6

Circular mRNA Encoding for Monomeric and Polymeric Green Fluorescent Protein

Rhonda Perriman

1. Introduction

Many proteins with unusual structural properties are comprised of multiple repeating amino acid sequences, and are often fractious to expression in recombinant systems. To facilitate recombinant production of such proteins for structural and engineering studies, the author has developed a method for producing messenger RNAs on circular RNA templates. This circularization process is derived from a rearranged group I intron, from which circular RNA is produced through the splicing activity of autocatalytic group I RNA elements (**Fig. 1**; ***1,2***). Because the only cofactors required for splicing of the group I intron are magnesium and guanosine, the process can take place in a variety of organisms, making it amenable to a wide variety of protein expression systems ***(1–4)***.

Fig. 1

This chapter details the design and construction of circular mRNAs containing the open reading frame (ORF) encoding for green fluorescent protein (GFP). Included on the circular GFP mRNA constructs are translation initiation sequences designed to recruit either prokaryotic or eukaryotic ribosomes. By removing in-frame stop codons, the author has also designed and tested circular, infinite mRNAs encoding GFP. The circular mRNAs produce extremely long protein chains of polyGFP, demonstrating that both prokaryotic and eukaryotic ribosomes can internally initiate, and repeatedly transit, a circular mRNA ***(3,5)***. The author has also analyzed fluorescence spectra from *E. coli* expressing the monomeric GFP or polyGFP from circular mRNA, and find that only the monomeric forms of GFP are fluorescent. The application of circular mRNA technology may provide a unique means of producing very long repeating sequence proteins (e.g., silks, mollusk shell framework) ***(6–8)***,

From: *Methods in Molecular Biology, vol. 176: Steroid Receptor Methods: Protocols and Assays*
Edited by: B. A. Lieberman © Humana Press Inc., Totowa, NJ

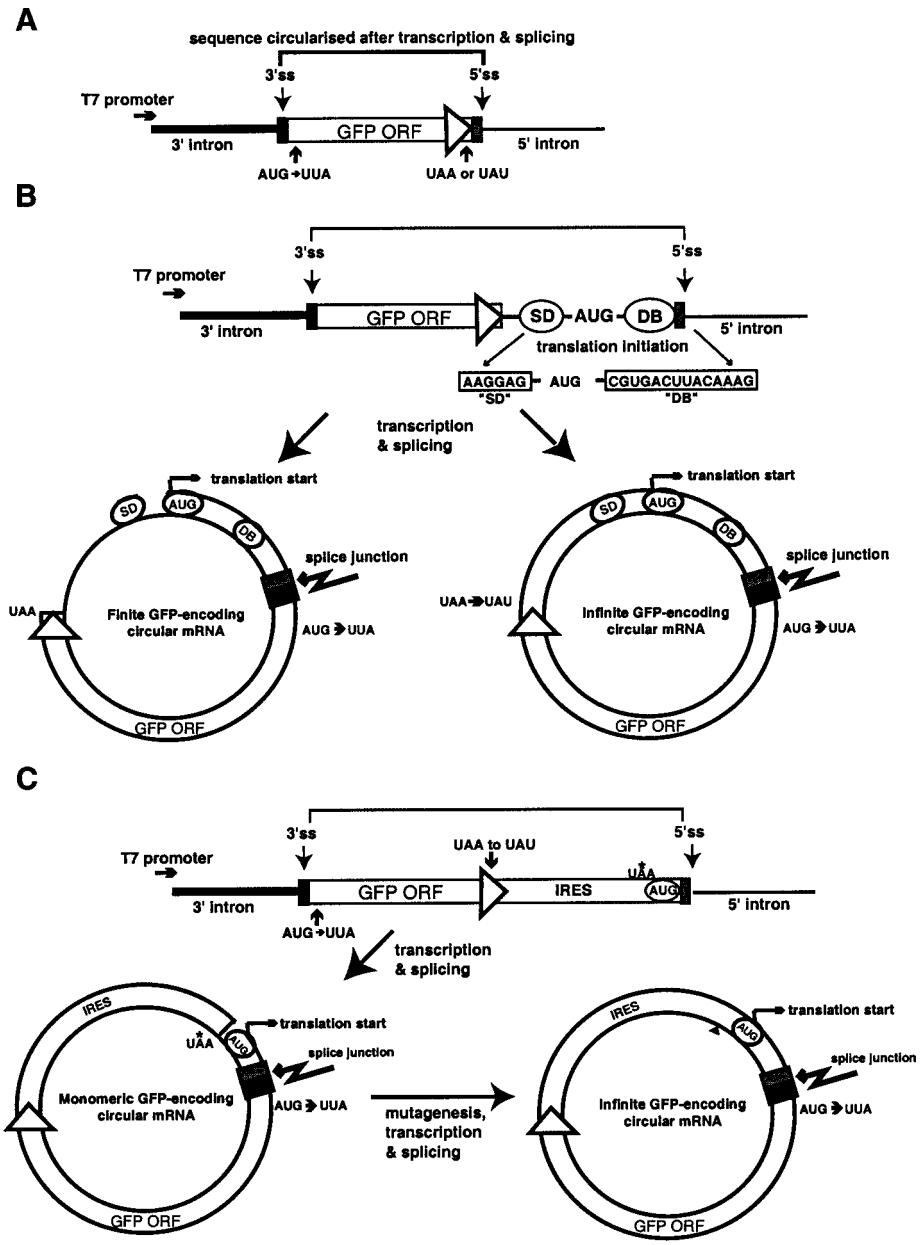

Fig. 1. Design features of plasmids containing rearranged group I intron elements for circular GFP mRNA expression of monomeric GFP or polyGFP in **(B)** *E. coli* or **(C)** rabbit reticulocyte lysates. Transcription and splicing results in circularization of the bracketed sequence, between the 3' (3'ss) and 5' (5'ss) splice sites shown in each figure. **(A)** Relevant region of circular GFP mRNA plasmid containing the GFP ORF.

Job: (MiMB 97–6.0) Hicks (905-6)
Chapter: 6/Perriman
Pub Date: 2001

Operator: MS
Date: 9/5&25/01
Revision: 1st Galley

opening the way for development of proteinaceous materials with novel properties (e.g., *see* **ref. *9*** for review). The methods described below can easily be adapted to the production of any desired protein sequence.

