base pair and is thus independent of the rearrangements that occur in the other double helix, the simultaneous movements of all four nucleotides complement each other in the sense that together they allow the maintenance of the close packing between the two helices. In particular, these movements preserve the dis-

### Table 1: Nucleotide sequences and characteristics of the selected clones

| Clone | N** | Randomized positions | GFPuv** | Plasmid-encoded 16 S rRNA* | Ribosome miscoding** |
|-------|-----|----------------------|---------|---------------------------|---------------------|
|       |     | 301 296 27 556       |         | 30 S                      |                     |
|       |     |                      | Ribosome miscoding** |
|       |     |                     | Highest | Misoding event |
| Group I | A1 (wt) | 25 G U G C | 100 | 37 ± 2 | 64 ± 2 | 1.0 | 1.0 | S → R |
|       | A8   | 3 G U C G | 81 ± 5 | 21 ± 2 | 61 ± 3 | 1.0 | 1.3 | Opal |
|       | A5   | 3 G C G U | 85 ± 8 | 39 ± 3 | 54 ± 4 | 1.2 | 1.7 | Ochre |
|       | A7   | 1 C G G U | 70 ± 4 | 44 ± 2 | 51 ± 1 | 1.1 | 1.7 | Amber |
| Group II | A2  | 7 G C G C | 79 ± 6 | 35 ± 1 | 60 ± 2 | 1.4 | 2.3 |                     |
| Group III | A12 | 1 G C G C | 49 ± 3 | 32 ± 2 | 17 ± 1 | 0.8 | 1.8 | S → R |
|         | A3   | 1 G C A G | 26 ± 3 | 26 ± 1 | 10 ± 1 | 5.8 | 28.3 | G → R |
|         | A9   | 1 G C C U | 11 ± 2 | 28 ± 2 | 7 ± 1 | 8.6 | 19.2 | −1 fs |
|         | A11  | 1 C G C A | 22 ± 3 | 10 ± 2 | 29 ± 3 | 1.6 | 2.7 | G → R |
|         | A4   | 1 C G G U | 78 ± 3 | 14 ± 4 | 17 ± 2 | 1.3 | 2.6 | Amber |
|         | A6   | 1 C C G G | 10 ± 2 | 11 ± 1 | 11 ± 2 | 3.0 | 3.9 | Amber |
|         | A10  | 1 G C C A | 4 ± 1 | 10 ± 2 | 11 ± 3 | 7.6 | 25.6 | −1 fs |
|         | A13  | 1 C C U G | 2 ± 1 | 8 ± 3 | 4 ± 1 | 6.8 | 21.8 | −1 fs |
|         | A15  | 1 C U U G | 2 ± 1 | 3 ± 1 | ND* | 7.6 | 17.5 | −1 fs |
|         | A14  | 1 G U C G | 2 ± 1 | 2 ± 1 | ND | ND | ND |                    |

**N is the number of times each clone has been isolated.

The ribosome activity was calculated as the mean ± S.D. of five to eight independent experiments.

The proportion of the plasmid-encoded 16 S rRNA in the 30 S and 70 S fractions was calculated as the mean ± S.D. of experiments performed with three independent RNA preparations. The determination of the level of specialized 30 S subunits within the 70 S ribosomes of 2% or less was not detectable.

Effect of the mutations in AGPM SU296 on the level of ribosome miscoding. A value of 1.0 was arbitrarily ascribed to the wt 16 S rRNA. For other clones, the level of miscoding was calculated as the mean ± S.D. of three independent experiments. The extended set of data is presented in supplemental Table S1. Here, only the average level of efficiency of seven different miscoding events as well as the highest level of miscoding efficiency and the corresponding miscoding event are provided (see text). Different miscoding events: ochre, amber, andopal: suppression of the corresponding stop codon; −1 fs, +1 fs: the −1 and +1 frameshifts; S → R: G → R: insertion of arginine in response to a serine or glycine codon, respectively. For clone A14, due to a very low activity, none of the levels of miscoding were reliably reproducible.

*ND, not detectable.
tance, and therefore, the H-bond between the O2' atoms of the internal nucleotides as well as between the N2 atom of G and the O3' atom of the opposite internal nucleotide (for reference, see Fig. 2).

