

## **Distinction of Individual Lanthanide Ions with a DNAzyme Beacon Array**

Po-Jung Jimmy Huang, Mahsa Vazin, Jennifer J. Lin, Rachel Pautler, and Juewen Liu\*

Department of Chemistry, Waterloo Institute for Nanotechnology

Waterloo, Ontario, Canada N2L 3G1

Email: [liujw@uwaterloo.ca](mailto:liujw@uwaterloo.ca)

## **Abstract**

Developing chemical probes to distinguish each lanthanide ion is a long-standing challenge. Aside from its analytical applications, solving this problem will also enhance our knowledge in metal ligand design. Using in vitro selection, we previously reported four RNA-cleaving DNazymes, each with a different activity trend cross the lanthanide series. We herein performed another eight in vitro selection experiments using each and every lanthanide from La<sup>3+</sup> to Tb<sup>3+</sup> but excluding the radioactive Pm<sup>3+</sup>. A new DNzyme named Gd2b was identified and characterized. By labeling this DNzyme with a fluorophore/quencher pair to create a catalytic beacon, a detection limit of 14 nM Gd<sup>3+</sup> was achieved. With the same beacon design, all the five lanthanide-specific DNzymes were used together to form a sensor array. Each lanthanide ion produces a unique response pattern with these five sensors, allowing a pattern-recognition-based linear discriminant analysis (LDA) algorithm to be applied, where separation was achieved between lanthanides and non-lanthanides, light and heavy lanthanides, and for the most part, each lanthanide. These lanthanide-specific DNA molecules are useful for understanding lanthanide coordination chemistry, designing hybrid materials, and developing related analytical probes.

**Keywords:** DNzymes, biosensors, lanthanides, sensor array, fluorescence, molecular beacons, SELEX

Lanthanides (Ln) refer to the 15 elements in the first row of the *f*-block in the periodic table (Figure 1A). For their electronic, optical, magnetic and catalytic properties, Ln are extremely important industrial vitamins.<sup>1,2</sup> Each year, ~100,000 metric tons of Ln are consumed,<sup>3</sup> thus demanding for cost-effective analytical tools for monitoring their pollution, recycling, and finding new mineral sources. In this regard, biosensors are attractive for providing fast and on-site information complementary to instrumentation methods.

In water, most Ln exist in the trivalent state (Ln<sup>3+</sup>). While a few optical and electrochemical sensors for Ln<sup>3+</sup> were reported based on small molecule chelators, they often suffer from poor selectivity and low sensitivity.<sup>4-7</sup> Chemical probing of Ln<sup>3+</sup> has been a long-standing challenge, since they only differ by the number of inner 4*f* electrons, which do not directly participate in bonding. In addition, all Ln<sup>3+</sup> ions have the same charge and similar sizes, displaying extremely similar chemical properties. With 15 of them in the group, discrimination of each Ln<sup>3+</sup> is a daunting task for rational ligand design.

Recent work indicates that DNA might be a useful platform for sensing Ln<sup>3+</sup>.<sup>8-13</sup> DNA contains a phosphate backbone that affords high Ln<sup>3+</sup> binding affinity, while the nitrogen containing nucleobases might help discriminate individual Ln<sup>3+</sup>.<sup>7</sup> Ln<sup>3+</sup> forms complexes with nucleotides, especially with adenosine and guanosine phosphates.<sup>14-16</sup> GT-rich DNA is particularly effective in sensitized Tb<sup>3+</sup> luminescence.<sup>17</sup> It is however still difficult to rationally design DNA sequences to selectively bind Ln<sup>3+</sup>.

DNAzymes are DNA-based biocatalysts, which require metal ion cofactors for activity.<sup>18-22</sup> In the past 20 years, RNA-cleaving DNAzymes have emerged to be a useful platform for metal sensing due to their fast cleavage rates and strong metal binding affinity.<sup>23-29</sup> Using in vitro selection, DNAzymes specific for Na<sup>+</sup>,<sup>30</sup> Ag<sup>+</sup>,<sup>31</sup> Pb<sup>2+</sup>,<sup>32</sup> Zn<sup>2+</sup>,<sup>33</sup> Cu<sup>2+</sup>,<sup>34</sup> UO<sub>2</sub><sup>2+</sup>,<sup>35</sup> Cd<sup>2+</sup>,<sup>36</sup> and Hg<sup>2+</sup> were isolated.<sup>37</sup> Free Ln<sup>3+</sup> at a few mM concentration can cleave RNA and thus Ln<sup>3+</sup> might be good DNAzyme cofactors as well. In fact, a small Pb<sup>2+</sup>-dependent ribozyme (leadzyme),<sup>38</sup> and the GR5 DNAzyme can

both be moderately activated by  $\text{Ln}^{3+}$ .<sup>39</sup> Aside from RNA cleavage,  $\text{Ln}^{3+}$  was also used for assisting other types of DNAzyme reactions.<sup>40,41</sup>

Over the past three years, we have reported four RNA-cleaving DNAzymes,<sup>9-12</sup> each has a distinct activity pattern cross the lanthanide series. We reason that by accumulating more such DNAzymes, a sensor array may be developed for pattern-recognition-based  $\text{Ln}^{3+}$  discrimination. In this work, we performed eight more independent selections, each using a different  $\text{Ln}^{3+}$  that was not employed previously. A new DNAzyme named Gd2b was identified and characterized. Combined with the previously reported four DNAzymes, an array containing five sensors was made, which can discriminate  $\text{Ln}^{3+}$  from other metals, light  $\text{Ln}^{3+}$  from the heavy ones, and for the most part, each  $\text{Ln}^{3+}$ .

## Materials and Methods

**Chemicals.** The DNA library for *in vitro* selection, related primers and fluorophore/quencher modified DNAs were from Integrated DNA Technologies (IDT, Coralville, IA). The sequences of DNA for *in vitro* selections are listed in Table S1. The trans-cleaving enzyme strands and their mutants were from Eurofins (Huntsville, AL). The  $\text{Ln}^{3+}$  and other metals were purchased from Sigma-Aldrich. Tris(hydroxymethyl)aminomethane (Tris), 2-(N-morpholino)ethanesulfonic acid (MES), 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES), EDTA disodium salt dihydrate, sodium chloride and ammonium acetate were from Mandel Scientific Inc (Guelph, Ontario, Canada). T4-DNA ligase, deoxynucleotide (dNTP) solution mix, Taq DNA polymerase with ThermoPol buffer, and low molecular weight DNA ladder were from New England Biolabs. All metal ions, buffers and gel stock solutions were prepared with Milli-Q water.

**In vitro selection.** The methods for *in vitro* selection, PCR, cloning, and sequencing were previously reported.<sup>9</sup> The only difference was that a different  $\text{Ln}^{3+}$  was used to induce cleavage in each selection. The selection conditions for each  $\text{Ln}^{3+}$  are in Tables S2-S9. The selections were stopped at round 5 or 6, where freshly prepared PCR1 products were cloned using the TA-TOPO Cloning Kit (Invitrogen). The

plasmid DNA was extracted and purified by using DirectPrep 96 MiniPrep Kit (QIAGEN) and submitted to TCAG DNA Sequencing Facility (Toronto, ON) for analysis. Each selection yielded approximately 40 sequences.

**Gel-based assays.** DNAzyme activity assays were performed with the FAM-labeled substrate (Sub-FAM, 0.7  $\mu$ M) and the Gd2b enzyme (or other enzymes, 1.1  $\mu$ M). The DNAzyme complex was annealed in buffer A (50 mM MES, pH 6.0, 25 mM NaCl) before adding a metal ion solution to induce cleavage. The reaction products were separated on a 15% denaturing polyacrylamide gel (dPAGE) and analyzed using a ChemiDoc MP imaging system (Bio-Rad).

