

Computational selection of nucleic acid biosensors via a slip structure model

Bradley Hall^a, Jay R. Hesselberth^b, Andrew D. Ellington^{a,*}

^a Department of Chemistry and Biochemistry, Institute for Cell and Molecular Biology, University of Texas at Austin, Austin, TX 78712, USA

^b Departments of Genome Sciences and Medicine, University of Washington, Box 357730, Seattle, WA 98195, USA

Received 24 April 2006; received in revised form 8 August 2006; accepted 14 August 2006

Available online 22 September 2006

Abstract

Aptamers have been shown to undergo ligand-dependent conformational changes, and can be joined to ribozymes to create allosteric ribozymes (aptazymes). An anti-flavin (FMN) aptamer joined to the hammerhead ribozyme yielded an aptazyme that underwent small, FMN-dependent displacements in the helix that joined the aptamer and ribozyme. This ‘slip structure’ model in which alternative sets of base-pairs are formed in the absence and presence of ligand proved amenable to energetic and computational modeling. Initial successes in modeling the activities of known aptazymes led to the *in silico* selection of new ligand-dependent aptazymes from virtual pools that contained millions of members. Those aptazymes that were predicted to best fit the slip structure model were synthesized and assayed, and the best-designed aptazyme was activated 60-fold by FMN. The slip structure model proved to be generalizable, and could be applied with equal facility to computationally generate aptazymes that proved to be experimentally activated by other ligands (theophylline) or that contained other catalytic cores (hairpin ribozyme). Moreover, the slip structure model could be applied to the prediction of a ligand-dependent aptamer beacon biosensor in which the addition of the protein vascular endothelial growth factor (VegF) led to a 10-fold increase in fluorescent signal.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Aptazyme; Aptamer beacon; Computational selection; Slip structure; Signaling aptamer

1. Introduction

Aptamers and ribozymes can be adapted to serve as biosensors by a variety of methods (Breaker, 2002; Nutiu and Li, 2005; Silverman, 2003; Verma et al., 2003). Most of these methods rely upon the aptamer or ribozyme undergoing a ligand-dependent conformational change. A number of authors have generated aptamer beacons in which a non-native conformation is established by base-pairing, and the cognate ligand then stabilizes the native conformation (Fang et al., 2003; Stojanovic et al., 2001). Ligand-dependent changes between non-native and native conformations can be read out by appending reporter molecules to the aptamer; for example, some aptamer beacons are hairpin stem structures in which a fluorophore and quencher are aligned in the non-native conformation, and split apart upon ligand-binding and conformational rearrangement, leading to fluorescence (Hamaguchi et al., 2001).

Ligand-dependent conformational changes can also be introduced into ribozymes by design or by selection, yielding allosteric ribozymes (so-called aptazymes). For example, an aptamer that undergoes a flavin mononucleotide (FMN) dependent conformational change was appended to a hammerhead ribozyme; ligand-binding yielded a coupled change in the conformation of the active site and a concomitant increase or decrease in catalytic activity (Soukup and Breaker, 1999a). Selections for RNA or DNA catalysis can also be designed in which the only successful catalysts will be those that can be activated by or utilize an introduced effector molecule (Roth and Breaker, 2004).

To date, though, almost all of these aptamer and aptazyme biosensors have been generated by empirical design (Tang and Breaker, 1997) or *in vitro* selection (Nutiu and Li, 2005; Robertson et al., 2004; Soukup and Breaker, 1999a). However, algorithms for the prediction of nucleic acid secondary structure have advanced to the point where nucleic acid secondary structures can be rapidly enumerated based on nucleic acid sequence (Hofacker et al., 1994; Zuker, 2000). Likewise, probabilistic methods have been developed to sample populations

* Corresponding author. Tel.: +1 512 232 3424; fax: +1 512 471 7014.
E-mail address: andy.ellington@mail.utexas.edu (A.D. Ellington).

of possible secondary structures and to predict structural stability (Ding and Lawrence, 2003; McCaskill, 1990). There have also been theoretical treatments of what computational design criteria will maximize sensitivity and specificity in nucleic acid sensors (Dirks et al., 2004). Therefore, it should be possible to develop algorithms for predicting the sequences of aptamers or aptazymes that will undergo ligand-dependent conformational changes and therefore function effectively as biosensors.

The Breaker lab has developed a computational method for developing allosteric ribozymes activated by nucleic acid sequences (Penchovsky and Breaker, 2005). In their work, a virtual pool of RNA molecules was created that contained a randomized region inserted into the middle of the hammerhead ribozyme. An algorithm was then used to predict the dominant secondary structures in the presence and absence of an oligonucleotide effector that was complementary to a given sequence in the randomized region. For example, using this algorithm it proved possible to predict the sequence of ribozymes that would be inactive in the absence of an oligonucleotide effector (due to internal pairings that disrupted the ribozymes), and then would undergo structural rearrangement in the presence of the oligonucleotide effector returning the ribozymes to their native structures. However, computational methods have not been applied to the more difficult problem of controlling nucleic acid conformational change by non-nucleic acid effectors.

In order to develop prediction methods for nucleic acid biosensors activated by analytes other than oligonucleotides, we wished to first explore a conformational transition that was reasonably well understood and that was computationally tractable. Breaker and co-workers had originally selected flavin-dependent hammerhead aptazymes from pools in which the joining region between an anti-flavin aptamer and the core catalytic domain of the hammerhead ribozyme were randomized (Soukup and Breaker, 1999a). At least one class of flavin-dependent hammerhead aptazymes was hypothesized to undergo a ‘slip structure’ conformational change in which the two strands of the joining region between the aptamer and the ribozyme realigned themselves (‘changed register’) upon the addition of flavin (Fig. 1a).

Herein, we describe the development of an algorithm for assessing the slip structure model, and the application of this algorithm to both designed aptazymes and designed aptamer beacons. Computational automation of the algorithm should allow the structures and minimum free energies of millions of different sequences to be computed in hours on a typical desktop computer. We have validated the utility of the algorithm by experimentally assaying a number of predicted constructs. In general, the model performs as well or better than empirical design, identifying aptazymes that are activated by up to 60-fold by small organic ligands and aptamer beacons that are activated by up to 10-fold by a protein ligand.

