Journal of Microbiological Methods 76 (2009) 181–187

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Rapid *in vivo* exploration of a 5S rRNA neutral network

Zhengdong D. Zhang [a,*], Madhavi Nayar [b], David Ammons [c], Joanne Rampersad [c], George E. Fox [b,*]

[a] Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA
[b] Department of Biology and Biochemistry, University of Houston, Houston, Texas 77204-5001, USA
[c] Department of Chemistry, The University of Texas Pan American, Edinburgh, TX 78539-2999, USA

ARTICLE INFO

Article history:
Received 12 August 2008
Received in revised form 13 October 2008
Accepted 15 October 2008
Available online 30 October 2008

Keywords:
5S rRNA
E. coli knockout strains
Neutral network
In vivo mutagenesis

ABSTRACT

A partial knockout compensation method to screen 5S ribosomal RNA sequence variants *in vivo* is described. The system utilizes an *Escherichia coli* strain in which five of eight genomic 5S rRNA genes were deleted in conjunction with a plasmid which is compensatory when carrying a functionally active 5S rRNA. The partial knockout strain is transformed with a population of potentially compensatory plasmids each carrying a randomly generated 5S rRNA gene variant. a The ability to compensate the slow growth rate of the knockout strain is used in conjunction with sequencing to rapidly identify variant 5S rRNAs that are functional as well as those that likely are not. The assay is validated by showing that the growth rate of 15 variants separately expressed in the partial knockout strain can be accurately correlated with *in vivo* assessments of the potential validity of the same variants. A region of 5S rRNA was mutagenized with this approach and nine novel variants were recovered and characterized. Unlike a complete knockout system, the method allows recovery of both deleterious and functional variants.. The method can be used to study variants of any 5S rRNA in the *E. coli* context including those of *E. coli*.

Published by Elsevier B.V.

1. Introduction

Computational studies of RNA evolution have revealed the importance of *neutral networks*—large sets of diverse sequences that fold into equivalent structures. For a typical RNA secondary structure, the sequences that fold into it as their ground state form a vast mutationally connected network that spans virtually all of sequence space (Eigen et al., 1988; Nino, 1983; Reidys et al., 1997). In such model systems, having the same structure is considered tantamount to having equal fitness. During evolution, populations are envisioned to explore such neutral networks through mutation, thereby experiencing long periods of structural stasis while accumulating significant silent genetic diversity. Evolutionary jumps are made rarely, when a mutation produces a sequence belonging to the neutral network of a better shape, it is then swept to fixation (Ancel and Fontana, 2000; Fontana and Schuster, 1998).

Much of this theoretical work is grounded in analyses of the RNA sequences with the potential to form equivalent secondary structures. Such genotype to phenotype maps are thus likely to only go part way in connecting genotype to fitness because real RNA structures are far more complex than envisioned by models that only consider canonical secondary structure. In addition to longer-range tertiary interactions, there are in many cases non-standard interactions associated with secondary structure regions. Also, real molecules can frequently tolerate 'imperfections' such as G–A oppositions that are not expected. That is to say, from a secondary structural perspective, a real neutral network has fuzzy boundaries the nature of which is not understood. In order to make further progress, we are seeking to characterize the real RNA neutral network associated with the *Vibrio proteolyticus* 5S rRNA and its nearby sequence space.

5S rRNA was selected as the model molecule to represent a realistic neutral network. It is a relatively small molecule (120 nucleotides) whose primary sequence and structure are sufficiently conserved that one can readily compare information between bacterial systems based on a universal numbering system (Fox, 1985). It has a well established secondary structure that contains five major helical stems whose relative orientation is defined and stabilized in part by a number of base-base tertiary interactions that have been identified at atomic resolution in *Haloarcula marismortui* 5S rRNA in the context of the 50S ribosomal subunit crystal structure (Ban et al., 2000). In addition, the RNA typically interacts with three ribosomal proteins (L5, L18, and L25) and considerable information about the nature of these interactions can also be gleamed from the crystal structure and earlier biochemical studies (Dinman, 2005; Szymanski et al., 2003). Another advantage offered by 5S rRNA is that unlike 16S rRNA, it does not appear to be directly involved in regulatory networks. Thus, neither 5S rRNA deficiency induced by gene deletions nor a plasmid encoded 5S rRNA addition seem to otherwise affect gene expression (Ammons et al., 1999; Tucker et al., 2005).

Earlier efforts sought to characterize the neutral network associated with *V. proteolyticus* 5S rRNA and its nearby sequence space using site-directed mutagenesis (Lee et al., 1993, 1997; Zhang et al., 2003). In order

* Corresponding authors. Zhang is to be contacted at Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA. Fox, Department of Biology and Biochemistry, University of Houston, Houston, Texas 77204-5001, USA.
E-mail addresses: zhengdong.zhang@yale.edu (Z.D. Zhang), fox@uh.edu (G.E. Fox).

