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

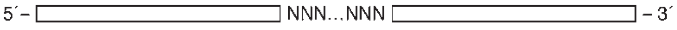
Compensatory plasmid pC5S (Ammons et al., 1999) which harbors an active *E. coli* 5S rRNA gene, was constructed from plasmid pCL1920 (Lerner and Inouye, 1990) and pKK5-1 (Brosius, 1984; Lerner and Inouye, 1990; Szeberenyi and Apirion, 1984). This plasmid was created in order to study the effects of deleting 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