Supporting information of the manuscript:

Crystal structure and fluorescence properties of the iSpinach aptamer in complex with DFHBI

Pablo Fernandez-Millan, Alexis Autour, Eric Ennifar, Eric Westhof & Michael Ryckelynck*

Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR 9002, F-67000 Strasbourg, France

TABLE S1. Constructions used in this study.

Construct name	RNA construct sequence*	Forward DNA primer **	Reverse DNA primer
iSpinach-WT	GCGACUACGGUGAGGGUCGGG UCCAGUAGCUUCGGCUACUGU UGAGUAGAGUGUGGGCUCCG UAGUCGC	GCATGC*TAATACGACTCACTATA*G CGAGTACGGTGAGGGTCGGG	GCGAGTACGGAGCCCACACTCTA C
iSpinach-D5	**GGGAG**UACGGUGAGGGUCGG GUCCAGUAGCUUCGGCUACUG UUGAGUAGAGUGUGGGCUCC GTA**CUCCC**	GCATGC*TAATACGACTCACTATA*G GGAGTACGGTGAGGGTCGGG	GGGAGTACGGAGCCCACACTCTA C
iSpinach-D5-D28	**GGGAG**UACGGUGAGGGUCGG GUCCAGUGCUUCGGCUACUGU UGAGUAGAGUGUGGGCUCCG UA**CUCCC**	GCATGC*TAATACGACTCACTATA*G GGAGTACGGTGAGGGTCGGG	GGGACTACGGAGCCCACACTCTA C
iSpinach-D5-C32	**GGGAC**UACGGUGAGGGUCGG GUCCAGUAGCU**C**GGCUACUGU UGAGUAGAGUGUGGGCUCCG UA**GUCCC**	GCATGC*TAATACGACTCACTATA*G GGAGTACGGTGAGGGTCGGG	GGGACTACGGAGCCCACACTCTA C
iSpinach-D5-G30-G32	**GGGAC**UACGGUGAGGGUCGG GUCCAGUAG**G**U**G**CG**C**CUACU GUUGAGUAGAGUGUGGGCUC CGUA**GUCCC**	GCATGC*TAATACGACTCACTATA*G GGAGTACGGTGAGGGTCGGG	GGGACTACGGAGCCCACACTCTA C
ISpinach-D5-G30-A32	**GGGAC**UACGGUGAGGGUCGG GUCCAGUAG**G**U**A**CG**C**CUACUG UUGAGUAGAGUGUGGGCUCC GUA**GUCCC**	GCATGC*TAATACGACTCACTATA*G GGAGTACGGTGAGGGTCGGG	GGGACTACGGAGCCCACACTCTA C

* Construct sequence resulting from the transcription

** T7 promoter sequence (*TAATACGACTCACTATA*) is italicized.

The differences with wild type sequence are bolded and underlined

TABLE S2. Data for the crystallogenesis experiments of Aptamer/DFHBI complexes.

Construct name	**Crystallographic conditions and results**	**Crystal shape**	**Data collected result-max resolution**	**Relative fluorescence***
iSpinach-WT	No crystal	No crystal	----	1
iSpinach-D5**	40 mM NaCl, 40 mM Sodium cacodylate, pH 7.0, 32% MPD, 12 mM Spermine	Large needles	No diffraction	0.98 ± 0.19
iSpinach-D5**	1.9 M Li_2SO_4, MES 50 mM, pH 5.6, MgCl2 5 mM	Lentils/Ball shape	2.4 Å / complete data set***	0.98 ± 0.19
iSpinach-D5-C32	No crystal	No crystal	----	0.97 ± 0.28
iSpinach-D5-D28	No crystal	No crystal	----	1.10 ± 0.05
iSpinach-D5-G30A32	40 mM NaCl, 40 mM Sodium cacodylate pH 7.0, 32% MPD, 12 mM Spermine	Laminar crystal	2.0 Å / complete data set	1.11 ± 0.11
iSpinach-D5-G30A32	1.9M Li_2SO_4, MES 50 mM, pH 5.6, $MgCl_2$ 5mM	Lentils/Ball	2.4 Å / complete data set***	1.11 ± 0.11
iSpinach-D5-G30G32	40 mM NaCl, 40 mM Sodium cacodylate pH 7.0, 32% MPD, 12 mM Spermine	Small Elliptic	15 Å / no data set collected	1.01 ± 0.11