2. Materials

2.1. Making Plasmids for Production of Circular GFP-mRNA

1. Multicopy plasmid vectors (the author uses pBluescript from Stratagene) and restriction enzymes from common vendors were used in buffers as directed by manufacturers.
2. *Escherichia coli* strains: CJ236 (*ung*-, *dut*-) and a standard *rec*- laboratory strain suitable for maintaining and amplifying plasmids (e.g., XL1-Blue from Stratagene).
3. Luria-Bertoni (LB) broth (1 L): 10 g Bacto-tryptone, 10 g NaCl, 5 g yeast extract, pH to 7.6. LB-ampicillin: LB broth containing 100 µg/mL ampicillin. LB agar: LB broth containing 15 g/L agar. LB-amp agar: LB agar containing 100 µg/mL ampicillin.

Au: Notes are renumbered as must cite in order

Site of T7 RNA polymerase promoter sequence used for in vitro and in vivo (*E. coli*) RNA production is indicated. The 3' and 5' group I intron sequences are shown. AUG to UUA is the mutation introduced by the GFP-AUG oligonucleotide, to remove the initiating GFP-AUG (*see* **Subheading 3.1.**). UAA to UAU is the mutation introduced by the GFP-stop oligonucleotide to remove the GFP-stop codon, thus allowing creation of infinite circular ORFs (see **Subheading 3.1.**). **(B)** Circular GFP mRNA plasmid containing the GFP ORF and *E. coli* protein expression cassette. Enlarged, boxed nucleotides are the SD and DB motifs of the translation initiation sequence. Circular species show either the monomeric or polyGFP mRNA created after transcription and splicing. Translated portion of circular mRNAs is shown as double circle. The monomeric GFP-encoding mRNA, created after transcription and splicing, encodes a ~30 kDa GFP species. The polyGFP-encoding mRNA has UAU in place of UAA termination codon, and is devoid of stop codons in the GFP reading frame. The initiating AUG (translation start) and fused 5'ss/3'ss (jagged arrow-splice junction) are shown on the circular mRNA species for both monomeric and polymeric constructs. Other abbreviations are: GFP, green fluorescent protein ORF; SD, Shine-Dalgarno sequence; AUG, initiating codon; DB, downstream box. **(C)** Same as **(B)**, except circular GFP mRNA plasmid containing the GFP ORF and IRES for mammalian protein expression cassette. Linear species shows relevant portions of circular GFP mRNA plasmids including both GFP-AUG and GFP-stop mutations (*see* **Subheading 3.1.**). The circular species show each GFP-encoding circular mRNA after transcription and splicing. The monomeric GFP-encoding mRNA contains a single UAA termination codon, as indicated by *, and encodes ~50 kDa protein. The polyGFP-encoding species has a two base insertion at this position (indicated by triangle), and is devoid of stop codons in the GFP reading frame. “IRES” is the internal ribosome entry sequence required for ribosome recruitment (*see* **Subheading 3.1.4.**).

4. M13KO7 helper phage (Pharmacia).
5. Kanamycin (10 mg/mL in LB-amp).
6. 20% Polyethylene glycol/2.5 *M* NaCl.
7. TES: 50 m*M* Tris-HCl, pH 7.5, 5 m*M* EDTA, 0.5% sodium dodecyl sulfate (SDS).
8. Phenol–chloroform–isoamyl alcohol (24:24:1 v/v/v).
9. 3 *M* sodium acetate, pH 6.0.
10. 70% Ethanol.
11. 1% Agarose gel containing 1 μg/mL ethidium bromide, 10X TBE (1 L): 108 g Tris, 55 g boric acid, 9.3 g EDTA.
12. 5X Annealing buffer: 200 m*M* Tris, pH 7.5, 100 m*M* $MgCl_2$, 250 m*M* NaCl; containing 1 m*M* each of deoxyribonucleoside triphosphate (dNTP)s, 1 m*M* dithiothreitol (DTT), 1 m*M* adenosine triphosphate (ATP), 400 U T4 DNA ligase (New England Biolabs), and 3 U Sequenase™ (U.S. Biochemicals, or other DNA polymerase; *see* **Note 1**).
13. 10X Filling buffer: 100 m*M* Tris, pH 7.5, 50 m*M* $MgCl_2$, 75 m*M* DTT.
14. dNTP mix containing 0.5 m*M* each deoxynucleotide.
15. Klenow DNA polymerase.
16. 0.5 *M* EDTA, pH 7.0.
17. 10X Dephosphorylation buffer: 1 *M* NaCl, 500 m*M* Tris, pH 7.9, 100 m*M* $MgCl_2$, 10 m*M* DTT.
18. Calf intestinal phosphatase.
19. 10X Ligation buffer: 500 m*M* Tris, pH 7.5, 100 m*M* $MgCl_2$, 100 m*M* DTT, 10 m*M* ATP, 0.25mg/mL bovine serum albumin (BSA).
20. Deoxyoligonucleotides: These can be obtained from many commercial sources, or, if available, an in-house facility (*see* **Note 2** for design tips). Oligonucleotides required for circular mRNA constructs in **Subheading 3.1.**
    a. GFP-AUG: alters the initiating AUG of the GFP ORF 5'-GGAGATATAA CCTTAAGTAA AGGAGA-3'.
    b. GFP-stop: alters the GFP stop codon 5'-AACTATACAA ATATTGAGCT CTCATGA-3'.
    c. *E. coli*-Shine-Dalgarno sequence (SD): adds ribosome recruitment signals to the circular mRNA plasmid for expression in *E. coli* 5'ATTGACCTGAGATC GCTTTTTGCTTTGTAAGTCACGTTAGAGCTAGCCATCTTGTGTCTC CTTGTGCAGACCTCTCGAGCTCCAT-3'.
    d. Internal ribosome entry sequence (IRES)-stop: adds two nucleotides to the circular RNA plasmid for polyGFP expression in rabbit reticulocyte lysates 5'-GGGACTAAGGCGGAAT TCTCGAGCTCCATG-3'.

2.2. Testing for Circular mRNA

All reagents for RNA work should be used for this purpose only to minimize contamination from RNase. If desired, H_2O used to prepare these solutions can be treated with diethyl pyrocarbonate (DEPC), by adding 1 mL DEPC/1 L water. Mix and stand, then incubate overnight at 55°C to inactivate DEPC (this is important, because active DEPC can chemically modify ribonucleic acids).