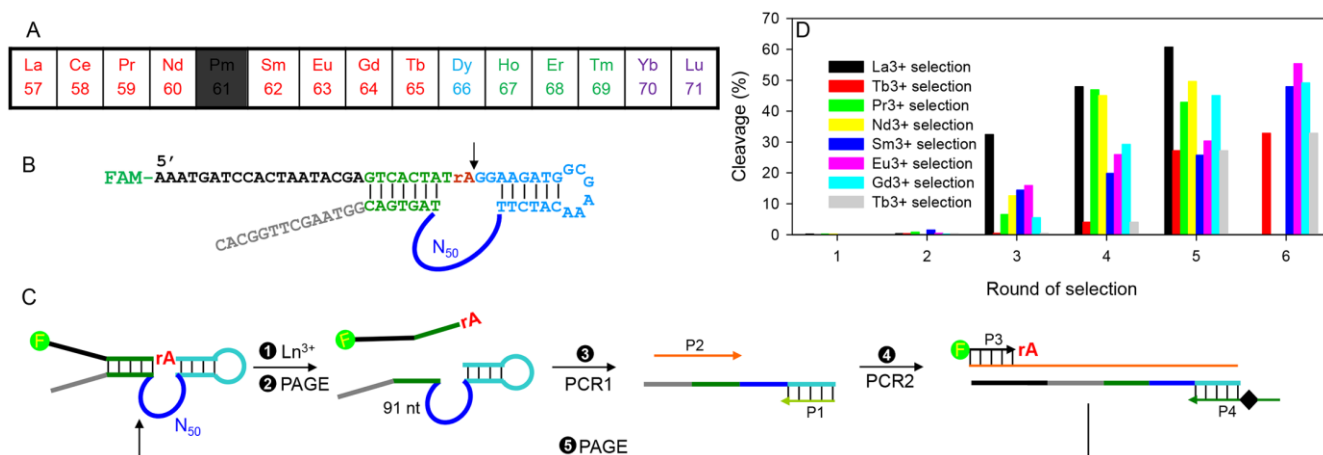
**DNAzyme sensor assays.** The kinetic studies were carried out using low binding half-area black 96 well plates using a microplate reader (SpectraMax M3). The stock complex was formed by annealing the FAM-labeled substrate and the quencher-labeled enzyme with molar ratio of 1:1.5 in buffer A. Each complex was further diluted with various concentrations of HEPES (pH 7.6) without additional NaCl. The HEPES buffer concentrations were individually optimized for each sensor: Lu12 (1 mM), Tm7 (10 mM), Gd2b (10 mM), Dy10a (10 mM), and Ce13d (50 mM). 100  $\mu$ L of DNAzyme complex (50 nM) was used for each well. 1  $\mu$ L of metal ion solution was added after 5 min of background reading followed by continuously monitored for at least 30 min with 25 s intervals. To test the sensor array, each DNAzyme was tested with 500 nM of each Ln<sup>3+</sup> (four replicates). The slope in the initial linear region was calculated for each kinetic trace. The data were analyzed using the Canonical Discriminant Analysis software from Origin.

## **Results and Discussion**

**In vitro selection.** Out of the 15 Ln (Figure 1A), Pm is radioactive and thus not studied. Within the remaining 14, we have already selected against seven ions so far. First, Ce<sup>4+</sup> was used with an N50 library (i.e. with 50 random nucleotides, Figure 1B), yielding the Ce13d DNAzyme that has a similar

activity for all the Ln<sup>3+</sup> (but inactive with Ce<sup>4+</sup>).<sup>9</sup> Then, we used Yb<sup>3+</sup> and Lu<sup>3+</sup> respectively with an N35 library, yielding the Lu12 DNAzyme.<sup>11</sup> Lu12 has a descending activity trend as the atomic number of Ln<sup>3+</sup> increasing, especially for the last a few Ln<sup>3+</sup>. Using the same N35 library but with Ho<sup>3+</sup>, Er<sup>3+</sup>, and Tm<sup>3+</sup>, respectively, the Tm7 DNAzyme was isolated, which is active only with the seven heavy Ln<sup>3+</sup>.<sup>10</sup> Finally, using Dy<sup>3+</sup> and the N50 library, we isolated the Dy10a DNAzyme that has the highest activity in the middle of the Ln<sup>3+</sup> series.<sup>12</sup>

In this work, we want to exhaust the sequence possibility, and thus used each of the remaining eight Ln<sup>3+</sup> ions (from La<sup>3+</sup> to Tb<sup>3+</sup> excluding Pm<sup>3+</sup>) with the N50 library. The initial library contained ~10<sup>14</sup> random DNA sequences. The randomized region (the blue loop in Figure 1B) is flanked by two short base paired duplexes, holding the single ribo-adenosine (rA) to its proximity. This RNA linkage is the putative cleavage site since it has much lower stability than the rest DNA linkages.<sup>42</sup> The library was respectively incubated with a Ln<sup>3+</sup> (Figure 1C, step 1). After incubation, the shorter cleaved sequences were harvested using gel electrophoresis (step 2). Then, two rounds of polymerase chain reactions (PCR) were performed to amplify the cleaved DNA, restore the original library length, and seed the next round of selection. We followed the cleavage yield of each round, and a gradually increased cleavage was observed in all the eight selections (Figure 1D). After 5 or 6 rounds of selection, cleavage reached >40% in most cases. At that point, the libraries were cloned and sequenced.



**Figure 1.** (A) Ln and their atomic numbers. The eight Ln<sup>3+</sup> ions in red were individually selected in this work. The ones in other colors were selected previously and the radioactive Pm<sup>3+</sup> was not studied. (B) The library sequence used here with 50 random nucleotides. A carboxyfluorescein (FAM) fluorophore was labeled to follow the selection progress. (C) A scheme for selecting Ln<sup>3+</sup>-dependent DNazymes. After metal incubation and gel electrophoresis (PAGE), the full-length library is re-generated by PCR1. In PCR2, P3 contains a FAM label on its 5'-end and a rA nearby its 3'-end. P4 has a polymer spacer to stop PCR extension. After PCR2, a duplex with an overhang is produced. The shorter strand is harvested to seed the next round of selection. (D) The progress of each selection characterized by the cleavage yield.