2. Methods and materials

2.1. Computational methods

The free energies of RNA secondary structures were calculated using the ViennaRNA package (Hofacker et al., 1994).

Scripts to automate the evaluation and comparison of RNA secondary structure profiles were coded in the scripting language Python (<http://www.python.org/>). These programs first generated nucleic acid sequences containing randomized regions: either a randomized joining region between an aptamer and a ribozyme, or a randomized region at one end of an aptamer. For each sequence in the randomized population, the free energy of folding was calculated along a series of ‘registers’, defined as sequential alignments of a randomized region with potential pairing partners in the native portions of an aptazyme or aptamer. Energetic profiles (free energies of folding versus register) were constructed based on these calculations. A flowchart that describes the computational selection of aptazymes is presented in Supplemental Fig. 1.

The aptazyme variants that best matched a defined profile were selected for further study. For example, the calculated energetic profiles were compared to the energetic profile of an experimentally validated reference aptazyme. Similarly, aptamer beacons were rank-ordered based on a number of criteria garnered from the *in silico* aptazyme selection (Supplemental Fig. 2).

2.2. *In vitro* transcription of aptazymes

Double-stranded DNA templates used for transcription were made by primer extension of complementary, synthetic DNA oligonucleotides (IDT, Coralville, IA). Radiolabeled RNA was generated by *in vitro* transcription using T7 RNA polymerase with trace amounts of [α - 32 P] UTP (Perkin-Elmer, Boston, MA), followed by purification on 10% denaturing polyacrylamide gels.

2.3. RNA synthesis of aptamer beacons

RNA was synthesized on an Expedite 8909 DNA synthesizer (PE Biosystems, Foster City, CA) using standard phosphoramidite chemistry. 3'-DABCYL (4-(4-dimethylamino-phenylazo)benzoic acid) and 5'-fluorescein (6-FAM) were added using a fluorescein phosphoramidite and a 3'-DABCYL controlled pore glass (CPG) column, respectively. All synthesis reagents were purchased from Glen Research (Sterling, Virginia). Products were purified by reverse phase chromatography or on denaturing polyacrylamide gels. The concentrations of aptamer beacons were determined based on their calculated extinction coefficients.

2.4. Aptazyme assays

Radiolabeled RNAs (20 pmol) were thermally equilibrated by denaturation at 70 °C in water followed by cooling to room temperature. Buffer (50 mM Tris-HCl, pH 7.5, 20 mM MgCl₂, final concentration) and ligand (FMN at 200 μ M, theophylline at 1 mM, or ATP at 1 mM). Reactions (20 μ L) were carried out at room temperature, and aliquots were removed at various time points and quenched in stop dye (85% formamide, 0.01% bromophenol blue, 0.01% xylene cyanol and 60 mM EDTA, pH 8.0). Uncleaved and cleaved ribozymes were separated on 8% denaturing polyacrylamide gels, and radiolabeled bands were

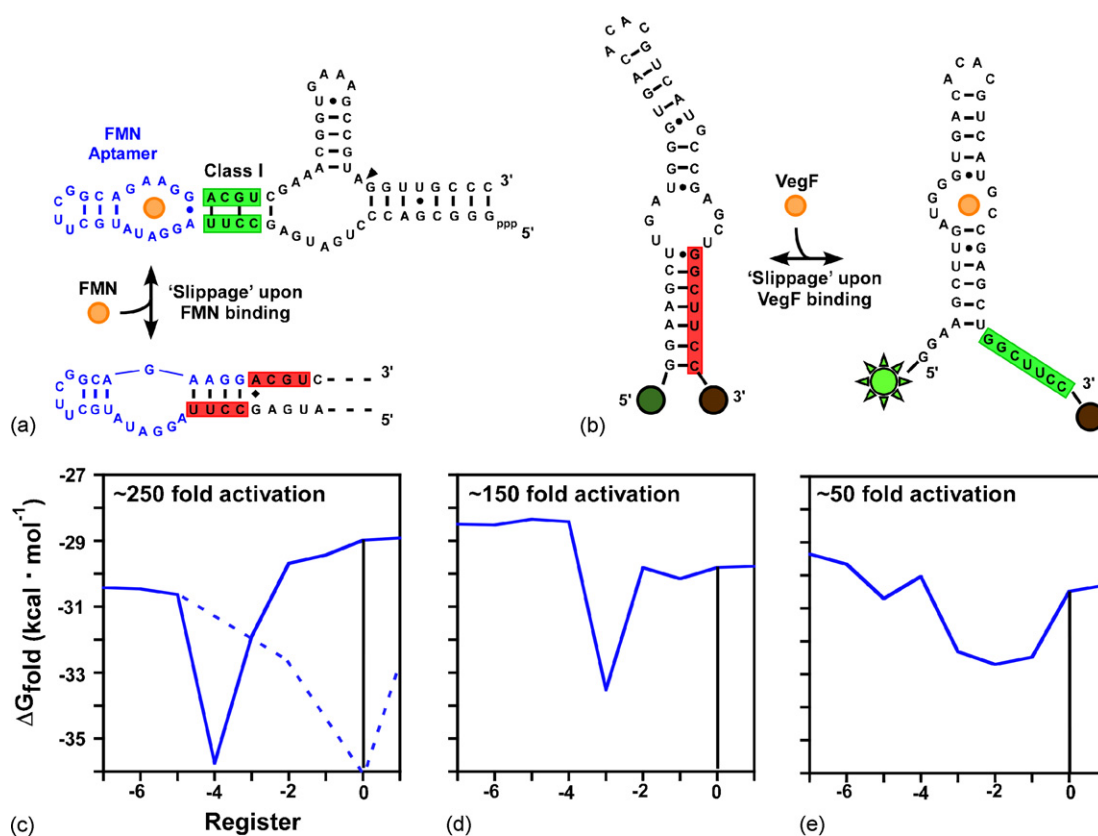


Fig. 1. Energetic profile model. (a and b) Schematic of 'slip structure' activation. In the absence of ligand (FMN or VegF), the joining region is trapped in an inactive conformation (red). A structural rearrangement occurs upon ligand binding, shifting the joining region to the 'active register' (green) and potentiating either catalysis (cleavage site denoted by filled arrow in (a) or binding). (c–e) Energetic profiles of three previously selected, FMN-responsive hammerhead aptazymes (Class I, Class III, and Class V respectively). The free energies of folding (ΔG_{fold}) in the absence of ligand are plotted versus the register of the joining region. Position '0' (vertical line) denotes the 'active' register of the ribozyme. The dashed line in (c) is a hypothetical energetic profile following ligand-binding.

quantitated using a PhosphorImager (Molecular Dynamics, Sunnyvale, CA). Rate constants were calculated by plotting the negative natural logarithm of the fraction of uncleaved ribozyme versus time, and determining the slope of a line fitted through at least three time points in the linear range of the assay. Assays were repeated twice or more; observed levels of activation differed by less than two-fold.