* These values are the mean ± SD of 3 independent experiments
** This variant possesses a wild type apical tetraloop
*** This data set did not allow building a reliable structure

TABLE S3. Data collection and refinement statistics of iSpinach-D5-G30A32/DFHBI complex.

	iSpinach-D5-G30A32 /K^+ * / **
Wavelength (Å)	1.800020
Resolution range	46.17 - 2.05 (2.12 - 2.05)
Space group	C 1 2 1
Unit cell	150.7 48.5 30.1 90 91.855 90
Total reflections	415162
Unique reflections	26531 (2470)
Multiplicity	15.6 (9.2)
Completeness (%)	0.99 (91.7)
Mean I/sigma(I)	17.6 (1.9)
Wilson B-factor	41.05
R-merge	0.103 (1.22)
R-meas	0.106 (1.290)
CC1/2	0.99(0.82)
Reflections used in refinement	13692 (1284)
Reflections used for R-free	694 (61)
R-work	0.1961 (0.2797)
R-free	0.2213 (0.2894)
Number of non-hydrogen atoms	1568
RNA	1483
ligands	36
RMS (bonds)	0.011
RMS (angles)	1.74
Clash score	17.48
Average B-factor* **	61.64
RNA	62.62
ligands	46.33
solvent	43.07

* Friedel mates were averaged when calculating reflection statistics.

** Statistics for the highest-resolution shell are given in parentheses.

*** We noticed that RNA B-factor exceeds that of the ligands and of the solvent. However, one should consider that, for the RNA, the value corresponds to an average value computed from individual B-factors spanning an order of magnitude (Supplemental Fig. S8). A local comparison of B-factors shows that the value for the RNA is lower than that of the ligands and of the solvent as expected (see Figure S8).

TABLE S4. RSMD of iSpinach against to Spinachs structures. The RMSD values were measured with full-length or core molecules as indicated.

Structure 1*	Structure 2*	RMSD (Å)
iSpinach	Spinach-P	1.74
iSpinach	Spinach-FdA	1.64
iSpinach-core	Spinach-P-core	0.58
iSpinach-core	Spinach-FdA-core	0.35
Spinach-P	Spinach-FdA	1.58
Spinach-P-core	Spinach-FdA-core	0.61

* Spinach-P corresponds to 4KZD (Huang et al. 2014) and Spinach-FdA corresponds to 4TS0 (Warner et al. 2014).

