A Catalytic DNA Activated by a Specific Strain of Bacterial Pathogen

Zhifa Shen*, Zaisheng Wu*, Dingran Chang*, Wenqing Zhang, Kha Tram, Christine Lee, Peter Kim, Bruno J. Salena, and Yingfu Li*

Abstract: Pathogenic strains of bacteria are known to cause various infectious diseases and there is a growing demand for molecular probes that can selectively recognize them. Here we report a special DNAzyme (catalytic DNA), RFD-CD1, that shows exquisite specificity for a pathogenic strain of Clostridium difficile (C. difficile). RFD-CD1 was derived by an in vitro selection approach where a random-sequence DNA library was allowed to react with an unpurified molecular mixture derived from this strain of C. difficile, coupled with a subtractive selection strategy to eliminate cross-reactivities to unintended C. difficile strains and other bacteria species. RFD-CD1 is activated by a truncated version of TcdC, a transcription factor, that is unique to the targeted strain of C. difficile. Our study demonstrates for the first time that in vitro selection offers an effective approach for deriving functional nucleic acid probes that are capable of achieving strain-specific recognition of bacterial pathogens.

Bacterial infections pose serious threats to public health and are responsible for many annual costly outbreaks.[1] The issue has been further compounded by the emergence of hyper-virulent and/or antibiotic-resistant strains, which has now become a serious global problem.[2] Early detection of specific pathogens has long been recognized as a vital strategy in the control of infectious diseases because it can lead to timely care of patients and prevent potential outbreaks. However, detection of specific bacteria represents a significant challenge because of the presence of many different species of bacteria in biological samples. Furthermore, for any given species of bacterium, only virulent strains are infectious while other strains of the same species may be harmless or even beneficial to human health. Therefore, there is a great need to develop molecular probes that are highly specific for pathogenic strains of bacteria.

DNAzymes refer to single-stranded DNA molecules with catalytic capabilities.[3] They have been widely explored as molecular tools for applications ranging from biosensing to gene regulation.[4] DNAzymes can be generated de novo by in vitro selection, a simple technique that allows for the isolation of rare functional DNA sequences from a random DNA pool.[5] We have previously developed an RNA-cleaving fluorogenic DNAzyme (RFD) as an indicator for E. coli.[6] RFD-EC1 has since been used in a colorimetric assay for the field detection of E. coli and in a microfluidic-based counting assay for the rapid detection of E. coli in blood.[7,8] Although RFD-EC1 is highly specific for E. coli with minimal cross-activity to other bacterial species,[9] it non-discriminatively recognizes all E. coli strains.[10] The goal of this study was to investigate whether it was possible to derive strain-specific RFDs by in vitro selection.

C. difficile, a gram-positive bacterium that has been identified as the major cause of the diarrheic disease known as Clostridium difficile infection (CDI), was chosen for the current study. The incidence and mortality of CDI have increased dramatically over the past 15 years and CDI has become one of the most common healthcare-associated infections in the Western hemisphere.[9] These have been linked to the emergence of a hyper-virulent clinical strain known as BI/027.[10] In addition to being more virulent, this strain is also more resistant to antibiotics that are used to treat CDI.[11] Therefore, developing a specific molecular probe for BI/027 has important clinical implications.

We used a locally isolated clinical strain of C. difficile for our investigation, which was confirmed to be a BI/027 strain by the typing experiment shown in Figure S1 in the Supporting Information. We refer this strain as BI/027-H (H: Hamilton, Ontario). We employed two strategies to derive the desired DNAzymes. First, we used a complex mixture (rather than a defined target) derived from BI/027-H for the
E. coli

Hafnia alvei

A) The activity of DNA pools. Each pool was incubated with CEM-

B. subtilis

and

C. difficile

Taking

C. genome encodes nearly 4000 proteins.

strains. Lane 1: Pediococcus acidilactici

and

(Figure 2A).

We first evaluated the reactivity of RFD-CD1 towards the CEM

of 13 bacterial species and found none of them was able to

activate RFD-CD1 (Figure 2A). We then tested the reactivity

of RFD-CD1 to CEM of 12 C. difficile strains. Only the

CEM from BI/027 was able to activate RFD-CD1 (lane 1,

Figure 2B). The results indicate that RFD-CD1 is both species-specific and strain-selective.

Next we turned our attention to the identification of the

target that activates RFD-CD1. We first treated CEM-CD

with three proteases, trypsin (TP) and proteinase K (PK) and

subtilisin (SL), to determine if the activator is a protein. We

found that CEM-CD treated with each protease was unable to

activate RFD-CD1 (Figure 3A), signifying that the target is

indeed a protein.

We evaluated the molecular weight of the target

using molecular sizing. CEM-CD was passed through

centrifugal filters with molecular weight (MW) cut-offs

from 3–100 kilodaltons (KDa). The 3 KDa and 10 KDa

filtrates did not induce the cleavage of RFD-CD1 while

the 30 KDa filtrate caused a small level of cleavage. In

contrast, the 50 K and 100 K filtrates prompted strong

cleavage (Figure 3B). This experiment suggests that the

protein target has a MW of about 30000 Daltons.

