In vitro Selection of Allosteric Ribozymes: Theory and Experimental Validation

Nicolas Piganeau1,2, Vincent Thuillier2* and Michael Famulok1*

1Kekule Institut für Organische Chemie und Biochemie, Rheinische Friedrich-Wilhelms Universität Bonn, Gerhard-Domagk-Strasse 1, 53121 Bonn, Germany
2Aventis Pharma, GenCell, 3825 Bay Center Place, Hayward CA 94545, USA

In vitro selection techniques offer powerful and versatile methods to isolate nucleic acid sequences with specific activities from huge libraries. We describe an in vitro selection strategy for the de novo selection of allosteric self-cleaving ribozymes responding to pefloxacin and other quinolone derivatives. Within 16 selection cycles, highly sensitive clones responding to drug levels in the sub-micromolar range were obtained. The morpholine moiety of the quinolone derivatives was required for inhibition of the self-cleavage of the selected ribozymes: modifications of the aromatic system were tolerated better than modifications of the morpholine ring. We also present a theoretical model that analyzes the predicted fraction of ribozymes with a given binding constant and cleavage rate recovered after each selection cycle. This model precisely predicts the actual experimental values obtained with the selection procedure. It can thus be used to determine the optimal conditions for an in vitro selection of an allosteric ribozyme with a desired dissociation constant and cleavage rate for a given application.

Keywords: ribozymes; aptamers; RNA/drug interaction; allosteric selection; selection modelization

Introduction

Gene expression may be regulated at every step of the pathway, from the initiation of transcription, to the activity of the protein itself. The most studied and seemingly major stage at which gene expression is regulated in the cell is the initiation of transcription. However, control of gene expression after this step can occur by regulating transcriptional rate, capping, splicing, polyadenylation, cellular localization, stability of the mRNA, translation initiation, translational rate, and finally protein modification. All of these events can be tightly regulated by external signals.1

Controlling the expression of a gene offers a broad range of potential applications, both in research and for clinical purposes.2 First, modulating the level of gene expression is an effective way to study its function. Both knockout and overexpression of genes of interest have long been powerful tools for understanding a gene’s biological role. However, in many cases these methods are limited by several factors: signaling molecules or pathways are often redundant and compensate for the under or over-activity of each other, while constitutive expression of many genes is cytotoxic or cytostatic. To circumvent such problems, the development of conditional gene expression systems that can be temporally or spatially regulated is needed.3

We have recently begun to develop systems that might allow the regulation of gene expression in response to an externally administered regulatory drug. The basis for such a system is allosteric ribozymes, which can specifically cleave their own RNA in the absence of the regulatory drug and can be inactivated in the presence of the regulatory drug.4 Thus, if such ribozymes were suitably inserted within mRNA transcripts, they would lead to its rapid degradation, unless the regulatory drug is present. Such catalytic RNAs can be rationally designed using known ribozymes such as the hammerhead ribozyme (HHR) and regulatory sequences, the so-called aptamers5,6 that bind small molecules by adaptive binding mechanisms.7,8 Alternatively, new effector-binding domains that activate or inhibit a ribozyme allosterically can be isolated by in vitro selection methods9,10 Using this method, Koizumi et al.9 identified allosteric ribozymes activated by cyclic nucleotide monophosphates at concentrations of around 100 μM. Upon
activation, the catalytic rates of these ribozymes were activated up to 5000-fold, and reached activities similar to the wild-type HHR. Conversely, we recently applied an in vitro selection scheme to obtain allosteric ribozymes that respond to the antibiotic doxycycline. While the level of inhibition was around 50-fold, the response could be achieved at very low concentrations with inhibition constants as low as 20 nM.

In the present study we applied an in vitro selection scheme similar to that described previously to select for an allosteric hammerhead ribozyme derivative inhibited by pefloxacin, a drug chosen a priori for its potential ability to control gene expression in vivo. Such a drug must meet several criteria: have a good bio-distribution and enter the cells properly, be as innocuous as possible, and present features suggesting good RNA binding properties. Pefloxacin and other quinolone derivatives contain unsaturated rings, as well as ketone or hydroxyl groups susceptible to interaction with RNA. Pefloxacin exhibits relatively low toxicity with an i.p. LD$_{50}$ in rats of 1.5 g/kg, is well distributed within the organism, and can permeate cells. Quinolones therefore seemed to be good candidates for both biological purposes and effective interaction with RNA.

The effective control of gene expression by an allosteric ribozyme that is inhibited by a specific drug relies on the self-cleavage rate of the ribozymes ($k$) and on the affinity of the drug towards the selected ribozyme ($K_D$). We reasoned that these two critical parameters result from the conditions applied during the in vitro selection of allosteric ribozymes, and therefore that we should be able to predict the efficiency of recovery of ribozymes with given $k$ and $K_D$. We report the description of such a mathematical model, discuss the fit between the theoretical predictions, and the actual characteristics of the selected ribozymes and show that this model can be used to enhance the effectiveness of the selection of allosteric ribozymes and aptamers.

