nature chemical biology

ARTICLES

https://doi.org/10.1038/s41589-021-00884-6

Check for updates

Structure-based investigation of fluorogenic Pepper aptamer

Kaiyi Huang[1,4], Xianjun Chen[2,3,4], Chunyan Li[1], Qianqian Song[1], Huiwen Li[2,3], Linyong Zhu[2]✉, Yi Yang[2,3]✉ and Aiming Ren[1]✉

Pepper fluorescent RNAs are a recently reported bright, stable and multicolor fluorogenic aptamer tag that enable imaging of diverse RNAs in live cells. To investigate the molecular basis of the superior properties of Pepper, we determined the structures of complexes of Pepper aptamer bound with its cognate HBC or HBC-like fluorophores at high resolution by X-ray crystallography. The Pepper aptamer folds in a monomeric non-G-quadruplex tuning-fork-like architecture composed of a helix and one protruded junction region. The near-planar fluorophore molecule intercalates in the middle of the structure and is sandwiched between one non-G-quadruplex base quadruple and one noncanonical G·U wobble helical base pair. In addition, structure-based mutational analysis is evaluated by in vitro and live-cell fluorogenic detection. Taken together, our research provides a structural basis for demystifying the fluorescence activation mechanism of Pepper aptamer and for further improvement of its future application in RNA visualization.

The spatiotemporal localization and tracing of biomacromolecules, such as proteins and RNAs, are essential for investigation of their functions, mechanisms and interplay in biology. The discovery and development of fluorescent proteins (FPs) have provided a marked increase in the applicability of live-cell protein imaging impacting function[1]. Live-cell RNA imaging, however, does not yet have such a powerful tag, given that no intrinsically fluorescent RNAs have ever been found in nature. Current state-of-the-art methods to detect the distribution and level of cellular RNA include fluorescence in situ hybridization (FISH)[2] and engineered RNA motifs tethering fluorescent protein fusions with specific RNA-binding proteins, for example, MS2, λ_N and even Cas[3–7]. However, FISH requires cell fixation and is not suitable for live-cell RNA imaging. By comparison, fluorescent protein tethering systems could allow real-time imaging of single RNA molecules in live cells, but still they suffer from high background signal due to unbound fluorescent proteins, and the dependence on many copies of tandem repeat RNA motifs, which might interfere with the biological functions of the target RNA[8,9].

Fluorescent RNAs (FRs), the complex of RNA aptamers with their cognate dye ligands, mimic the fluorogenic properties of fluorescent proteins and enable genetically encoded labeling of target RNA[10–29]. Despite RNA tagging and imaging being achieved in live cells, many FRs still suffer from limitations, such as high background[11–15,21–25], weak cellular brightness[15–19], limited thermostability or photostability[15–19,30,31] or their multimeric nature[19,32]. To this end, the recently reported Pepper fluorescent RNAs (Extended Data Fig. 1a) have emerged as a powerful tool for live-cell RNA imaging, as they are monomeric, bright, stable and exhibit multiple spectral properties in vitro and in live cells, thus allowing robust and no-background imaging of diverse RNA in live cells[33]. Compared to other fluorescent RNAs, Pepper tags have one order of magnitude enhanced cellular brightness and fluorescence turn-on ratio, one or two orders of magnitude enhanced fluorophore affinity, an increase in melting temperature (T_m) of ~20 °C and expanded pH tolerance[33]. Sequence analysis of Pepper aptamer shows that a 43-nucleotide, non-G-tetrad RNA module is essential for HBC ((4-((2-hydroxyrthyl)(methyl) amino)-benzylidene)-cyanophenyl-acetonitrile) (**1**) binding and activation, which is composed of two short duplexes flanked by two invariant intermediate junctions and one terminal loop. To elucidate the mechanism underlying the fluorescence activation of this fluorogenic RNA aptamer, we determined the tertiary structures of Pepper aptamer in complex with HBC and its related analogs using X-ray crystallography, which was further supplemented with structure-guided in vitro and live-cell mutational analysis. Our research thus establishes thorough structural principles underlying the photophysics of Pepper fluorescent RNA and provides insights into the rational design of HBC complexes with improved properties.

