CHEMBIOCHEM

DOI: 10.1002/cbic.200700546

Random Walks to Synthetic Riboswitches—A High-Throughput Selection Based on Cell Motility

Shana Topp and Justin P. Gallivan*[a]

A major goal of chemical and synthetic biologists is to create ligand-dependent genetic-control systems to report on cellular metabolism, to construct synthetic gene circuits, or to reprogram cellular behavior. Because designing genetic switches de novo is challenging, many groups have employed directed-evolution strategies to achieve these goals.[1,2] Our laboratory recently reported a high-throughput robotic screen that was used to identify synthetic riboswitches that displayed low background levels of gene expression in the absence of a ligand, and strongly activated gene expression in the presence of the ligand.[3] We subsequently demonstrated that these riboswitches could control bacterial motility in a ligand-dependent fashion.[4] Because the differences in cell motility at different ligand concentrations were easy to distinguish by using only a ruler, we asked whether motility differences could be the basis of a high-throughput selection to discover new synthetic riboswitches from large libraries. We envisioned that this method could equal, if not exceed, the throughput of our previously reported robotic screen, and could be performed at a fraction of the cost. Here we report an inexpensive and operationally simple selection method based on cell motility that not only approaches the throughput of a genetic selection, but also provides the quantitative nature of a genetic screen. We further show that this selection quickly identifies synthetic riboswitches that display low background levels of gene expression in the absence of a ligand, and robust increases in the presence of a ligand. We anticipate that motility-based selections will be generally useful in the discovery of rare events from large genetic libraries.

There is good precedent for using motility to select for rare events. For decades, microbiologists have identified rare mutants by spotting cells at the center of a Petri dish containing semisolid media, and manually selecting cells that migrate abnormally.[5] Capitalizing on these successes, Goulian and coworkers recently developed a motility-based selection to discover mutant chemoreceptors that could recognize a new ligand.[6] Because a riboswitch displays two different phenotypes, depending on whether or not the ligand is present, a selection process must be able to quantitatively assay both the "on" and "off" states. Such counter-selections, which select for one phenotype and against another, are powerful, but can be challenging to implement. We envisioned that a motility-based selection could be used to discover "on" switches by selecting for mutants that do not move in the absence of the ligand, and then further selecting for mutants that move in the presence of the ligand. Of course, there are many ways that one can imagine setting up the experiment (for example, looking for "off" switches), but the key is to be able to rapidly and inexpensively determine the phenotype under two sets of conditions.

To compare our motility selection to a more traditional screening technique, we assayed combinatorial libraries that were similar to those previously reported by our laboratory,[3] which were comprised of four to eight randomized base pairs flanked by the mTCT8-4 theophylline aptamer[7] and a fixed ribosome-binding site upstream of the β-galactosidase reporter gene. *E. coli* expressing these libraries were assayed by using a multistep process that used a robotic colony picker and an automated liquid-handling system.[3] While our previous assay identified several outstanding synthetic riboswitches, it relied on expensive robotics, in addition to large quantities of consumables. We anticipated that a motility-based selection could eliminate the need for specialized capital equipment, and identify synthetic riboswitches with only standard consumables.

To select for riboswitches by using motility as the readout, we used *cheZ* as a reporter gene. CheZ plays a critical role in *E. coli* chemotaxis by dephosphorylating the CheY-P protein,[8] which binds to the flagellar motor and causes cells to tumble. Optimal levels of CheZ are necessary for *E. coli* cells to migrate on semisolid media. If too little CheZ is present, the level of CheY-P will increase, and the cells will tumble incessantly and not migrate.[9] If cells have excess CheZ, they will swim very smoothly and rarely tumble. Because cells that swim extremely smoothly can become embedded in the semisolid media, they cannot migrate.[10] Thus it is critical to ensure that CheZ is not over-expressed in these assays.