0167-7012/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.mimet.2008.10.010

to determine whether particular variants belong to the neutral network, their phenotype was examined and classified using two measures. They are the fitness, as determined by comparing the relative growth rate of each mutant 5S rRNA when competed against an otherwise identical strain carrying a wild-type 5S rRNA, and the accumulation level of the mutant 5S rRNA in both the cytoplasm and the 70S ribosomes. Using this approach most variants can be classified into one of three categories. These are: high accumulation in the cells with high incorporation into 70S ribosome (Type I), high accumulation in cells with no accumulation into ribosomes (Type II), and no accumulation in cells or ribosomes (Type III). Most mutations had little or no effect on growth rate regardless of category. Occasional deleterious mutants were observed in all three categories with the most extreme examples in variants that are otherwise Type II. Excluding the deleterious mutants, the Type-I variants are considered to belong to the neutral network, the Type-II do not and the Type-III are uncertain but probably do not in most cases. This method is theoretically well established and also experimentally well tested (Lee et al., 1993, 1997; Zhang et al., 2003). However, because this approach is labor-intense and time-consuming, it is inappropriate for a large scale exploration of the neutral network and the sequence space.

Previous studies on 5S rRNA gene numbers in *Escherichia coli* have shown that the loss of multiple genes through sequential genomic deletion decreases cell fitness and increases cell doubling time (Ammons and Rampersad, 2001; Ammons et al., 1999). The deletion of five of the eight 5S rRNA genes in the *E. coli* genome results in a ~60% increase in the doubling time. The same studies also showed that a functional plasmid-borne 5S rRNA gene was capable of compensating for the deleted 5S rRNA genes. Thus, the success or failure of a plasmid-borne 5S rRNA to compensate for the reduced level of 5S rRNA in such a partial knockout strain should depend on the validity of the 5S rRNA gene that is carried. One would therefore expect a valid 5S rRNA (e.g. a Type-I variant) to compensate for the deleted genomic 5S rRNA genes while an invalid sequence should not compensate as much or not at all (Ammons and Rampersad, 2001).

In the work described here, we show that the ability of plasmid-borne 5S rRNAs to compensate the partial knockout strain can in fact be used to rapidly and reliably determine the validity/invalidity of random 5S rRNA variants by simply examining the growth rate and sequencing the plasmid insert to identify the variant. The approach was validated by using it to reexamine 15 representative variants belonging to each of the three categories described above. In addition, a small set of new variants were characterized. The system based on the partial knockout strain is a substantial improvement because measurement of growth rate is much easier and faster than the 5S rRNA quantification and growth competition assays used previously (Lee, D'souza and Fox, 1993). The new methodology also offers an advantage with regard to *in vitro* selection methods because one explicitly recovers both good and bad variants.

2. Materials and methods

2.1. pCV251

Expression vector pCV251 (Hedenstierna et al., 1993) carries the *V. proteolyticus* 5S rRNA gene including the *E. coli* promoters and terminator. This plasmid was used for the expression of *V. proteolyticus* 5S rRNA in *E. coli* HB101-MO400 cells (See Table 1 in the Appendix) in order to study the incorporation of variant 5S rRNA in the ribosomes and in the cytoplasm.

2.2. Construction of pCV5S

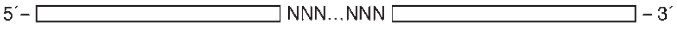
Compensatory plasmid pC5S (Ammons et al., 1999) which harbors an active *E. coli* 5S rRNA gene, was constructed from plasmid pCL1920 (Lerner and Inouye, 1990) and pKK5-1 (Brosius, 1984; Lerner and Inouye, 1990; Szeberenyi and Apirion, 1984). This plasmid was created in order to study the effects of deleting 5S rRNA genes from the *E. coli* strain EMG2 genome. In order to study the phenotypes of *V. proteolyticus* 5S rRNA mutants, a compensatory plasmid, which has an active *V. proteolyticus* 5S rRNA gene instead of the *E. coli* 5S rRNA, was needed. This new plasmid pCV5S was constructed from pC5S via mutagenesis. *V. proteolyticus* and *E. coli* 5S rRNA each has 120 nucleotides, but differs at 19 different positions. In order to construct pCV5S (See Table 2 in the Appendix), the bases at these 19 positions were changed cumulatively in three steps using the QuikChange Site-Directed Mutagenesis Kit. Thus, the resulting plasmid, pCV5S, has the desired sequence and features.

2.3. Construction of deletion strain EMG2ΔBDHG

In *E. coli*, there are seven rRNA operons (operons A, B, C, D, E, G, and H). All these operons contain one gene for each of the rRNAs except for operon D, which contains two 5S rRNA genes. Operons B, D, H, and G, which are scattered in the *E. coli* genome, were deleted to construct the 5S rRNA-deficient *E. coli* strain used in our study. As described previously (Ammons and Rampersad, 2001; Ammons et al., 1999), the construction of the deletion strain began with EMG2ΔBDH. Confirmation of the EMG2ΔBDHG genotype was obtained from observation of change in growth rate and PCR using the operon-specific primer (5'-agt tga gtc gtc gtg ttc cc-3'). Relative growth was ascertained by streaking deletion strains EMG2ΔBDH, and EMG2ΔBDHG to single colonies on the same YT plate and incubating at 37 °C. The five gene deletion strain used here was preferable to the four gene deletion strain because its growth is more significantly impacted.