A different situation would occur if only one of the two replacements GU → WC or WC → GU takes place, while the other base pair remains unchanged. In this case, the nucleotide displacements in the mutated base pair will not be matched by the corresponding displacements in the other base pair, which will result in the loss of the close helix packing. For example, replacement GU → WC would displace the external nucleotide of this base pair for several angstroms farther from the opposite helix, thus creating a crack in the arrangement clearly seen in the known exceptional cases (Fig. 2C). From this point of view, the GU ↔ WC interchange constitutes a structural compensation in which a negative effect of one base pair replacement is compensated by the other base pair replacement. Because GU and WC base pairs co-exist in all Group I clones, and only in them, we can conclude that the close packing of helices 3 and 12 is critical for the normal ribosome function.

Because the GU ↔ WC interchange does not affect the packing of the helices, it has a chance to occur in evolution and thus could be detected through comparison of available nucleotide sequences of rRNA. Indeed, as recently pointed out by Mokdad and co-workers (28), the GU ↔ WC co-variation of the central base pairs exists in some AGPMs of both ribosomal subunits. Interestingly, SU296 is not included in these AGPMs and, on the contrary, demonstrates a strong preference of GU for helix 12 and of WC base pair for helix 3 (Table 2 and supplemental Table S2). In fact, the latter observation was one of the reasons that determined our choice of motif SU296 for experimental studies (31). The results presented here show that a possibility for the GU ↔ WC interchange is a general property of AGPM, which goes beyond the natural variability of rRNA.

On a more detailed level, however, one can see differences in the performance of the Group I clones before and after the GU ↔ WC interchange. In particular, comparison of clones A5 to A1 and A7 to A8 shows that a displacement of the GU base pair from helix 12 to helix 3 results in a drop of the overall activity and of the 70 S level by 15%. Although these changes are not dramatic, they could be sufficient for driving the evolution toward the wt sequence. The differences in the performance of the Group I clones demonstrate that the GU ↔ WC interchange is not, in fact, completely adequate, even though it preserves the close helix packing. A probable explanation of this phenomenon deals with the discussed earlier fact that the GU ↔

**TABLE 2**

Statistical spectrum of the identities of the central base pairs in AGPM SU296

The data were obtained from the available rRNA alignments (29). Only those cases were taken into consideration where the identities of all four nucleotides were known. For SU296 both in eubacteria and in archaebacteria, combinations of the RY-RY type dominate among non-GU-WC cases.

| Eubacteria | No. of cases | Archaeabacteria | No. of cases |
|------------|--------------|----------------|--------------|
|            | 301 296 27 556 | 301 296 27 556 |              |
| A) GU-WC type |               |               |              |
| G U G C | 5968 | G U G C | 286 |
| G U A U | 3 | G U C G | 61 |
| G U U A | 6 | G U U A | 6 |
| G U A U | 6 | G U A U | 6 |
| Total | 5971 | 359 |
| B) RY-RY type |               |               |              |
| G C G C | 26 | G U G U | 31 |
| G U A C | 26 | A U G C | 4 |
| A U G C | 25 | A U G C | 4 |
| G U G U | 4 | G U G U | 4 |
| A C G C | 1 | A C G C | 1 |
| Total | 82 | 35 |
| C) Others |               |               |              |
| G U U C | 18 | G U U G | 1 |
| G U C C | 14 | A U C G | 1 |
| U G C C | 3 | G U C U | 1 |
| C U G C | 2 | G U A G | 1 |
| G U G A | 1 | C U U C | 1 |
| C U U C | 1 | A U C C | 1 |
| Total | 40 | 4 |
WC interchange forces a displacement of all four nucleotides of the central base pairs. This displacement could affect the positions of the neighboring nucleotides and, indirectly, the conformations of more distant regions of the 30 S subunit. A potentially negative effect of such conformational changes could explain why the GU ↔ GC interchange does not occur naturally in SU296.

Preferences for the WC Base Pair—Further analysis shows that the identity of the WC base pair is also important for the ribosome function. Thus, one can see that the WC base pair is neither AU nor UA in any of the Group I clones. Although at least partly this fact can be explained by the mentioned above bias of the original library against A, it also fits well to what is known about AGPM in general. Indeed, the GC/CG base pairs have an obvious advantage compared with AU/UA because of the H-bond between the amino group of G and O3’ of U from the opposite GU base pair (Fig. 2, A and B), which does not exist in the AU/UA case. This additional H-bond provides a stabilizing effect on the AGPM, which could be critical for the normal ribosome function. The preference toward GC/CG observed in our selection goes along with the fact that, in the wt rRNA sequences, GC/CG as central base pairs of AGPM occur overwhelmingly more often than AU/UA. For most AGPMs, the content of AU/UA is lower than 10% and only in one case exceeds 20% (29) (supplemental Table S2).