### Results

#### Pool design

Allosteric hammerhead ribozymes have been rationally designed in previous studies by fusing an aptamer RNA onto helix II. The stability of helix II was previously reported as being crucial for hammerhead ribozyme activity. In addition, this helix is advantageous over positions located downstream of the hammerhead ribozyme sequence, since the putative binding site of the potential inhibitor would be very close to the hammerhead catalytic core (see Figure 1). We therefore chose helix II for positioning the randomized sequence. Since the desired ribozymes need to be active in the absence of the drug, it was important to conserve as much ribozyme activity as possible in the initial pool. This requirement is in contention with the fact that the potential drug-binding pocket should be as close as possible to the hammerhead catalytic core for inhibition to occur. In keeping with these conflicting requirements, we reduced helix II to a two base-pair stem and substituted the remainder of stem-loop II by 40 randomized nucleotides (Figure 1).

The randomized pool was constructed by employing two successive large-scale PCRs using a Reaction...

---

**Figure 1.** (a) Secondary structure of the transcripts from the initial pool. Helix II was shortened to two base-pairs, and loop II was replaced with a 40 nt random region. The arrow indicates the cleavage site. The chemical link between the biotin moiety and the RNA used during the selection is also shown. (b) Pool amplification scheme. A first large-scale PCR was applied to the template Pnp-pool using the amplification primers Pnp-1 and Pnp-rev. In a second step, the full-length transcriptable pool was produced by PCR with primers Pnp-3 and Pnp-rev. The sequence of the primers is given in Materials and Methods.
the primers shown in Figure 1(b). After synthesis of the double-stranded DNA pool, its complexity was evaluated as 6 nmol of initial single-stranded oligonucleotides, i.e. $10^{15}$ different molecules. This pool was transcriptionally competent, producing RNA molecules of the expected size, and showed a cleavage efficiency of 11% after overnight transcription as quantified by denaturing gel electrophoresis (data not shown). Subsequently, the selection process will actually involve only 11% of the molecules of the initial pool, since our procedure selects only ribozymes that are active in the absence of the drug (see below). No effect of 1.0 mM inhibitory molecules on the initial pool could be detected after transcription under selection conditions.

**In vitro selection**

**In vitro** selection was performed as described. Briefly, the RNA was transcribed in the presence of the inhibitory drug to avoid cleavage, and was biotinylated at its 5' end after T7-transcription. After immobilization on streptavidine agarose, the pool was incubated in the presence of the inhibitory molecule. At the end of this incubation, the bound drug ligand and cleaved ribozyme products were removed by denaturing washing steps. The remaining ribozymes on the column were subjected to a second incubation, without inhibitor. Finally, the cleaved RNA molecules were eluted, purified, reverse-transcribed, and amplified by PCR before use in the next selection cycle. The design of the PCR primers allowed the restoration of the 5' cleaved fraction of those HHR that were left bound to the column plus the T7 promoter. The resulting DNA was used for the next selection cycle.

Figure 2 shows the course of the selection in direct comparison to a previously performed selection that used doxycycline as the inhibitory drug. Remarkably similar results were obtained with pefloxacin. The elution profile remained constant during the first three cycles. However, in the fourth cycle a significant decrease in the amount of RNA eluted during the first incubation plus inhibitor was observed. At the same time, the percentage of RNA eluted during the second incubation increased. This result was repeated during cycles 5 and 6. However, a control selection cycle in the absence of inhibitor molecule revealed that the selected pools responded only poorly to their respective ligand. For example, in the presence of doxycycline, 14% of the RNA was eluted in selection step 4, versus 10% in the negative control without regulatory drug. To confirm this result, we assessed the time-course of self-cleavage with the pools after cycle 5 with or without the inhibitor molecules. In the case of the pefloxacin pool, a slight inhibition was observed, as illustrated in Figure 3.

In contrast, no inhibition at all could be seen in the pool selected for doxycycline, a phenomenon that can be explained by the enrichment of parasitic sequences. These are sequences that can fold into different conformations, some cleavage-competent, some not. After the denaturation-renaturation step to remove bound ligand and cleaved RNA, some parasitic sequences refold into active cleavage conformations and are eluted together with ligand-inhibited species.

![Graph](image)

**Figure 2.** Percentages of eluted radioactivity at the end of each selection cycle. The radioactivity eluted after the second incubation for the doxycycline (filled circles) or the pefloxacin (filled squares) selections are shown. From cycle 6, a negative control experiment of the respective RNA without regulatory drug in the first incubation step was carried out to distinguish between enrichment of inhibited and "parasitic" ribozymes (open symbols). The table below the graph relates each selection cycle to the respective selection conditions. Arrows indicate selection cycles amplified under mutagenic PCR conditions.
To eliminate parasitic species, we modified the selection procedure from cycle 6 onwards by interrupting the first incubation step by repetitive washing steps under denaturation/renaturation conditions. This slight modification of the in vitro selection scheme was based on a theoretical analysis (see Theory). Indeed, after round 7 the enriched pool showed much lower parasitic activity. After rounds 10 and 13, the complexity of the enriched pool was increased by mutagenic PCR. In parallel, the selection stringency was also increased to promote the enrichment of allosteric ribozymes with improved values of $K_i$ and catalytic rate.