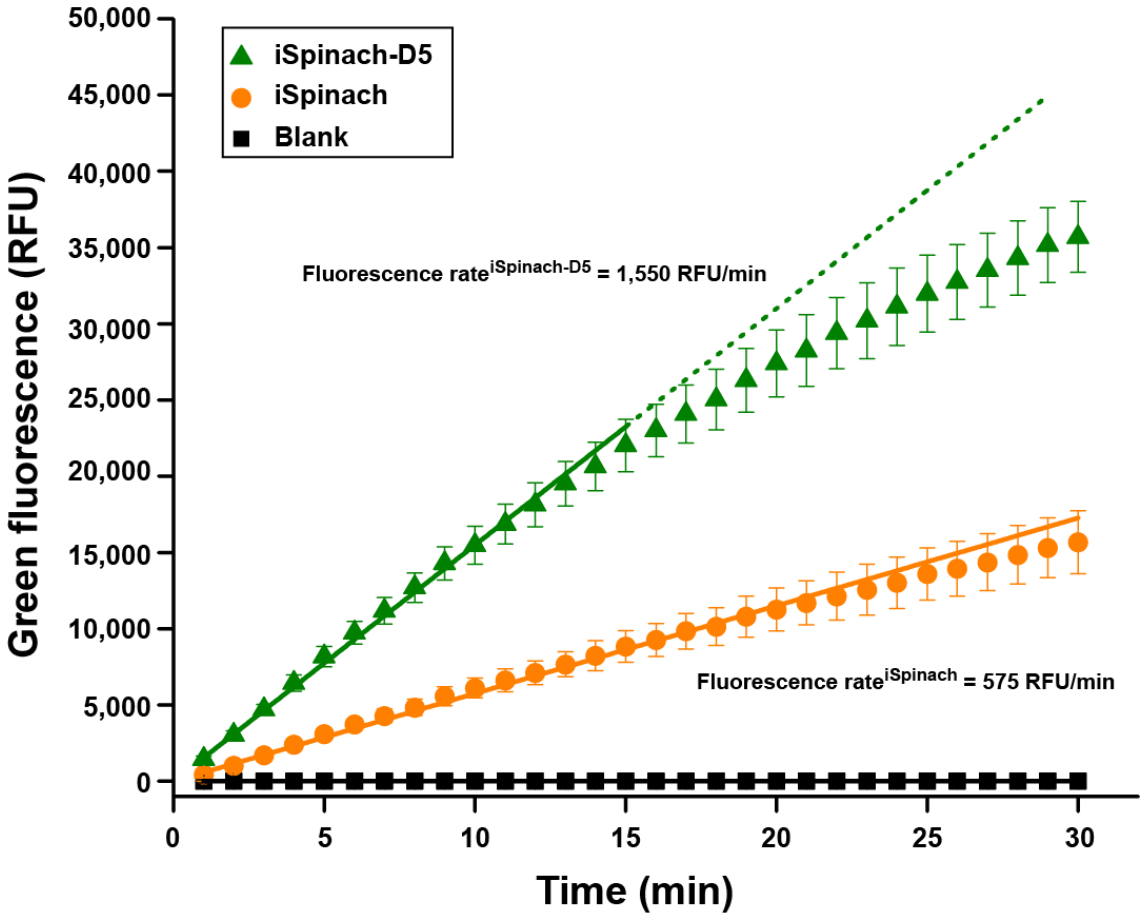

FIGURE S1. Transcription efficiency of the genes coding for iSpinach and its D5 variants. An identical amount of DNA coding for iSpinach or iSpinach-D5 placed under the control of the T7 RNA polymerase promoter was *in vitro* transcribed in the presence of 20 μM of DFHBI at 37°C. The green fluorescence (ex: 492 nm/em: 516 nm) was monitored every minute on a real-time thermocycler (Mx 3005P, Agilent). The transcription efficiency can be evaluated from the fluorescence apparition rate (value given for each construct) during the linear phase. Since both RNAs have the same size and the same relative fluorescence (Supplemental Table S2), the 3-time higher rate observed with iSpinach-D5 can be assigned to a better transcription rate as compared with iSpinach. The values are the mean of three independent experiments and the error bars correspond to ± 1 standard error.