2.5. Aptamer beacon assays

Aptamer beacons (50 μ M) in buffer (PBS [137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂PO₄·7H₂O, 1.4 mM KH₂PO₄; pH 7.4] or TMK [100 mM Tris pH 7.5; 80 mM KOAc; 10 mM Mg(OAc)₂]; 30 μ L) were denatured at 80 °C for 3 min and cooled to room temperature over 5–15 min. Either an MJ Research DNA Engine Opticon (Waltham, MA) with LED fluorescence excitation at 495 nm and photomultiplier tube (PMT) emission detection at 518 nm or a Tecan Safire (San Jose, CA) fluorescent monochromator plate reader with excitation at 480/2 nm and emission detection at 520/2 nm were used to assay aptamer beacons. The thermal denaturation profiles for aptamer beacons were determined in the Opticon by obtaining fluorescence readings at temperatures from 25 to 90 °C in 1 °C increments and with equilibration at each temperature for 1 min.

In order to measure analyte activation, aptamer beacons were denatured as described above, then allowed to cool to room temperature over 10 min. Aptamer beacons were then heated to 30 °C in the Safire for 10 min, protein was added, and the fluorescence data was taken at 2-min intervals for up to 2 h. The proteins assayed included recombinant human VegF (R&D Systems, Minneapolis, MN), recombinant human bFGF (Bachem, King of Prussia, PA), recombinant human PDGF-AB (R&D Systems, Minneapolis, MN), recombinant human angiogenin (Bachem), recombinant human MEK-1 (Santa Cruz Biotechnology, Santa Cruz, CA), glycosylated ricin A chain (Sigma–Aldrich, St. Louis, MO), and bovine serum albumin (New England Biolabs, Ipswich, MA). All proteins were initially reconstituted in the buffers described by their manufacturers.

3. Results and discussion

3.1. Development and computational assessment of the slip structure model

Aptazymes have been designed and selected in which ligand-binding aptamer domains are connected to the catalytic domain via a stem or 'joining region'. Breaker and co-workers have previously postulated that for some aptazymes the mechanism of activation involves a 'slip structure' in which ligand-binding

induces a local reorientation of a stem and a corresponding change in ribozyme activity. The slip structure model tacitly assumes that there are at least two conformational states for an aptazyme: an unliganded, inactive conformer in which the joining region was ‘out-of-register’, and a liganded, active conformer in which the joining region was ‘in register’. In this model, the actual conformational change between the two structures could involve sequential changes in register: the stepwise movement of one strand of a joining region relative to the other (Fig. 1a).

In order to assess whether there might be one or more stable, out-of-register conformers for previously selected FMN-aptazymes, we constructed a plot of the Gibbs energy of folding (ΔG_{fold}) for each possible base-pair register within their joining regions (Fig. 1c–e). An ‘active register’ (denoted ‘0’) was first defined based on the known secondary structure of the hammerhead ribozyme. All conformations near the active register were then computationally generated by simply shifting the lower (5'-most) strand of the joining region relative to the upper (3'-most) strand, and determining the predicted folding energy at each register (see also Supplemental Fig. 1). Conceptually, the energetic profiles that were generated describe the changes in free energy that the aptazymes would undergo as they ‘slipped’ from one register to the next.

The free energy profiles generated for the different flavin-activated hammerhead aptazymes are consistent with activation via a slip structure mechanism. In the previously selected Class I aptazyme one strand of the joining region (CCUU) is predicted to base-pair with a complementary portion of the anti-FMN aptamer (Fig. 1a). The assumption of this ‘inactive register’ effectively disrupts the structure of both the hammerhead and the anti-FMN aptamer (Fan et al., 1996). The free energy of flavin-binding would stabilize the formation of a different stem structure containing only two Watson–Crick base-pairs in the joining region. In this ‘active register’ aptamer residues important for flavin-binding (such as the terminal A–G pairing, Fan et al., 1996) would no longer be sequestered and the hammerhead ribozyme could in turn assume an active conformation. The energetic profiles of two other flavin-activated hammerhead aptazymes with different joining regions (Classes III and V) were also constructed. As was the case with the Class I aptazyme, out-of-active-register free energy minima were observed (Fig. 1d, e). The free energy of FMN binding to the aptamer (-8.6 kcal/mol) is more than enough to overcome the predicted energetic differences between the active and inactive registers.

The prediction that there are active and inactive registers is consistent with the available structural data. A hammerhead aptazyme that contained an anti-FMN aptamer and a Class I communication module was probed using an ‘in-line assay’ that exploited conformation-dependent differences in hydrolysis (Jose et al., 2001; Soukup and Breaker, 1999b). Flexible residues, such as those not ensconced in helices, show greater cleavage activity than residues that retain a fixed conformation. The residues that we predict to be unpaired in the unliganded, inactive register were in fact cleaved (see Fig. 4 in Jose et al., 2001). Conversely, these residues should assume a more stable

conformation in the liganded, active register, and the cleavage patterns were found to be correspondingly suppressed.

The predicted free energy profiles are not only consistent with activation via a slip structure mechanism, but could also account for the relative levels of activation. The Class I aptazyme was predicted to have the greatest energetic separation between the stable active and inactive registers at nearly 7 kcal/mol (Fig. 1c), and also showed the greatest effector-activation. Class III and Class V have smaller predicted energy barriers of 4 and 2.5 kcal/mol respectively, and exhibited correspondingly smaller activations. Both the height of the energy barrier and the distance between the stable inactive register and the less stable active register may affect the level of aptazyme activation.