Results

Crystallization of the Pepper-HBC complex. The chemical structure of the fluorophore molecule HBC is shown in Fig. 1a, together with the SELEX-selected initial RNA sequence, which can specifically recognize HBC, as shown in Extended Data Fig. 1a (ref. [33]). To facilitate the crystallization of the complex of Pepper bound with HBC, we changed the length of the stem P1 and P3 and replaced the terminus variable loops of stem P3 and P1 with other optional sequences. After screening a large number of different Pepper aptamer constructs, we obtained high diffraction-quality crystals of one 49-nucleotide Pepper construct in complex with HBC, in which the UUCG tetraloop was introduced into the terminus of stem P3 as a closed loop, L3, and the terminus of stem P1 was mutated to a blunt end (Fig. 1b). We solved the Pepper-HBC complex structure with the single-wavelength anomalous diffraction (SAD) method by using the anomalous signal of three $Ir(NH_3)_6^{3+}$ heavy atoms from $Ir(NH_3)_6^{3+}$-soaked crystals (Extended Data Fig. 2a). The structure of the Pepper-HBC complex was refined at a high resolution of 1.64 Å with R_{work} / R_{free} values of 0.20 / 0.25 (Supplementary Table 1).

[1]Life Sciences Institute, Zhejiang University, Hangzhou, China. [2]Optogenetics & Synthetic Biology Interdisciplinary Research Center, State Key Laboratory of Bioreactor Engineering, Shanghai Collaborative Innovation Center for Biomanufacturing Technology, East China University of Science and Technology, Shanghai, China. [3]CAS Center for Excellence in Brain Science and Intelligence Technology, Institute of Neuroscience, Chinese Academy of Sciences, Shanghai, China. [4]These authors contributed equally: Kaiyi Huang, Xianjun Chen. ✉e-mail: linyongzhu@ecust.edu.cn; yiyang@ecust.edu.cn; aimingren@zju.edu.cn

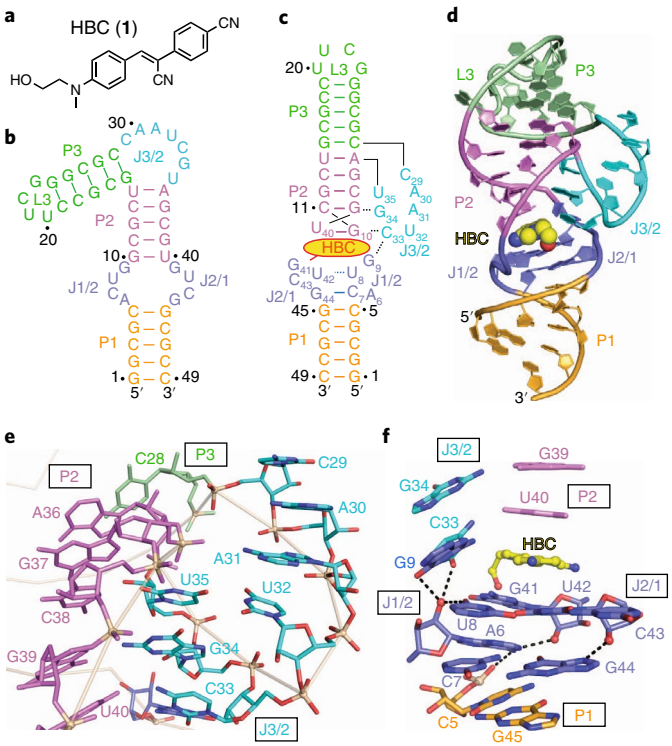


Fig. 1 | Secondary and tertiary structure of Pepper aptamer in complex with HBC. a, Chemical structure of HBC. **b**, Predicted secondary structure of the Pepper aptamer construct used in crystallization. **c**, Schematic of the secondary structure of the Pepper aptamer depicted on the basis of the tertiary structure determined by X-ray crystallography. **d**, Tertiary structure of the Pepper aptamer (shown in cartoon representation) in complex with HBC (shown in sphere representation). **e**, Tertiary conformation of the downward-pointing junction J3/2, in which two stacking alignments are formed by C29-A30-A31-U32 and C33-G34-U35, respectively. **f**, Tertiary interaction of junction J1/2 and J2/1, in which a non-G tetrad, U8-G41-U42-C43, and a base triple, G44-C7-A6, are formed beneath the bound HBC. It is noted that A6, U42 and C43 adopted 2′-*endo* sugar pucker conformations and formed hydrogen bonds with adjacent residues.