Another observation pertains to the preference of GC over CG. In particular, comparison of clones A8 to A1 and A7 to A5 shows that the replacement of GC by CG in any of the two helices results in a drop of the overall activity by ~20%. As one can see in Fig. 2B, nucleotide C in the external position, unlike G, does not form any interaction with the opposite base pair. The absence of such an interaction will make the arrangement less cooperative and consequently, less stable, which can explain the preference of GC versus CG. It is important that this preference is not specific to SU296 and is also observed in other cases of AGPM. For example, in eubacteria, GC as a central base pair is, on average, about three times as frequent as CG (28, 29) (supplemental Table S2). In the case of SU296, the local context can also contribute to the GC preference over CG in helix 3. Indeed, the presence of C in position 27 would allow an alternative base pairing of this nucleotide with the unpaired G in position 557 (Fig. 3C). We may speculate that a potential formation of such an alternative base pair in clone A8 results in the drop of the 30 S level by more than 40% compared with A1 (Table 1).

A Suboptimal Close Packing Arrangement for Clone A2—Clone A2, the only one composing Group II, represents a special case. In this clone, both central base pairs are GC, which, according to our findings, does not allow the close packing of the two helices. Although all other clones without a close helix packing are functionally damaged, A2 functions almost as well as the Group I clones. This ability of A2 is not shared by other clones with two WC central base pairs. Indeed, a replacement of a central GC base pair in A2 by either CG or UA, which happens in clones A12 and A10, respectively, results in a substantial drop of the ribosome activity. For A12, which contains CG compared with GC in A2, the drop of activity is much stronger than in the previously discussed pairs of clones A1–A8 and A5–A7, experiencing the same kind of base pair replacement. Interestingly, the GC–GC combination can occur in different AGPM naturally, without any alarming consequence for the survival of the organism. In particular, for SU296 in eubacteria, combination GC–GC is one of two most popular alternatives to GU–GC (29) (Table 2). Based on our results and on the mentioned above phylogenetic data, one can consider the GC–GC combination as a possible exception of the GU-WC rule able to make the ribosome functional.

Further analysis revealed an unusual feature of the GC–GC combination that may relate to its ability to maintain the ribosome function. We found that, if the geometry of H-bonds within the two central base pairs is allowed to be somewhat deformed, a closely packed suboptimal arrangement will still be possible (Fig. 5). Despite the deformed hydrogen bonds, this arrangement corresponds to an energy minimum and is thus characterized by a certain level of stability. Essentially, this arrangement can be built only if both internal positions are occupied by pyrimidines, i.e. for GC–GC and not for GC–CG, CG–GC, or CG–CG. We can thus speculate that the ability of the GC–GC combination to form such a closely packed arrangement is responsible for the strong performance of clone A2. Because for the GC–GC combination, the suggested arrange-
Along-groove Packing Motif in 16 S rRNA

...ment would compete with the arrangement seen in Fig. 2C, some level of conformational flexibility for AGPM containing this combination seems to be unavoidable. This flexibility could be responsible for the increased level of miscoding observed for clone A2 compared with Group I clones.

Interestingly, a similar close packing arrangement with deformed H-bonds can also be formed for any other combination of two RC base pairs, where R and Y stand for a purine and a pyrimidine, respectively. Fig. 5 demonstrates how such an arrangement is formed for two GC base pairs. Other RC dinucleotide combinations like GU, AU, and AC could also fit to this arrangement, although the two latter combinations would make it weaker due to the absence of an amino group in the minor groove. In connection with this, it is worth mentioning that, in the case of SU296, combinations R-Y-RY-RC are accountable for most exceptions from the standard GU-WC pattern (Table 2).