Since the strength of the promoter will ultimately dictate the maximum expression level of the *cheZ* gene, we began with two different promoters: a "weak" *IS10* promoter[11] and the *tac* promoter,[12] which is 60- to 100-fold stronger (S.T. unpublished results). We anticipated that the motility selections would readily reveal which promoter provides the appropriate CheZ expression level. Using cassette-based PCR mutagenesis, we constructed a library in which the mTCT8-4 theophylline aptamer[7] was followed by ten consecutive randomized base pairs (N_{10}) in the 5′ untranslated region (5′ UTR) of *cheZ* (Figure 1). We chose to assay an N_{10} library that lacked a preset ribosome-binding site for these experiments rather than our previously described "N_8" library, for two reasons. First, we anticipated that this selection could more effectively sample the additional sequence space; and second, having the additional

[a] S. Topp, Prof. J. P. Gallivan
Department of Chemistry and Center for Fundamental and Applied Molecular Evolution
Emory University
1515 Dickey Drive, Atlanta, GA 30322 (USA)
Fax: (+1) 404-727-6586
E-mail: justin.gallivan@emory.edu

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

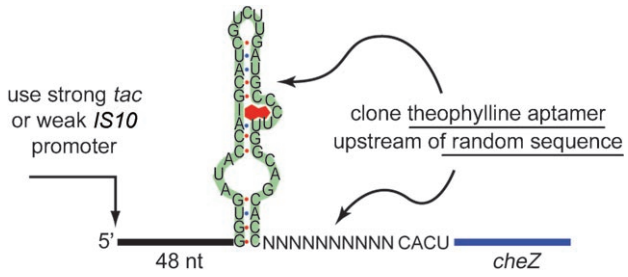

Figure 1. Cloning strategy to create libraries.

randomized bases would allow us to optimize the strength of the ribosome-binding site to ensure proper levels of CheZ expression.

E. coli strain JW1870[13] (*ΔcheZ*) was transformed separately with the N_{10} plasmid libraries under the control of either the *IS10* or *tac* promoter. After each transformation, ~600 000 cells were spotted on semisolid selection media with or without theophylline, and grown at 30 °C for 13 h (Figure 2). Cells that

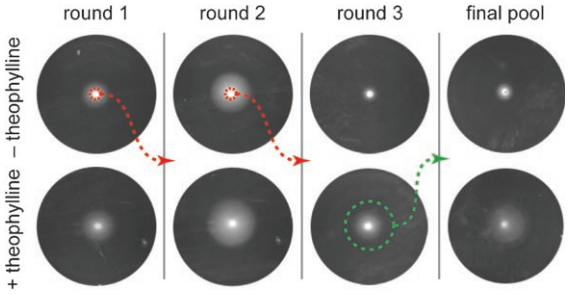

Figure 2. Progress of selections for synthetic riboswitch activity by using the *IS10* promoter. In rounds 1 and 2, cells were chosen from the center of plates without theophylline. In round 3, we picked cells from the outside of the plate that contained theophylline (1 mM). The final pool showed a significant increase in motility when theophylline was present. Plates from rounds 1, 2, 3, and the final pool were incubated at 30 °C for 13, 18, 14, and 14 h, respectively. The longer incubation in round 2 ensured that cells had time to migrate; this allowed us to pick the least motile cells.

did not migrate in the absence of theophylline were isolated and subjected to a second round of selection. After 18 h, immotile cells on the theophylline-free plate were isolated and spotted on plates with and without theophylline. Populations of cells from both libraries displayed increased motility in the presence of theophylline, and cells were collected by scraping the region of the plates that extended beyond the outermost visible ring of cells. Cultures derived from the weaker *IS10* promoter library grew normally, and when spotted on semisolid media in the presence or absence of theophylline (1 mM), showed clear differences in motility. In contrast, cells derived from the stronger *tac* promoter library failed to grow; this suggests that very few cells migrated far across the plate.

We suspected that this was a result of the strong *tac* promoter, which led to the over-expression of *cheZ* and caused the majority of the cells not to move, regardless of whether theophylline was present. To confirm this, we cloned the 5′ UTRs from each library after the second round of negative selection upstream of the β-galactosidase reporter gene. These cells were plated onto selection media containing X-gal, but no theophylline. As expected, the majority of the cells from the weak *IS10* promoter library were white or light blue; this indicates that gene expression was low in the absence of theophylline (see Figure S1 in the Supporting Information). However, the majority of cells from the *tac* promoter library were dark blue; this suggests that they expressed substantial amounts of β-galactosidase (and by extension, CheZ) in the absence of theophylline. Thus, many of the cells in the *tac* library were immotile because CheZ was over-expressed and the cells became embedded in the media. This result underscores the need to control for promoter strength in motility assays.