1. rNTPs (10 m*M* each A, C, and G, 1 m*M* U).
2. 5X Transcription buffer: 200 m*M* Tris, pH 7.5, 30 m*M* $MgCl_2$, 10 m*M* spermidine, 50 m*M* NaCl.
3. RNasin (Promega).
4. a-^{32}P-uridine triphosphate (UTP) (3000 Ci/mmol, Amersham).
5. T7 RNA polymerase.
6. RQ1 DNase I (Promega).
7. 4.5% polyacrylamide gel containing 7 *M* urea.
8. Autoradiography film (Kodak).
9. Formamide dye mix: 100% formamide, 1% bromophenol blue, 1% xylene cyanol.

2.3. Testing Protein Expression in E. coli

1. *E. coli* strain BL21-DE3 (or other T7-expressing *E. coli* strain; *see* **Note 3**).
2. Isopropyl-β-D-thiogalactopyranoside (IPTG).
3. 50 m*M* HEPES, pH 7.5.
4. 8 *M* urea/0.1 *M* NaH_2PO_4/10 m*M* Tris, pH 6.
5. Bradford assay kit (Bio-Rad).
6. SDS-PAGE solutions: 30 polyacrylamide: 0.8 *bis*-acrylamide solution (a premade 30% stock can be purchased from Bio-Rad); 4X stacking buffer (100 mL): 6.06 g Tris, 4 mL 10% SDS (pH 6.8); 4X separating buffer (100 mL): 18.17 g Tris, 4 mL 10% SDS (pH 8.8); 10X protein running buffer (1 L): 30 g Tris, 144 g glycine, 10 mL 10% SDS; protein running dye: 2 mL glycerol, 2 mL 10% SDS, 0.25 mg bromophenol blue, 2.5 mL 4X stacking buffer, 0.5 mL β-mercaptoethanol.
7. Western blotting materials: nitrocellulose membrane; BSA; polyclonal GFP antibody (Clontech); immunoglobulin G-goat-antirabbit secondary antibody; 5-bromo-4-chloro-3-indolyl phosphate (BCIP), dissolved at 25 mg/500 µL in dimethyl formamide; nitroblue tetrazolium (dissolved at 25 mg/350 µL in dimethyl formamide + 150 µL water); Western transfer buffer (4 L): 9.6 g Tris, 45.2 g glycine, 800 mL methanol, to 4 L with H_2O; TBS-T (500 mL): 0.45 g NaCl, 50 mL, 1 *M* Tris, pH 7.5, 500 µL Tween; Western substrate buffer (10 mL): 250 µL, 4 *M* NaCl, 50 mL 1 *M* $MgCl_2$, 1 *M* Tris-HCl, pH 9.5.
8. Autoradiography materials: ^{35}S-methionine (Amersham); Flexi™ rabbit reticulocyte lysate in vitro translation kit (Promega); 10 mg/mL RNase A solution; film (Kodak).

3. Methods

3.1. Making Plasmids for Production of Circular GFP-mRNA

A general-purpose plasmid, containing self-splicing introns, which produces circular mRNA, must be created (**Fig. 1**; ***1,3***). The ORF encoding GFP ***(10)***, is inserted, as an end-filled *BstB*1-*Sac*I fragment, into an end-filled *Nco*I-*Sac*I-digested circular mRNA plasmid (**Fig. 1A**). This plasmid can be used after subsequent engineering to allow expression of either monomeric or poly GFP in either eukaryotic or prokaryotic systems (*see* **Note 4**). The translation initia-

Job: (MiMB 97–6.0) Hicks (905-6)
Chapter: 6/Perriman
Pub Date: 2001

Operator: MS
Date: 9/5&25/01
Revision: 1st Galley

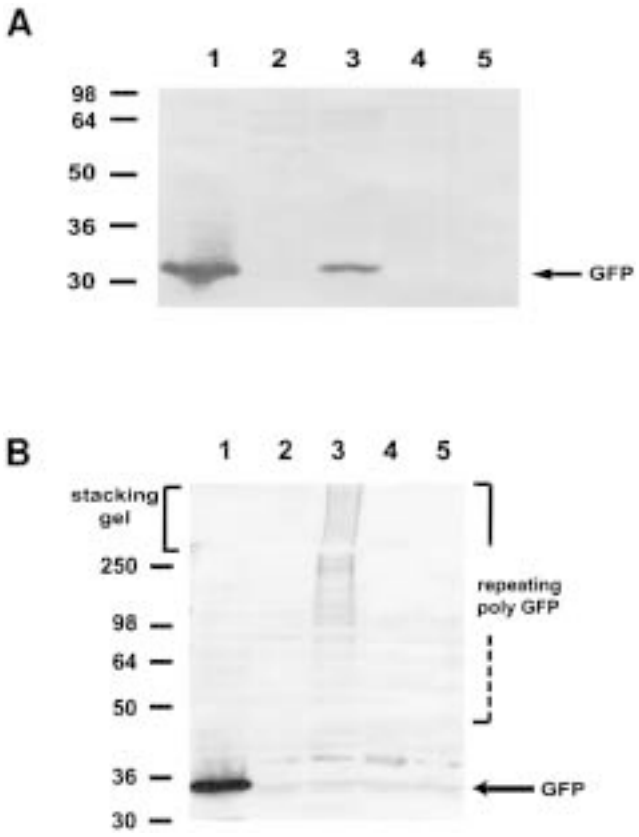

Fig. 2. Translation of GFP from circular mRNAs in *E. coli*. Western blot for GFP expression from strains expressing various monomeric **(A)** and polymeric **(B)** GFP-encoding circular mRNAs, and linear control mRNA. **(A)** Lane 1, linear GFP mRNA control; lane 2, vector control; lane 3, monomeric GFP-encoding circular mRNA; lane 4, monomeric GFP-encoding circular mRNA without translation initiation sequence; lane 5, monomeric GFP-encoding lacking the 5' half group I intron element. Arrowed "GFP" is protein species. Protein size markers are indicated. **(B)** As for (A), except that infinite-encoding circular mRNAs are being expressed.

tion sequences are inserted in-frame with the GFP ORF (**Fig. 1**), so that continuous translation will be possible, when the GFP stop codon is removed. For the prokaryotes, this produces a circular mRNA transcript containing 795 nucleotides, from which a ~30-kDa monomeric GFP (**Fig. 1B** and **Fig. 2**) or polyGFP (the GFP stop codon is mutated to code for isoleucine) can be expressed. For eukaryotic expression using rabbit reticulocyte lysate, the DNA construct produces a circular mRNA transcript containing 1311 nucleotides