To further validate the utility of using energetic descriptions for slip structure mechanisms, energetic profiles for previously selected hammerhead aptazymes (Soukup and Breaker, 1999a) that were inhibited rather than activated by FMN were constructed (data not shown). As expected, there were energetic minima at the active registers and less stable structures in adjacent registers; in other words, the predicted energy profiles were the converse of the energetic profiles predicted for flavin-activated aptazymes. Presumably, FMN-binding leads to a reorganization of the joining region such that an inactive register becomes favored relative to the active register.

3.2. Computational selection of allosteric ribozymes

To the extent that the slip structure model can be used to explain the activation parameters of known hammerhead aptazymes, it should prove possible to use this model to predict the sequences of new hammerhead aptazymes. To this end, we developed a simple algorithm to select slip structure aptazymes from computationally generated random sequence populations. Energetic profiles similar to those shown in Fig. 1 were computed for each member of a pool in which the hammerhead ribozyme was joined to the anti-FMN aptamer via a symmetric N3–N3 joining region ($4^6 = 4096$ variants; Fig. 2a). The energetic profiles of each candidate in the pool were compared to a reference energetic profile of a known hammerhead aptazyme (Class VI; Fig. 2a, blue line). Aptazyme variants that had similar energy profiles to the Class VI hammerhead aptazyme were selected for further analysis. A total of 23 candidate aptazyme sequences were chosen that most closely matched the energy profile Class VI aptazyme; the energetic profiles of two of these aptazymes are shown in Fig. 2a. These energy profiles differed substantially from the norm. The black line in Fig. 2a corresponds to the average energetic composition of the un-selected pool.

All 23 designed aptazymes were synthesized and the aptazyme variants were assayed *in vitro*. Since these aptazymes were predicted to structurally rearrange from an inactive to an active conformation only in the presence of effector, they should show substantially greater cleavage activity only in the presence of FMN. Three variants exhibited greater than 4-fold activation in the presence of FMN (Table 1), and two of these, c23 and c52, exhibited activations ($k_{\text{obs}}^+/k_{\text{obs}}^-$) of 21- and 60-fold,

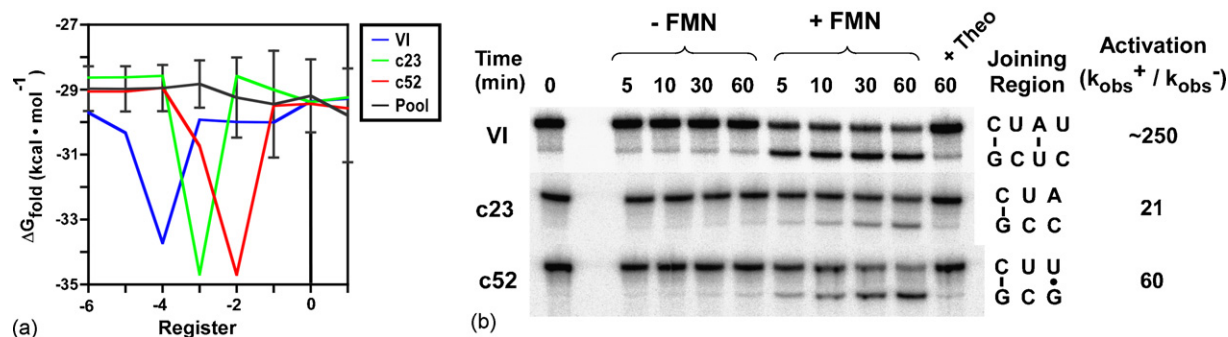


Fig. 2. Computationally selected hammerhead aptazyme variants. (a) The energetic profiles of every member of a pool in which an anti-flavin aptamer was joined to the hammerhead ribozyme via a symmetric, N3–N3 stem were compared to a reference energy profile (Class VI hammerhead aptazyme, blue line). The average energetic profile and distribution (error bars) of the pool are shown in black. The energetic profiles of two candidate aptazymes are shown in green (aptazyme c23) and red (aptazyme c52). (b) Activity assays with selected variants. Cleavage as a function of time following the addition of flavin was analyzed by polyacrylamide gel electrophoresis. Non-specific activation by theophylline (Theo) was also monitored. The sequences of the joining regions for the aptazymes and the calculated levels of activation in the presence of FMN are listed to the right.

respectively (Fig. 2b). Of the remaining aptazymes 8 were constitutively inactive and 12 were constitutively active.

A control was carried out in which 23 aptazyme variants were selected at random from the population. Three were slightly activated by FMN, although the best variant was activated only 5-fold (Table 1), as compared with over 20-fold for the computationally selected aptazymes. In addition, five of the randomly selected aptazyme variants were found to be inactivated by FMN, whereas no computationally selected aptazymes were inactivated.

3.3. Varying the aptamer sensor domain

Some joining regions identified by *in vitro* selection experiments have been shown to act as ‘communication modules’ that can lead to the activation of ribozymes by many different aptamers (Soukup and Breaker, 1999a). We therefore replaced the anti-flavin aptamer on the two best computationally selected aptazymes with an anti-theophylline aptamer which should in turn alter the aptazyme’s effector specificity from FMN

to theophylline. One of the computationally selected joining regions (c23) conferred roughly the same activation by theophylline (17-fold); the other joining region (c52) showed no activation.

Given this success and the fact that theophylline-activated hammerhead aptazymes had previously been selected (Jenison et al., 1994; Soukup et al., 2000), the anti-theophylline aptamer was computationally appended to a N3–N3 pool, and the pool was sieved using energy parameters identical to those that had been used to identify flavin-dependent aptazymes. Four other sequences were identified as candidate aptazymes, and were assayed for theophylline responsivity (Table 1). A single variant from the four exhibited a 14-fold increase in rate upon the addition of theophylline.