The self-cleavage activity of the selected pools from cycles 10, 13, and 16 was monitored by time-course experiments with different concentrations of the inhibitor. Figure 4 directly compares the pool-activities for the previously described doxycycline and the now performed pefloxacin selections.

While the sensitivity of the pooled ribozymes for doxycycline increased continuously up to cycle 16, where the pool was completely inhibited in the presence of $1 \mu M$ doxycycline, the pefloxacin selection showed no improvement in the pool response between cycles 13 and 16 but showed inhibition levels comparable to those of doxycycline after cycle 16. Therefore, the majority of the sequenced clones that resulted from the pefloxacin-selection were taken from cycle 13.

Figure 3. Cleavage of the pools selected after cycle 5. The cleavage reaction was performed in the absence (filled circles) or in the presence (empty circles) of 1.0 mM regulatory drug ((a) doxycycline selection and (b) pefloxacin selection). The pool cleavage efficiency is severely reduced compared to the wild-type ribozyme (less than 25% cleaved after one hour). The response to doxycycline is not detectable, and the response to pefloxacin is very low.

Figure 4. Responsiveness of pools from different selection cycles to doxycycline ((a) top row) or pefloxacin ((b) lower row). The selected pools were incubated with the indicated amount of the respective drug after cycle 10, 13 and 16, and the cleavage activity was monitored. The sensitivity to doxycycline increased gradually during the selection. The sensitivity to pefloxacin however did not seem to increase between cycles 13 and 16.
Cloning and sequencing

The selected pools were cloned into the vector pNPG1, and 40 clones were sequenced (Figure 5). Sequence analysis showed that 95% of the sequences were unique. However, these sequences could be grouped into different families as shown in Figure 5. The sequences in each group most probably originated from one unique sequence in the initial pool, and the mutagenic PCR gave rise to the intra-group variability. Some groups were significantly different and were unlikely to have evolved from the same sequence but nevertheless showed some common motifs. However, among these groups, standard RNA-folding algorithms could not predict any common secondary structure.18,19

Whereas a certain bias towards T could be observed in the sequenced pool for doxycycline (660 T compared to 443 A, 473 G and 461 C), this bias was not present in the pefloxacin selection (393 T, 367 A, 482 G and 357 C).10

Clone analysis and inhibition of clones selected by pefloxacin

A total of 12 clones were chosen on the basis of their representation and were functionally tested. The results obtained are summarized in Table 1. The levels of self-cleavage inhibition observed in the presence of 1 μM pefloxacin were somewhat lower than those observed with the doxycycline selection, ranging from no inhibition (two clones) to 15-fold inhibition.10 It is also noteworthy that the cleavage rates in the absence of pefloxacin are substantially lower than the cleavage rate of the wild-type ribozyme from which they are derived (3 min⁻¹, data not shown).

Sensitivity of the clones

Four clones with the highest ratios between uninhibited and inhibited rates were chosen for Ki determination. Among them, one yielded irreproducible kinetics. The other three showed very similar properties, with Ki values ranging from 130 to 250 nM (Table 2).

These values are about two- to tenfold higher than those determined for doxycycline.10 A possible explanation is that the clones isolated after the doxycycline selection were mainly from cycle 16, whereas the clones isolated after the pefloxacin selection were from the 13th, and therefore less stringent, cycle.

Specificity of the inhibition and important features of the pefloxacin molecule

The two clones showing the lowest Ki, P13-15 and P13-28, were tested in the presence of 2 μM different derivatives of 1 (Figure 6). These two clones exhibited the same pattern of recognition for the different derivatives, as summarized in Figure 6 and Table 3. The different modifications in the

Table 1. Single clone analysis after pefloxacin selection

<table>
<thead>
<tr>
<th>Class</th>
<th>Clones per class</th>
<th>Uninhibited k (min⁻¹)</th>
<th>Factor of inhibition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P13-28</td>
<td>5</td>
<td>0.34</td>
<td>4.4</td>
</tr>
<tr>
<td>P16-04</td>
<td>-</td>
<td>0.03</td>
<td>3.5</td>
</tr>
<tr>
<td>P13-25</td>
<td>2</td>
<td>0.11</td>
<td>0.8</td>
</tr>
<tr>
<td>P13-06</td>
<td>2</td>
<td>0.15</td>
<td>4.5</td>
</tr>
<tr>
<td>P13-27</td>
<td></td>
<td>0.15</td>
<td>1.8</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P13-07</td>
<td>3</td>
<td>0.26</td>
<td>6.1</td>
</tr>
<tr>
<td>P13-30</td>
<td>2</td>
<td>0.30</td>
<td>5.3</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P13-03</td>
<td>2</td>
<td>0.26</td>
<td>2.8</td>
</tr>
<tr>
<td>Orphans</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P10-01</td>
<td></td>
<td>0.30</td>
<td>1.1</td>
</tr>
<tr>
<td>P13-15</td>
<td></td>
<td>0.33</td>
<td>13.5</td>
</tr>
<tr>
<td>P13-19</td>
<td></td>
<td>0.06</td>
<td>4.1</td>
</tr>
<tr>
<td>P13-24</td>
<td></td>
<td>0.42</td>
<td>3.4</td>
</tr>
</tbody>
</table>

Single clone analysis after pefloxacin selection. The clones analyzed are shown grouped in classes. The cleavage rate (k) was determined during the first three minutes of reaction. The cleavage rate in the presence of 1 μM pefloxacin was also determined. The ratio between the uninhibited and inhibited reaction rates is shown (factor of inhibition).