Fig. 2

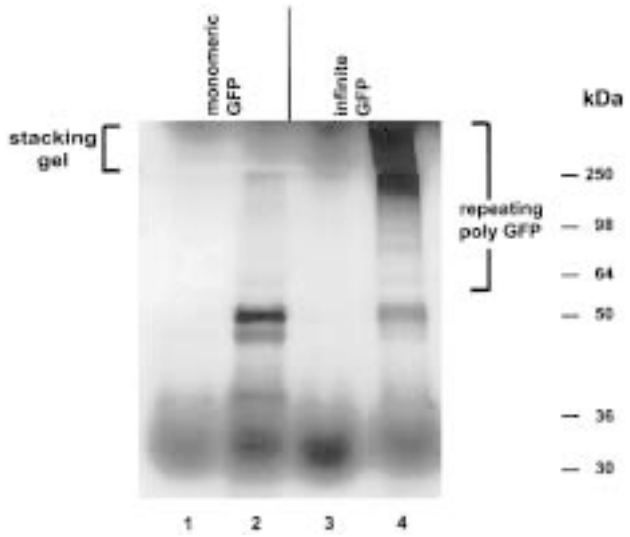

Fig. 3. Translation of GFP from circular mRNAs in rabbit reticulocyte lysates. Autoradiograph of in vitro translation products from monomeric (lanes 1 and 2) and polymeric (lanes 3 and 4) GFP-encoding circular RNA plasmids. Lane 1, monomeric GFP-encoding RNA lacking the 5' half intron; lane 2, monomeric GFP-encoding circular mRNA; lane 3, as for lane 1 except the polyGFP-encoding construct was translated; lane 4, polyGFP-encoding circular mRNA (as is the case in *E. coli*, a heterogeneous collection of proteins is seen, ranging in size from ~50 to >300 kDa).

encoding for a ~50-kDa monomeric GFP, or a polyGFP (after mutation of the GFP stop codon) (**Fig. 1C** and **Fig. 3**).

Fig. 3

3.1.1. Site-Directed Mutagenesis of GFP Start and Stop Codons

Prior to addition of ribosome entry signals, site directed mutagenesis is used to create two sequence alterations in the circular GFP mRNA plasmid. The first mutation (GFP-AUG) changes the existing initiating GFP-AUG, and is necessary to ensure that no translation of the GFP ORF occurs from unspliced linear RNAs. The author changed the AUG to UUA, which encodes for leucine. This construct is used to make monomeric GFP in *E. coli*, and it acts as a template for the second mutation. The second mutation (GFP-stop) alters the GFP stop codon, to code for isoleucine (UAA to AUU).

1. Transform CJ236 (*ung*$^-$, *dut*$^-$) *E. coli*-competent cells with cylase-GFP plasmid. Select transformed colonies by spreading dilutions of transformed cells onto LB-amp agar plates after incubating overnight at 37°C.
2. Pick a single colony, and inoculate 2 mL LB-amp in a sterile test tube. Shake at 37°C for >6 h.

3. Transfer 60 µL to a fresh 3-mL aliquot of LB-amp liquid in a new tube, then add 2×10^7 pfu M13KO7 helper phage, and shake at 37°C for 2 h. Add kanamycin to 10 µg/mL, and continue shaking overnight at 37°C.
4. Pellet the cells, and remove 2 mL supernatant. Add 0.54 mL 20% polyethylene glycol/2.5 *M* NaCl solution to the supernatant. Incubate for 15 min at room temperature, then spin at 10,000*g* for 10 min to pellet phage particles.
5. Resuspend pellet in 100 µL TES. Add an equal volume of phenol–chloroform–isoamyl alcohol, then vortex to mix. Spin at 10,000*g* for 5 min, to separate phases. Transfer top phase to a new tube and repeat the phenol–chloroform–isoamyl alcohol extraction.
6. Add 0.1 vol 3 *M* NaAc (pH 6.0) and 2.5 vol ethanol, to precipitate single-stranded DNA. Spin at 10,000*g* for 10 min. Remove the supernatant, and wash the pellet with 70% ethanol (be careful not to disrupt pellet), and dry remaining solvent under vacuum. Resuspend pellet in 30 µL water. About 3 µL of this single-stranded DNA solution should be enough to use as template for site-directed mutagenesis (as a guide, 3 µL should be easily visible by ethidum bromide staining, after electrophoresis through a 1% agarose gel).
7. Mix 3 µL single-stranded DNA with 50 ng phosphorylated oligonucleotide containing mutagenic sequence designed to alter GFP-AUG or GFP-stop (*see* **Note 2**) in a total volume of 10 µL. Incubate at 65°C for 10 min.
8. Add 10 µL 5X annealing buffer containing 1 m*M* each dNTP, 1 m*M* DTT, 1 m*M* ATP, 400 U (New England Biolabs units; *see* Materials) T4 DNA ligase, and 3 U DNA polymerase (*see* **Note 1**). Incubate at 37°C for 1 h.
9. Transform XL1-Blue (or equivalent *rec*$^-$ lab strain) *E. coli* with 0.1 vol mutagenesis reaction from **step 8**. Select for transformed bacteria on LB-amp agar after overnight incubation at 37°C.
10. Isolate plasmid DNA by alkaline lysis from six or more cultures grown in liquid LB-amp, and analyze for mutant sequences by restriction digest and DNA sequencing (*see* **Note 5**).

3.1.2. Mutagenesis of Plasmids for Monomeric GFP Expression in E. coli

1. A sequence designed for translation initiation in *E. coli* is inserted, by site-directed mutagenesis (*see* **Subheading 3.1.1.**), into the circular mRNA construct from **Subheading 3.1.1.** in which only the GFP-AUG is mutated. The oligonucleotide for this is called "*E. coli*-SD" (*see* **Subheading 2.1.** and **Note 1**) and contains an SD ***(10–13)***, an AUG codon for initiation, and a downstream box (DB) (***8***; **Fig. 1B**). Although this sequence does recruit ribosomes in vivo (***3***; **Fig. 3**), it is possible that alterations could enhance its activity (*see* **Note 4**).
2. The *E. coli* translation sequence is inserted in-frame and downstream of the GFP ORF in the cyclic RNA transcription unit (**Fig. 1A,B**).
3. To create this construct, follow the site directed mutagenesis procedure outlined in **Subheading 3.1.1.** substituting the *E. coli*-SD oligonucleotide for the GFP-AUG or GFP-stop oligonucleotides (**Fig. 1B**).