3.4. Varying the length of the joining region

Since computational selections with short joining regions had proved successful, we were interested whether longer, more complex joining regions could also be sieved. The anti-

Table 1
Summary of computationally identified and experimentally tested aptazymes

Pool	Ribozyme	Aptamer	State	Register ^a	Off ^b	On ^c	Activated	Inhibited	Identified ^d	Tested ^e
N3–N3	HH	FMN	Random	n/a	3	12	3	5	–	23
N3–N3	HH	FMN	Selected	–4	8	12	3	0	23	23
N3–N3	HH	Theo	Selected (1)	–4	0	3	1	0	4	4
N3–N3	HH	Theo	Selected (2)	–8	4	19	1	0	44	24
N4–N4	HH	Theo	Random	n/a	3	7	0	0	–	10
N4–N4	HH	Theo	Selected	–8	17	8	2	0	48	27
N5–N5	HH	Theo	Random	n/a	4	6	0	0	–	10
N5–N5	HH	Theo	Selected	–8	23	1	1	0	11,637(68)	25
N3–N3	Hairpin	FMN	Selected	–4	12	0	1	0	13	13
N5–N5	HH	FMN	Selected	n/a	12	1	2	0	8607	15

HH: Hammerhead.

^a ‘Register’ indicates the required distance between the stable active and inactive registers. This distance could vary up to two positions in the algorithm, as stated in Fig. 2.

^b ‘Off’ indicates ribozymes that were found to be constitutively inactive.

^c ‘On’ indicates ribozymes that were found to be constitutively active.

^d ‘Identified’ gives the number of ribozymes picked by the algorithm.

^e ‘Tested’ gives the number of ribozymes synthesized and assayed.

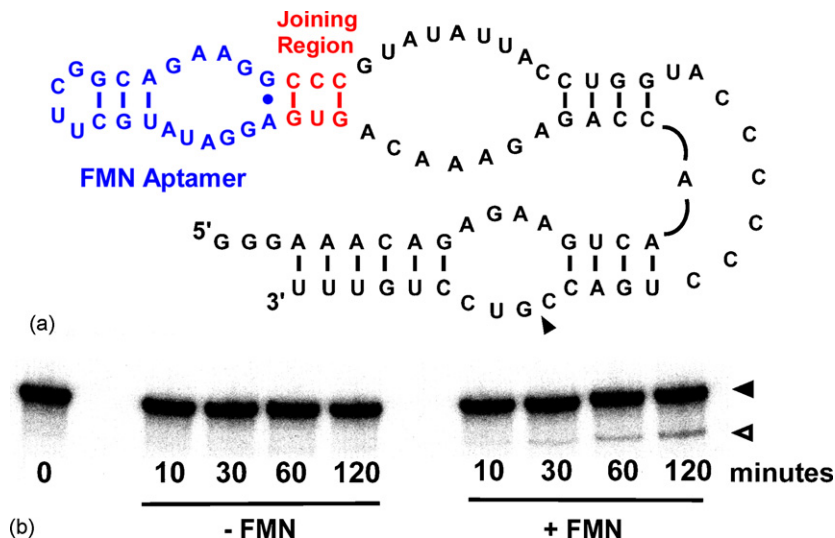


Fig. 3. Computationally selected hairpin aptazyme variants. (a) The hairpin ribozyme (black, cleavage site marked with filled arrow) was modified with a randomized stem region (red) and an anti-FMN aptamer (blue). Selection identified an aptazyme whose joining region facilitated activation in the presence of FMN. (b) Activity assay with the selected variant. The filled and open arrows mark the positions of uncleaved and cleaved products, respectively.

theophylline aptamer was computationally joined to the hammerhead ribozyme by both N4–N4 or N5–N5 stems and the resultant pools were sieved using the slip structure energy model. However, the distance between the active and inactive registers was changed from 2–4 to 4–8, reflecting the best-activated variants from the experimental N5–N5 selection (Table 1). From the pool with a N4–N4 joining region 48 candidates were identified and 27 were experimentally assayed for self-cleavage in the presence of theophylline. However, from the pool containing a N5–N5 joining region, 11,637 sequences (representing 1.1% of the original pool) were identified, despite the fact that the same parameter sets were utilized. The complexity of this population was reduced by taking only one variant from selected sets whose 5'-most N5 regions were identical. This left 68 candidate sequences, and 25 were experimentally assayed. Two variants from the N4–N4 pool were found to be activated by theophylline (9- and 20-fold), and one clone from the N5–N5 pool was theophylline-responsive (16-fold).

To again ensure that we had not identified aptazymes by chance alone, 10 designs from the N4–N4 and N5–N5 pools were picked at random from these pools, synthesized, and assayed for activation in the presence of theophylline. Not only were no effector-dependent variants found, but the distribution of active and inactive variants was very different between the randomly selected and computationally sieved sequences (Table 1). The randomly selected sequences favored constitutively active variants, while the computationally sieved sequences contained many fewer constitutively active variants than inactive variants.

3.5. Varying the ribozyme response domain

While various aptamer domains could be easily interchanged using the slip structure model, we also wanted to test whether the same model could be generalized to other catalytic platforms. The hairpin ribozyme was chosen as a candidate for engineer-

ing because like the hammerhead its structure also consisted of stacked helices (Rupert and Ferre-D'Amare, 2001).

The anti-FMN aptamer was appended to the end of loop B of the hairpin ribozyme via a symmetric N3–N3 random region, and 13 candidate aptazymes were identified using the same slip structure energy model that had proven successful for hammerhead aptazymes. All 13 sequences were then assayed *in vitro* for their ability to undergo cleavage in the presence of FMN. A single sequence was found to be activated 10-fold in the presence of FMN (Fig. 3).

3.6. Computational selection of signaling aptamers

Following up on our success using the slip structure energy model to computationally predict aptazyme sequences, we sought to use the model to generate conformation-switching aptamers that could be used as optical biosensors. A non-binding or inactive register was generated by manipulating the terminal stem of a hairpin-like aptamer. In the model, the addition of ligand should shift the register of the aptamer stem from the non-binding conformation to the original binding conformation. Instead of detecting ligand-activated ribozyme cleavage, ligand-binding was coupled to structural changes that induced fluorescent signaling, much like a molecular beacon (Fig. 1b).