In addition to A2, there are two more selected clones A12 and A10 in which both central base pairs are WC. The efficiency of these clones, however, was notably lower than that of A2, which was the reason for us to put them in Group III. A poorer performance of these clones correlates with the fact that the central pairs fit to the RY-RC pattern neither in A12 nor in A10, and, therefore, a close helix packing is reachable in none of these clones even as a suboptimal arrangement. The most probable non-close arrangement of two WC base pairs is seen in Fig. 2D. Compared with the close packing, in this arrangement, one inter-helix hydrogen bond will be lost, which can explain a lower activity of clones A12 and A10 relative to A2. Also, the presence of a UA base pair instead of GC or CG would deprive the helix-helix contact of another inter-helix hydrogen bond, which could be critical for the function and would explain the low efficiency of A10 compared not only to A12, but also to most Group III clones. In all other Group III clones, at least one central base pair is different from GU and WC, which suggests stronger rearrangements in the area of contact between helices 3 and 12. Correspondingly, none of these clones functions properly. Moreover, these clones demonstrate a variety of problems related to all aspects of the ribosome function from the formation of the 30 S subunit structure and the subunit association to low elongation efficiency and a compromised accuracy. Altogether, these results clearly show that a normal function of the ribosome requires helices 3 and 12 be closely packed accordingly to the standard GU-WC pattern, or, at least, to the RY-RC pattern.

DISCUSSION

In this study, we used in vivo selection of functional clones of 16 S rRNA for analysis of general constraints imposed on AGPM SU296 in working ribosomes. The selection was based on the ability of the mutated ribosomes to synthesize CAT that would allow the cells to survive in the presence of chloramphenicol. It is important, however, that the fact of survival per se has not been able to guarantee a high efficiency of protein synthesis. Indeed, only an extensive functional characterization of the selected clones, which included the measurement of the activity of protein synthesis, of the proportion of the specialized 30 S subunits, and of specialized 70 S ribosomes within the whole population of, respectively, 30 S subunits and 70 S ribosomes, and, finally, a systematic control of the accuracy of the specialized ribosomes has allowed us to distinguish the clones with efficiency approaching wt from those clones in which some aspects of functionality were severely damaged. A systematic comparison of the nucleotide sequences of more and less effective clones helped us to reveal some important characteristics of AGPM SU296.

Most characteristics of SU296 elucidated in this study pertain to close packing and stability of the inter-helix arrangement and as such are expected to be shared by all or at least most AGPMs. The close packing of the helices can be achieved if one of the central base pairs is GU, whereas the other one is WC. To make the arrangement sufficiently stable, the WC base pair should be either GC or CG. The Group I clones satisfying both requirements demonstrate the activity comparable to wt. The GC base pair is expected to provide a more stable arrangement than CG, and, correspondingly, the clones containing combinations GU-GC and GC-GU demonstrate a higher level of activity than the clones containing GU-CG and CG-GU. On average, these four combinations occupy the central base pairs in different AGPMs more often than any other combination, thus reflecting the general importance of stable helix-helix packing for the ribosome structure and function (supplemental Table S2). Although one successful A2 clone does not follow the GU-WC sequence pattern, even here, a close helix packing scheme can be suggested as a suboptimal arrangement. Even a replacement of a GC or CG base pair by UA, which deprives the helix-helix interaction of an important hydrogen bond, is critical for the ribosome function. The latter conclusion is made based on comparison of clone A10 with other clones and is independent of the fact that, as mentioned above, there was some bias in the randomized positions toward C and against A.

The importance of the close packing between helices 3 and 12 becomes even more evident if one compares clones of the first two groups with those of Group III. In no clone of the latter group, the standard packing of helices 3 and 12 is reachable; correspondingly, none of these clones works properly. It would be interesting to know which aspects of the ribosome functional cycle are affected the most if helices 3 and 12 are not closely packed. Based on the results presented here, we can say that the absence of close packing would affect both the integrity of the tertiary structure of the 30 S subunit and its function. Indeed, as it is seen in Table 1, for most Group III clones the proportion of the specialized 30 S subunits in the population of 30 S subunits is substantially lower than for wt. However, even when this proportion is comparable to that of wt, which happens in clones A12, A3, and A9, the level of 70 S ribosomes is still severely affected. In addition, in most Group III clones, the ribosomes are essentially less accurate than in wt. In some clones, the miscoding level for specific events exceeds the corresponding level for the wt ribosomes more that 20 times. Altogether, these data demonstrate that AGPM SU296 is positioned at the intersection of different important processes related to the ribosome function, and, depending on the nature of the interaction between helices 3 and 12 within this motif, different aspects of the ribosome function can be affected differently. Additional studies are required to elucidate the particular
mechanism of how imperfections in packing of helices 3 and 12 affect the 30 S and 70 S formations, as well as the efficiency and accuracy of translation.