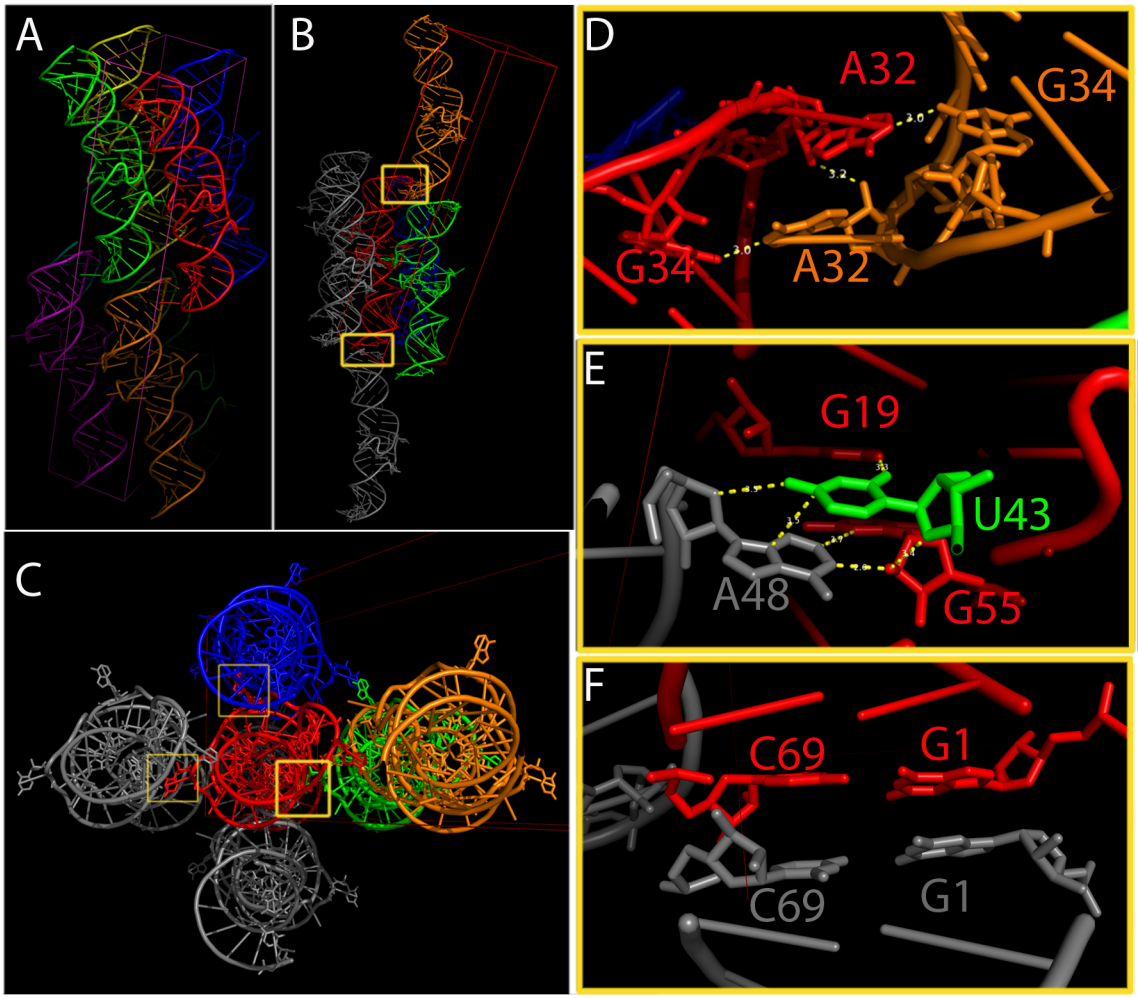

FIGURE S2. Crystal packing and molecular interactions. (A) Crystal packing. The application of the crystallographic symmetry to the red molecule generated the seven other molecules contained in the unit cell, each molecule is shown in a different color. (B) Molecular interactions established with the red molecule. The regions where interactions occur are squared in yellow. (C) Top view of the unit cell. (D) to (E) Zoom on the regions where interactions occur. The residues are shown in the same color than the molecule they originate from.