Structure of Pepper bound to HBC. The secondary structure schematic and cartoon representation of the tertiary structure of the Pepper-HBC complex are shown in Fig. 1c,d. The tertiary fold of Pepper aptamer contains three stems, P1 (gold), P2 (violet) and P3 (green), connected by three junctions, J1/2 (slate), J2/1 (slate) and J3/2 (cyan), which conform to the proposed secondary structure (Fig. 1b and Extended Data Fig. 1a). Stems P1, P2 and P3, and the zipped junctions J1/2 and J2/1, stack continuously, thereby forming a long stable helix (Extended Data Fig. 1b). The bulge region J3/2 located between stems P2 and P3 extrudes from the helix and is positioned in the major groove of stem P2, where it forms long-distance interactions with stem P2 and the zipped junction J2/1 (Fig. 1c,d). HBC is bound at the intersection of stem P2, zipped junctions J1/2 and J2/1 and the bulge region J3/2 (Fig. 1c,d and Extended Data Fig. 1b).

The downward-projecting bulge region J3/2 is composed of nucleotides from C29 to U35 (Fig. 1c). From top to bottom, two continuously stacking alignments, C29-A30-A31-U32 and U35-G34-C33, are aligned towards the junctional HBC-binding pocket (Fig. 1e). On the top of J3/2, C29 and A30 stack on A31. In the middle of J3/2, A31 forms one hydrogen bond with G37p (the phosphate of G37) from stem P2 (Fig. 1e and Extended Data Fig. 3a,b). U32 forms a noncanonical pairing interaction with U35 between the two stacking alignments C29-A30-A31-32 and U35-G34-C33, wherein U35 forms two additional hydrogen bonds with the base of G37 and C38p from stem P2 (Fig. 1e and Extended Data Fig. 3a,b). Below the noncanonical U35-U32 base pair, G34 from J3/2 forms one base triple with G39·C11 from stem P2 in the tertiary structure of the complex (Fig. 1f and Extended Data Fig. 3c). In the bottom of J3/2, C33 forms hydrogen-bond interactions with G9 from J1/2 and U40·G10, the terminal base pair of stem P2. This results in the formation of a sheared base tetrad G10-U40-C33-G9 at the intersection of J3/2, stem P2 and J1/2, which is positioned to accommodate the bound ligand HBC (Fig. 1f and Extended Data Fig. 3d).

There are eight nucleotides distributed in the junction region between stem P1 and P2, in which the consecutive nucleotides A6-C7-U8-G9 and G41-U42-C43-G44 in two strands are designated as J1/2 and J2/1, respectively (Fig. 1b,c). Except for G9, which is involved in the formation of the base tetrad above the bound HBC, all the other nucleotides participate in the formation of two base tiers below the bound HBC (Extended Data Fig. 3d). U8 from J1/2 and G41, U42 and C43 from J2/1 are positioned in the same plane and form the first non-G-tetrad tier (Fig. 1f and Extended Data Fig. 3e). C7 from J1/2 forms a canonical Watson–Crick base pair with G44 from J2/1 (Fig. 1f and Extended Data Fig. 3f). A6 from J1/2 is positioned in the major groove side of C7·G44 and hydrogen bonds with the base of C7, resulting in the formation of a base triple G44-C7-A6, constituting formation of a second tier above stem P1 (Fig. 1f and Extended Data Fig. 3f). Between the two tiers, it should be noted that both G41 and U42 in the first tier adopt a 2′-*endo* sugar pucker conformation and form hydrogen-bonding interactions with the second-tier G44-C7-A6 base triple (Fig. 1f). At the same time, A6 in the second tier also adopts a 2′-*endo* sugar pucker and interacts with the base of U8 in the first tier (Fig. 1f). In addition, the 2′-OH of A6 interacts with the bases of G9 and C33 in the base tetrad above the bound HBC (Fig. 1f), while the base of A6 forms a hydrogen-bond interaction with the phosphate of C5 from stem P1 (Extended Data Fig. 3f,g). Taken together, these stacking interactions and the extensive tertiary contacts stabilize the conformation of the junction region, including J1/2, J2/1 and J3/2, therefore assisting in defining the tertiary structure of the Pepper aptamer.

Metal ions mediate tertiary structure formation. It is notable that several hydrated Mg^{2+} ions are identified participating in the tertiary folding of the HBC-bound Pepper aptamer, which have been confirmed by the observed anomalous signal collected from divalent cation Mn^{2+}-soaked crystals of the Pepper-HBC complex, which contrasts with the absence of an anomalous signal for monovalent cation, Cs^{+}-soaked crystals of the complex (Extended Data Fig. 2b,c). Except for M5, which is coordinated to the major groove of stem P1, all other metal ions, including M1, M2, M3 and M4, are coordinated to the junction regions J1/2, J2/1 and J3/2 (Fig. 2a).