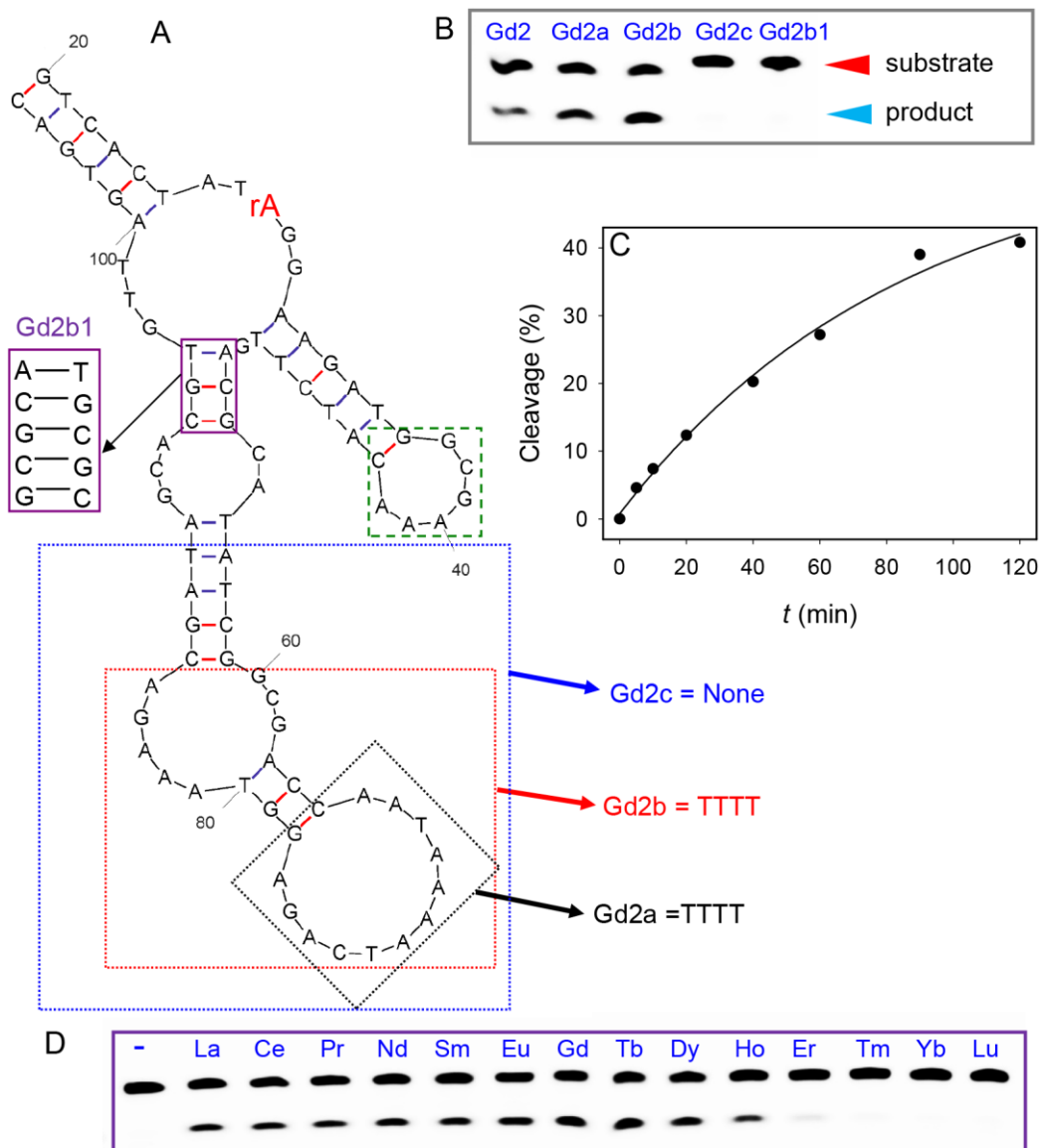
**Identification of the Gd2 DNzyme.** Nearly 40 well-aligned sequences were obtained for each selection (Tables S10-S17). We noticed sequence similarity across the different selections, prompting us to align them together. To better identify DNazymes that are already known, we also included the sequences from the previous six selections (using Dy<sup>3+</sup> to Lu<sup>3+</sup>) and aligned all of them together. There are 336 sequences from the N50 library (Table S18) and 135 from the N35 library (Table S19). The sequences from different selections are well mixed, strongly indicating the chemical similarity across the Ln<sup>3+</sup> ions.

Based on extensive alignment and folding analysis of these sequences, we identified only one new DNzyme from our eight selections in this work, while the rest active sequences belong to one of the four DNazymes reported previously. This new DNzyme was named Gd2. In fact, Gd2 appeared only in the Gd<sup>3+</sup> and Tb<sup>3+</sup> selections (highlighted in pink in Table S18). The classification of DNzyme sequences is based on their folded structures and also activity trend cross the Ln<sup>3+</sup> series. For each potential new sequence, we designed a trans-cleaving DNzyme to test its cleavage activity. Table S20 shows the sequences tested, and some of the examples related to their analysis are shown in Figures S1-S9. Below, the Gd2 is analyzed in detail.

**Analysis of the Gd2 DNAzyme.** The secondary structure of the cis-cleaving Gd2 DNAzyme is shown in Figure 2A (generated by Mfold).<sup>43</sup> The rA in red is the cleavage site and the overall folding agrees with the library design in Figure 1B. To characterize this enzyme, we first truncated the hairpin loop in the green box to generate a trans-cleaving system: the enzyme is the bottom half and is still named Gd2, and the substrate strand is the top half containing the rA. For activity assays, the substrate was labeled with a FAM on its 3'-terminus. With this trans-cleaving Gd2, we observed Gd<sup>3+</sup>-dependent cleavage activity, suggesting this is an active enzyme (Figure 2B, the first lane). To analyze the catalytic core, a few truncations were made. For example, we first replaced the nucleotides in the black box with 'TTTT' (now named Gd2a), and it has slightly enhanced activity (Figure 2B). The Gd2b replaces the nucleotides in the blue box with 'TTTT', and the activity further increased, suggesting the truncated nucleotide are redundant. Further shortening the DNAzyme to make Gd2c by deleting the nucleotides in the red box has however abolished the activity. We further made another mutant named Gd2b1 on the basis of Gd2b, where we elongated the three-base-pair stem to five base pairs. This mutant also failed to show activity, arguing against a stem structure at this position.

Overall, Gd2b appears to be an optimal enzyme, and its structure is drawn in Figure 3A. It has a cleavage rate of 0.012 min<sup>-1</sup> with 10 μM Gd<sup>3+</sup> (Figure 2C). We next measured the activity of Gd2b with different Ln<sup>3+</sup> ions (Figure 2D). The highest cleavage was observed with Gd<sup>3+</sup> and Tb<sup>3+</sup>. The activity goes down quickly beyond Ho<sup>3+</sup>, and this trend was not observed previously. Therefore, Gd2b represents another example to Ln<sup>3+</sup>-dependent DNAzymes.