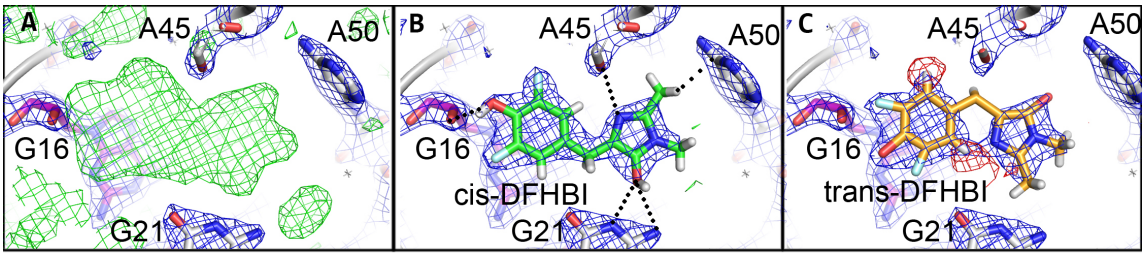

FIGURE S3. DFHBI is present in the *cis* isomeric form in the crystal. (A) Density map in the DFHBI-binding pocket. The Fo-Fc map (green) is shown in the absence of any ligand modeled in the DFHBI-binding pocket. (B) The 2Fo-Fc map (blue) is shown with the DFHBI (green) modeled in its *cis* isomeric form. Interactions (dash lines) are represented and no incompatibility was observed. (C) The 2Fo-Fc map (blue) is shown with the DFHBI (orange) modeled in its *trans* isomeric form. The recalculation of the electron density maps indicates the wrong position of the molecule with 2Fo-Fc maps (blue) and the negative values (incompatibilities) of Fo-Fc maps are represented (red).

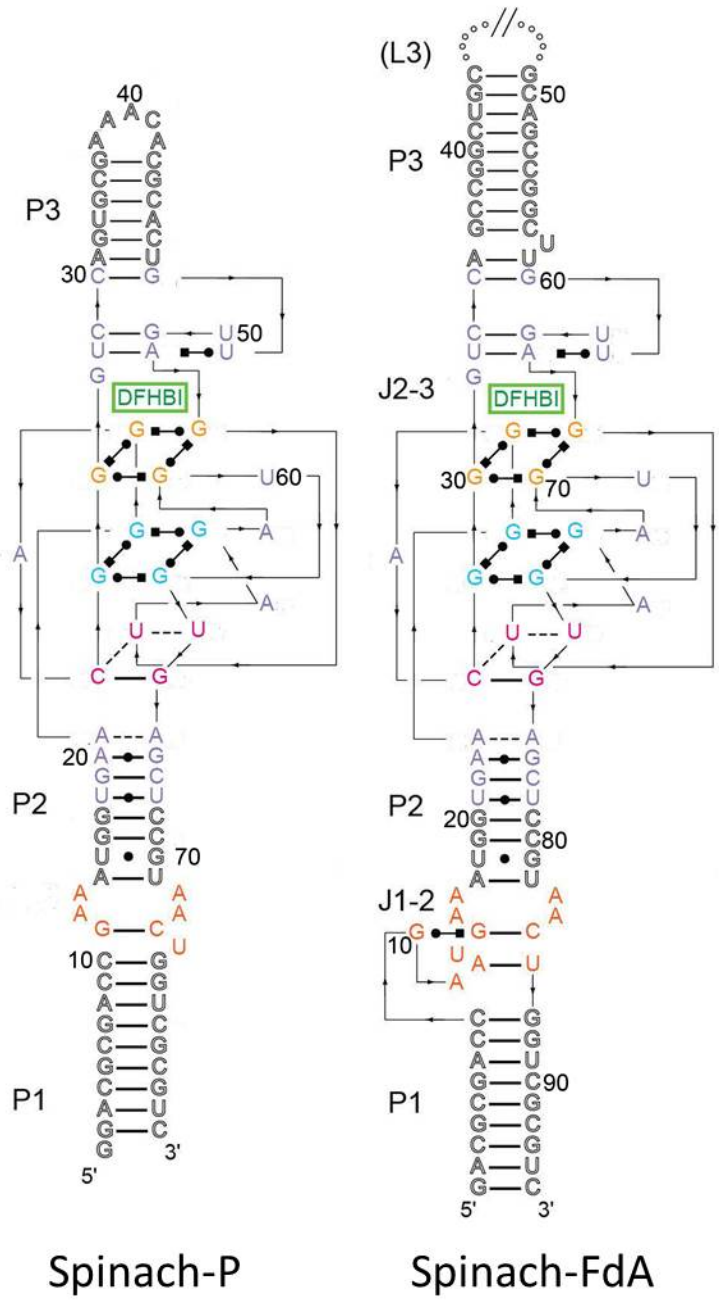

FIGURE S4. Sequences and secondary structures of Spinach-P and Spinach-FdA. The molecules are shown as initially reported (Huang et al. 2014; Warner et al. 2014).