Table 2. Sensitivity of inhibition for three clones after the pefloxacin selection

<table>
<thead>
<tr>
<th>Clone</th>
<th>Uninhibited k (min⁻¹)</th>
<th>Pefloxacin Ki (nM)</th>
<th>Factor of inhibition</th>
</tr>
</thead>
<tbody>
<tr>
<td>P13-07</td>
<td>0.26 ± 0.12</td>
<td>250 ± 170</td>
<td>6</td>
</tr>
<tr>
<td>P13-15</td>
<td>0.33 ± 0.06</td>
<td>220 ± 100</td>
<td>14</td>
</tr>
<tr>
<td>P13-28</td>
<td>0.34 ± 0.13</td>
<td>130 ± 70</td>
<td>5</td>
</tr>
</tbody>
</table>

Ki values for pefloxacin, rate constants of the non-inhibited reaction (k), and maximal levels of inhibition are shown.
derivatives had different effects on the inhibition of the ribozymes.

When the fluorine residue in position 6 was substituted with a chlorine atom (3) or the carbon atom in position 8 was substituted with a nitrogen atom (2) or the shortening of ethyl at position 1 to methyl (8) did not show any effect. However, modifications of the piperazine group, such as addition of an oxygen atom either 3' (6) or 4' (4), totally abolished the inhibition effect. Two other derivatives, where the piperazine group was either substituted for a chlorine residue (5) or displaced (7) showed no inhibitory activity.

Taken together, these results suggest that the piperazine ring plays a central role in the inhibitory function of pefloxacin. However, it is not the only motif involved in recognition, as displacement of this group on the molecule abolishes activity. The face of the molecule containing the carboxylic acid function (atoms 2 to 5) presents interesting features, which could be involved in the binding motif. However, the lack of available derivatives for this part of the molecule prevented further investigation.

Theory
Elimination of parasitic ribozymes
The impact of denaturation/renaturation steps on the recovered fraction \( r(x) \) of ribozymes insensitive to pefloxacin but folding into both active and inactive conformations can be computed for a given ribozyme sequence as:

![Figure 5](image)

Sequences could be grouped into classes, whose consensus sequence is shown in bold. Some classes showed some homology between each other (class A and B); the conserved regions among class A or B members are underlined. Orphans are sequences that could not be assigned to any class. The clones chosen for further analysis are shown in bold.

### Table 3. Effect of the pefloxacin derivatives 1-9 on clones P13-15 and P13-28

<table>
<thead>
<tr>
<th>Entry</th>
<th>Compound name</th>
<th>Effect</th>
<th>Compound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H2O</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>Pefloxacin</td>
<td>++ +</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Norfloxacin</td>
<td>++ +</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>RP54328</td>
<td>++ +</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>RP44092</td>
<td>-/+</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>RP54240</td>
<td>-/+</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>RP57606</td>
<td>-/+</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>RP105509</td>
<td>-</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>RP130687</td>
<td>++ +</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>RP41983</td>
<td>++ +</td>
<td>9</td>
</tr>
</tbody>
</table>

Pefloxacin (1) and the modification derivatives (2 to 9) refer to the molecules in Figure 6.
where $x$ is the fraction of the molecules folding into a self-cleaving conformation and $n$ is the number of denaturation-renaturation steps:

$$r(x) = x(1 - x)^n$$  \hspace{1cm} (1)

is obtained for:

$$r'(x) = [1 - (n + 1)x](1 - x)^{n-1} = 0$$  \hspace{1cm} (2)

Therefore the recovered fraction of parasitic ribozymes, $r(x)$, is maximal for the species that fold into a self-cleaving conformation with a probability of:

$$x = \frac{1}{n+1}$$  \hspace{1cm} (3)

The recovered fraction of such ribozymes is expressed by:

$$r\left(\frac{1}{n+1}\right) = \frac{1}{n+1} \left(1 - \frac{1}{n+1}\right)^n$$  \hspace{1cm} (4)

The maximum fraction of parasitic ribozymes that could be eluted after a selection cycle using 12 denaturation-renaturation steps is 3% (equation (4)). This value is an overestimate of the proportion of parasitic ribozymes recovered, because they do not all fold into a self-cleaving conformation with a probability of $1/(n+1)$. 

Figure 6. Structure of pefloxacin (1) and its derivatives (2-9).
Therefore, the theoretical analysis supports denaturation-renaturation as an efficient means to eliminate parasitic sequences.

**Outcome of a selection cycle**

In order to better understand the outcome of the selection of allosteric ribozymes, we developed a theoretical model, which calculates the enrichment of allosteric ribozymes with different catalytic and binding efficiencies during a selection cycle.