2.4. Site-directed and random mutagenesis

The QuikChange Site-Directed Mutagenesis Kit from Stratagene Cloning Systems (La Jolla, CA) was used to introduce site-directed mutations into the *V. proteolyticus* 5S rRNA gene. In this study, random mutagenesis was employed to rapidly generate sequence variants. To make random mutations, the standard protocol for the QuikChange Site-Directed Mutagenesis Kit was modified to make use of pairs of degenerate oligonucleotides. The region in the 5S rRNA to be randomly mutated is surrounded on each side by 15 nucleotides, which form base pairs with the 5S rRNA template. The degenerate oligonucleotide usually has the form (N is either A, G, C, or T; the box represents the base-pairing region):



The randomized regions in the oligonucleotide can be commercially synthesized to lengths that considerably greater than the practical limitation set by the experimental approach. If the length of the randomized region is n, theoretically there will be (4^n-1) mutagenic oligonucleotides in the pool of synthesized oligonucleotides. For a 5-nucleotide randomized region, 1023 (4^5-1) different mutations in the 5S rRNA could be made in one step. However, this number of colonies cannot be readily accommodated on one 100×15 mm Petri dish and hence not all variants are expected to be present. The stability of each 5S rRNA mutant was assessed by sequencing plasmid-borne 5S rDNAs, harvested from transfected *E. coli* culture. Only the mutant nucleotide at the intended position was seen unambiguously.

2.5. Growth rate determination

In all cases, 60 ml of YT medium were inoculated with 100 μl of an overnight (16 h) culture also grown in YT medium and bacterial growth rates were determined by monitoring the optical density of the culture at 610 nm (OD_{610}) with respect to time. Strain EMG2 (wild-type, See Table 1 in the Appendix), EMG2ΔBDHG (knockout), and EMG2ΔBDHG-pCV5S

Table 1
The phenotypes of fifteen variants of *V. proteolyticus* 5S rRNA, which were chosen to test the partial knockout variant system

Mutation	Growth rate difference (min^{-1})[a]	% Incorporation			
		ΣRNA[b]	70S	50S[c]	ΣPOLY[d]
Type I: high accumulation and high incorporation (potentially valid)					
U35C	0.041	45	47	47	46
U36C	0.019	>50	32	29	26
G41U	Ne[e]	>50	>50	>50	>50
U52A	Ne	>50	>50	>50	>50
A62U	0.069	22	21	20	19
Type II: high accumulation and low incorporation (not incorporated)					
G44C	−0.208	>50	tr	tr	tr
A53G	−0.029	>50	tr	tr	tr
U55A	−0.114	42	tr	tr	tr
A58C	−0.038	34	tr	tr	tr
U64A	0.047	23	tr	tr	tr
Type III: low accumulation and low incorporation (not stable)					
C17U	0.062	6	7	7	8
A19U	0.021	tr[e]	nd[e]	nd	nd
U21A	−0.024	5	tr	tr	tr
G44U	−0.099	6	tr	tr	tr
C70U	0.045	9	8	6	10

[a] Growth rate difference between the mutant and the wild type.
[b] ΣRNA: total cellular RNA (cytosolic RNA and ribosomal RNA).
[c] The fraction of unassembled large ribosomal subunits.
[d] ΣPOLY: polysomal pool (i.e. 2x, 3x, 4x). An mRNA can be translated by several ribosomes simultaneously and this forms a polyribosome. In a typical spectrum of ribosomal fractions, we see substantial fractions of polysomes with up to four (4x) ribosomes. The polysomal pool is composed of these fractions. The ΣPOLY column gives the extended measurement of the incorporation of 5S rRNA into ribosomes.
[e] Abbreviations: ne, no effect; nd, not detected; tr, trace (<5%).

(the knockout strain transformed with plasmid pCV5S harboring wild-type or mutant *V. proteolyticus* 5S rRNA gene) were grown at 37 °C and ~200 rpm in an incubator shaker (New Brunswick Scientific Co., Edison, NJ). One-milliliter samples were removed every ~20 min for OD_{610} determination. The exponential phase of the growth curve is specified by

$$N_t = N_{t_0} \times 2^{(t-t_0)/\tau} \quad (1)$$

where N_t is the cell number at time t, N_{t0} the cell number at time t_0, and τ the doubling time of the culture. The doubling time τ and the growth rate r—the slope of the regression line for the exponential portion of the growth curve of $\log_{10}(OD_{610})$ versus time—have the relation: $\tau \cdot r = \log_{10}2$ (Monod, 1949).

2.6. 70S ribosome extraction

70S ribosomes were extracted using the method of Staehelin *et al.* (1969) with modifications. In all cases, *E. coli* HB101-MO400 cells (wild type or the knockout strain transformed with plasmid pCV251 harboring mutant *V. proteolyticus* 5S rRNA gene) in log phase was thawed and suspended in equal volume of Buffer A (10 mM $MgCl_2$/ 10 mM Tris–HCl, pH7.3/50 mM NH_4Cl/0.1 mM Na_2EDTA/6 mM 2-mercaptoethanol) and lysed with lysozyme (1 mg/ml). The mixture was incubated a 4 °C and ~150 rpm for 20 min. Sodium deoxycholate (0.1%) and DNAse I (0.5 μg/ml) was added to the mixture and incubated at 4 °C and ~150 rpm for 5 min. Cell debris was removed by centrifuging at 23,000 g for 2 h and ribosomes were pelleted from the cell extract by centrifugation for 7 h at 105,000 g. The pellet was resuspended in Buffer A and stored at −80 °C.