3.1.3. Mutagenesis of Plasmids for PolyGFP Expression in E. coli

1. For poly GFP translation in *E. coli*, the author inserted the *E. coli*-SD oligonucleotide containing the sequence used in **Subheading 3.1.2.**, but into plasmid from **Subheading 3.1.1.**, in which both the GFP-AUG and the GFP-stop had been mutated.
2. Follow site-directed mutagenesis protocols in **Subheading 3.1.1.**, remembering to verify the exact sequence of this construct, very carefully, by dideoxynucleotide sequencing. For success of the polyGFP translation system, it is critical that each nucleotide within the circularized mRNA is accounted for (*see* **Note 5**).

3.1.4. Plasmids for Monomeric GFP Translation from Rabbit Reticulocyte Lysates

For translation in rabbit reticulocyte lysates, the IRES from the picornavirus, encephelomyocarditis virus ***(14)***, is introduced downstream of the GFP ORF in the DNA construct from **Subheading 3.1.1.**, in which both the GFP-AUG and the GFP-stop are mutated. The IRES fragment is inserted, by subcloning, into the circular GFP mRNA plasmid (**Fig. 1C**). The author used an IRES fragment that had been modified ***(5)***, to remove all potential in-frame internal stop codons, so that polyGFP mRNA constructs could eventually be created (*see* **Subheading 3.1.5.**).

1. Digest 1–5 µg picornavirus IRES-containing plasmid ***(5)*** with restriction enzymes *Xho*I and *Bgl*II, using standard protocols and instructions from the manufacturers. Likewise, digest 1–5 µg circular GFP mRNA plasmid with restriction enzymes *Eco*RI and *Nco*I.
2. Digestion with these restriction enzymes leaves incompatible ends, so both the IRES fragment and the circular GFP mRNA vector are repaired, using Klenow DNA polymerase to generate blunt ends. For a 20-µL Klenow reaction, mix restricted DNA, 1 µL of 0.5 m*M* dNTPs, 2 µL 10X filling buffer, 1–5 U Klenow, and water to 20 µL. Incubate for 15 min at 30°C, and stop the reaction by adding 1 µL 0.5 *M* EDTA.
3. **Steps 3** and **4** for digested GFP DNA only. Increase volume to 50 µL, extract with phenol–chloroform–isoamyl alcohol, precipiate with ethanol, and isolate DNA as in **Subheading 3.1.1.** Resuspend the DNA at 50–100 ng/µL in water.
4. To prevent self-ligation, add 2 µL of 10X dephosphorylation buffer, 0.5 U calf intestinal phosphates/pmol DNA, and water to 20 µL. Incubate for 60 min at 37°C and terminate by addition of 1 µL 0.5 *M* EDTA. Repeat **step 3** above.
5. **Steps 5** and **6** for IRES DNA fragment only. To prevent religation into the original vector, the author uses a "freeze-squeeze" procedure to isolate the IRES-containing fragment. Run digested DNA on a 1% agarose gel containing 1 µg/mL ethidium bromide in 1X TBE. Visualize DNA using an ultraviolet (UV) light-box.
6. Poke a small hole in a 0.5-mL tube and plug with small ball of tissue (e.g., Kimwipe); rest this tube in a 1.5-mL tube. Excise the gel containing the IRES fragment from the gel and place it in the 0.5-mL tube. Freeze on dry ice for ~5

min. Spin the 0.5-mL tube (within 1.5-mL tube) at 10,000*g* for 10 min, and collect DNA-containing supernatant now in the 1.5-mL tube. Repeat ethanol precipitation, and resuspend the isolated fragment in water at 100–200 ng/μL.

7. For ligation of vector and fragment, the author uses a molar ratio of ~1 vector: 2–5 fragment. Mix 50–100 ng GFP vector, 200–500 ng IRES fragment, 1 μL 10X ligation buffer, 400 U New England Biolabs T4 DNA ligase (*see* Materials), and add water to 10 μL. Incubate 4–12 h at 16°C.
8. Isolate the ligated DNA by ethanol precipitation, and resuspend the pellet in 10 μL water. Transform *E. coli* XL1-Blue with 2–5 μL or the resuspended DNA. Analyze for the IRES insert, by restriction digest and/or sequencing (*see* **Note 5**). This plasmid is now ready for monomeric protein expression in rabbit reticulocyte lysate.

3.1.5. Plasmids for PolyGFP Translation in Rabbit Reticulocyte Lysates

1. An additional round of site-directed mutagenesis is required to create circular GFP mRNA encoding plasmids designed for translation in rabbit reticulocyte lysates.
2. Despite using the circular GFP mRNA plasmid lacking the GFP-stop, insertion of the IRES creates an in-frame stop codon at exactly one complete revolution of the 1311 nucleotide circle (*see* **Fig. 1C** and **4**).
3. A two nucleotide insertion by site directed mutagenesis using IRES-stop oligonucleotide is required to create the circular GFP-ORF. Follow the site directed mutagenesis protocol using the cyclic GFP plasmid containing the IRES (**Fig. 1C**) as the template for mutagenesis.

Fig. 4

3.2. Testing Circularization of In Vitro RNA Transcripts

Prior to introduction into *E. coli* or rabbit reticulocyte lysates, the circularization efficiency for each of the constructs can be tested by making in vitro RNA transcripts using T7 RNA polymerase and analyzing RNA products on denaturing polyacrylamide gels. All the requirements for the circular mRNA production are within a transcription reaction.