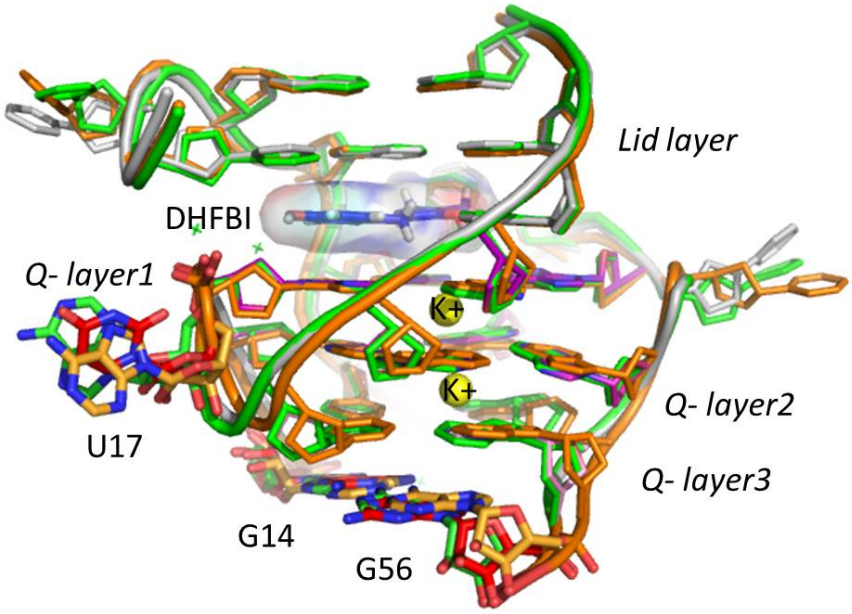

FIGURE S5. Superposition of the region involved in DFHBI binding. The structures of the core region of the three molecules (iSpinach-D5-G30-A32, Spinach-P and Spinach-FdA) are superposed. This region includes the GoG pair, the three layers of G-quadruplex, the DFHBI pocket, the lid layer and the first base pair of the apical stem. iSpinach-D5-G30-A32 is shown in gray with the mutations colored in red while Spinach-P and Spinach-FdA are shown in green and orange respectively.