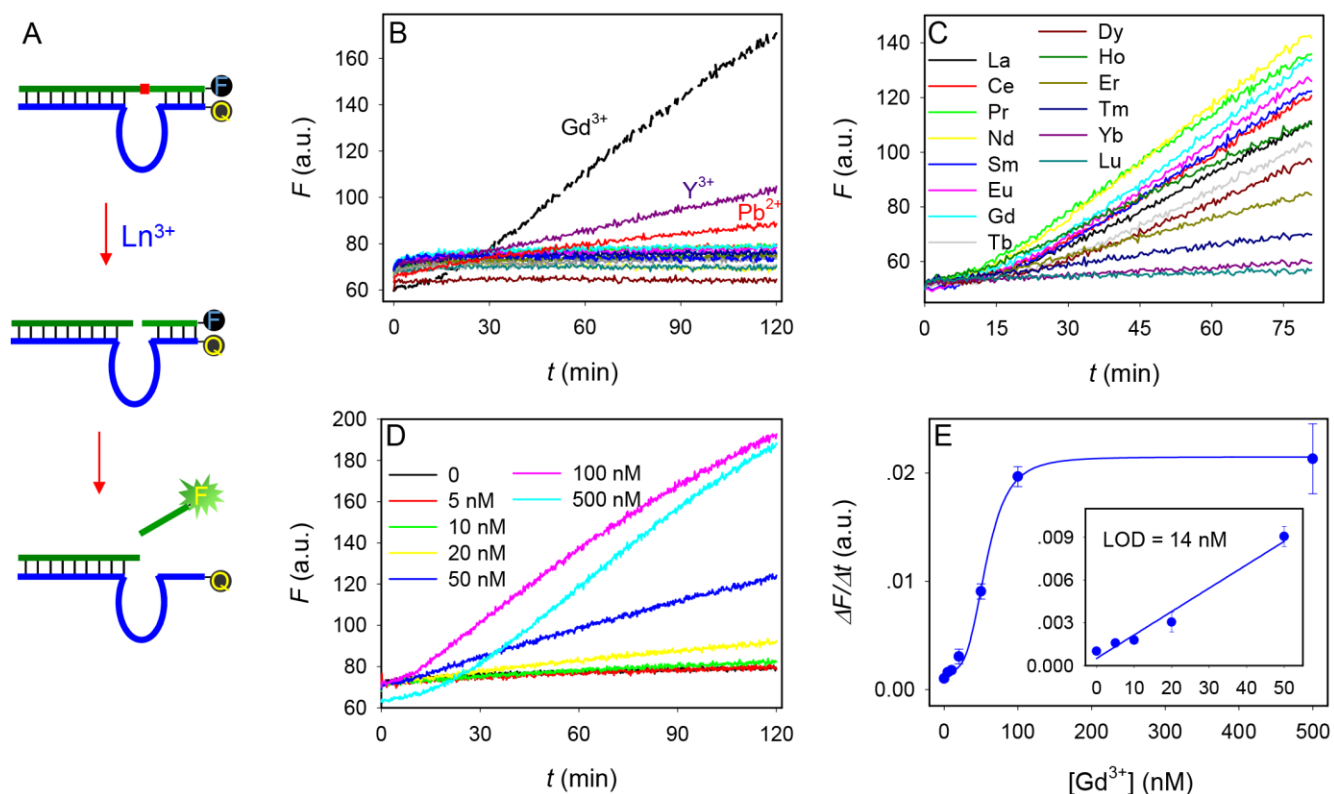




**Figure 2.** (A) The secondary structure of the Gd2 DNAzyme predicted by Mfold. The green dashed box cuts this cis-cleaving DNAzyme to the trans-cleaving version. Three truncated enzymes were tested. In each case, the sequences in the boxes were replaced by the nucleotides next to the arrowhead. (B) Activity of the truncated DNAzymes. (C) Cleavage kinetics of Gd2b in the presence of 10  $\mu\text{M}$   $\text{Gd}^{3+}$ . (D) Cleavage of the Gd2b DNAzyme by 10  $\mu\text{M}$  different  $\text{Ln}^{3+}$  for 2 h.

**A catalytic beacon for Ln<sup>3+</sup> using Gd2b.** After understanding the structure and activity of Gd2b, we next studied its analytical performance. We labeled the 3'-terminus of the substrate with a FAM and the 5'-terminus of the enzyme with a dark quencher (Figure 3A, and sequence in Figure S10A).<sup>44</sup> After cleavage and release of the cleaved fragment, the fluorescence is unmasked. To test whether Ln<sup>3+</sup> can be distinguished from other metals, we measured the sensor response to 16 divalent and trivalent metals. Other than the positive control Gd<sup>3+</sup>, only Y<sup>3+</sup> and Pb<sup>2+</sup> produced a moderate signal (Figure 3B). It is not surprising that Y<sup>3+</sup> is also active since it has the same size as Ho<sup>3+</sup>. Pb<sup>2+</sup> is a known interfering ion for RNA-cleaving DNazymes, likely due to the unique pK<sub>a</sub> value of the Pb<sup>2+</sup> bound water.<sup>9, 45</sup>

Next, the sensor response to each Ln<sup>3+</sup> was studied (Figure 3C). Most Ln<sup>3+</sup> showed a signal except for the last few, which is consistent with the result from the gel-based assay (Figure 2D). Therefore, the sensor signal is the direct result of the cleavage reaction. Next, we used Gd<sup>3+</sup> as a representative target metal since it has the fastest cleavage rate. We measured the Gd<sup>3+</sup> concentration-dependent response (Figure 3D). Below 100 nM Gd<sup>3+</sup>, we observed a linear kinetics increase. From these kinetic traces, their initial slopes were plotted (Figure 3E). Saturated signal was observed with just 100 nM Gd<sup>3+</sup>, and a detection limit of 14 nM Gd<sup>3+</sup> was achieved based on the signal greater than three times of background variation.



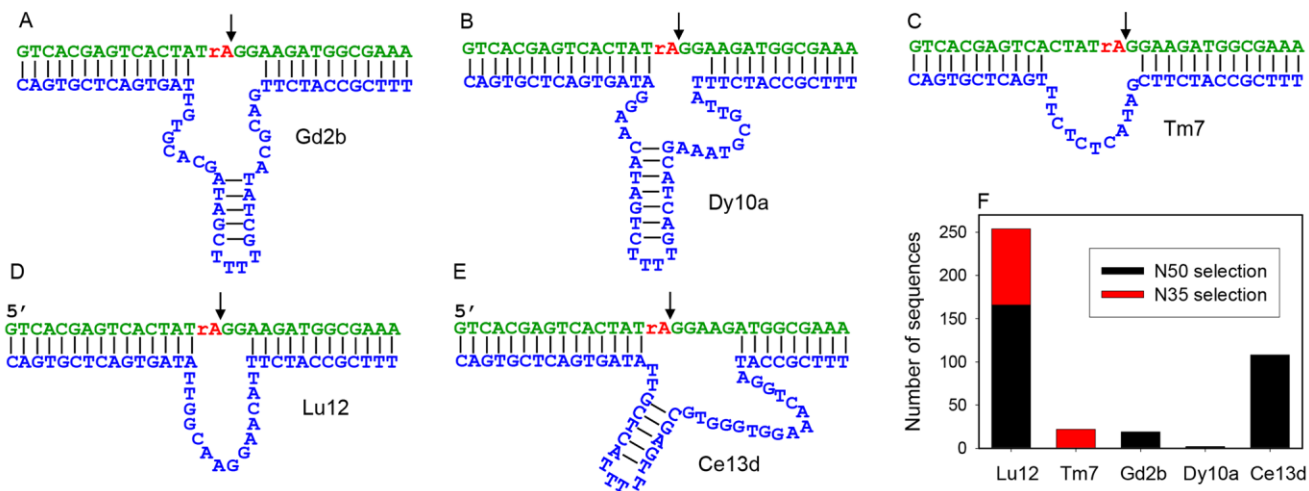
**Figure 3.** (A) A scheme of DNAzyme catalytic beacon. After  $\text{Ln}^{3+}$ -induced cleavage and release of the cleaved fragment, fluorescence signal is produced. (B) The Gd2b-based sensor selectivity test. A positive control ( $\text{Gd}^{3+}$ ) was also included. The tested competing metal ions (500 nM) include  $\text{Mg}^{2+}$ ,  $\text{Mn}^{2+}$ ,  $\text{Co}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Sr}^{2+}$ ,  $\text{Cd}^{2+}$ ,  $\text{Hg}^{2+}$ ,  $\text{Pb}^{2+}$ ,  $\text{Al}^{3+}$ ,  $\text{Sc}^{3+}$ ,  $\text{Fe}^{3+}$ ,  $\text{Ga}^{3+}$ ,  $\text{Y}^{3+}$ ,  $\text{In}^{3+}$ , and  $\text{Tl}^{3+}$ . (C) Sensor response to the 14  $\text{Ln}^{3+}$  ions (500 nM). (D) Sensor signaling kinetics with various concentrations of  $\text{Gd}^{3+}$ . (E) The rate of fluorescence enhancement as a function of  $\text{Gd}^{3+}$  concentration. Insets: the linear region at low  $\text{Gd}^{3+}$  concentrations.

**Five different DNAzymes.** By now, we have completed selection against each of the 14  $\text{Ln}^{3+}$  ions plus  $\text{Ce}^{4+}$ . A total of five independent DNAzymes were identified. In addition to the Gd2b discovered in this paper (Figure 4A), the secondary structures of the rest four are also presented (Figure 4B-E). These five DNAzymes account for ~90% of all the sequences we obtained from these selections. Their

distributions are color coded in Tables S18-19. The rest sequences are either inactive or cannot fold into a reasonable secondary structure.