M1 is located in the upper segment of J3/2, near the intersection of stems P2 and P3, where it forms inner-sphere coordination with the nonbridging phosphate oxygen of A30 and forms outer-sphere coordination with the nonbridging phosphate oxygen of C29, A31 and G37, as well as the bases of A30, A31 and U32 (Fig. 2b). M2 is coordinated to the lower segment of J3/2, where it forms inner-sphere coordination with the nonbridging phosphate oxygen of U32 and forms outer-sphere coordination with the nonbridging phosphate oxygen of A31, C33 and G34 (Fig. 2b).

M3 is located adjacent to the HBC-binding site, where it is coordinated directly to the residues in junction J1/2, including N7 of G9, the phosphate between A6 and C7 and the phosphate between C7 and U8 (Fig. 2c). M4 is located between junction J2/1 and stem P1 and forms inner-sphere coordination with the nonbridging phosphate oxygen of U42, while forming outer-sphere coordination with the 2′-OH of G41 and the bases of G4, C5 and G45 (Fig. 2d). The

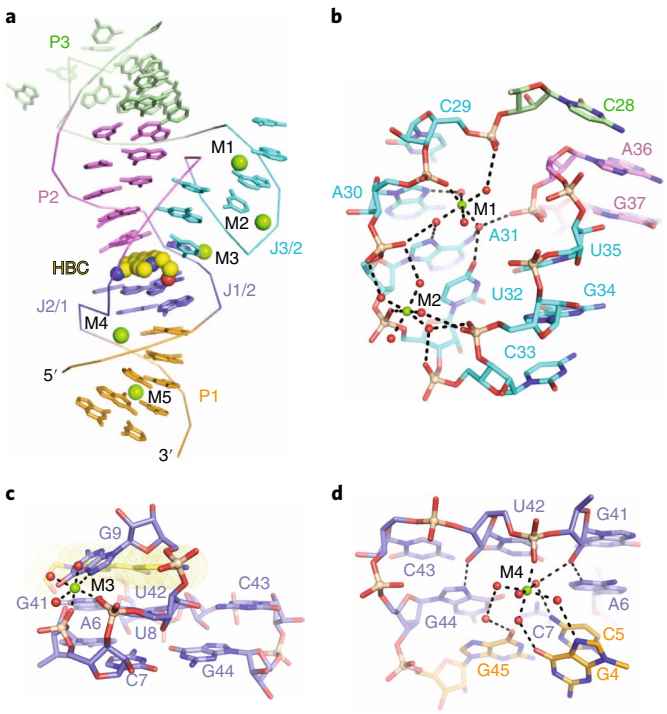


Fig. 2 | Metals involved in interaction in the tertiary structure of Pepper aptamer in complex with HBC. a, Five divalent metal cations, M1–M5, are identified in the tertiary structure of Pepper aptamer in complex with HBC, four of which (M1–M4) are located in the junction region and one of which (M5) is bound to the stem P1. **b**, M1 and M2 are located in the bulge junction region J3/2. M1 forms direct coordination with A30p (the phosphate of A30), while it forms indirect coordination with phosphates of C29p, A31p and G37p and the bases of A30, A31 and U32. M2 forms direct and indirect coordination with A31p, U32p, C33p and G34p. **c**, M3 forms inner-sphere coordination with U8p, C7p and the base of G9. **d**, M4 is located in the intersection of stem P1 and junction region J1/2 and J2/1, which forms direct coordination with U42p and indirect coordination with the sugar of G41 and the bases of G4 and C5.

composite omit electron density map (contoured at level 1.0σ) of M1 to M4 and the coordinated residues are shown in Extended Data Fig. 4a. We identified one additional metal, M5, that is located in the major groove of stem P1, which forms outer-sphere coordination with the main chain of G45, C46 and G47 and the bases of G47 and G2 (Fig. 2a and Extended Data Fig. 4b).