We also attempted to extend the range of analytes that could activate computationally predicted biosensors from small organic molecules (such as theophylline and FMN) to proteins (such as vascular endothelial growth factor (VegF)). A series of random sequence libraries of four to nine residues in length (349,440 variants total) were created *in silico* and appended to the 3'-end of an anti-VegF aptamer 44T (Jellinek et al., 1994). We chose this aptamer because of the structural similarities to the FMN aptamer including a stem loop and internal bulge. Minimum free energies were then calculated for each register of each member of the library.

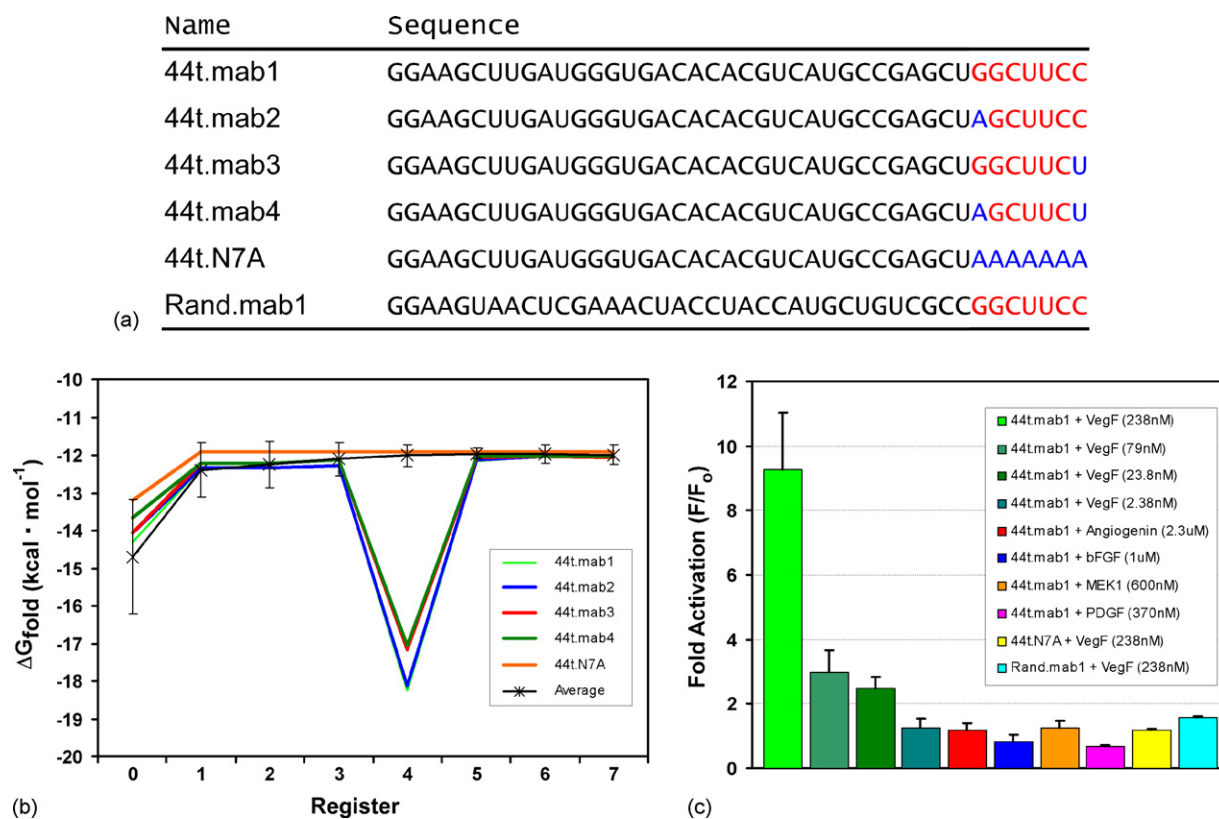


Fig. 4. Design of a VegF-responsive aptamer beacon. (a) Following computational selection (as described in the text), four variants (44t.mab1-mab4) were identified from a N7 pool that had an energy minima at register 4. The sequences of these variants are shown, along with two negative controls (44t.N7A and Rand.mab1). The aptamer beacon 44t.mab1 was used in experimental studies. The randomized region of 44t.mab1 is shown in red, and residues in other aptamer beacons that differ from 44t.mab1 are shown in blue. (b) Energetic profiles for selected aptamer beacons. In addition to the energetic profiles for 44t.mab1-mab4 and 44t.N7A, the average energetic profile of all 16,384 variants is shown in black with error bars representing the standard deviation. (c) Sensitivity and selectivity assays of the designed aptamer beacon. The aptamer beacon 44t.mab1 (2 μ M) was assayed at four VegF concentrations and with four other proteins involved in angiogenesis (concentrations shown). Also, negative controls were assayed with VegF (238 nM). These experiments were performed in 30 μ L reactions on an MJ Opticon at 37 $^{\circ}$ C. VegF-responsive aptamer beacons were at 2 μ M final concentration and readings were taken after 60 min.

Since no VegF-activated aptamer beacon existed, no reference energy profile was available to sieve putative sequences from the library. Therefore, we devised a set of criteria that could be sequentially applied (see also Supplemental Fig. 2): (i) we first rank-ordered the sequences based on how well-defined the local minima were at the inactive register. Those aptamer beacons in which the free energy of folding at the inactive register differed most from the free energies of folding for the registers to either side were ranked most highly (see also Fig. 4b). By choosing well-defined local minima we hoped to reduce folding transitions in the absence of ligand and hence background signal (Flamm et al., 2001). (ii) We then chose an energetic difference between the active and inactive registers that we hoped would lead to an increase in signal in the presence of ligand. An arbitrary cutoff value of at least 3.5 kcal mol $^{-1}$ energy difference between the active and inactive registers ($\Delta G_{\text{fold,inactive}} - \Delta G_{\text{fold,active}}$) was initially chosen. This arbitrary choice can be further refined by experimental feedback, as described below.

Based on these criteria, no variants from the N4, N5 or N6 pools were predicted to function as aptamer beacons. A total of 7 sequences were selected from the N7 pool (Fig. 4a), 58 from the N8 pool, and 514 from the N9 pool. The increase in predicted

active variants roughly corresponds to the increase in pool complexity. In general, we found that the largest energy differences between the active and inactive registers (criterion (ii), above) did not exceed 7 kcal mol $^{-1}$ ($\Delta G_{\text{fold,inactive}} - \Delta G_{\text{fold,active}}$). This energy difference should have been more than adequate for conformational transduction, as the binding energy for VegF has been measured to be -12.7 kcal mol $^{-1}$ (Jellinek et al., 1994).