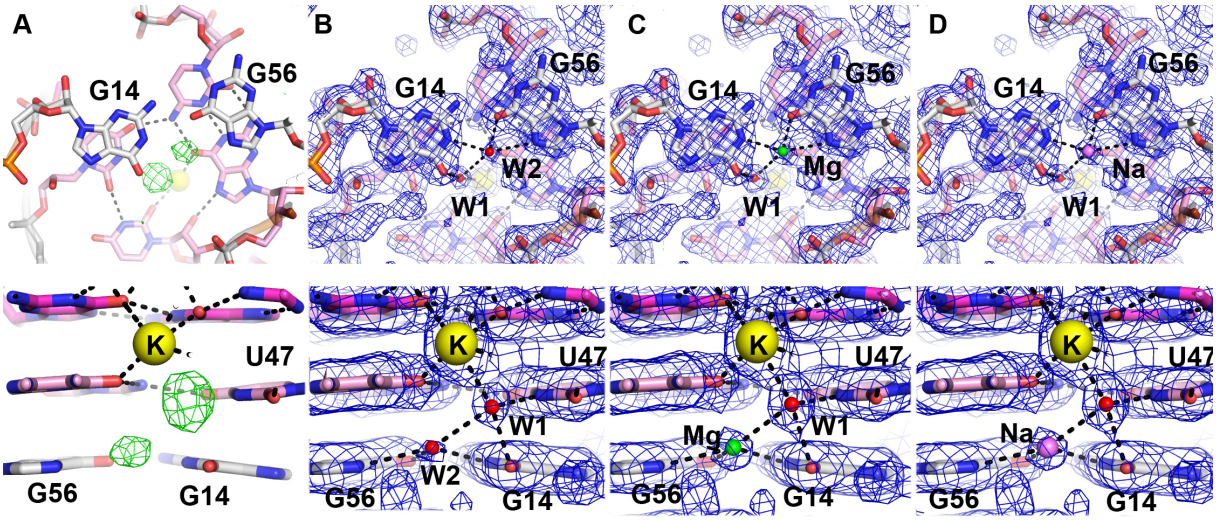

FIGURE S6. The GoG base pair water network. The Fo-Fc map (green) was generated at 5 σ (A). Then, either (B) two water (W1 and W2, red spheres) molecules, or (C) one water molecule (W1, red sphere) and one magnesium ion (Mg, green sphere), or (D) one water molecule (W1, red sphere) and one sodium ion (Na, violet sphere) were fit in the 2Fo-Fc maps. On the top, the GoG pair is shown parallel to the helical axis of the molecule and, on the bottom, a perpendicular view is shown.

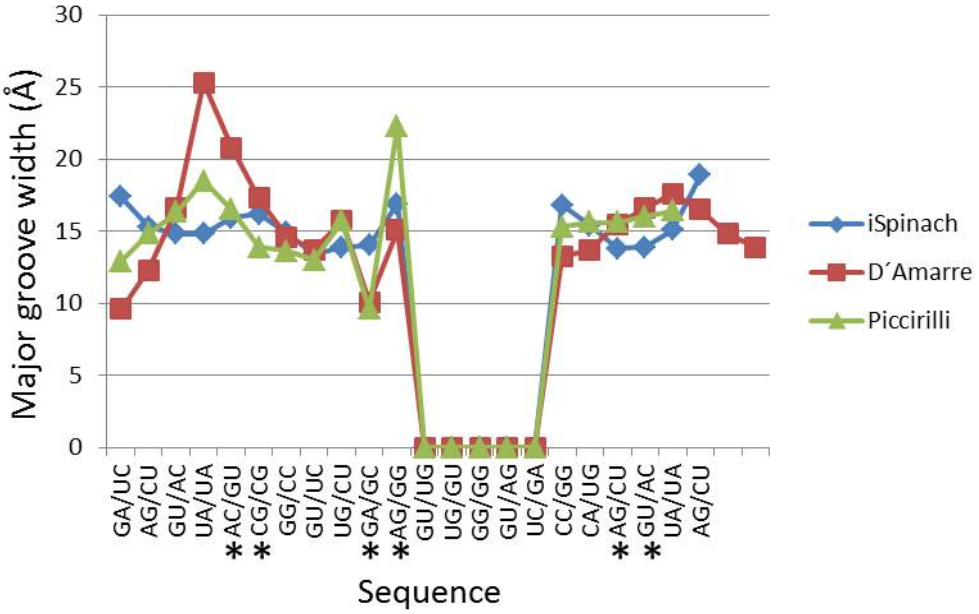

FIGURE S7. Major groove distances between the phosphates were calculated with 3DNA. The blue line corresponds to iSpinach structure, whereas the red and the green lines correspond respectively to Spinach-FdA and Spinach-P. As iSpinach structure is shorter than Spinach, the supplementary base pairs of Spinach-FdA and Spinach-P structures were not plotted. The stars indicate the steps with a mutation in iSpinach molecule.