2.7. RNA extraction

Total RNA from whole cells and ribosomal RNA from the ribosomes was extracted (Chomczynski and Sacchi, 1987). Cells or ribosomes were resuspended in 0.5 M sodium acetate pH5, followed by trizol

Table 2
Growth rates and doubling time of the knockout strain compensated with fifteen *V. proteolyticus* 5S rRNA variants, whose phenotypes have been determined

Geno-type[a]	WT	Δ	Δ-*p*	Type I variants					Type II variants					Type III variants				
				U35C[b]	U36C	G41U	U52A	A62U	G44C	A53G	U55A	A58C	U64A	C17U	A19U	U21A	G44U	C70U
Growth rate (r, in min^{-1})																		
1°	0.0134	0.0054	0.0107	0.0095	0.0100	0.0109	0.0103	0.0090	0.0063	0.0071	0.0063	0.0063	0.0073	0.0077	0.0082	0.0081	0.0081	0.0082
2°	0.0129	0.0067	0.0096	0.0098	0.0099	0.0100	0.0098	0.0093	0.0058	0.0071	0.0065	0.0063	0.0065	0.0077	0.0086	0.0081	0.0080	0.0085
3°	0.0130	0.0046	0.0094	0.0090	0.0108	0.0100	0.0094	0.0091	0.0063	0.0069	0.0070	0.0078	0.0064	0.0080	0.0084	0.0085	0.0070	0.0089
Doubling time (τ, in min)																		
1°	22.46	55.75	28.13	31.69	30.10	27.62	29.23	33.45	47.78	42.40	47.78	47.78	41.23	39.09	36.71	37.16	37.16	36.71
2°	23.34	44.93	31.36	30.72	30.41	30.10	30.72	32.37	51.90	42.40	46.31	47.78	46.31	39.09	35.00	37.16	37.63	35.42
3°	23.16	65.44	32.02	33.45	27.87	30.10	32.02	33.08	47.78	43.63	43.00	38.59	47.04	37.63	35.84	35.42	43.00	33.82
Ave.[c]	22.99	55.37	30.51	31.95	29.46	29.27	30.66	32.97	49.16	42.81	45.70	44.72	44.86	38.61	35.85	36.58	39.27	35.32
Max-Δ.[d]	0.350	10.07	1.520	1.500	0.950	0.830	1.370	0.480	2.750	0.820	2.080	3.060	2.170	0.490	0.860	0.580	3.740	1.390
Δ.-Min[e]	0.520	10.44	2.370	1.230	1.590	1.660	1.430	0.600	1.370	0.410	2.700	6.130	3.620	0.980	0.850	1.170	2.100	1.490

[a] Type I, high accumulation and high incorporation. Type II, high accumulation and low incorporation. Type III, low accumulation and low incorporation. WT, wild type strain EMG2 of *E. coli*. Δ, knockout strain EMG2ΔBDHG. Δ-*p*, knockout strain EMG2ΔBDHG compensated with the plasmid pCV5S.
[b] The nature of the variant is identified by its original base, its position in 5S rRNA, and the new base. Thus, U35C is a variant in which U at position 35 is changed to C.
[c] Average of the three measurements. Average=(1°+2°+3°)/3.
[d] Distance between the average and the maximum of the three measurements, calculated by Max(1°, 2°, 3°)−Ave.
[e] Distance between the minimum of the three measurements and the average, calculated by Ave.−Min(1°, 2°, 3°).