The occurrence of the DNAzymes was plotted in Figure 4F. Lu12 is a dominating sequence (Figure 4D), occupying more than 50% of the final selection libraries. It appeared in both the N35 and N50 selections.<sup>11</sup> Dy10a (Figure 4B) appeared only twice in the Dy<sup>3+</sup> selection.<sup>12</sup> Tm7 (Figure 4C) has a simple loop structure and was only observed from the N35 selections.<sup>10</sup> Ce13d is a very popular sequence in the N50 library selections (Figure 4E). Ce13d has a similar activity with each Ln<sup>3+</sup>,<sup>9</sup> explaining its widespread existence.

It is interesting to note that out of the 471 sequences, only these five distinct DNAzyme families were identified. It needs to be noted that our classification was based on the conserved sequences and activity trend. For each family, there are also sequence variations in non-critical positions. Some examples of our analysis are shown in Figure S1-S9. It is likely that these five have covered most of the possible Ln<sup>3+</sup>-dependent RNA-cleaving DNAzymes, since we have selected against each Ln<sup>3+</sup>.

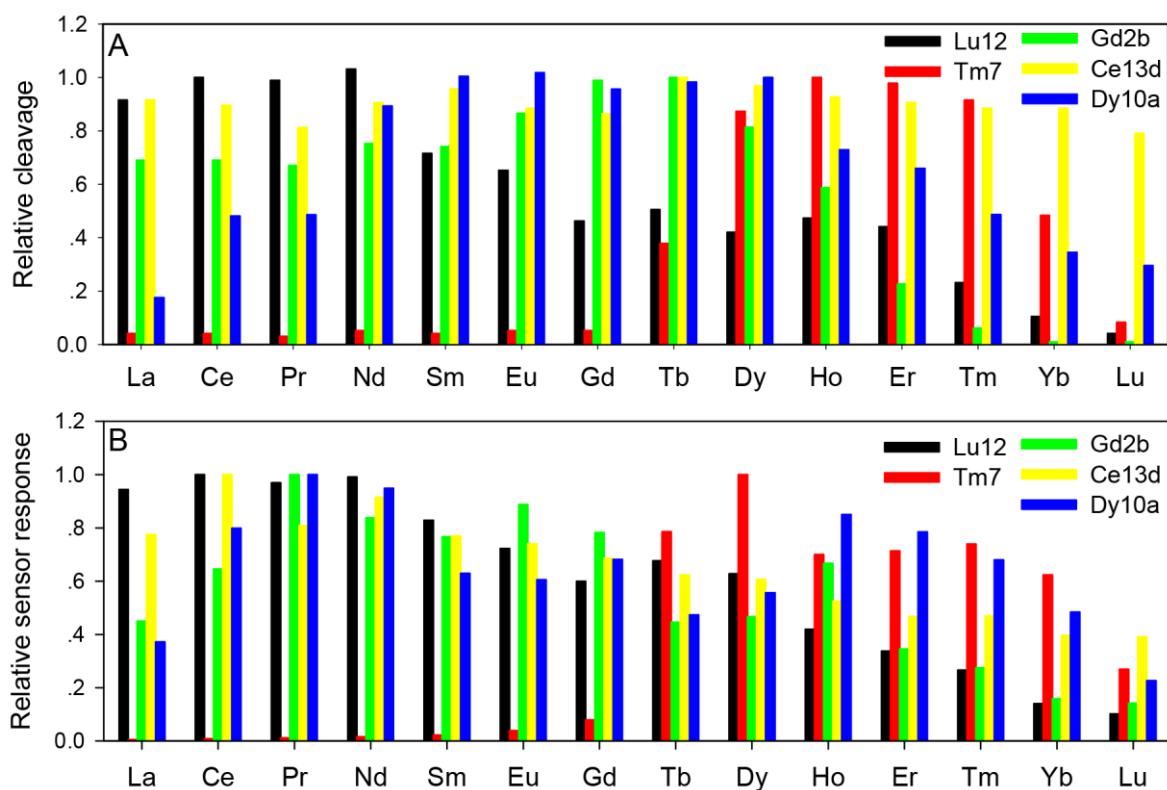


**Figure 4.** The secondary structures of the (A) Gd2b, (B) Dy10a, (C) Tm7, (D) Lu12, and (E) Ce13d DNAzymes. They all share the same substrate sequence in green and the cleavage sites are indicated by

the arrowheads. (F) The number of DNA sequences belonging to each family including all our previous  $\text{Ln}^{3+}$  selections.

**Activity trend cross the  $\text{Ln}^{3+}$  series.** These five DNAzymes have the following properties. 1) They are all quite active for  $\text{Ln}^{3+}$  as a group, but with almost no activity outside the  $\text{Ln}^{3+}$  (except for  $\text{Pb}^{2+}$  and  $\text{Y}^{3+}$  in some DNAzymes). Therefore, they are able to distinct  $\text{Ln}^{3+}$  from other metals. 2) However, none of them have an absolute selectivity for any particular  $\text{Ln}^{3+}$ . This is understandable since  $\text{Ln}^{3+}$  ions are extremely similar to each other. As a result, identification of each  $\text{Ln}^{3+}$  has to rely on the subtle differences. 3) They all have a similar secondary structures, allowing a common catalytic beacon signaling method to be used as shown in Figure 3A. With these properties, we aim to prepare a sensor array for pattern-recognition-based detection. Compared with other array strategies for  $\text{Ln}^{3+}$  discrimination,<sup>8</sup> the choice of the sensor molecules here is more targeted since they were selected based on their specific activity, which may facilitate the distinction.

With these five DNAzymes, we might have enough sensors to discriminate each  $\text{Ln}^{3+}$ . To achieve efficient discrimination, each DNAzyme should have a different activity pattern cross the series, and these enzymes can satisfy this requirement. Their gel-based cleavage activities are in Figure 5A. We normalized the data such that for each DNAzyme the most active metal gives a relative yield of 1. The Tm7 activity peaks with  $\text{Ho}^{3+}$  and  $\text{Er}^{3+}$ , descending on each side; while Gd2b peaks with  $\text{Gd}^{3+}$  and  $\text{Tb}^{3+}$ . The Dy10a has the highest activity from  $\text{Sm}^{3+}$  to  $\text{Dy}^{3+}$ .<sup>9</sup> Ce13d has a similar activity for all the  $\text{Ln}^{3+}$ , which is less useful for distinction within this  $\text{Ln}^{3+}$  group but important for separating  $\text{Ln}^{3+}$  from other metals.



**Figure 5.** Relative activities of the five DNAzymes with different Ln<sup>3+</sup> using (A) gel-based and (B) sensor-based assays. In (A), the reaction conditions were 1  $\mu$ M Ln<sup>3+</sup> for 5 min for Lu12 and Dy10a. For the rest, 10  $\mu$ M Ln<sup>3+</sup> was used with the Ce13d reaction time of 1 h, the Gd2b 2 h, and the Tm7 1 h. The reactions in (B) were all with 500 nM Ln<sup>3+</sup>. For each DNAzyme, the metal induced fastest cleavage was assigned a value of 1 and the other metals were normalized linearly accordingly. The signaling rate for Gd2b was collected at 40 min, while the other sensors were measured at 10 min.