Since most of the cells in the *IS10* library showed low levels of gene expression in the absence of theophylline, and significant increases in motility in the presence of theophylline, we anticipated that this library contained functional synthetic riboswitches. To confirm this, the plasmids were extracted from the population of cells collected after the third round of selection, the DNA pool was amplified by PCR, and this library was cloned in the 5′ UTR of a *cheZ–lacZ* translational fusion that was used to transform *E. coli*. Ninety five random transformants were assayed for β-galactosidase activity in the absence and presence of theophylline (1 mM) by using a multichannel pipettor. Of the 95 clones assayed, 12 contained riboswitches that activated gene expression more than fivefold in the presence of theophylline (1 mM), and of these, four exhibited activation ratios greater than 15-fold (Figure 3). Sequencing revealed that two of the four switches were identical in the N_{10} region (see the Supporting Information). Another had 11 nucleotides in this region, which might have resulted from errors in oligonucleotide synthesis, or from random mutations introduced by replication in *E. coli*; regardless of the cause, this result highlights the ability of the motility selection to discover rare events. All of the sequences could adopt folds[14] similar to riboswitches identified by using our previously described

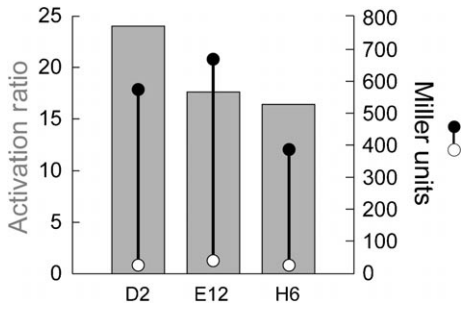

Figure 3. Measures of the activities of the synthetic riboswitches identified in the motility selections. Each measurement was performed in triplicate. Right-hand *y* axis: β-galactosidase activity in the absence (○) or presence of theophylline (1 mM; ●). Activities are expressed in Miller units, and the standard errors of the mean are less than the diameters of the circles. Left-hand *y* axis: activation ratios of the synthetic riboswitches (■), which were determined by using the ratio of the Miller units in the presence and absence of theophylline.

screen,[3] and each contained a region of purines that could function as a ribosome-binding site (Figure S2).

Although our selection method identified new riboswitches, we wanted to confirm whether the motility of the cells on semisolid agar corresponded to the absolute levels of gene expression as measured by β-galactosidase expression. If this were the case, our motility selection, which can be performed with millions of clones in parallel by using only semisolid agar and a ruler, could be used in place of the β-galactosidase assay, which requires significantly more resources to perform. To address this question, we cloned several riboswitches that were discovered by robotic screening into the 5′ UTR of *cheZ*, and performed motility assays in the presence and absence of theophylline. We correlated the β-galactosidase expression in the presence and absence of theophylline with migration on semisolid agar for these previously described clones,[3] as well as for the newly-discovered synthetic riboswitches (Figure 4).

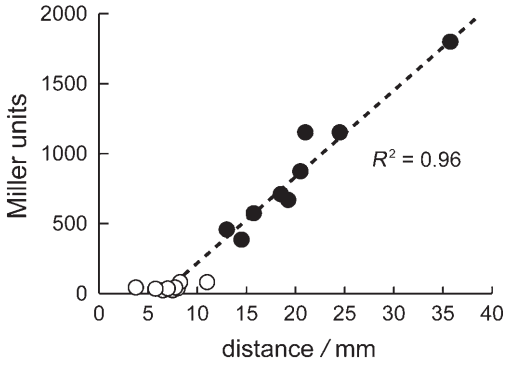

Figure 4. β-Galactosidase activity (in Miller units) as a function of cell motility (distance migrated on semisolid agar) for synthetic riboswitches. Measurements in the absence of theophylline (○); measurements in the presence of theophylline (1 mM; ●). There is a strong correlation between migration distance and β-galactosidase activity.