To determine the importance of this Mg^{2+}-involved coordination in the tertiary structure folding of Pepper aptamer and the fluorescence activation of HBC, we performed the fluorescence activation assay of Pepper aptamer in the presence of different concentrations of Mg^{2+}, as well as other divalent cations, such as Mn^{2+}, Ca^{2+} and Ba^{2+} (Extended Data Fig. 5a–d and Supplementary Table 3). It was found that in the absence of Mg^{2+}, the Pepper has minimal activated fluorescence. Increasing the concentration of Mg^{2+} leads to a significant increase of the fitted binding affinity of Pepper with HBC, and a lower HBC concentration to reach the maximal fluorescence intensity (Extended Data Fig. 5a and Supplementary Table 3). Pepper can also activate the fluorescence of HBC in the presence of Mn^{2+}, but showed lower dependence on Mn^{2+} and lower maximal fluorescence intensity compared to Mg^{2+} (Extended Data Fig. 5b). Other divalent cations, such as Ca^{2+} and Ba^{2+}, do not support the binding of Pepper with HBC and result in minimal fluorescence (Extended Data Fig. 5c,d and Supplementary Table 3). Beside these divalent cations, we also tested the dependence of Pepper on the monovalent cation K^+ and found that the activated fluorescence is not affected by K^+ concentrations from 0 mM to 50 mM (Extended Data Fig. 5e and Supplementary Table 3).

Architecture of the HBC-binding pocket. As shown in Fig. 1c,d, the fluorophore molecule HBC is bound in the intersection region of stem P2, internal junctions J1/2 and J2/1 and the bulge junction J3/2. To investigate the composition of the HBC-binding site, we show the schematic secondary structure of Pepper RNA in Fig. 3a, with emphasis on the junction region architecture. A surface representation reveals that a planar and compact binding pocket is formed in the center of the overall structure (Fig. 3b), in which two alignments of consecutive residues U8-G9-G10 and U40-G41-U42-C43, supplemented with one terminal residue C33 in J3/2, form a four-sided box that accommodates the bound HBC molecule (Fig. 3c and Extended Data Fig. 6a,b). HBC intercalates in the ligand-binding pocket and becomes stacked between the top and bottom faces of the box (Fig. 3b). The noncanonical base pair formed by G10 and U40 at the terminus of stem P2 constitutes the top face of the box, while the non-G-tetrad G41-U42-C43-U8 comprises the bottom face of the box (Figs. 1f and 3c–e). The face at the left-hand side of the box consists of G9 and C33, which is fastened to the top face by the hydrogen-bonding interaction between C33 and U40 and the linked phosphate between G10 and G9, and to the bottom face by the linked phosphate between U8 and G9 (Figs. 1f and 3c). The face at the right-hand side of the box consists of the linked phosphate between U40 and G41, which stabilizes both the top and bottom faces of the box and anchors HBC firmly in the box (Fig. 3c). Within the binding box, it is notable that one hydrogen bond is formed between the terminal group –OH of HBC and N7 of G41 (Fig. 3d,e). Outside the binding box, three faces of the box are wrapped additionally on the top by the base triple G39-C11-G34 (Extended Data Fig. 6c,d) and on the bottom by the base layer formed by A6, C7 and G44 (Fig. 1f and Extended Data Fig. 6c,d), followed by the continuous stacking interaction in opposite directions from stem P1 and P2 (Fig. 3a). The composite omit electron density maps (contoured at 1.0σ level) of the residues involved in the binding interaction and the bound HBC are shown in Extended Data Fig. 6e,f.

Fluorescence assay of Pepper aptamer mutants. To evaluate the impact of the residues involved in HBC binding, we performed nucleotide substitutions of the binding pocket and measured their ability to enhance the fluorescence of HBC in vitro and in live mammalian cells (Fig. 4 and Extended Data Figs. 7–9). G10 and U40 form a wobble base pair above the bound HBC in the binding site (Fig. 3c–e). The individual U40 to C (M01) and G10 to A (M02) mutations, which resulted in replacement of the wobble base pair with canonical Watson–Crick base pair variants, partly decreased the activated fluorescence intensity to 40% and 60%, respectively, in vitro (Fig. 4a,b). Accordingly, the population of the cells expressing M01 and M02 Pepper mutants with green fluorescence became noticeably smaller, and the fluorescence signals of these cells with green fluorescence were also markedly decreased when compared to the cells expressing wild-type (WT) Pepper (Fig. 4c). Additionally, the residue switched mutant U40G·G10U (M03) was also prepared and expressed in mammalian cells; however, the fluorescence was no longer detectable (Fig. 4). G9 from J1/2 forms a long-distance interaction with C33 from J3/2 in the tertiary structure lining the HBC-binding pocket (Figs. 1c,f and 3c). Substitutions of G9 to A (M04) or C (M05) completely abolished the activated fluorescence, while substitutions of C33 to U (M06) resulted in a reduction in fluorescence by more than a half, and of C33 to A (M07) decreased the fluorescence by more than tenfold (Fig. 4b). These results were further confirmed by the expression of these mutants (M04, M05, M06 and M07) in mammalian cells (Fig. 4c), which was in agreement with our observations in the tertiary structure (Fig. 1f). Mutations