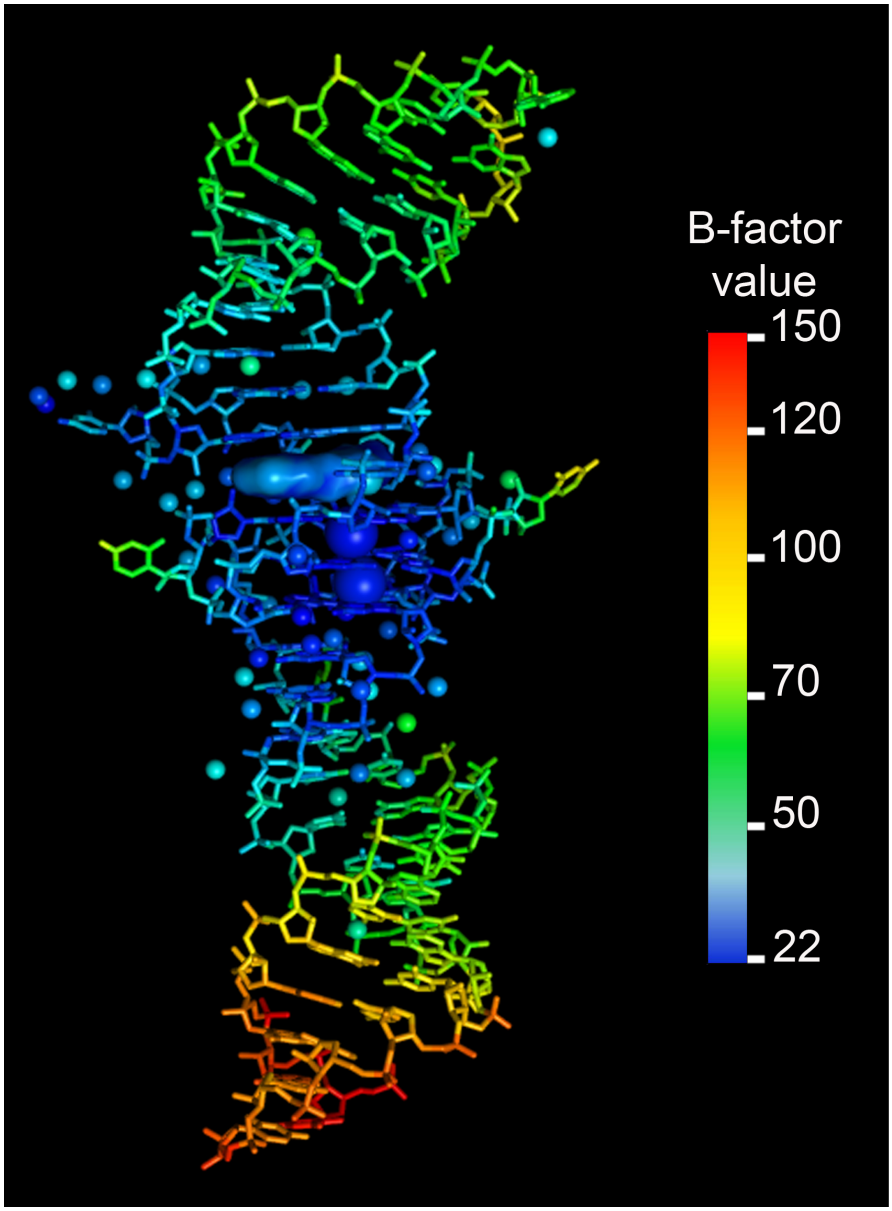

FIGURE S8. The distribution of the B-factors in the structure. The B-factor of each atom has been color-coded with the RNA shown as sticks, water molecules as small spheres, and potassium ions as large spheres and the DFHBI as a surface. Note that, whereas B-factors tend to be low in the middle of the structure, they increase when moving towards the extremities.

REFERENCES

Huang H, Suslov NB, Li NS, Shelke SA, Evans ME, Koldobskaya Y, Rice PA, Piccirilli JA. 2014. A G-quadruplex-containing RNA activates fluorescence in a GFP-like fluorophore. *Nature chemical biology* **10**: 686-691.

Warner KD, Chen MC, Song W, Strack RL, Thorn A, Jaffrey SR, Ferre-D'Amare AR. 2014. Structural basis for activity of highly efficient RNA mimics of green fluorescent protein. *Nature structural & molecular biology* **21**: 658-663.