In standard selections (aptamers), the enrichment in “winner” RNA sequences is explained by competition between the RNA species for binding to the target molecule at the equilibrium. However, in the case of the selection described here, there is no such competition between RNA molecules because the concentration of the regulatory drug is always kept much higher than the RNA concentration. Rather, the dynamic nature of the enzymatic reaction of self-cleavage drives the selection.

To illustrate this, a simple dynamic selection occurs if one selects for self-cleaving ribozymes with a cleavage rate over 0.69 min\(^{-1}\) (i.e. a half-life of less than one minute), assuming a simple exponential decrease in the reaction kinetics. If one uses a cleavage time of one minute, the ribozymes with cleavage rates of exactly 0.69 min\(^{-1}\) will be recovered at 50%, because one half of the ribozyme will cleave during one minute and the other half will not. Ribozymes with cleavage rates of 2 min\(^{-1}\) would be recovered at 86% and ribozymes with cleavage rates of 0.2 min\(^{-1}\) will be recovered at 18%. The inactive ribozymes will be recovered at the same level as the background. If the background recovery is 1%, the enrichment will therefore be 86-, 50-, 18-, or onefold over the background for ribozymes with the respective cleavage rates of 2, 0.69, 0.2, or 0 min\(^{-1}\). The selection will therefore preferentially enrich for ribozymes with higher cleavage rates.

In the case of the present selection, the same kind of reasoning can be applied, with more complex equations. The theoretically recovered fraction is the percentage of the input RNA that will be eluted at the end of the selection cycle for RNA sequences exhibiting a given dissociation constant \((K_D)\) and cleavage rate \((k)\) after one selection cycle was calculated. For modeling purposes the selection cycle can be reduced to three steps: transcription, incubation with drug, incubation without drug (Figure 7(a) and (b)). The recovered fraction at the end of the selection cycle is the product of the recovered fractions at each step.

The kinetics of the ribozyme self-cleavage can be expressed by \(f\) that is the fraction of uncleaved ribozymes after a time \(t\):

\[
f = \frac{[U]}{[U] + [C]} = e^{-kt}
\]

Uncleaved RNA \( \xrightarrow{k} \) Cleaved RNA

The recovered fraction at the end of the transcription (step 0, Figure 7(a)) is the fraction of the ribozyme molecules which remain uncleaved. With a constant transcription rate \(k\), a transcription time \(T_{0}\), and a drug concentration \(c_{0}\) this fraction can be calculated:

\[
f_0 = \frac{[U]}{[U] + [C]} = 1 - e^{k_{app}T_{0}}
\]

Matrix \( \xrightarrow{K_D} \) Uncleaved RNA \( \xrightarrow{k} \) Cleaved RNA

\[
k_{app} = \frac{k K_D}{c_0 + K_D}
\]

Unc. Drug \( \xrightarrow{k'} \) Cleav. Drug

If one makes the assumption that \(k' = 0\) (the ribozyme complexed with the drug is totally inactive), the expression of \(k_{app}\) can be simplified:
Equations (6) and (8) show that inhibited ribozymes behave like wild-type ribozymes with an apparent cleavage rate \(k_{\text{app}}\). Another consequence of this model is that the inhibition constant, \(K_i\), and the binding constant, \(K_D\), are equal (equation (8)).

Similarly, the uncleaved fraction of the ribozyme after the incubation with the effector (Figure 7, step 1) with an incubation time \(T_1\) and a drug concentration \(c_1\), is:

\[
f_1 = \frac{[U]}{[U] + [C]} = e^{-k_{\text{app}}T_1}
\]

Uncleaved RNA \(\rightarrow\) Cleaved RNA

\[
K_D \Downarrow \quad \Uparrow
\]

Unc. Drug \(\rightarrow\) Cleav. Drug
$k_{app}$ is calculated as in equations (7) and (8) with $c_1$ instead of $c_0$.

The uncleaved fraction of the ribozyme after incubation without effector (Figure 7, step 2) with an incubation time $T_2$, is:

$$f_2 = \frac{[U]}{[U] + [C]} e^{-kT_2} \quad (10)$$

Uncleaved RNA → Cleaved RNA

For this step the recovered fraction corresponds to the cleaved ribozyme fraction that is $(1 - f_2)$. Consequently, the recovered fraction ($r$) at the end of one selection cycle (Figure 7) will be for this particular ribozyme species:

$$r = f_0 f_1 (1 - f_2) \quad (11)$$

It is intuitive that for a given cleavage rate, the molecules having the strongest affinity for their ligand will be selected. Thereby, $r$ is a growing function of $K_D$. For a given $K_D$, however, there will be an intermediate cleavage rate for optimal recovery. Indeed, the ribozyme cleavage rate in the presence of the ligand ($k_{app}$) is proportional to the cleavage rate in the absence of ligand (equation (8)). If $k$ is too high, $k_{app}$ will be so high that the majority of the RNA molecules will cleave during step 1 (Figure 7(a)) and thus will be discarded. Conversely, if the cleavage rate is too low, only a very small fraction of the ribozymes will cleave during step 2 (Figure 7(a)), and the eluted amount will be very low. As a consequence, $r$ shows a maximum as a function of $k$ (see below).

