Supporting Information

for

Fragment screening against the thiamine pyrophosphate riboswitch *thiM*

Elena Cressina\textsuperscript{a}, Liuhong Chen\textsuperscript{b}, Chris Abell\textsuperscript{b}, Finian J. Leeper\textsuperscript{b} and Alison G. Smith*,\textsuperscript{a}

\textsuperscript{a} Department of Plant Sciences, University of Cambridge, Downing Street, CB2 3EA, Cambridge United Kingdom.
\textsuperscript{b} University Chemical Laboratory, University of Cambridge, Lensfield Road, CB2 1EW, Cambridge, United Kingdom

Constructs

RNA constructs

Sequence of RNAs used in the biophysical experiments; underlined regions correspond to the aptamers as defined elsewhere.\textsuperscript{1,2}

\textit{E. coli thiM}:

GGGCGAAUUGGGGCCCAGUCGCAUGCUCCCGGCAGCGCGCCGAUGGCGGGCCGCCGCGGGAAUUCGAUUGAUCAUGAA
UUCGCAACCAAACGACUCGGGGUGCCCUUCUGCGUGAAGGCUGAGAAAUACCCGUAUCACCUGAUCUGGA
UAUAGCCAGCGUAGGGAAG

\textit{B. subtilis lysC}:

GGGCGAAUUGGGGCCCAGUCGCAUGCUCCCGGCAGCGCGCCGAUGGCGGGCCGCCGCGGGAAUUCGAUUGAUCAUGAA
UUCGCAACCAAACGACUCGGGGUGCCCUUCUGCGUGAAGGCUGAGAAAUACCCGUAUCACCUGAUCUGGA
UAUAGCCAGCGUAGGGAAG

DNA constructs used for \textit{in vitro} transcription translation reporter assays

Primary sequences of the DNA templates used for the \textit{in vitro} transcription translation (IVTT) assay. Both plasmids were constructed from the vector pBluescript II KS (-) and were transcribed from the T7 promoter. The Renilla \textit{luc} gene was codon optimized for \textit{C. reinhardtii}.\textsuperscript{3}

Construct 1 (pKS- ThiM luc):

The sequence of the T7 promoter is highlighted in yellow, the \textit{thiM} RS aptamer and expression platform are highlighted in green, the ribosome binding site (RBS) is typed in red and the translation start codon is highlighted in red. The restriction sites are highlighted in grey.

\textbf{T7 promoter}:

\textbf{ThiMRiboswitch}:

\textbf{SacI} Renilla \textit{luc}
Construct 2 (pKS- luc): Control construct were the Renilla luc gene is not under control of the thiM riboswitch. The colour coding is the same described for the previous construct, except for the sequence in green, now contains only the last 41 nt of thiM-RS expression platform.

```
T7 promoter
GTAATACGACTCACTATAGGGCGATTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAAGGAG

SacI   Renilla luc
CAAGTCGACCTGCTGGGTTCAGCGCAAGAGCTCATGGCCAGCAAGGTGTACGCCCCCGAGCAGCGCAAGCG
CATGATCACCGGCCCTCAGTGGTGGGCTCGCTGCAAGCAGATGAACGTGCTGGACAGCTTCATCAACTACT
ACGACAGCGAAGACGCGCAAGCGCAGCCGCAAGCGCAGCTACCGCCTGCTGGACCACTACAAGTACCTGACCGCCTGGTTCGAGCTGCTGAACCTGCCCAAGAAGAT
CATCTTCGTGGGCCACCGACCTTGCCGCTGGGATTCGAGCTGCTGAACCTGCCCAAGAAGATCATCTTCGTGGGCCACGACTGGGGCGCCTGCCTGGCCTTCCACTACAGCTACGAGCNCCAGGACAAGATCAAGGCCATCGTGCACGCCGAGAGCGTGGTGGACGTGATCGAGAGCTGGGACGAGTGGCCCGACATCGAGGAGGACATCGCCCTGATCAAGAGCGAGGAGGGCGAGAAGATGGTGCTGGAGAACAACTTCTTCGTGGAGACCATGCTGCCCAGCAAGATCATGCGCAAGCTGGAGCCCGAGGAGTTCGCCGCCTACCTGGAGCCCTTCAAGGAGAAGGGCGAGGTGCGCCGTCCCACCCTGAGCTGGCCTCGCGAGATCCCCCTGGTGAAAGGGCGGCAAGCCCGACGTGGTGCAGATCGTGCGCAACTACAACGCCTACTTGCGCGCCAGCGACGACC
```