Regardless of the method by which the riboswitch was identified, the results correlate well; this suggests that motility on semisolid agar is an inexpensive and effective proxy for high-throughput assays for β-galactosidase expression, which require expensive reagents and equipment.

In summary, we have developed a rapid, efficient, and inexpensive method to discover synthetic riboswitches. This method can be used to perform both positive and negative selections with very large libraries ($>10^5$ clones) and does not require robotics: only standard consumables, and a minimal amount of human intervention. The riboswitches identified by using this assay have low background levels of gene expression in the absence of ligand and display good activation ratios. We expect that this method will be useful for discovering riboswitches starting from a variety of RNA aptamers, and that the resulting synthetic riboswitches will be generally useful for controlling bacterial gene expression. Furthermore, we suggest that current genetic screens that detect changes in the production of proteins, such as green fluorescent protein or β-galactosidase, can be adapted to use cell motility as an inexpensive high-throughput readout.

Experimental Section

General: Synthetic oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA, USA). Culture media was obtained from EMD Bioscience (San Diego, CA, USA). Ampicillin was purchased from Fisher. Theophylline was obtained from Sigma. DNA polymerase, restriction enzymes, and the pUC18 cloning vector were purchased from New England BioLabs. Purifications of plasmid DNA, PCR products, and enzymatic digestions were performed by using kits from Qiagen. All new plasmids were verified by DNA sequencing provided by Agencourt, MWG Biotech, or the Center for Fundamental and Applied Molecular Evolution at Emory University.

Construction of libraries for motility selection: Cassette-based PCR mutagenesis was used to construct the libraries shown in Figure 1, which were subcloned into the *Kpn*I and *Sac*I sites of a vector derived from pUC18 (Ap^R). Plasmid manipulations were performed with *E. coli* TOP10F′ cells (Invitrogen) that were transformed by electroporation. The theoretical library size was 4^{10} (~10^6) clones. On a single Petri dish (85 mm), we typically assayed ~6×10^5 clones. Approximately five Petri dishes could provide 95% coverage of the library.[15]

Motility selection: Motility experiments were performed with *E. coli* strain JW1870.[13] For all motility plates, selection media (tryptone broth with 0.25% agar, 50 μg mL^{-1} ampicillin, and theophylline: 0 or 1 mM) was prepared in Petri dishes (85 mm diameter). Diluted cell suspensions from mid-log-phase cultures (3 μL, ~6×10^5 cells) were applied to the center of the plates, which were dried in air for 15 min, and incubated at 30 °C. Selected cells were suspended in liquid media (tryptone broth, 50 μg mL^{-1} ampicillin) and grown to mid-log-phase to begin the next round of selection.

To begin selection, JW1870 cells were transformed with both plasmid libraries and ~600000 cells from each library were spotted at the center of motility plates with and without theophylline. After 13 h, immotile cells remaining at the center of the plates without theophylline were suspended in liquid media, grown to mid-log phase, and spotted on plates with and without theophylline. After 18 h, cells that remained at the center of the plates without theophylline were transferred to liquid media, and again spotted on motility plates with and without theophylline. After 14 h, cells were collected from the region beyond the visible migration edge of the theophylline (1 mM) plates. A diluted suspension of the final pool derived from the *IS10* promoter library was spotted on motility plates to confirm the theophylline-dependent migration of this population.

Construction of *cheZ–lacZ* translational fusions: Plasmid DNA was extracted from cell cultures grown after the second and third rounds of selection. These pools were amplified with the forward primer JPG-007 (which anneals 5′ to the aptamer) and the reverse primer ST-278 (which anneals to the 3′ end of *cheZ*, removes the stop codon, adds a GGSSAA linker, and introduces a *Hind*III restriction site to construct a translation fusion to *lacZ*). The pools of PCR products were subcloned into the *Kpn*I and *Hind*III sites of SAL172 (*IS10* promoter) or SAL109 (*tac* promoter), both of which contain *lacZ* and are derivatives of pUC18.[3]

β-Galactosidase promoter comparison: TOP10F′ cells were transformed with the *cheZ–lacZ* pools constructed for both the *IS10* and *tac* promoter populations after two rounds of motility selection. Cells were grown at 37 °C on LB agar (25 mL) supplemented with ampicillin (50 μg mL^{-1}) and X-gal (2.1 mg dissolved in 83.3 μL dimethyl formamide).