1. Linearize 1–2 mg plasmid template DNA with appropriate restriction enzyme and resuspend in 20 mL sterile RNase-free water.
2. Assemble the transcription mix in the following order: 5 μL rNTP mix (10 m*M* rATP, rCTP, rGTP + 1 m*M* rUTP), 10 μL 5X transcription buffer, 5 μL 0.1 m*M* DTT, 6 μL dH_2O, 1 μL RNasin, 2 μL α-^{32}P-UTP (3000 Ci/mmol), 1–2 μg linear DNA in 20 μL water, 1 μL T7 RNA polymerase.
3. Incubate at 37°C for 1 h. This reaction should be set up at room temperature (not ice) to avoid precipitation of rNTPs or linear DNA template, which can occur in the presence of spermidine (found in the transcription buffer). Add 1 μL of RQ1 DNase I to remove DNA template from the mixture. Terminate DNase I treatment by addition of 50 μL formamide dye mix.
4. Load 1–2 μL of the transcription mix onto a 40-cm, 4.5% denaturing (7 *M* urea) polyacrylamide gel. Run until bromophenol blue has reached the bottom of the gel. Dry the gel and analyze RNA products by exposing using either autoradi-

Job: (MiMB 97–6.0) Hicks (905-6)
Chapter: 6/Perriman
Pub Date: 2001

Operator: MS
Date: 9/5&25/01
Revision: 1st Galley

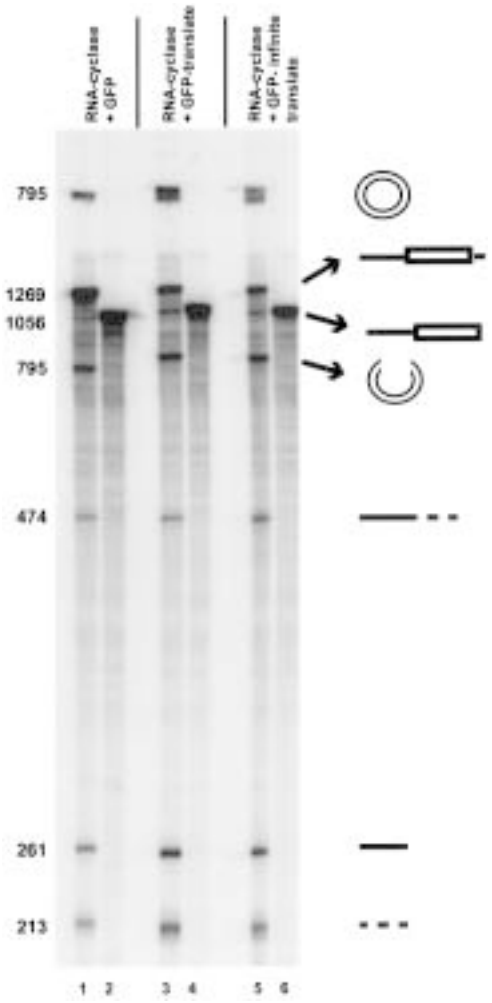

Fig. 4. Example of in vitro transcription of circular GFP mRNA from plasmids run on 4% PAG containing GFP (lanes 1 and 2), monomeric GFP + *E. coli* initiation cassette (lanes 3 and 4), polyGFP + *E. coli* initiation cassette (lanes 5 and 6). Lanes 2, 4, and 6 are negative controls, in which the circular GFP mRNA plasmids have been linearized to exclude the 5' intron sequence from the transcribed RNA. These RNA species represent the 3' intron +internal circularizing GFP ORF, which is also the first step of the group I splicing reaction. Lanes 1, 3, and 5 are complete circular GFP mRNA plasmids, and show products of the group I splicing reaction. Product designations, and their approximate length, are indicated to the right and left of the panel, respectively, and, from top to bottom, are: circularized RNA; unspliced precursor; 3' intron + internal circularizing GFP ORF; nicked circular RNA; ligated 5' and 3' introns; 5' intron; 3' intron. Note the difference in migration of the nicked circular species and the closed circular species.

ographic film or a Phosphoimager screen. Circularized RNA can be detected as a species distinct from other self-splicing RNA products in a variety of ways (e.g., **Fig. 4**; **Note 6**).

3.3. Testing for Protein Expression

3.3.1. Testing for Expression by Electrophoretic Analysis

3.3.1.1. Protein Extraction from *E. coli*

1. Transform *E. coli* strain BL21(DE3), or another *E. coli* strain expressing T7 RNA polymerase and designed for protein expression, with plasmids containing monomeric or poly GFP, and select on LB-AMP agar plates (*see* **Note 7**).
2. Grow 1-mL cultures of a single colony from each transformant in LB-AMP overnight at 37°C with shaking.
3. Seed 5 mL LB-AMP to A_{600} of ~0.05, with overnight cultures.
4. Grow at 37°C, with shaking, to A_{600} ~0.4. Add IPTG (1 m*M*) and grow cultures for another 4 h.
5. Pellet the cells, and resuspend them in 400 µL of 50 m*M* HEPES, pH 7.5. Lyse cells by 2–3 15-s bursts using a sonicator. Clarify by centrifugation at 10,000*g* for 10 min at 4°C. Decant, and save the supernatant. Protein extracts can be stored at –70°C.
6. Sonication destroys the repeating polyGFP multimers from the infinite encoding constructs (data not shown). Protein from these can be isolated as follows. Pellet cells and resuspend in 8*M* urea–0.1 *M* NaH_2PO_4–10 m*M* Tris, pH 6.0. Freeze on dry ice, then heat to 90°C for 5 min. Clarify by centrifugation at 10,000*g* for 10 min at 4°C. Aliquots can be stored at –70°C.

3.3.1.2. Detection of Monomeric and PolyGFP by Western Blots

1. Determine protein concentrations of extracts with Bradford assay kit per manufacturer's specifications.
2. For detection of polyGFP, equivalent total protein amounts (as determined in **step 1**) are separated on 8% separating–4% stacking polyacrylamide–SDS gels. For detection of monomeric GFP, 12.5% separating–4% stacking polyacrylamide–SDS gels are used. The author finds that 5–10 µL total cell lysate is sufficient for Western analysis and visualization of both monomeric GFP and polyGFP products (**Fig. 2A,B**).
3. Proteins are electrophoretically transferred to nitrocellulose membrane at 4°C, 50 V for ≥ 2 h in Western transfer buffer.
4. The nitrocellulose membrane is blocked by incubating in 3% BSA in TBS-T for 30 min at room temperature. Polyclonal GFP antibody is added at 1:1000 dilution in TBS-T–3% BSA, and incubated for ~2 h at room temperature. The membrane is washed in TBS-T 3 × 15 min, then incubated in TBS-T–3% BSA containing IgG-goat-antirabbit secondary antibody at 1:1000 dilution for 1 h at room temperature. The membrane is then washed in TBS-T 3 × 15 min, and GFP products visualized via conjugated alkaline phosphatase, in a solution of 17 µL BCIP and 33 µL nitroblue tetrazolium in 5 mL substrate buffer.