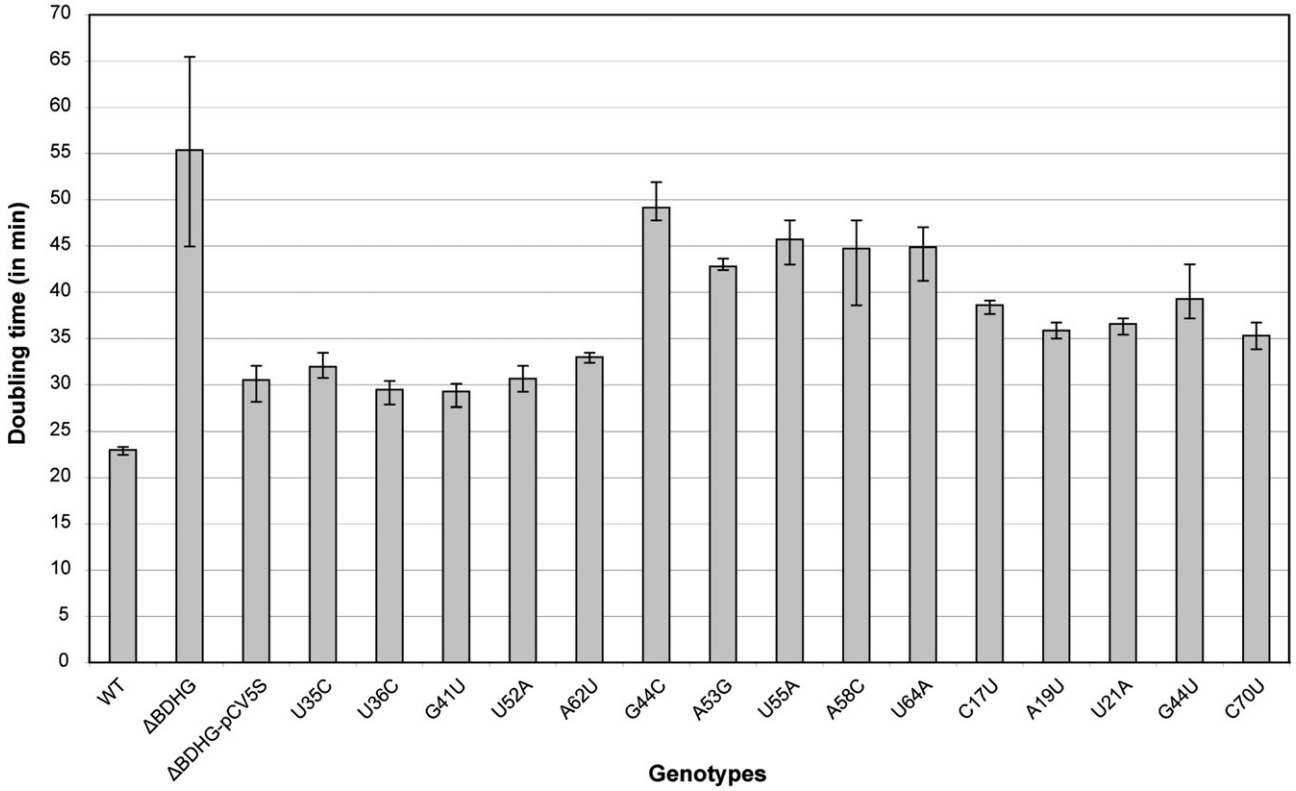


Fig. 1. Average doubling time of the knockout strain EMG2ΔBDHG compensated with fifteen previously categorized *V. proteolyticus* 5S rRNA variants. The range of the maximum and the minimum measurements is shown as the error bar at the top of each column.

reagent and chloroform. The mixture was centrifuged for 15 min at 6000 rpm to separate the chloroform layer and aqueous layer. The aqueous layer that contains the RNA was separated from the chloroform layer and RNA precipitated using isopropanol and 0.5 M NaOAc by incubating at −80 °C for 24 h. Precipitated RNA was collected by centrifugation at 5000 rpm for 15 min without thawing.

2.8. 5S rRNA incorporation studies

The ratio of variant *V. proteolyticus* 5S rRNA to total 5S rRNA in the cells and in the 70S ribosomes was determined as described else where (Ammons and Rampersad, 2001). In brief, membrane bound RNA was hybridized to a probe which is complementary to the variant *Vibrio* 5S rRNA sequence (HV2: gtc caa atc gct atg gtc gc), and an autoradiogram was obtained. The membrane was stripped and rehybridized with a probe complementary to *E. coli* 5S rRNA (HE2: gac cac cgc gct act gcc gc) (Sambrook and Russel, 2001) and a second autoradiogram was obtained. Autoradiograms of samples containing known ratios of *V. proteolyticus* 5S rRNA to *E. coli* 5S rRNA was also made. The imaging plates were read using image Ready V1.4E software package and image data analyzed using Image Gauge V3.0 on a Power Macintosh G3.

3. Results

3.1. The new variant-phenotyping system can determine the genotypes of the 5S rRNA mutations

Before the partial knockout compensation system can be employed to classify large numbers of 5S rRNA variants, its accuracy needed to be confirmed and a basis for comparison with earlier methodology established. This important step will make it possible to link the phenotypes obtained using partial knockout compensation with results obtained using site-directed mutagenesis. For this verification step, fifteen variants of *V. proteolyticus* 5S rRNA were selected. This included five variants (Table 1) from each of the three major phenotypes that were observed previously (Ammons et al., 1999; Lee et al., 1997; Tucker et al., 2005). The growth rate of each variant was measured in triplicate with three control strains, namely EMG2 (the wild type), EMG2ΔBDHG (the knockout strain), and EMG2ΔBDHG–pCV5S (the knockout compensated by plasmid pCV5S).

The results are detailed in Table 2. The wild type EMG2 strain has a typical *E. coli* doubling time of 23 min. With five 5S rRNA genes deleted from its genome, the knockout strain EMG2ΔBDHG grew much more slowly under the same conditions—its doubling time rises to 55 min, a 250% increase. Compensation of 5S rRNA loss by the expression of a 5S rRNA gene from plasmid pCV5S decreases the doubling time of EMG2ΔBDHG to 30 min. The growth rates and the doubling times for each variant are also listed in Table 2. A column is included in Table 2 for each variant. The nature of the variant is identified by its original base followed by its position in 5S rRNA followed by the new base. Thus, U35C is a variant in which U at position 35 is changed to C. For each variant, the average, as well as the range of the fluctuation among the three measurements, of the doubling time of each strain were also calculated and included in the table. The bar graph in Fig. 1 gives a quick and intuitive comparison of the doubling time among the eighteen strains (3 controls + 15 tests). As Fig. 1 shows, the *V. proteolyticus* 5S rRNA variants of Type I, II, and III each demonstrate distinguishable effects on the growth of EMG2ΔBDHG. The average doubling time of EMG2ΔBDHG compensated with Type-I (U35C–A62U), Type-II (G44C–U64A) or Type-III (C17U ~ C70U) 5S rRNA variants is 31 min, 45 min or 37 min respectively.