A specific characteristic of SU296 pertains to the fact that, for this motif, unlike for several other AGPMs, the GU ↔ WC interchange has never been observed among natural rRNA. The conservation of the positions of the GU and WC base pairs demonstrates the asymmetry between the two helices, which has been confirmed by a somewhat lower activity of the clones having GU in helix 3 compared with those having this base pair in helix 12. The existence of such an asymmetry shows that the stability and the close helix packing are not the only aspects of SU296 that are important for the ribosome function. It also demonstrates the asymmetry between the two helices, which affects the 30 S and 70 S formations, as well as the efficiency and accuracy of translation.

Along-groove Packing Motif in 16 S rRNA

REFERENCES

1. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) Science 289, 905–920
2. Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., and Yonath, A. (2000) Cell 102, 615–623
3. Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Jr., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) Nature 407, 327–339
4. Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001) Cell 107, 679–688
5. Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H., and Noller, H. F. (2001) Science 292, 883–896
6. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. (2005) Science 310, 827–834
7. Gagnon, M. G., and Steinberg, S. V. (2002) RNA (N. Y.) 8, 873–877
8. Luria, S. E., and Burrous, J. W. (1957) J. Bacteriol. 74, 461–476
9. Lee, K., Varma, S., SantaLucia, J., Jr., and Cunningham, P. R. (1997) J. Mol. Biol. 269, 732–743
10. Morosyuk, S. V., Lee, K., SantaLucia, J., Jr., and Cunningham, P. R. (2000) J. Mol. Biol. 300, 113–126
11. Lee, K., Holland-Staley, C. A., and Cunningham, P. R. (2001) J. Nutr. 131, 2994S–3004S
12. Belanger, F., Gagnon, M. G., Steinberg, S. V., Cunningham, P. R., and Brakier-Gingras, L. (2004) J. Mol. Biol. 338, 683–693
13. Morosyuk, S. V., SantaLucia, J., Jr., and Cunningham, P. R. (2001) J. Mol. Biol. 307, 213–228
14. Rackham, O., and Chin, J. W. (2005) Nat. Chem. Biol. 1, 159–166
15. Abdii, N. M., and Fredrick, K. (2005) RNA (N. Y.) 11, 1624–1632
16. Hui, A., and de Boer, H. A. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 4762–4766
17. Hui, A., Juhrani, P., and de Boer, H. A. (1987) Methods Enzymol. 153, 432–452
18. Brosius, J., Ulrich, A., Raker, M. A., Gray, A., Dull, T. J., Gutell, R. R., and Noller, H. F. (1981) Plasmid 6, 112–118
19. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Gene (Amst.) 77, 51–59
20. Ito, Y., Suzuki, M., and Husimi, Y. (1999) Biochem. Biophys. Res. Commun. 264, 556–560
21. Jensen, P. R., and Hammer, K. (1998) Appl. Environ. Microbiol. 64, 82–87
22. Ron, E. Z., Kohler, R. E., and Davis, B. D. (1966) Science 153, 1119–1120
23. Sigmund, C. D., Ettayebi, M., Borden, A., and Morgan, E. A. (1988) Methods Enzymol. 164, 673–690
24. Lee, K., Holland-Staley, C. A., and Cunningham, P. R. (1996) RNA (N. Y.) 2, 1270–1285
25. Dyer, B. W., Ferrer, F. A., Klindinst, D. K., and Rodriguez, R. (2000) Anal. Biochem. 282, 158–161
26. Viviani, V. R., Uchida, A., Viviani, W., and Ohmiya, Y. (2002) Photochem. Photobiol. 76, 538–544
27. Rodriguez-Correa, D., and Dahlberg, A. E. (2004) RNA (N. Y.) 10, 28–33
28. Mokdad, A., Krasovska, M. V., Sponer, J., and Leontis, N. B. (2006) Nucleic Acids Res. 34, 1326–1341
29. Wuyts, J., Perriere, G., and Van De Peer, Y. (2004) Nucleic Acids Res. 32, D101–D103
30. Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D’Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Muller, K. M., Pande, N., Shang, Z., Yu, N., and Gutell, R. R. (2002) BMC Bioinformatics 3, 2
31. Gagnon, M. G., Mukhopadhyay, A., Butorin, Y., Chen, J., and Steinberg, S. V. (2005) in RNA 2005, Tenth Annual Meeting of the RNA Society, Banff, Alberta, Canada, May 24–29, 2005, p. 380

Acknowledgments—We thank Dr. Harry F. Noller for providing plasmid pKK1192U and Dr. Léa Brakier-Gingras for critical reading of the manuscript.