HindIII
GGGCGTATTTGAAGCGAATTCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTT
GTTCCCTTTAGT...
Supplementary Figures and Tables

**Figure S1**: Fluorescence melting curve of *thiM*-RS (0.5 µM), with the dye EvaGreen ± 50 µM TPP.

![Fluorescence melting curve](chart)

**Table S1**: Structures and equilibrium dialysis [³H]-thiamine displacement percentages of fragments S1-S3 for which $K_D$ could not be determined by ITC.

<table>
<thead>
<tr>
<th>Fragment</th>
<th>Structure</th>
<th>% thiamine displacement</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td><img src="image" alt="Structure S1" /></td>
<td>41</td>
</tr>
<tr>
<td>S2</td>
<td><img src="image" alt="Structure S2" /></td>
<td>31</td>
</tr>
<tr>
<td>S3</td>
<td><img src="image" alt="Structure S3" /></td>
<td>31</td>
</tr>
</tbody>
</table>
Table S2: WaterLOGSY (left panels) and $T_2$ relaxation-edited (right panels) NMR spectra of fragments 1-17, S1-S3, 24-26. The chemical shift $\delta$ (ppm) scale is indicated for each spectrum.
Ligand only

waterLOGSY

+ RNA

Ligand only

$T_2$-relaxation edited NMR

+ RNA

Ligand only

Ligand only

6

N

H

N

H

waterLOGSY

+ RNA

Ligand only

$T_2$-relaxation edited NMR

+ RNA

Ligand only

Ligand only

N

N

NH

NH

7

N

H

N

H

waterLOGSY

+ RNA

Ligand only

$T_2$-relaxation edited NMR

+ RNA

Ligand only

Ligand only
Supplementary Material (ESI) for Chemical Science
This journal is (c) The Royal Society of Chemistry 2010

**Chemical Structures and Spectra**

1. **Compounds:**
   - **11**
   - **12**
   - **13**

2. **Spectra Types:**
   - WaterLOGSY
   - T$_2$-relaxation edited NMR

3. **Spectra Details:**
   - Each compound shows spectra with and without RNA present.
   - Spectra are labeled as **Ligand only** when RNA is not present.

4. **Chemical Shifts:**
   - The spectra show chemical shifts in the range of 8.0 to 2.0.

5. **Conclusion:**
   - The spectra provide evidence for the interaction of the ligands with RNA, highlighting the potential binding modes.

---

**Figures:**

1. **Figure 1:** WaterLOGSY spectra for compounds 11 and 12 with and without RNA.
2. **Figure 2:** T$_2$-relaxation edited NMR spectra for compounds 13 with and without RNA.

---

**Additional Observations:**

- The waterLOGSY spectra display differences in peak integration with the presence of RNA, suggesting specific interactions.
- The T$_2$-relaxation edited NMR spectra show altered relaxation times, indicating potential changes in the conformation or dynamics of the ligands in the presence of RNA.
Supplementary Material (ESI) for Chemical Science
This journal is (c) The Royal Society of Chemistry 2010

$\text{NH}_2$
$\text{CN}$

waterLOGSY

$T_2$-relaxation edited NMR

+ RNA

Ligand only

$\text{NH}_2$
$\text{CO}_2\text{CH}_3$

waterLOGSY

$T_2$-relaxation edited NMR

+ RNA

Ligand only

$\text{NH}_2$
$\text{S}$
$\text{OH}$

waterLOGSY

$T_2$-relaxation edited NMR

+ RNA

+ Ligand

+ TPP

+ Asterix denote TPP peaks

+ RNA

Ligand only

Supplementary Material (ESI) for Chemical Science
This journal is (c) The Royal Society of Chemistry 2010
Table S3: Isothermal titration thermograms of fragments 1-3, 5-17, 24, 25. Isothermal titration thermogram of fragment 4 is shown in Figure 4a. The thermodynamic parameters for each titration are indicated in the boxes. N is the calculated number of ligand molecules that bind to one RNA molecule, K (M⁻¹) is the thermodynamic association constant, ΔH and ΔS are the enthalpy and entropy variation respectively.