3.3.1.3. Testing for Protein Expression in Rabbit Reticulocyte Lysates by Autoradiography

To test for protein expression from circular mRNA constructs encoding monomeric GFP or polyGFP from the mammalian expression cassette (**Fig. 1**), mRNAs are transcribed in vitro, then incubated in a rabbit reticulocyte lysate translation system (*see* Materials and **Note 8**) using ^{35}S-methionine for autoradiographic detection of proteins (**Fig. 4**).

1. In vitro transcription can be done as described in **Subheading 3.2.**, **steps 1** and **2**, except radiolabeled α-^{32}P-UTP is omitted. Instead, substitute 10 m*M* rUTP for 1 m*M* rUTP. Include RQ1 DNase treatment, to remove template DNA.
2. Add an equal volume (50 µL) of phenol–chloroform–isoamyl alcohol; mix, and spin at 10,000*g* for 5 min. Transfer top phase to new tube, and precipitate the RNA with ethanol. Resuspend the precipitated RNA in RNase-free water at ~1–2 µg/µL (usually requires ~10 µL of water).
3. The author typically uses 2–3 µg in vitro transcribed RNA in a 25-µL in vitro translation reaction, and follows a proportionally scaled-down version of the standard 50-µL protocol found in the technical bulletin supplied with rabbit reticulocyte lysates from Promega (with exceptions listed in **Note 8**). The author uses ^{35}S-methionine as the radiolabeled amino acid. Generally, one-tenth of this reaction (2–3 µL), is sufficient for electrophoresis in 10% separating–4% stacking SDS-polyacrylamide gel (both products can be analyzed using 10% polyacrylamide as the separating gel).
4. Following electrophoresis, the gel is dried and used to expose film for autoradiography or, if available, phosphoimager analysis (**Fig. 3**). In vitro translation products of GFP are visualized via incorporated ^{35}S-methionine.

3.3.2. Detection of GFP by Fluorimetry

GFP expression from the circular mRNAs in *E. coli* can also be analyzed by looking for emission of green light, upon photoexcitation ***(11)***. This can be done by observing *E. coli* colonies on a UV lightbox after growth on LB agar containing IPTG, or, for more sensitive and quantitative detection, by fluorimetry. The author can detect fluorescence from *E. coli* expressing the monomeric circular GFP mRNA constructs using either method, but The author is unable to detect any fluorescence from the polyGFP expressing constructs (*see* **Note 3**).

1. Fluorimetry, if available, is the method of choice for detecting GFP expression from circular mRNAs. The author uses a Perkin-Elmer (LS50B) luminescence spectrometer to do an emission scan with a 5 nm slit width at 240 nm/min. Excitation is fixed at 397 nm and emission is scanned between 350 and 550 nm (peak emission for the GFP derivative is 508 nm).
2. Cultures of GFP-expressing *E. coli* (2–3 mL) are grown overnight at 37°C. The cells are pelleted and resuspended in 1 mL of water to analyze. A four-sided glass cuvet, with a 10-mm path length works well in these assays (*see* **Note 3**).

Job: (MiMB 97–6.0) Hicks (905-6)
Chapter: 6/Perriman
Pub Date: 2001
Operator: MS
Date: 9/5&25/01
Revision: 1st Galley

4. Notes

1. Several other DNA polymerases work well in this protocol. The author has used T7 DNA polymerase (from which Sequenase™ is a derivative), T4 DNA polymerase, and Klenow. Buffer requirements may vary, so be sure to check if using a different enzyme.
2. Deoxyoligonucleotides for site-directed mutagenesis using the method described in this chapter should be designed such that they "band-aid" the region of insertion. For point mutations, such as those required to remove start or stop codons in the circular GFP mRNA constructs, potential hybridization 5' and 3' of the mutagenic region should be at least 10–12 bases, to ensure efficient annealing. For larger sequence alterations, such as the insertion of the *E. coli* expression cassette, The author recommends extending this potential hybridization to at least 17 nucleotides. Also, synthetic deoxyoligonucleotides do not have phosphates at their 5' termini. For any ligation of the mutagenic strand, these oligonucleotides must be phosphorylated, using standard molecular biology protocols involving T4 DNA kinase.

   <Au: More specific than "band-aid"?

   The author carried out sequential mutagenesis reactions (i.e., first, the author altered the GFP-AUG) then this was used as a template in a second mutagenic reaction using the GFP-stop oligonucleotide), but there is no reason why both could not be attempted simultaneously, by including both oligonucleotides in the annealing step. If this is done, one needs to acquire the GFP-AUG as a single mutation as well as in combination with the GFP-stop.
3. Prokaryotic *in vivo* translation of the monomeric GFP from the circular mRNA constructs is relatively inefficient, compared with the identical linear constructs ***(3)***. Thus, reduced fluorescence makes visualization by UV illumination with a light-box difficult. To be certain that one is not observing autofluorescence (which can be remarkably deceiving), it is essential to compare with non-GFP expressing *E. coli*. The author could convincingly observe fluorescence from the monomeric GFP constructs in *E. coli* using a UV light-box, and verified this by fluorimetry. In contrast, the author was unable to distinguish any fluorescence above background using the UV light-box method or fluorimetry when looking at the polyGFP mRNA construct ***(3)***. Presumably, this means that multimeric linked units of GFP are unable to fluoresce. The author has not analyzed why this is the case, but the polyGFP may be inefficiently posttranslationally modified ***(17)***, and/or not able to fold into the correct conformation to allow formation of the cyclic tripeptide chromophore. Possibly the addition of a linker region between the repeating GFP units would allow enough space for correct processing, folding, and fluorescence. Additionally, because a heterogeneous collection of proteins is observed, ranging in size from ~50 kDa (the size of the monomer) to >300 kDa, possibly much of the polyGFP is made from incomplete GFP units that might interfere with posttranslational processing.
4. Although the author has not tested other sequences, our previous data ***(3)*** show that both the SD and DB sequence motifs, in the translation initiation sequence of the circular mRNAs, are necessary to produce maximal protein levels, as meas-

ured in this experimental system. The author finds that the SD sequence contributes more than the DB motif ***(3)***. Since very little is known about whether there is a role for 5' ends of mRNA contributing to translation in *E. coli*, possibly the SD and DB used in this study may not be best for exclusively internal initiation. Other RNA sequence motifs may enhance direct ribosome recruitment, and thus improve circular mRNA translation. Data also show that the apparent effects of removing either the SD or the DB motif is minimized during expression from the polyGFP ORF, compared to the monomeric GFP ORF ***(3)***, which indicates that initiation is the rate-limiting step in translation in *E. coli* from circular mRNAs, at least in the context of sequences contained in our circular GFP mRNA constructs.