Because of the difficulties inherent in the synthesis of long, derivatized RNA molecules, we wished to choose a single variant that would have the best chance of showing ligand-dependent signaling. Of the seven sequences chosen from the shortest, most synthetically accessible pool (N7), only four could form terminal base-pairings that would place fluorophores and quenchers adjacent to one another. Of these four, 44t.mab1 had the best-defined local minimum (criterion (i), $\Delta G_{\text{fold}} = -5.97$ kcal mol $^{-1}$; Fig. 4b); and had the second largest energetic differences between the active and inactive registers (criterion (ii), $\Delta G_{\text{fold,inactive}} - \Delta G_{\text{fold,active}} = 3.91$ kcal mol $^{-1}$); and also formed beacon like terminal base pairs in the inactive conformer.

We synthesized 44t.mab1 as an aptamer beacon with fluorescein and DABCYL incorporated at the 5' and 3' ends,

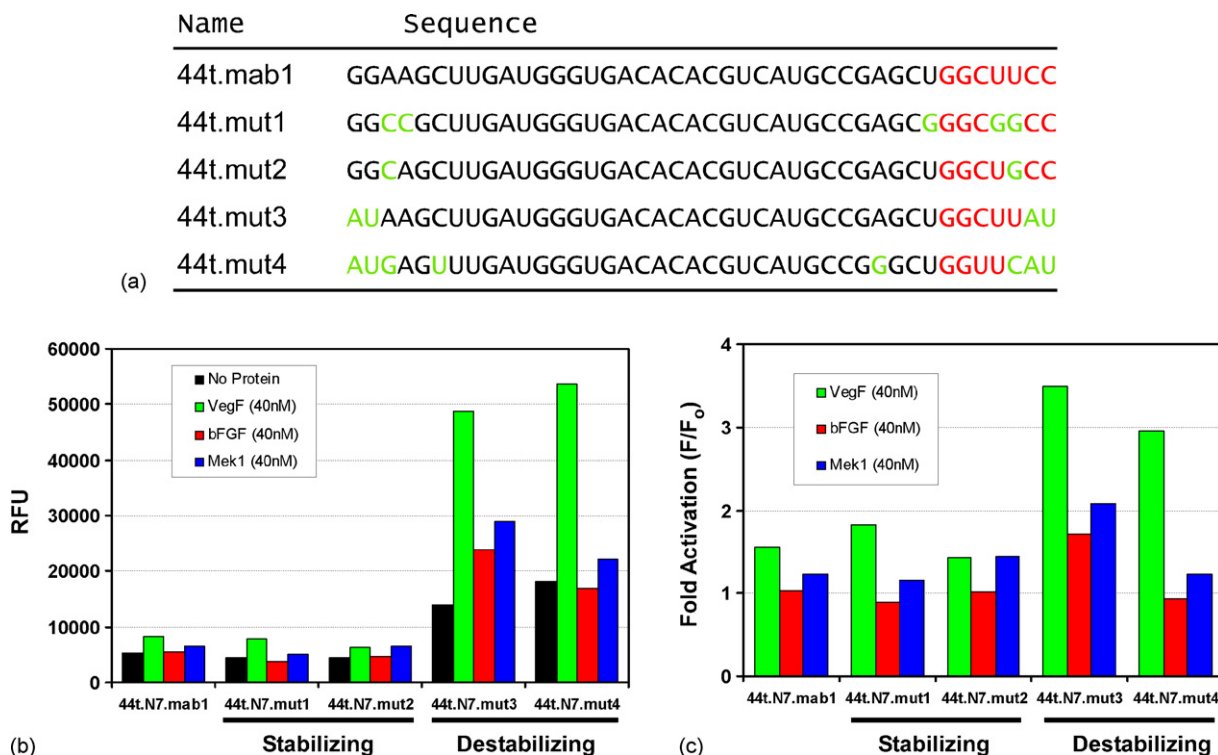


Fig. 5. Modulating the performance of the VegF-responsive aptamer beacon. (a) A series of predicted stabilizing (44t.mut1-2) or destabilizing mutations (44t.mut3-4) of 44t.mab1 were designed. The randomized region of 44t.mab1 is red, and mutational changes are green. (b) Observed fluorescence (RFU) of 44t.mab1 and its mutations with various angiogenesis proteins. (c) Same results as (b), but represented in terms of fold fluorescence increase in the presence of protein. Reactions were carried out in TMK buffer on a Tecan Safire at 40 °C. Aptamer beacons were at 200 nM final concentration and readings were taken after 100 min.

respectively. In the absence of VegF, the aptamer should form a structure in which DABCYL quenches the fluorescein, while in the presence of VegF the aptamer should form a structure in which fluorescence is not quenched. The computationally predicted signaling aptamer was activated 10-fold by 238 nM VegF (Fig. 4c), and displayed optimum activity at 37 °C (data not shown). These results were especially impressive considering that other experimentalists have observed only three- to four-fold activation at similar protein concentrations when examining similar constructs designed ‘by hand’ (Nutiu and Li, 2003). The aptamer beacon was both sensitive and selective for VegF. The beacon displayed a two-fold fluorescence increase at analyte concentrations as low as 24 nM ($1 \mu\text{g ml}^{-1}$), and was not activated by non-cognate growth factors including PDGF, bFGF, or angiogenin (Fig. 4c).

The kinetics of conformational change and activation were, however, relatively slow: it took 15 min to reach 50% activation and about 75 min to reach full activation (data not shown). The kinetics of activation are consistent with a model in which the rearrangement of the engineered stem precedes protein binding. The observed responsivity is comparable to the typical time required for the development of signals from protein-binding aptamer microarrays and much faster than the time required for antibody microarrays that typically require complicated post-binding procedures (such as the binding of secondary reagents in sandwiches and wash steps (Collett et al., 2005)). Thus designed signaling aptamers should be compatible with high-throughput, chip-based analyses.