β-Galactosidase assay to confirm activity: TOP10F′ cells were transformed with the *cheZ–lacZ* pool that was constructed after the final round of motility selection. Cells were grown at 37 °C on LB agar supplemented with ampicillin (50 μg mL^{-1}). Ninety five clones were picked by hand, and were assayed for β-galactosidase activity in the absence and presence of theophylline (1 mM) by using a multichannel pipettor as previously described.[3] Clones with activation ratios greater than 15 were subcultured and assayed in triplicate, as previously described.[3] Miller units were calculated by using the following formula (1):

$$\text{Miller units} = \frac{OD_{420}}{OD_{600} \times \text{hydrolysis time} \times (\text{volume of cell lysate}/\text{total volume})} \quad (1)$$

Motility assay to correlate migration to Miller units: Forward primer JPG-007 and reverse primer SKD-056 were used to clone several riboswitches (which were discovered previously by using the robotic assay)[3] into the 5′ UTR of a pUC18 derivative containing the *cheZ* gene. JW1870 (*ΔcheZ*) cells were transformed with these riboswitches. Cells were grown to mid-log phase in tryptone broth supplemented with ampicillin (50 μg mL^{-1}), and diluted suspensions (~600 000 cells) were spotted at the center of motility plates with or without theophylline (1 mM). Plates were imaged after 13 h incubation at 30 °C. The Miller units for each riboswitch were measured, and these values were plotted as a function of the corresponding migration radii.

Acknowledgements

This research was supported by the NIH (GM070740 to J.P.G.). J.P.G. is a Beckman Young Investigator and a Camille Dreyfus Teacher–Scholar. S.T. is a G. W. Woodruff Fellow and an ARCS Scholar.

Keywords: high-throughput screening · riboswitches · RNA · selection · synthetic biology

[1] Y. Yokobayashi, R. Weiss, F. H. Arnold, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16587–16591.
[2] K. Chockalingam, Z. L. Chen, J. A. Katzenellenbogen, H. M. Zhao, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5691–5696.
[3] S. A. Lynch, S. K. Desai, H. K. Sajja, J. P. Gallivan, *Chem. Biol.* **2007**, *14*, 173–184.
[4] S. Topp, J. P. Gallivan, *J. Am. Chem. Soc.* **2007**, *129*, 6807–6811.
[5] J. Adler, *Annu. Rev. Biochem.* **1975**, *44*, 341–356.
[6] P. Derr, E. Boder, M. Goulian, *J. Mol. Biol.* **2006**, *355*, 923–932.
[7] R. D. Jenison, S. C. Gill, A. Pardi, B. Polisky, *Science* **1994**, *263*, 1425–1429.
[8] R. Zhao, E. J. Collins, R. B. Bourret, R. E. Silversmith, *Nat. Struct. Biol.* **2002**, *9*, 570–575.
[9] A. J. Wolfe, M. P. Conley, T. J. Kramer, H. C. Berg, *J. Bacteriol.* **1987**, *169*, 1878–1885.
[10] A. J. Wolfe, H. C. Berg, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 6973–6977.
[11] C. Jain, N. Kleckner, *Mol. Microbiol.* **1993**, *9*, 233–247.
[12] H. A. De Boer, L. J. Comstock, M. Vasser, *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 21–25.
[13] T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, H. Mori, *Mol. Syst. Biol.* **2006**, *2*, 1–11.
[14] M. Zuker, *Nucl. Acids Res.* **2003**, *31*, 3406–3415.
[15] A. E. Firth, W. M. Patrick, *Bioinformatics* **2005**, *21*, 3314–3315.

Received: September 12, 2007
Published online on December 20, 2007