3.2. The phenotype of a variant can be frequently be recognized by colony size

Measurement of the growth rate of a 5S rRNA variant in the knockout system is the quantitative, accurate, and thus the definitive

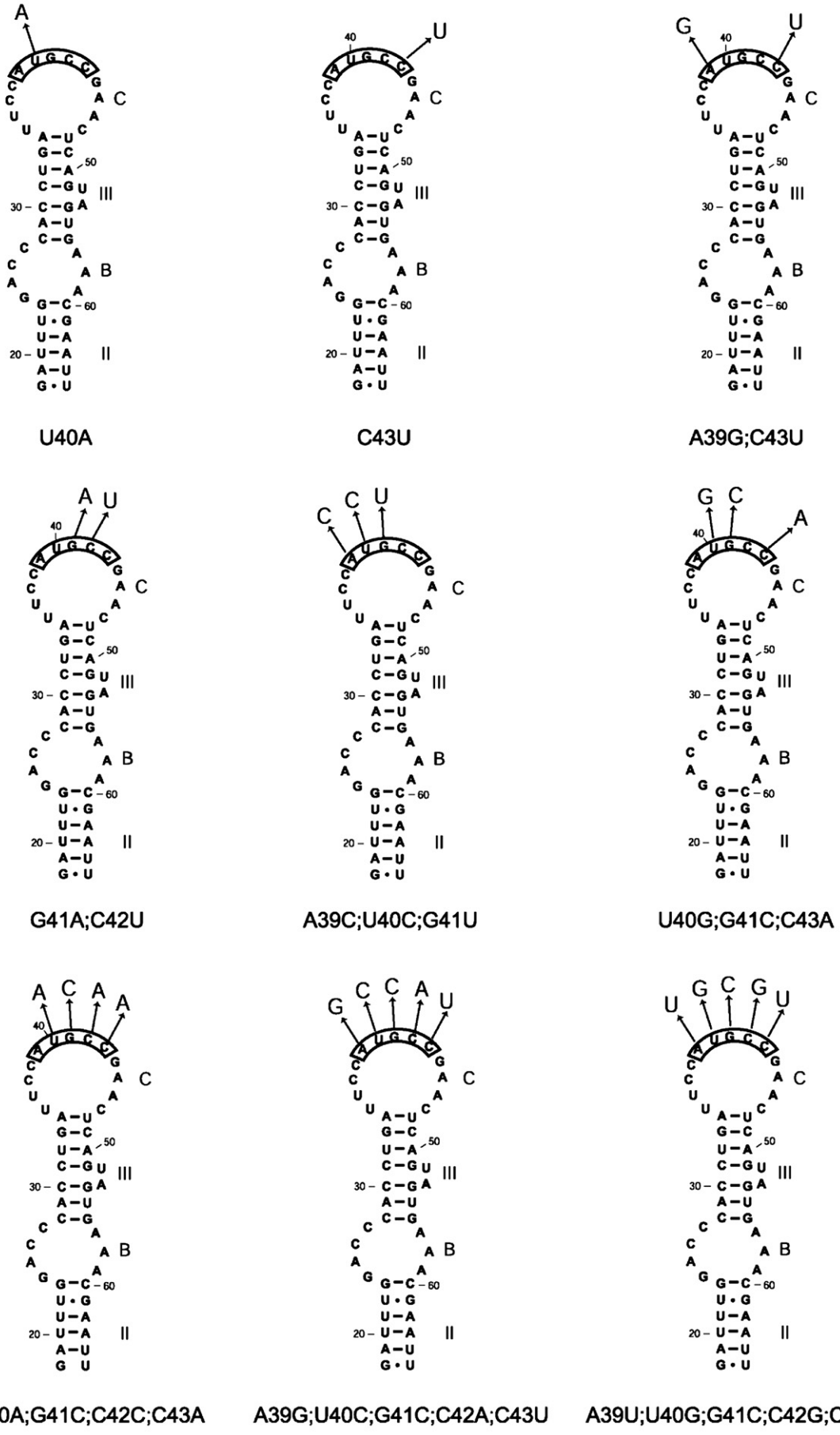


Fig. 2. Location of nine random mutations made between positions 39 and 43 in the C loop of *V. proteolyticus* 5S rRNA. The randomized region is boxed.

Table 3
The doubling time and growth rates of nine randomly created mutations

Genotype	r[a]	τ[b]	Phenotype[c]
EMG2	0.0134	22.46	Wild type
EMG2ΔBDHG	0.0054	56.74	Knockout
EMG2ΔBDHG – pCV5S	0.0097	31.03	Compensated
U40A	0.0102	29.51	I
C43U	0.0096	31.36	I
A39G;C43U	0.0098	30.72	I
G41A;C42U	0.0078	38.59	III
A39C;U40C;G41U	0.0066	45.61	II
U40G;G41C;C43A	0.0079	38.11	III
U40A;G41C;C42C;C43A	0.0066	45.61	II
A39G;U40C;G41C;C42A;C43U	0.0065	46.31	II
A39U;U40G;G41C;C42G;C43U	0.0064	47.34	II

Each of the variants is assigned to one of the three primary phenotypes based on this information.
[a] : growth rate in min^{-1}..
[b] τ: doubling time in min.
[c] Major phenotypes: I, potentially valid; II, not incorporated; III, not stable.

way to reveal its phenotype. However, when a large pool of 5S rRNA variants generated by random mutagenesis need to be screened and categorized, a method that can rapidly distinguish variants of different types will allow one to readily select variants likely to have desired properties. The ability to conduct such pre-screening of variants dramatically decreases the number of growth studies needed to fully characterize variants.