3.7. Mutational analysis of the computationally designed signaling aptamer

To better assess how the selection criteria related to experimental performance, we rationally introduced mutations into the VegF signaling aptamer that were predicted to either stabilize or destabilize the inactive conformer (Fig. 5a). As can be seen in Fig. 5b, stabilizing mutations had no effect on activation, suggesting that the criteria used for the selection of the inactive register overpredicted stability. Consistent with this interpretation, the destabilizing mutations increased specific activation from approximately 1.5-fold at low (40 nM) protein concentrations to 3-fold (Fig. 5c). As expected, the destabilizing mutations also increased background fluorescence (pre-formation of the active conformer) prior to protein addition. The initial successes of the slip structure model now provide a basis for further determining how systematically varying the free energy parameters for computationally derived biosensors affects their overall performance.

4. Conclusion

While *in vitro* selection methods are extremely robust, the ability to use computational design methods for the generation of biosensors would be extremely valuable for a variety of reasons. First, the ability to design nucleic acid biosensors further validates the utility of methods used for the computational prediction of nucleic acid secondary structures. In particular, by showing

that a slip structural model accords with both computational predictions and experimental data it may prove possible to use similar models to better understand conformational transitions in both natural and selected functional nucleic acids. Second, computational selection should be inherently faster than experimental selection methods, and thus the time and effort required for the development of biosensors may be greatly decreased. Based on the methods described in this paper, only a few sequences had to be synthesized and assayed to identify functional biosensors, while following selection experiments numerous different sequences typically have to be selected, characterized, and screened for their ability to function as biosensors. Ultimately, the computational sieving algorithm can be employed within hours to quickly generate a set of aptazymes or aptamer beacons corresponding to any of the large number of aptamers that have already been found. Coupled with high-throughput, chip-based nucleic acid synthesis (McGall and Fidanza, 2001; Nuwaysir et al., 2002) it may therefore prove possible to rapidly generate biosensor chips for the detection of multiple analytes. Lastly, to the extent that computational models match experimental results, it should be possible to finely control the design process, generating biosensors with optimal sensitivities, signal-to-noise ratios, and dynamic ranges.

Acknowledgements

The authors would like to thank Lauren Ancel Myers for helpful discussions and knowledgeable insights. J.H. and B.H. would like to acknowledge partial financial support from an Integrative Graduate Education and Research Traineeship (IGERT) and the National Science Foundation. This work was funded by a grant from the NIH.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bios.2006.08.019.

References

- Breaker, R.R., 2002. *Curr. Opin. Biotechnol.* 13, 31–39.
- Collett, J.R., Cho, E.J., Ellington, A.D., 2005. *Methods* 37, 4–15.
- Ding, Y., Lawrence, C.E., 2003. *Nucleic Acids Res.* 31, 7280–7301.
- Dirks, R.M., Lin, M., Winfree, E., Pierce, N.A., 2004. *Nucleic Acids Res.* 32, 1392–1403.
- Fan, P., Suri, A.K., Fiala, R., Live, D., Patel, D.J., 1996. *J. Mol. Biol.* 258, 480–500.
- Fang, X., Sen, A., Vicens, M., Tan, W., 2003. *ChemBiochemistry* 4, 829–834.
- Flamm, C., Hofacker, I.L., Maurer-Stroh, S., Stadler, P.F., Zehl, M., 2001. *Rna* 7, 254–265.
- Hamaguchi, N., Ellington, A., Stanton, M., 2001. *Anal. Biochem.* 294, 126–131.
- Hofacker, I.L., Fontana, W., Stadler, P.F., Bonhoeffer, L.S., Tacker, M., Schuster, P., 1994. *Monatsh. Chem.* 125, 167–188.
- Jellinek, D., Green, L.S., Bell, C., Janjic, N., 1994. *Biochemistry* 33, 10450–10456.
- Jenison, R.D., Gill, S.C., Pardi, A., Polisky, B., 1994. *Science* 263, 1425–1429.
- Jose, A.M., Soukup, G.A., Breaker, R.R., 2001. *Nucleic Acids Res.* 29, 1631–1637.
- McCaskill, J.S., 1990. *Biopolymers* 29, 1105–1119.
- McGall, G.H., Fidanza, J.A., 2001. *Meth. Mol. Biol.* 170, 71–101.
- Nutiu, R., Li, Y., 2003. *J. Am. Chem. Soc.* 125, 4771–4778.
- Nutiu, R., Li, Y., 2005. *Angew. Chem. Int. Ed. Engl.* 44, 1061–1065.
- Nuwaysir, E.F., Huang, W., Albert, T.J., Singh, J., Nuwaysir, K., Pitas, A., Richmond, T., Gorski, T., Berg, J.P., Ballin, J., et al., 2002. *Genome Res.* 12, 1749–1755.
- Penchovsky, R., Breaker, R.R., 2005. *Nat. Biotechnol.* 23, 1424–1433.
- Robertson, M.P., Knudsen, S.M., Ellington, A.D., 2004. *Rna* 10, 114–127.
- Roth, A., Breaker, R.R., 2004. *Meth. Mol. Biol.* 252, 145–164.
- Rupert, P.B., Ferre-D'Amare, A.R., 2001. *Nature* 410, 780–786.
- Silverman, S.K., 2003. *Rna* 9, 377–383.
- Soukup, G.A., Breaker, R.R., 1999a. *Proc. Natl. Acad. Sci. U.S.A.* 96, 3584–3589.
- Soukup, G.A., Breaker, R.R., 1999b. *Rna* 5, 1308–1325.
- Soukup, G.A., Emilsson, G.A., Breaker, R.R., 2000. *J. Mol. Biol.* 298, 623–632.
- Stojanovic, M.N., de Prada, P., Landry, D.W., 2001. *J. Am. Chem. Soc.* 123, 4928–4931.
- Tang, J., Breaker, R.R., 1997. *Chem. Biol.* 4, 453–459.
- Verma, S., Jager, S., Thum, O., Famulok, M., 2003. *Chem. Rec.* 3, 51–60.
- Zuker, M., 2000. *Curr. Opin. Struct. Biol.* 10, 303–310.