**Fit of the model to experimental data**

The experimental analysis of the individual clones allowed a verification of the pertinence of the hypothesis underlying the theoretical model. (i) The kinetics performed on the different clones at various concentrations of the regulatory drug showed a good fit to the first-order kinetic reaction (simple exponential decay of the uncleaved fraction during at least two half-lives of the reaction). (ii) for each clone the observed apparent cleavage rate ($k_{app}$) in the presence of various concentrations of the regulatory drug fitted equation (8) nicely, justifying the hypotheses that each selected ribozyme binds a unique regulatory molecule in a reversible fashion and that this association-dissociation is at equilibrium. Furthermore, this also demonstrates the hypothesis that $k'$ in equation (7) can be neglected.

Under the conditions used during the final rounds of selection, for a determined $K_D$, an optimal $k$ can be calculated (equation (11)). It is striking that for the seven clones analyzed from the doxycycline and the pefloxacin selections, the actual cleavage rate is in excellent accordance with the predicted optimal $k$ (Figure 8) at which maximal recovery is observed. Therefore, our model seems to be valid, and explains well the slow self-cleavage obtained for the selected ribozymes.

**Use of the model for setting the parameters of the selection**

Next, we show how the model could be used to determine the optimal conditions for a selection protocol, knowing the desired $K_D$ and $k$ for a given application. First, we want to ensure that the selection pressure remains low during the transcription reaction and is applied only during the later two

![Figure 8](image-url)

**Figure 8.** Theoretical recovery after cycles 16 and 13. The curves showing the fractional function of recovery of the ribozymes with cleavage rates along the x-axis, for the $K_i$ observed for selected clones. The experimentally measured $k$ values for these clones are indicated by open circles. Except for clone D13-01 ($K_i = 70$ nM) the measured cleavage rates ($k$ values, open circles) very closely correspond to the cleavage rate at which the theoretical fraction of recovery is maximal. Thus, for these clones the theory almost perfectly reflects the experimental results.
incubations; therefore we maintained the target drug concentration high (1 mM) during the whole transcription stage. As a consequence, we consider that no selection is imposed on aptazymes with dissociation constants lower than 1 μM during the transcription reaction. For these molecules, the selection concentrates on the two later incubations and the expression of the recovered fraction \( r \) can be simplified with the approximation that \( f_0 = 1 \).

With this approximation, the catalytic rate with the maximum level of recovery can be calculated (maximum of \( r(k) \)):

\[
k_{\text{opt}} = \frac{1}{T_2} \ln \left( \frac{1 + \frac{c_1}{K_D}}{\left( 1 + \frac{c_1}{K_D} \right) \frac{T_2}{T_1} + 1} \right)
\]

(12)

For this value of \( k \) the fraction of recovery becomes:

\[
r(k_{\text{opt}}) = \exp \left[ -\ln \left( \frac{1 + \frac{c_1}{K_D}}{\left( 1 + \frac{c_1}{K_D} \right) \frac{T_2}{T_1} + 1} \right) \right] \left( 1 + \frac{c_1}{K_D} \right) \frac{T_2}{T_1} + 1
\]

(13)

The observation of equations (12) and (13) allows us to draw a strategy for the determination of optimal selection parameters in a selection cycle. \( r(k_{\text{opt}}) \) is a function of:

\[
\left( 1 + \frac{c_1}{K_D} \right) \frac{T_2}{T_1}
\]

To get a fraction of recovery \( r(k_{\text{opt}}) = 0.2 \) for the desired aptazymes:

\[
\left( 1 + \frac{c_1}{K_D} \right) \frac{T_2}{T_1}
\]

should be approximately 0.7 (numerical resolution). When this parameter is fixed, \( k_{\text{opt}} \) is only depending on \( T_2 \).

If the desired aptazymes are to have catalytic efficiencies in the range of the wild-type ribozyme (2 min\(^{-1}\)), the parameter \( T_2 \) should be fixed at 16 seconds. Now, with \( T_2 \) and:

\[
\left( 1 + \frac{c_1}{K_D} \right) \frac{T_2}{T_1}
\]

fixed, and a desired \( K_D \) of 50 nM it is possible to determine \( T_1 \) and \( c_1 \). For example one could use a concentration of drug \( c_1 = 1 \) μM and an incubation time \( T_2 = 8 \) minutes. Another combination giving the same results could be \( c_1 = 10 \) μM and \( T_1 = 1.28 \) hours.

It should be noted that the parameters used here are not optimal for the first round of selection, because a recovered fraction of 20 % would discard many unique molecules with the appropriate properties.

**Discussion**

We report the selection of allosteric ribozymes that are inhibited by pefloxacin, an antibiotic of the family of fluoroquinolones. This class of compounds target bacterial DNA gyrase and topoisomerase IV and stabilize the complexes formed by these enzymes and DNA. Moreover quinolones have some affinity for DNA (\( K_D \) in the low millimolar range) in the presence of magnesium ions. Hammerhead ribozymes are dependent on Mg\(^{2+}\) for self-cleavage, therefore it is tempting to propose that inhibition of the selected ribozymes by pefloxacin occurs through Mg\(^{2+}\) displacement.