**Pattern recognition based sensing.** Based on the above assays, it might be possible to distinguish different Ln<sup>3+</sup> using a sensor array strategy.<sup>46</sup> This ‘DNA tongue’ relies on the combined response of each sensor to the different ions. For this purpose, we used the five sensors with sequences shown in Figure S10. Four of them have been previously fully characterized, and fifth one was studied in this work. We chose to use a metal ion concentration of 500 nM to achieve a fast and stable signaling rate.

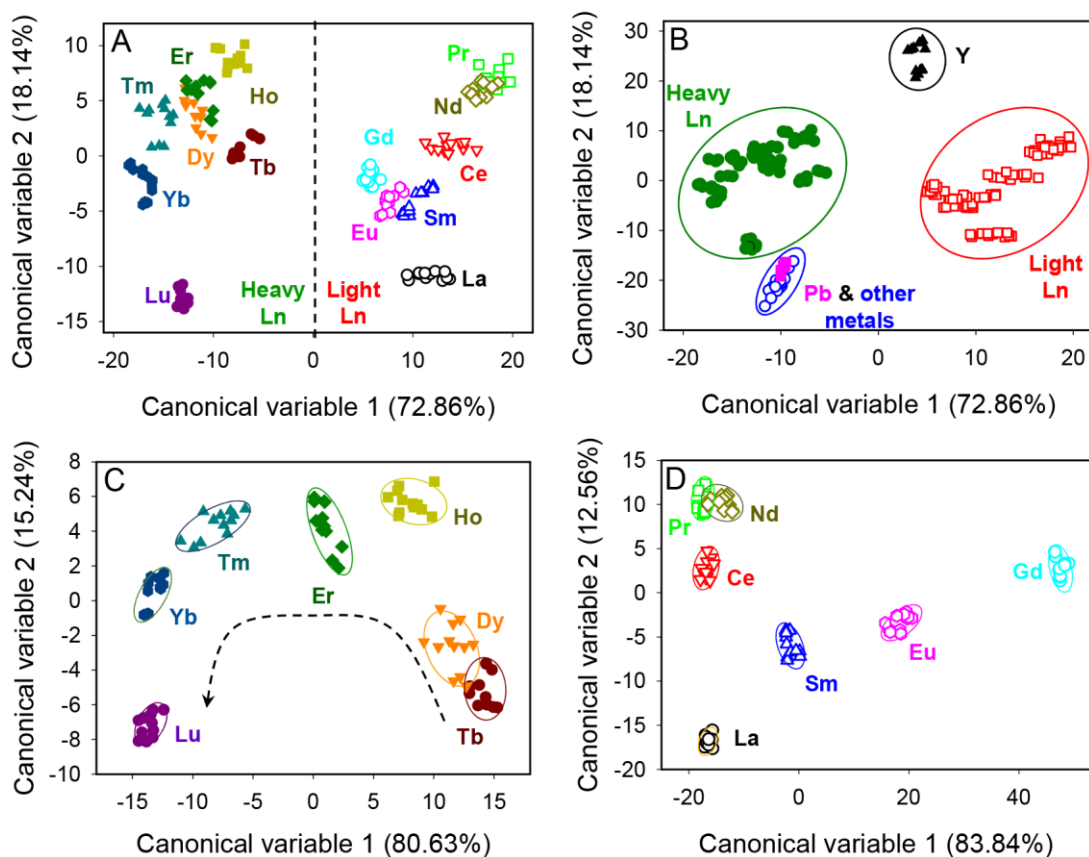
Since all the sensor responses are saturated at 500 nM  $\text{Ln}^{3+}$ , the sensor response is less affected by slight variations in the  $\text{Ln}^{3+}$  concentration near this region.

For each sensor, we picked the linear region around 10 min to obtain the rate of fluorescence enhancement (except for the Gd2b sensor whose rate was calculated at 40 min). Eventually, all the  $\text{Ln}^{3+}$  will produce a similar final fluorescence after cleavage is completed. Our choice of the detection time was to maximize the difference between different  $\text{Ln}^{3+}$ . The sensor responses are plotted in Figure 5B. In general, the trends agree with that in Figure 5A. There are a few inconsistencies though, which is likely due to the different assay conditions. For example, the gel-based assays in Figure 5A were based on a single time point, while the sensor response in Figure 5B was from the combined kinetic response of DNAzyme cleavage followed by release of the cleaved fragment. In addition, the DNAzyme and lanthanide concentrations were also different. Our array analysis was based on the sensor data in Figure 5B.

The raw data have five dimensions based on the five sensors (Table S21). To reduce dimension, linear discriminant analysis (LDA) was performed on the 14  $\text{Ln}^{3+}$  using 12 sets of data. With this, we obtained a canonical score plot (Figure 6A). Interestingly, the  $\text{Ln}^{3+}$  ions are separated into two groups: the first seven light  $\text{Ln}^{3+}$  and the last seven heavy ones reside on the two sides of the dashed line. At the dashed line, the canonical variable 1 is zero. Using these LDA parameters, we calculated the positions of other metals (Figure 6B).  $\text{Y}^{3+}$  is in a unique spot, while  $\text{Pb}^{2+}$  and other metals stand at the lower left corner, all well separated from the  $\text{Ln}^{3+}$ .

Separation for each  $\text{Ln}^{3+}$ , however, is not ideal based on Figure 6A. This can be improved by reducing the group size. A reasonable way to achieve this is to first judge whether it is a light or heavy  $\text{Ln}^{3+}$  based on Figure 6A. Then a second analysis is used for its further identification. For example, by just using the seven heavy  $\text{Ln}^{3+}$ , we achieved effective separation except for  $\text{Dy}^{3+}$  and  $\text{Tb}^{3+}$ , which are right next to each other (Figure 6C). It is interesting to note that the spots are arranged in a counter-

clockwise order according to the atomic number of the  $\text{Ln}^{3+}$ . Similarly, the light  $\text{Ln}^{3+}$  were also separated much better in the second analysis except for  $\text{Pr}^{3+}$  and  $\text{Nd}^{3+}$  (Figure 6D).



**Figure 6.** (A) Canonical score plot based on the sensor array for the 14  $\text{Ln}^{3+}$  ions. Light and heavy  $\text{Ln}^{3+}$  are separated nicely by the dashed line. (B) Using the parameters obtained in (A), the positions of other metal ions are calculated and plotted. Canonical plots of the seven heavy  $\text{Ln}^{3+}$  (C) and light  $\text{Ln}^{3+}$  (D) based on each subgroup. In these two plots, instead of using the 14  $\text{Ln}^{3+}$  ions altogether, only 7 of them were used each time.

**Further discussion.** Biosensors often rely on highly specific ligands such as antibodies and aptamers, where the goal is to achieve the highest possible selectivity towards specific target analytes.



Alternatively, less selective ligands can form an array for pattern recognition, which has the advantage of being more adaptive. Our noses and tongues are such examples. Artificial pattern recognition sensors have been demonstrated with many receptors such as polymers,<sup>47</sup> small molecule dyes,<sup>48</sup> functionalized nanomaterials,<sup>49, 50</sup> and DNA.<sup>8, 51</sup> In this work, we used Ln<sup>3+</sup>-specific DNAzymes for pattern-recognition-based detection. We reason this approach might be more efficient than using random DNA sequences or other rationally designed polymers.