5. The author recommends always sequencing mutagenic clones, particularly in the region containing the mutation(s) because site-directed mutagenesis can sometimes introduce sequenced changes other than those intended. In addition to verifying them by restriction digest. The author also strongly recommends sequencing when fragments and/or vector have been end-filled. This is particularly critical for the circular GFP mRNA constructs, in which each nucleotide must be accounted for, to ensure the reading frame remains correct.
6. In vitro transcribed RNAs from the circular GFP mRNA plasmids can be analyzed for circularized products in several ways. Analysis of circular GFP mRNA expression products on high- and low-% denaturing acrylamide gels leads to a shift in mobility of nonlinear (i.e., circular) molecules, whereby circles run proportionally faster than their linear counterparts on a low % gel, but proportionally slower on a high-% gel ***(12)***. Only the migration of nonlinear, circularized RNA species are affected in this way. Two dimensional denaturing gel electrophoresis ***(1)*** is another way of visualizing circular RNAs, which is an extension of the first, but here a diagonal of linear RNA molecules is produced. The circular molecules appear above the diagonal, because of the reduction in mobility of nonlinear species in higher-% polyacrylamide gels.
7. There are several commercially available strains of *E. coli* expressing T7 RNA polymerase from a chromosomal gene copy. The author uses BL21(DE3) from Novagen, in which T7 is under control of a regulated *lacUV5* promoter that can be induced using IPTG. The Novagen catalog (for 2000) and web site ***(18)*** contain excellent overviews on protein expression and some detail regarding the advantages and disadvantages of each strain variant. Because the author has not tested strains other than BL21(DE3), comment cannot be made on their suitability for expressing protein from our circular GFP mRNA constructs. If desired, circularized mRNAs can also be analyzed by purifying total RNA from IPTG-induced *E. coli* and doing Northern hybridization ***(3)***.
8. The author found the following modifications necessary to optimize in vitro translation products from our circular GFP mRNA plasmids. The author titrated addition of 25 m*M* magnesium acetate and found 0.5/25 μL reaction optimal for monomeric and polyGFP translation from IRES circular GFP mRNA constructs. Similarly, the author found 1.2/25 μL reaction of 2.5 m*M* potassium chloride increases translation efficiency. DTT was not added. An incubation time of 60

min is sufficient. Following incubation, and prior to loading on polyacrylamide gel, The author finds it necessary to add 0.5 µL 10 mg/mL RNase A to a 3-µL aliquot of the reaction, to digest aminoacyl tRNAs, which produce background bands. The GFP produced in the rabbit reticulocyte lysate system does not fluoresce in our hands. The author was unable to obtain any GFP translation products using the coupled transcription–translation kits from Promega.

Acknowledgments

This work was done in the laboratory of Professor Manuel Ares, under partial support by grant GM40478 from the National Institutes of Health. Rhonda Perriman was also supported by an American Cancer Society Postdoctoral fellowship. The author thanks Roland Nagel for review and M.A. for stimulating discussion, support, and critical review of this manuscript.

<Au: Please define M. A. (author)?

References

1. Ford, E. and Ares, M. (1994) Synthesis of circular RNA in bacteria and yeast using RNA cyclase ribozymes derived from a group I intron of phage T4. *Proc. Natl. Acad. Sci. USA* **91,** 3117–3121.
2. Puttaraju, M. and Been, M. D. (1996) Circular ribozymes generated in *Escherichia coli* using group I self-splicing permuted intron-exon sequences. *J. Biol. Chem.* **271,** 26,081–26,087.
3. Perriman, R. and Ares, M. (1998) Circular mRNA can direct translation of extremely long repeating sequence proteins *in vivo*. *RNA* **4,** 1047–1054.
4. Long, M. B. and Sullenger, B. A. (1999) Evaluating group I intron catalytic efficiency in mammalian cells. *Mol. Cell Biol.* **19,** 6479–6487.
5. Chen, C. Y. and Sarnow, P. (1995) Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* **268,** 415–417.
6. Prince, J. T., McGrath, K. P., DiGirolamo, C. M., and Kaplan, D. L. (1995) Construction cloning and expression of synthetic genes encoding spider dragline silk. *Biochemistry* **34,** 10,879–10,885.
7. Oshimi, Y. and Suzuki, Y. (1977). Cloning of the silk fibroin gene and its flanking sequences. *Proc. Natl. Acad. Sci. USA* **74,** 5363–5367.
8. Sudo, S., Fujikawa, T., Nagakura, T., et al. (1997) Structures of mollusc shell framework proteins. *Nature* **387,** 563–564.
9. Heslot, H. (1998) Artificial fibrous proteins: a review. *Biochemie* **80,** 19–31.
10. Shine, J. and Dalgarno, L. (1974) The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. USA* **71,** 1342–1346.
11. Steitz, J. A. and Jakes, K. (1975) How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli. Proc. Natl. Acad. Sci. USA* **72,** 4734–4738.
12. Gold, L. (1988) Posttranscriptional regulatory mechanisms in *Escherichia coli. Annu. Rev. Biochem.* **57,** 199–233.

13. Sprengart, M. L., Fuchs, E., and Porter, A. G. (1996) The downstream box: an efficient and independent translation initiation signal in *Escherichia coli. EMBO J.* **15,** 665–674.
14. Hellen, C. U. and Wimmer, E. (1995) Translation of encephalomyocarditis virus RNA by internal ribosomal entry.*Curr. Top. Microbiol. Immunol.* **203,** 31–63.
15. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994) Green fluorescent protein as a marker for gene expression. *Science* **263,** 802–805.
16. Puttaraju, M. and Been, M. D. (1992) Group I permuted intron-exon (PIE) sequences self-splice to produce circular exons. *Nucleic Acids Res.* **20,** 5357–5364.
17. Heim R., Prasher D. C., and Tsien R. Y. 1994 Wavelength mutations and post-translational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci. USA* **91,** 12,501–12,504.
18. http://www.novagen.com; URL verified Oct. 10, 2000.