C. difficile genome encodes nearly 4000 proteins.[12]

Given the estimated MW of the target, we tailored our search

to 986 proteins with an MW between 25000 and

35000 Daltons (Figure 3C). Based on the fact that the

RFD-CD1 is a DNA molecule, we hypothesized that the

target might have an intrinsic ability to interact with DNA.

Therefore, we decided to narrow down our search
to gene expression regulators. After applying this

collection strategy using Procom software, the potential

candidates were further reduced to 78 (Figure 3C).

The strain specificity data presented in Figure 2B indicates that RFD-CD1 is only active with BI/027-H but not with other strains including CD630. Previous studies have shown that there are numerous genetic differences between CD-BI/027 and CD630.[13] Taking these into account, we decided to compare the genomic

removal using centrifugation. This CEM, named CEM-CD,
served as the positive selection target. A CEM was also

preprepared similarly for each control bacterium. These CEMs

were combined and incubated with the DNA library for

five hours. The uncleaved DNA molecules in this step were

purified by 10% denaturing polyacrylamide gel electropho-

resis (dPAGE), and then incubated with CEM-CD for

30 minutes. The cleaved DNA products in this step were

purified by dPAGE, amplified by polymerase chain reaction

(PCR), and used for the next round of selection (experimen-
tal details are described in the Supporting Information).

In total, 19 iterations were conducted. Figure 1A depicts the
reactivity of representative DNA pools when incubated with

CEM-CD. By round 19, approximately 20% cleavage was
observed. The 19th DNA pool was cloned and sequenced.

Five DNAzyme classes were discovered (Figure 1B) and the

DNAzyme with highest cleavage activity (class 2) was named

RFD-CD1 (Figure 1C) and chosen for further investiga-

Figure 1. A) The activity of DNA pools. Each pool was incubated with CEM-

for 30 minutes, followed by dPAGE analysis. Unclv: uncleaved DNA pool;

Clv: cleaved DNA pool; % Clv = (Fclv/F0) × 100. B) Five DNAzyme classes were discovered (Figure 3C). Based on the fact that the

RFD-CD1 is a DNA molecule, we hypothesized that the target might have an intrinsic ability to interact with DNA. Therefore, we decided to narrow down our search
to gene expression regulators. After applying this filtration strategy using Procom software, the potential candidates were further reduced to 78 (Figure 3C).

The strain specificity data presented in Figure 2B indicates that RFD-CD1 is only active with BI/027-H but not with other strains including CD630. Previous studies have shown that there are numerous genetic differences between CD-BI/027 and CD630.[13] Taking these into account, we decided to compare the genomic

removal using centrifugation. This CEM, named CEM-CD,
served as the positive selection target. A CEM was also

prepared similarly for each control bacterium. These CEMs

were combined and incubated with the DNA library for

five hours. The uncleaved DNA molecules in this step were

purified by 10% denaturing polyacrylamide gel electropho-

resis (dPAGE), and then incubated with CEM-CD for

30 minutes. The cleaved DNA products in this step were

purified by dPAGE, amplified by polymerase chain reaction

(PCR), and used for the next round of selection (experimen-
tal details are described in the Supporting Information).

In total, 19 iterations were conducted. Figure 1A depicts the
reactivity of representative DNA pools when incubated with

CEM-CD. By round 19, approximately 20% cleavage was
observed. The 19th DNA pool was cloned and sequenced.

Five DNAzyme classes were discovered (Figure 1B) and the

DNAzyme with highest cleavage activity (class 2) was named

RFD-CD1 (Figure 1C) and chosen for further investiga-

Figure 1. A) The activity of DNA pools. Each pool was incubated with CEM-

for 30 minutes, followed by dPAGE analysis. Unclv: uncleaved DNA pool;

Clv: cleaved DNA pool; % Clv = (Fclv/F0) × 100. B) Five DNAzyme classes were discovered (Figure 3C). Based on the fact that the

RFD-CD1 is a DNA molecule, we hypothesized that the target might have an intrinsic ability to interact with DNA. Therefore, we decided to narrow down our search
to gene expression regulators. After applying this filtration strategy using Procom software, the potential candidates were further reduced to 78 (Figure 3C).

The strain specificity data presented in Figure 2B indicates that RFD-CD1 is only active with BI/027-H but not with other strains including CD630. Previous studies have shown that there are numerous genetic differences between CD-BI/027 and CD630.[13] Taking these into account, we decided to compare the genomic

removal using centrifugation. This CEM, named CEM-CD,
served as the positive selection target. A CEM was also

prepared similarly for each control bacterium. These CEMs

were combined and incubated with the DNA library for

five hours. The uncleaved DNA molecules in this step were

purified by 10% denaturing polyacrylamide gel electropho-

resis (dPAGE), and then incubated with CEM-CD for

30 minutes. The cleaved DNA products in this step were

purified by dPAGE, amplified by polymerase chain reaction

(PCR), and used for the next round of selection (experimen-
tal details are described in the Supporting Information).