- **Fragment 1:**
  - N: 0.632 ± 0.0268 Sites
  - K: 4.54E4 ± 4.83E3 M⁻¹
  - ΔH: -1.412E4 ± 762.9 cal/mol
  - ΔS: -26.1 (cal/mol)/deg

- **Fragment 2:**
  - N: 0.654 ± 0.0511 Sites
  - K: 2.01E4 ± 1.82E3 M⁻¹
  - ΔH: -1.317E4 ± 1202 cal/mol
  - ΔS: -24.5 (cal/mol)/deg

- **Fragment 3:**
  - N: 1.00 ± 0.0671 Sites
  - K: 1.77E4 ± 1.94E3 M⁻¹
  - ΔH: -7397 ± 601.1 cal/mol
  - ΔS: -5.38 (cal/mol)/deg

- **Fragment 4:**
  - N: 0.632 ± 0.0651 Sites
  - K: 9.68E3 ± 629 M⁻¹
  - ΔH: -3966 ± 185.7 cal/mol
  - ΔS: 4.93 (cal/mol)/deg

- **Fragment 5:**
  - N: 5.85 ± 0.367 sites
  - K: 3.04E3 ± 261 M⁻¹
  - ΔH: -1004 ± 79.87 cal/mol
  - ΔS: 12.6 (cal/mol)/deg

- **Fragment 6:**
  - N: 1.72 ± 0.0651 Sites
  - K: 2.01E4 ± 1.82E3 M⁻¹
  - ΔH: -1.317E4 ± 1202 cal/mol
  - ΔS: -24.5 (cal/mol)/deg

- **Fragment 7:**
  - N: 1.00 ± 0.0671 Sites
  - K: 1.77E4 ± 1.94E3 M⁻¹
  - ΔH: -7397 ± 601.1 cal/mol
  - ΔS: -5.38 (cal/mol)/deg

- **Fragment 8:**
  - N: 0.632 ± 0.0268 Sites
  - K: 4.54E4 ± 4.83E3 M⁻¹
  - ΔH: -1.412E4 ± 762.9 cal/mol
  - ΔS: -26.1 (cal/mol)/deg

- **Fragment 9:**
  - N: 1.00 ± 0.0671 Sites
  - K: 1.77E4 ± 1.94E3 M⁻¹
  - ΔH: -7397 ± 601.1 cal/mol
  - ΔS: -5.38 (cal/mol)/deg

- **Fragment 10:**
  - N: 4.90 ± 0.340 Sites
  - K: 3.10E3 ± 279 M⁻¹
  - ΔH: -3124 ± 386.9 cal/mol
  - ΔS: 6.79 (cal/mol)/deg

This journal is (c) The Royal Society of Chemistry 2010

Supplementary Material (ESI) for Chemical Science
Figure S2: Linear correlation between the luminescence signal and the amount of IVTT reaction mix. A complete IVTT reaction was incubated at 37°C for 1 hour, using the construct 1 as DNA template. The mix was then diluted 4 fold, and aliquots of different volumes were taken, diluted to 50 μL and assayed for luciferase activity as described in the Experimental Section. The x-axis shows the volume of the undiluted IVTT reaction mix used in the assays. The experimental values (red squares) were fitted with a line (black).

\[ y = 4E+06x + 406472 \]
\[ R^2 = 0.9761 \]

Figure S3: TPP (0.1 mM) effect on Renilla luciferase expression in IVTT reactions with construct 1, containing the thiM-RS and construct 2 (pKS- luc) not containing the riboswitch thiM-RS. Red bars indicate refer to IVTT performed with 0.1 mM TPP, green bars to IVTTS without TPP. The experiments were performed in duplicate and the error bars were calculated from the experimental error on the duplicates.
**Figure S4:** Effect of the compounds 1-5 and 22-25 on *Renilla* luciferase expression in IVTT systems. The compounds at 100 μM concentration were added to IVTT reactions in presence or absence of 100 μM TPP and incubated for 2 hours at 37°C. The plot shows the normalised luminescence (obtained by dividing the luminescence of each incubation by that of the control reaction without any compound and without TPP). The incubations were performed in duplicate and the error bars were calculated from the experimental error on the duplicates.

![Graph showing effect of compounds on Renilla luciferase expression](image_url)

**References**