When bacteria are grown in liquid media the difference in the growth rates of different strains is reflected by the speed at which the turbidity of the media increases. This relationship is the foundation of the methods used to measure growth rate. When bacteria are grown on solid media agar plates after synchronization, the growth rate difference is marked by the colony size. (The one-side Wilcoxon rank sum test shows the increase in the growth rate of the 5S rRNA mutants from large colonies over the ones from small colonies is statistically significant with the test statistic $W = 415.5$ and p-value $= 1.678 \times 10^{-13}$.) To test the ability of the new system to distinguish 5S rRNA variants of different types by their colony size, equal amounts of the knockout strain EMG2ΔBDHG , the knock out strain with wild type pCV5S and previously studied Type- I, II, and III variants were mixed in equal amounts and plated on YT agar plates. After 16 h of incubation at 37 °C several colonies were picked and the plasmid DNA sequenced. It was observed that the colony size has direct correlation with the doubling time and the type of mutation. Later, the knockout strain EMG2ΔBDHG was transformed with a mixture of equal amount of all 15 pCV5S variants listed in Table 1. After 15 h incubation at 37 °C, 20 small and 20 big colonies were picked. The plasmid, pCV5S, from these 40 colonies was extracted and the variant 5S rDNA was sequenced in each case. Of the fifteen variants added, 14 were recovered at least once among the 40 samples. 75% of the small colonies and 80% of the large colonies were identified as Type-II/III and Type-I *V. proteolyticus* 5S rRNA variants respectively.

3.3. Novel random 5S rRNA mutants have been characterized by the new phenotyping method

In the illustrative study reported here, positions 39 through 43 in the loop C region of *V. proteolyticus* 5S rRNA were randomized. Approximately 600 colonies appeared on the LB/streptomycin agar plate after 50 μl of XL1-Blue competent cells transformed with 1 μl mutagenesis reaction product and incubated over night. Thirty of these were selected and the 5S rRNA genes carried by plasmid pCV5S were sequenced. Nine (35% with 4 unreadable sequencing results discarded) had mutations in the target region, which varied from single mutations to quintuple mutations (Fig. 2).

The growth rates of the nine random mutations were measured in one set of experiments, which included the three control strains: the wild type (EMG2), the 5S knockout (EMG2ΔBDHG), and the knockout compensated by plasmid pCV5S (EMG2ΔBDHG-pCV5S). The results are listed in Table 3 and the phenotype of each mutant was derived from the comparison between its doubling time and the characteristic doubling times of the known Type- I, II and III variants as listed in Table 2. Of the nine mutations, three have a doubling time of approximately 30 min and thus are classified as Type-I variants. The remaining six are either Type-II or Type-III variants. The three Type-I variants include two single and one double mutation (Table 3). Fortuitously, one of these, C43U, was also previously categorized as a Type-I variant (high accumulation and high incorporation), which confirms the result of the previous accumulation/incorporation study conducted on this mutation (Zhang et al., 2003).

4. Discussion

This study tested the hypothesis that the partial 5S rRNA knockout strain EMG2ΔBDHG can be used to rapidly explore the *Vibrio* region of 5S rRNA sequence space *in vivo*. The stability of each 5S rRNA mutant was assessed by sequencing plasmid-bourn 5S rDNAs, harvested from transfected *E. coli* cultures. In each case, only the intended mutant nucleotide was unambiguously seen. The approach was validated using 15 5S rRNA variants previously shown to exhibit one of three distinct phenotypes. The partial knockout strain was transformed with a plasmid carrying each variant and the effect on growth rate was determined. The results show that the three types of variants have characteristic growth rates and doubling times in the partial knockout strain and thus are distinguishable. The observation that the average doubling time of EMG2ΔBDHG compensated with Type-I variants is essentially identical to that of EMG2ΔBDHG compensated with the wild type *V. proteolyticus* 5S rRNA show that the Type-I variants and wild type 5S rRNA behave similarly in the cell. Type-II variants have the slowest growth rates. The Type-III variants exhibit intermediate growth rates confirming the speculation that Type-III 5S rRNA variants are frequently partially functional and thus comparatively less deleterious.

Above all, it has been demonstrated that the partial knockout system is capable of readily distinguishing the Type-I and the non-Type-I variants. This distinction is essential to ongoing efforts to understand how allowable (Type-I phenotype) sequences are distributed in the 5S rRNA neutral network. It was also shown that simple observation of colony size can be used to quickly enrich for variants belonging or not belonging to the Type-I phenotype. Once validated with the mutants of known types, the approach was used to examine variants in the loop C region of 5S rRNA were identified by sequencing and subsequently classified using the new system. An alternative to the partial knockout system utilized here would be the use of a strain in which all the 5S rRNA genes are deleted as was done previously for the larger RNA (Asai et al., 1999). However, in such a system only variants that were functional could be characterized. Thus, if a particular mutation did not occur, one would not know whether it was missing from the pool or deleterious. Thus, a major advantage of the partial knockout system described here is its ability to recover both deleterious and viable mutations. Now that the approach has been validated by the use of incorporation data and competitive growth rate studies, one can use it to study other systems, e.g. *E. coli* 5S rRNA variants in E. coli. The validation was much easier using *V. proteolyticus* variants because the sequence differences made it easy to use hybridization to quantitatively examine expression levels of 5S rRNA.