Compared to other arrays, this study highlights some advantages of DNAzymes. Using combinatorial selection, the molecular recognition property of DNA can reach its full potential. For synthetic polymers, it is sometimes difficult to reach true multi-dimensional recognition based on rational design. The potential of sequence-defined polymers is more and more appreciated, and DNA is such an example.<sup>52</sup> Starting with 10<sup>14</sup> random but well-defined sequences, the subtle difference in the size of Ln<sup>3+</sup>, *pK<sub>a</sub>* of bound water, and other properties (we may not yet aware) can be maximally probed and expressed in a general final readout platform: DNAzyme beacon. Finally, the DNAzymes allow for ultrahigh (low nM and even pM) sensitivity. With these five DNAzymes, we have made a preliminary demonstration at one metal concentration of 500 nM. Identification of individual Ln<sup>3+</sup> in more complex situation remains challenging. This might be solved by incorporating more sensor components and other properties of Ln<sup>3+</sup>, such as their magnetic and luminescent properties. By using Ln<sup>3+</sup> for other DNAzyme reactions, we may also achieve further distinction.

The current discussion is centered on Ln<sup>3+</sup>, which is probably the most difficult part of distinction metal ions in the periodic tables. Outside Ln<sup>3+</sup>, many metal analytes are also of environmental and biological importance. DNAzymes sensitive to many of these metals have already been reported.<sup>24</sup> Combined with these metal-specific DNAzymes and aptamers, it might be possible to construct a much larger sensor array to discriminate more metals in the periodic table.

## Conclusions

In summary, we performed eight in vitro selections with eight different Ln<sup>3+</sup> ions in this work. A new RNA-cleaving DNzyme named Gd2b was obtained. It has the highest activity with Gd<sup>3+</sup> and Tb<sup>3+</sup>, and the activity descends on either side. This DNzyme was made into a catalytic beacon with a detection limit of 14 nM Gd<sup>3+</sup>. Combined with the previously reported four Ln<sup>3+</sup>-dependent DNzymes, a sensor array with 5 components was made. These sensors have Ln<sup>3+</sup>-size dependent activity, allowing distinction of each Ln<sup>3+</sup>. Overall, DNA can detect subtle differences of these very similar ions, suggesting the possibility of DNA-based separation of Ln<sup>3+</sup> and other applications.

## Acknowledgements

Funding for this work was from the Natural Sciences and Engineering Research Council of Canada (NSERC) Grant Application Number: 386326.

**Supporting Information Available:** The following files are available free of charge.

DNA sequences, in vitro selection conditions, sequence alignment, additional DNA folding and activity assays, and DNA sensor array data (PDF)

## References

- (1) Kobayashi, S.; Sugiura, M.; Kitagawa, H.; Lam, W. W. L., Rare-Earth Metal Triflates in Organic Synthesis. *Chem. Rev.* **2002**, 102, 2227-2302.
- (2) Dong, H.; Du, S.-R.; Zheng, X.-Y.; Lyu, G.-M.; Sun, L.-D.; Li, L.-D.; Zhang, P.-Z.; Zhang, C.; Yan, C.-H., Lanthanide Nanoparticles: From Design toward Bioimaging and Therapy. *Chem. Rev.* **2015**, 115, 10725-10815.
- (3) Atwood, D. A., Sustainability of Rare Earth Resources. In *The Rare Earth Elements: Fundamentals and Applications*, Atwood, D. A., Ed. Wiley: 2012; pp 21-25.
- (4) Vargas-Zuniga, G. I.; Sessler, J. L., Sensors for Lanthanides and Actinides. In *The Rare Earth Elements: Fundamentals and Applications*, Atwood, D. A., Ed. Wiley: 2012; pp 561-573.
- (5) Saito, S.; Hoshino, H., Highly-Sensitive Simultaneous Detection of Lanthanide(III) Ions as Kinetically Stable Aromatic Polyaminocarboxylato Complexes via Capillary Electrophoresis Using Resolution Enhancement with Carbonate Ion. *Anal. Bioanal. Chem.* **2004**, 378, 1644-1647.
- (6) Dam, H. H.; Reinhoudt, D. N.; Verboom, W., Multicoordinate Ligands for Actinide/Lanthanide Separations. *Chem. Soc. Rev.* **2007**, 36, 367-377.
- (7) Kolarik, Z., Complexation and Separation of Lanthanides(III) and Actinides(III) by Heterocyclic N-Donors in Solutions. *Chem. Rev.* **2008**, 108, 4208-4252.
- (8) Yuen, L. H.; Franzini, R. M.; Tan, S. S.; Kool, E. T., Large-Scale Detection of Metals with a Small Set of Fluorescent DNA-Like Chemosensors. *J. Am. Chem. Soc.* **2014**, 136, 14576-14582.
- (9) Huang, P.-J. J.; Lin, J.; Cao, J.; Vazin, M.; Liu, J., Ultrasensitive DNzyme Beacon for Lanthanides and Metal Speciation. *Anal. Chem.* **2014**, 86, 1816-1821.
- (10) Huang, P.-J. J.; Vazin, M.; Matuszek, Z.; Liu, J., A New Heavy Lanthanide-Dependent DNzyme Displaying Strong Metal Cooperativity and Unrescuable Phosphorothioate Effect. *Nucleic Acids Res.* **2015**, 43, 461-469.

- (11) Huang, P.-J. J.; Vazin, M.; Liu, J., In Vitro Selection of a New Lanthanide-Dependent DNAzyme for Ratiometric Sensing Lanthanides. *Anal. Chem.* **2014**, 86, 9993-9999.
- (12) Huang, P.-J. J.; Vazin, M.; Liu, J., In Vitro Selection of a DNAzyme Cooperatively Binding Two Lanthanide Ions for RNA Cleavage. *Biochemistry* **2016**, DOI: 10.1021/acs.biochem.6b00132.
- (13) Zhou, W.; Ding, J.; Liu, J., An Efficient Lanthanide-Dependent DNAzyme Cleaving 2'-5' Linked RNA. *ChemBioChem* **2016**, DOI: 10.1002/cbic.201500690.
- (14) Nishiyabu, R.; Hashimoto, N.; Cho, T.; Watanabe, K.; Yasunaga, T.; Endo, A.; Kaneko, K.; Niidome, T.; Murata, M.; Adachi, C.; Katayama, Y.; Hashizume, M.; Kimizuka, N., Nanoparticles of Adaptive Supramolecular Networks Self-Assembled from Nucleotides and Lanthanide Ions. *J. Am. Chem. Soc.* **2009**, 131, 2151-2158.
- (15) Wang, F.; Liu, B.; Huang, P.-J. J.; Liu, J., Rationally Designed Nucleobase and Nucleotide Coordinated Nanoparticles for Selective DNA Adsorption and Detection. *Anal. Chem.* **2013**, 85, 12144-12151.
- (16) Liang, H.; Jiang, S.; Yuan, Q.; Li, G.; Wang, F.; Zhang, Z.; Liu, J., Co-Immobilization of Multiple Enzymes by Metal Coordinated Nucleotide Hydrogel Nanofibers: Improved Stability and an Enzyme Cascade for Glucose Detection. *Nanoscale* **2016**, 8, 6071-6078.
- (17) Zhang, M.; Le, H.-N.; Jiang, X.-Q.; Yin, B.-C.; Ye, B.-C., Time-Resolved Probes Based on Guanine/Thymine-Rich DNA-Sensitized Luminescence of Terbium(III). *Anal. Chem.* **2013**, 85, 11665-11674.
- (18) Breaker, R. R., DNA Enzymes. *Nat. Biotechnol.* **1997**, 15, 427-431.
- (19) Santoro, S. W.; Joyce, G. F., A General Purpose RNA-Cleaving DNA Enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, 94, 4262-4266.
- (20) Navani, N. K.; Li, Y., Nucleic Acid Aptamers and Enzymes as Sensors. *Curr. Opin. Chem. Biol.* **2006**, 10, 272-281.