The success of obtaining any biological molecule with desired characteristics or function depends to a great extent on the initial diversity of the population and on the power of selection. The advantage of generating the starting material in vitro, as with functional RNA molecules is that the pool size and complexity are not limited by biological constraints. Pool size and complexity can routinely reach \( 10^{15} \) molecules representing orders of magnitude more than can possibly be obtained in vivo. However, this offers no advantage if suitable selection parameters cannot be applied, this is why we built a mathematical model that describes our in vitro selection with a restricted set of parameters, which recapitulate the selection parameters. We show that once the desired \( k \) and \( K_D \) of the selected ribozymes are known, the model enables one to optimize the incubation times and the ligand concentration used at the different steps of the selection. Additionally, with simple calculations we were able to improve our selection strategy to remove parasitic molecules and obtain better fractions of molecules with desired characteristics. The experimental selection produced doxycycline and pefloxacin-sensitive aptazymes with self-cleavage kinetics that correlated remarkably well with theoretical predictions.

We have demonstrated previously that the selected allosteric ribozymes derived from hammerhead ribozyme can be converted into aptamers by a single mutation (cytosine residue 5 to the cleavage site into guanosine) that abolishes the self-cleavage activity. Therefore, we believe that the selection procedure we report here can be applied more generally for selecting aptamers, in particular towards small molecules that should otherwise be crosslinked to a solid phase during the selection. This gives widespread application to the predictive model of the selection described here.

**Materials and Methods**

**Pool synthesis**

The five primers were synthesized at a micromolar scale on an Expedite-Nucleic Acid Synthesis System (Millipore) following classical phosphoramidite chemistry. Special care was taken that the amplification pri-
mers were synthesized and purified before the template primer was synthesized, to avoid any contamination by amplification primer.

The sequences of the primers used were: Pnp-1, 5'-CGC GTT GTG TTT ACG CGT CTG ATG-3'; Pnp-2, 5'-AGC TGG TAC CTA ATG AGT-N40-AGT-3'‡, Pnp-3, 5'-AGC TGG TAC CTA ATG AGT-N40-AGT-3'‡, Pnp-rev, 5'-AGC TCT CGA GTT TTC GT-3'.

After synthesis primers were cleaved from the resin and deprotected overnight in 1 ml of 32 % ammonia at 55 °C. Following precipitation with butanol-1, pellets were washed with diethyl ether and precipitated with 1 ml of diethyl ether. The pellets were resuspended in 300 μl of water. Concentration was determined by measuring absorption at 260 nm.

After primer synthesis, large-scale PCR with the primer pool as template plus primers 1 and Rev was performed as follows: total volume 80 ml, template (Pnp-pool) 6.6 nmol, primers (Pnp-1 and Pnp-rev) 80 nmol, MgCl2 1.5 mM, dNTP 0.2 mM, Gold star reaction buffer (Eurogentec) 250 units, five cycles (94 °C five minutes, 55 °C five minutes, 72 °C seven minutes, in 8 ml aliquots with shaking every two minutes). Aliquots (5μl) were taken after each cycle to verify amplification. After PCR, DNA was purified by phenol/chloroform-extraction, ethanol-purification and chromatography on a G50 column (0.7 cm × 15 cm) before phenol/chloroform-extraction and ethanol-precipitation. Purified primers were resuspended in 300 μl of water. Concentration was determined by measuring absorption at 260 nm.