In total, 19 iterations were conducted. Figure 1A depicts the
reactivity of representative DNA pools when incubated with

CEM-CD. By round 19, approximately 20% cleavage was
observed. The 19th DNA pool was cloned and sequenced.

Five DNAzyme classes were discovered (Figure 1B) and the

DNAzyme with highest cleavage activity (class 2) was named

RFD-CD1 (Figure 1C) and chosen for further investiga-
The A) Responses of RFD-CD1 to protease-treated CEM-CD. TP: trypsin; PK: proteinase K; SL: subtilisin. B) Estimation of the molecular weight of the protein target. CEM-CD was passed through centrifugal filters with specified molecular weight cut-offs, and the filtrates were then tested for reactivity with RFD-CD1. C) The Venn diagram showing the number of proteins in CD630 in each listed category. D) Cumulative distribution of proteins in CD630 and CD196 having lowest to highest identity ratio. Insert shows the first 100 proteins with the lowest identity ratio. E) RFD-CD1 treated with the CEM prepared from E. coli cells transformed with a plasmid expressing one of the five candidate targets: CdtR, MerR, RpiR, TrmB, and TcdC. NC: the negative control made of the CEM prepared from E. coli containing an empty plasmid. PC: positive control made of the CEM prepared from B/I 027-H. F) RFD-CD1 treated with purified TcdC-WT and TcdC-24. SB: Selection buffer only. Reaction time in panels A, B, E and F: 30 minutes.

Table 1: Typing of tcdC genes of the C. difficile strains used in Figure 2B for strain specificity test.

<table>
<thead>
<tr>
<th>Strain names</th>
<th>tcdC version</th>
<th>Size of TcdC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI/027-H</td>
<td>TcdC-24</td>
<td>226</td>
</tr>
<tr>
<td>BAA-1382 (CD630)</td>
<td>TcdC-WT</td>
<td>232</td>
</tr>
<tr>
<td>43 255</td>
<td>TcdC-WT</td>
<td>232</td>
</tr>
<tr>
<td>43 594</td>
<td>TcdC-WT</td>
<td>232</td>
</tr>
<tr>
<td>BAA-1804</td>
<td>TcdC-WT</td>
<td>232</td>
</tr>
<tr>
<td>BAA-1814</td>
<td>TcdC-WT</td>
<td>232</td>
</tr>
<tr>
<td>BAA-1871</td>
<td>TcdC-WT</td>
<td>232</td>
</tr>
<tr>
<td>BAA-1872</td>
<td>TcdC-WT</td>
<td>232</td>
</tr>
<tr>
<td>BAA-1870</td>
<td>TcdC-1</td>
<td>65</td>
</tr>
<tr>
<td>BAA-1875</td>
<td>TcdC-5</td>
<td>61</td>
</tr>
<tr>
<td>BAA-1801</td>
<td>no TcdC</td>
<td>0</td>
</tr>
</tbody>
</table>

We next sequenced the tcdC genes from other strains used for the strain specificity test in Figure 2B and the results are summarized in Table 1. BAA-1801 is a non-pathogenic strain and does not have a tcdC gene. BAA-1870 and BAA-1875 produce significantly truncated TcdC proteins containing only 65 and 61 amino acids, respectively, due to sequence mutations. All other C. difficile strains produce wild-type TcdC. Taken together, the experiments above not only result in the discovery of TcdC as the target for RFD-CD1, but also explain the origin of the strain specificity. That is: RFD-CD1 has a high specificity for TcdC-24 and recognizes neither full-length TcdC nor the truncated TcdC variants.

In summary, using in vitro selection, we have obtained an RNA-cleaving fluorescent DNAzyme (RFD) that can recognize an infectious strain of C. difficile. This DNAzyme not only exhibits cross-reactivity to other bacterial species, but also is highly strain-selective for C. difficile. To our knowledge, our work reports the first example of a functional nucleic acid that can recognize a specific strain of bacterium. Moreover, our approach also represents an unconventional strategy to molecular probe engineering as it does not begin with the identification of a specific biomarker for the probe development, which can be a lengthy and possibly unproductive process. Instead, our approach relies on a test-tube selection process to come up with a solution. This is achieved through tandem selection steps where a complex biological mixture from the targeted bacterial strain is used as the positive-selection target and similar mixtures from control bacteria are used as the subtractive selection target. Even though this approach does present the challenge of identifying the target
upon the completion of in vitro selection, we can exploit simple bioinformatic tools and experimental techniques to progressively narrow down target candidates and ultimately identify the target. Taken together, our work here demonstrates that in vitro selection can be an effective solution to engineering functional DNA probes that are able to recognize specific strains of bacterial pathogens. This sets up the stage to exploit synthetic DNAzyme probes for infectious disease diagnosis.

Acknowledgements

Funding is provided by Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canadian Institute of Health Research (CIHR) through a CHRP grant.

Keywords: aptamers · bacterial detection · biosensors · DNAzymes · in vitro selection

How to cite: Angew. Chem. Int. Ed. 2016, 55, 2431–2434


Received: October 30, 2015
Published online: December 16, 2015