- (21) Lu, Y., New Transition Metal-Dependent DNazymes as Efficient Endonucleases and as Selective Metal Biosensors. *Chem. Eur. J.* **2002**, 8, 4588-4596.
- (22) Wilson, D. S.; Szostak, J. W., In Vitro Selection of Functional Nucleic Acids. *Annu. Rev. Biochem.* **1999**, 68, 611-647.
- (23) Liu, J.; Cao, Z.; Lu, Y., Functional Nucleic Acid Sensors. *Chem. Rev.* **2009**, 109, 1948–1998.
- (24) Zhang, X.-B.; Kong, R.-M.; Lu, Y., Metal Ion Sensors Based on DNazymes and Related DNA Molecules. *Annu. Rev. Anal. Chem.* **2011**, 4, 105-128.
- (25) Liu, J.; Lu, Y., A Colorimetric Lead Biosensor Using DNAzyme-Directed Assembly of Gold Nanoparticles. *J. Am. Chem. Soc.* **2003**, 125, 6642-6643.
- (26) Xiao, Y.; Rowe, A. A.; Plaxco, K. W., Electrochemical Detection of Parts-Per-Billion Lead via an Electrode-Bound DNAzyme Assembly. *J. Am. Chem. Soc.* **2007**, 129, 262.
- (27) Liu, Z.; Mei, S. H. J.; Brennan, J. D.; Li, Y., Assemblage of Signaling DNA Enzymes with Intriguing Metal-Ion Specificities and Ph Dependences. *J. Am. Chem. Soc.* **2003**, 125, 7539-7545.
- (28) Lu, L. M.; Zhang, X. B.; Kong, R. M.; Yang, B.; Tan, W. H., A Ligation-Triggered DNAzyme Cascade for Amplified Fluorescence Detection of Biological Small Molecules with Zero-Background Signal. *J. Am. Chem. Soc.* **2011**, 133, 11686-11691.
- (29) Wang, H.; Kim, Y.; Liu, H.; Zhu, Z.; Bamrungsap, S.; Tan, W., Engineering a Unimolecular DNA-Catalytic Probe for Single Lead Ion Monitoring. *J. Am. Chem. Soc.* **2009**, 131, 8221-8226.
- (30) Torabi, S.-F.; Wu, P.; McGhee, C. E.; Chen, L.; Hwang, K.; Zheng, N.; Cheng, J.; Lu, Y., In Vitro Selection of a Sodium-Specific DNAzyme and Its Application in Intracellular Sensing. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, 112, 5903-5908.
- (31) Saran, R.; Liu, J., A Silver DNAzyme. *Anal. Chem.* **2016**, 88, 4014–4020.
- (32) Breaker, R. R.; Joyce, G. F., A DNA Enzyme That Cleaves RNA. *Chem. Biol.* **1994**, 1, 223-229.

- (33) Li, J.; Zheng, W.; Kwon, A. H.; Lu, Y., In vitro Selection and Characterization of a Highly Efficient Zn(II)-Dependent RNA-Cleaving Deoxyribozyme. *Nucleic Acids Res.* **2000**, 28, 481-488.
- (34) Huang, P.-J. J.; Liu, J., An Ultrasensitive Light-up Cu<sup>2+</sup> Biosensor Using a New DNAzyme Cleaving a Phosphorothioate-Modified Substrate. *Anal. Chem.* **2016**, 88, 3341–3347.
- (35) Liu, J.; Brown, A. K.; Meng, X.; Cropek, D. M.; Istok, J. D.; Watson, D. B.; Lu, Y., A Catalytic Beacon Sensor for Uranium with Parts-per-Trillion Sensitivity and Millionfold Selectivity. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, 104, 2056-2061.
- (36) Huang, P.-J. J.; Liu, J., Rational Evolution of Cd<sup>2+</sup>-Specific Dnazymes with Phosphorothioate Modified Cleavage Junction and Cd<sup>2+</sup> Sensing. *Nucleic Acids Res.* **2015**, 43, 6125-6133.
- (37) Hollenstein, M.; Hipolito, C.; Lam, C.; Dietrich, D.; Perrin, D. M., A Highly Selective DNAzyme Sensor for Mercuric Ions. *Angew. Chem., Int. Ed.* **2008**, 47, 4346 - 4350.
- (38) Sugimoto, N.; Ohmichi, T., Site-Specific Cleavage Reaction Catalyzed by Leadzyme Is Enhanced by Combined Effect of Lead and Rare Earth Ions. *FEBS Lett.* **1996**, 393, 97-100.
- (39) Geyer, C. R.; Sen, D., Lanthanide Probes for a Phosphodiester-Cleaving, Lead-Dependent, DNAzyme. *J.Mol.Biol.* **1998**, 275, 483-489.
- (40) Dokukin, V.; Silverman, S. K., Lanthanide Ions as Required Cofactors for DNA Catalysts. *Chem. Sci.* **2012**, 3, 1707-1714.
- (41) Javadi-Zarnaghi, F.; Hobartner, C., Lanthanide Cofactors Accelerate DNA-Catalyzed Synthesis of Branched RNA. *J. Am. Chem. Soc.* **2013**, 135, 12839-12848.
- (42) Li, Y.; Breaker, R. R., Kinetics of RNA Degradation by Specific Base Catalysis of Transesterification Involving the 2'-Hydroxyl Group. *J. Am. Chem. Soc.* **1999**, 121, 5364-5372.
- (43) Zuker, M., Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction. *Nucleic Acids Res.* **2003**, 31, 3406-3415.

- (44) Li, J.; Lu, Y., A Highly Sensitive and Selective Catalytic DNA Biosensor for Lead Ions. *J. Am. Chem. Soc.* **2000**, 122, 10466-10467.
- (45) Saran, R.; Liu, J., A Comparison of Two Classic Pb<sup>2+</sup>-Dependent RNA-Cleaving Dnazymes. *Inorg. Chem. Front.* **2016**, 3, 494-501
- (46) Lavigne, J. J.; Anslyn, E. V., Sensing a Paradigm Shift in the Field of Molecular Recognition: From Selective to Differential Receptors. *Angew. Chem., Int. Ed.* **2001**, 40, 3118-3130.
- (47) Albert, K. J.; Lewis, N. S.; Schauer, C. L.; Sotzing, G. A.; Stitzel, S. E.; Vaid, T. P.; Walt, D. R., Cross-Reactive Chemical Sensor Arrays. *Chem. Rev.* **2000**, 100, 2595-2626.
- (48) Rakow, N. A.; Suslick, K. S., A Colorimetric Sensor Array for Odor Visualization. *Nature* **2000**, 406, 710-713.
- (49) De, M.; Rana, S.; Akpınar, H.; Miranda, O. R.; Arvizo, R. R.; Bunz, U. H. F.; Rotello, V. M., Sensing of Proteins in Human Serum Using Conjugates of Nanoparticles and Green Fluorescent Protein. *Nat Chem* **2009**, 1, 461-465.
- (50) Pei, H.; Li, J.; Lv, M.; Wang, J.; Gao, J.; Lu, J.; Li, Y.; Huang, Q.; Hu, J.; Fan, C., A Graphene-Based Sensor Array for High-Precision and Adaptive Target Identification with Ensemble Aptamers. *J. Am. Chem. Soc.* **2012**, 134, 13843-13849.
- (51) Yuen, L. H.; Franzini, R. M.; Wang, S.; Crisalli, P.; Singh, V.; Jiang, W.; Kool, E. T., Pattern-Based Detection of Toxic Metals in Surface Water with DNA Polyfluorophores. *Angew. Chem. Int. Ed.* **2014**, 53, 5361-5365.
- (52) Lutz, J. F.; Ouchi, M.; Liu, D. R.; Sawamoto, M., Sequence-Controlled Polymers. *Science* **2013**, 341, 1238149.