During the selection process, the transcription reaction was carried out under classical conditions to produce internally (α-32P)-labeled RNA for scintillation counting and quantification of bound and eluted RNA. For selection purposes, 1 mM inhibitor molecule and an 8 M excess (20 mM final) of guanosine monophosphothioate (GMPS) were used. An 8 M excess of GMPS was adopted because it gave the best yields of primed molecules. With smaller amounts, the priming efficiency was lower while with larger amounts transcription was partially inhibited. Quantification of priming was achieved either through mercury PAGE analysis or by linking to thiopropyl Sepharose. After classical preparation including DNase I treatment, precipitation, PAGE purification (8 %) and photometric quantification, primed RNA was coupled to iodoacetyl-LC-biotin (Pierce). A 2 μM solution of primed RNA in 50 mM Tris, 5 mM EDTA (pH 8.3) was incubated with a 200-fold excess of iodoacetetyl-LC-biotin (stock 4 mM in DMF) for 90 minutes at room temperature protected from light with occasional shaking. This labeling method was used because we observed that direct GMPS linking on thiopropyl Sepharose was very unstable. The reaction products were precipitated (sodium acetate), purified with PAGE and the final amount of RNA recovered was quantified on a photometer. Biotinylated RNA (1 nmol) was linked for 30 minutes at room temperature to 250 μl of streptavidin agarose equilibrated in coupling buffer (PBS, 150 mM NaCl, 50 % slurry). The amount of RNA linked on the column typically ranged between 20 and 40 %. To eliminate unlinked species the column was washed thoroughly (six times with alternatively 1 ml WA (25 mM Hepes (pH 7.4); 1 M NaCl, 5 mM EDTA) and 1 ml WB (3 M urea, 5 mM EDTA)) and rinsed with water (five times 500 μl). For the first selection incubation, the column material was incubated in 500 μl selection buffer (40 mM Tris (pH 8.0, 25 °C), 50 mM NaCl, 2 mM spermidine, 8 mM MgCl2) with the appropriate amount of inhibitor molecule at 37 °C with gentle shaking for the appropriate time. The incubation was initiated upon addition of magnesium. The column washing and rinsing sequence was then repeated as before and the column was incubated under similar conditions as before, but without inhibitor molecule. Finally, the cleaved RNA was eluted with WB (two times 500 μl). The different fractions (washes and elution) were quantified in a scintillation counter to determine the percentage of eluted RNA. The eluted RNA was purified with two or three phenol/chloroform-extractions, one chloroform extraction and precipitated (sodium acetate) in the presence of glycogen (1 %, v/v). The pellets were washed two additional times with 70 % ethanol before resuspension in 20 μl of water with 200 pmol of Pnp-rev primer. The RNA-oligonucleotide mix was denatured (one minute, 95 °C) and mixed with 80 μl of reverse transcriptase mix (5 μl of dNTP 4 mM, 10 μl of MmOAc 25 mM, 20 μl of 5x RT-PCR buffer (250 mM bicine/KOH (pH 8.2, 25 °C), 575 mM K-acetate; 40 % (v/v) glycerol), 43 μl of water, 2 μl of Th DNA polymerase (two to ten units)). The total mix was incubated for 30 minutes at 72 °C (at the same time a control without RNA was performed). The reverse transcription mix was then diluted into a 500 μl of PCR reaction under standard conditions with primers Pnp-3 and Pnp-rev (including a negative control). The number of cycles was calculated using the following equation:

\[
 n \geq 1 + \ln \left(\frac{1000}{x}\right) / \ln(2)
\]

where \(x\) is the amount of eluted RNA in pmol.

The PCR products were analyzed on a 2 % (w/v) agarose gel, before phenol/chloroform-extraction and precipitation (sodium acetate), 25 % of the resulting DNA was used for the next selection cycle.

**Mutagenic PCR**

At certain points during the selection (after cycles 10 and 13), we performed mutagenic PCR using a protocol derived from that described by Cadwell & Joyce. On the total DNA molecule, 43 bases could be subject to mutagenic PCR (the other bases are covered by the primers). According to mutation distribution, the primers would yield an average number of 1.13 mutations per molecule after 60 cycles of mutagenic PCR. The number of mutations

\[
 n \geq 1 + \ln \left(\frac{1000}{x}\right) / \ln(2)
\]
per individual sequence should follow a Poisson distribution. Hypothetically, the average frequency of unmutagenized molecules should be 32 %.

The PCR reaction (100 μl) contained 10 mM Tris (pH 8.3), 50 mM KCl, 0.001 % gelatin, 0.2 mM dATP-tagenized molecules should be 32 %.

Kinetische

The DNA templates were transcribed, kinased and diluted to 1 nM in selection buffer. For the shotgun experiment, self-cleavage time-course experiments were performed without or with 1 μM of the inhibitory drug in selection buffer with 8 mM magnesium at 37 °C (selection conditions). The reaction was initiated upon addition of magnesium, and aliquots were taken every 20 seconds. The aliquots were immediately quenched by mixing with an equal volume of PAGE loading buffer on ice. The different aliquots were loaded on a sequencing gel and the uncleaved fraction (radioactivity of the uncleaved band/radioactivity of uncleaved band + radioactivity of the cleaved band) was determined for each sample after quantification on a PhosphorImager. The cleavage rates (k) were then determined, using an exponential fitting of the uncleaved ribozyme fraction during the first two minutes of reaction. For the k_i determination, 1 nM kinased ribozyme was incubated with increasing amounts of the regulatory drug ranging from 20 nM to 1 μM, and the reaction rates were calculated as above. The k_i was calculated after fitting the cleavage-rate values obtained for different drug concentrations with the following equation (where k is the uninhibited cleavage rate):

$$k_{obs}(\text{Drug}) = \frac{k}{1 + \frac{[\text{Drug}]}{k_i}}$$  \hspace{1cm} (15)

Each experiment was repeated at least two times, and the values given in the text are the mean and standard deviation of the observed results.

For templates poorly or not inhibited by the drug, the reaction rate was assayed co-transcriptionally as described previously.16,26

Acknowledgements

We gratefully acknowledge Aventis Corporation for providing compounds 2-9. We thank Michael Blind, Andreas Jenne, Gerhard Sengle and Günter Mayer for helpful discussions and suggestions. We thank J. Crouzet, J.F. Mayaux and M. Finer for their support. This work was supported from the Volkswagenstiftung, the Fonds der Chemischen Industrie, and Aventis Gencecell.

References


Edited by R. Huber

(Received 8 May 2001; received in revised form 24 July 2001; accepted 24 July 2001)