The method of random mutagenesis described here places a limit of approximately 50 nucleotides on the length of the region that can be randomized. This congenital limitation is preset by the usage of a pair of artificially synthesized oligonucleotides. Therefore, this method cannot to be used to obtain random mutations throughout the entire 5S rRNA at present. One solution to this might be to include a unique restriction enzyme digestion site at each end of the 5S rRNA gene in the plasmid and then use random mutagenesis PCR to

randomize the whole 5S rRNA sequence. The restriction sites introduced into the RNA operon in the plasmid for this purpose should not have any deleterious effects on 5S rRNA transcription and downstream processing. This concern is minor, however, as in practice the main limitation is the number of colonies that will be produced.

Acknowledgements

This work was supported in part by grants to GEF from the Robert A. Welch Foundation (E-1451), NASA (NAG5-8140 and its successor NNG05GN75G), and the Institute of Space Systems Operations.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.mimet.2008.10.010.

References

Ammons, D., Rampersad, J., 2001. An *E. coli* 5S rRNA deletion mutant useful for the study of 5S rRNA structure/function relationships. Curr. Microbiol. 43, 89–92.
Ammons, D., Rampersad, J., Fox, G.E., 1999. 5S rRNA gene deletions cause an unexpectedly high fitness loss in *Escherichia coli*. Nucleic Acids Res. 27, 637–642.
Ancel, L.W., Fontana, W., 2000. Plasticity, evolvability, and modularity in RNA. J. Exp. Zool. 288, 242–283.
Asai, T., Condon, C., Voulgaris, J., Zaporojets, D., Shen, B., Al-Omar, M., Squires, C., Squires, C.L., 1999. Construction and initial characterization of *Escherichia coli* strains with few or no intact chromosomal rRNA operons. J. Bacteriol. 181, 3803–3809.
Ban, N., Nissen, P., Hansen, J., Moore, P.B., Steitz, T.A., 2000. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science 289, 905–920.
Brosius, J., 1984. Toxicity of an overproduced foreign gene product in *Escherichia coli* and its use in plasmid vectors for the selection of transcription terminators. Gene 27, 161–172.
Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162, 156–159.
Dinman, J.D., 2005. 5S rRNA: structure and function from head to toe. Int. J. Biomed. Sci. 1, 2–7.
Eigen, M., Winkler-Oswatitsch, R., Dress, A., 1988. Statistical geometry in sequence space: a method of quantitative comparative sequence analysis. Proc. Natl. Acad. Sci. U. S. A. 85, 5913–5917.
Fontana, W., Schuster, P., 1998. Shaping space: the possible and the attainable in RNA genotype-phenotype mapping. J. Theor. Biol. 194, 491–515.
Fox, G.E., 1985. The structure and evolution of archaebacterial RNA. In: Gunsalus, I.C. (Ed.), The bacteria, vol. 8. Academic Press, New York, pp. 267–310.
Hedenstierna, K.O., Lee, Y.H., Yang, Y., Fox, G.E., 1993. A prototype stable RNA identification cassette for monitoring plasmids of genetically engineered micro-organisms. Syst. Appl. Microbiol. 16, 280–286.
Lee, Y.H., D'souza, L., Fox, G.E., 1993. Experimental investigation of an RNA sequence space. Orig. Life Evol. Biosph. 23, 365–372.
Lee, Y.H., D'Souza, L., Fox, G.E., 1997. Equally parsimonious pathways through an RNA sequence space are not equally likely. J. Mol. Evol. 45, 278–284.
Lerner, C.G., Inouye, M., 1990. Low copy number plasmids for regulated low-level expression of cloned genes in *Escherichia coli* with blue/white insert screening capability. Nucleic Acids Res. 18, 4631.
Monod, J., 1949. The growth of bacterial cultures. Ann. Rev. Microbiology 3, 371–394.
Nino, J., 1983. Molecular approach to evolution. Princeton University Press, Princeton, NJ.
Reidys, C., Stadler, P.F., Schuster, P., 1997. Generic properties of combinatory maps: neutral networks of RNA secondary structures. Bull. Math. Biol. 59, 339–397.
Sambrook, J., Russel, D.W., 2001. Molecular cloning. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Staehelin, T., Maglott, D.M., Monro, R.E., 1969. On the catalytic center of peptidyl transfer: a part of the 50 S ribosome structure. Cold. Spring. Harb. Symp. Quant. Biol. 34, 39–48.
Szeberenyi, J., Apirion, D., 1984. Synthesis and processing of 5 S rRNA from an rrnB minigene in a plasmid. Biochim. Biophys. Acta. 783, 15–25.
Szymanski, M., Barciszewska, M.Z., Erdmann, V.A., Barciszewski, J., 2003. 5 S rRNA: structure and interactions. Biochem. J. 371, 641–651.
Tucker, D.L., Karouia, F., Wang, J., Luo, Y., Li, T.B., Willson, R.C., Fofanov, Y., Fox, G.E., 2005. Effect of an artificial RNA marker on gene expression in *Escherichia coli*. Appl. Environ. Microbiol. 71, 4156–4159.
Zhang, Z., D'Souza, L.M., Lee, Y.H., Fox, G.E., 2003. Common 5S rRNA variants are likely to be accepted in many sequence contexts. J. Mol. Evol. 56, 69–76.