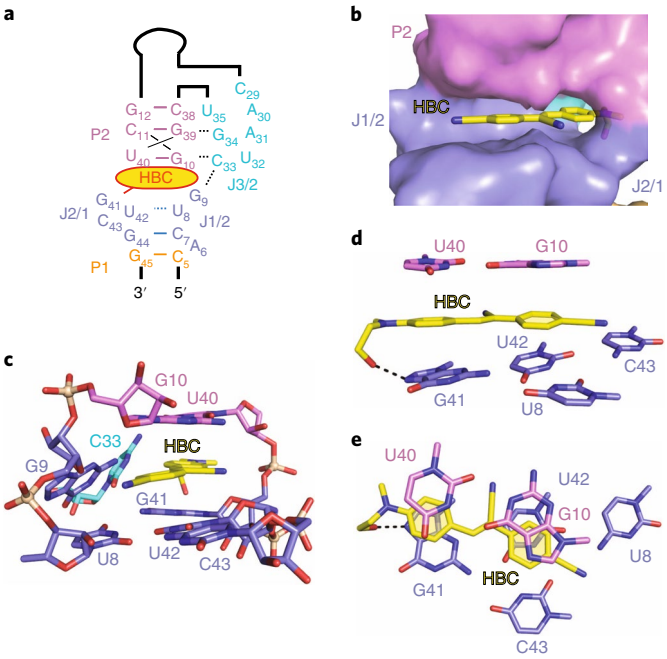


Fig. 3 | Architecture of the HBC-binding pocket. a, Close-up view of the schematic secondary structure of Pepper aptamer with emphasis on the junction region J1/2, J2/1 and J3/2. **b**, HBC (shown in stick representation) intercalates into the ligand-binding pocket (shown in surface representation), resulting in stacking and burial of the chromophore. **c**–**e**, HBC is encapsulated in a four-sided box arrangement. The top and bottom faces are constituted by one noncanonical base pair, G10·U40, and one non-G-tetrad, G41-U42-C43-U8, respectively. G9 and C33 and the linked phosphate between U40 and G41 comprise the faces on the left-hand and right-hand sides (**c**). HBC is stacked between the top face G10·U40 and the bottom face G41-U42-C43-U8 inside the box, where the terminal –OH group of HBC forms one hydrogen bond with N7 of G41 (**d**,**e**).

(M08–M13) based on the formation of the U8-G41-U42-C43 tetrad positioned beneath the ligand reduced the fluorescence by varying levels (Fig. 4). Except for the mutant U42A (M09), which generated less than half the fluorescence, the remaining mutants, such as U8A (M08), U42C·U8G (M10), U42C·U8C (M11), G41A (M12) and C43U (M13), generated either low fluorescence (around 10%) or no fluorescence (Fig. 4b). Mutations of G44, C7 and A6 (M14–M17) also showed diminished fluorescence compared to WT Pepper aptamer (Fig. 4b). Live-cell expression of these mutants (M08–M17) in mammalian cells generated coincident results with the in vitro fluorescence activation experiments (Fig. 4c). Together, all these mutation results highlight the importance of the related nucleotides lining the binding pocket.

We noticed that A6, G41 and U42 all adopt 2′-*endo* sugar pucker conformations in the tertiary structure and participate in the formation of the platform in the HBC-binding pocket (Fig. 1f and Extended Data Fig. 3f–g). Thus, we also investigated the significance of the 2′-OH of A6, G41 and U42, by replacing those residues with the cognate deoxynucleotides. As shown in Extended Data Fig. 7, A6dA, G41dG and U42dU exhibited markedly low HBC-binding affinity and weak capacity to activate HBC fluorescence, even in the presence of high concentration of ligand (5 μM HBC) (Extended Data Fig. 7). These results agree with the observed interaction in the crystal structure and highlight the functional importance of the 2′-*endo* sugar pucker and 2′-OH groups in RNA molecules.

To further explore the composition tolerance of the Pepper aptamer, we tried truncation variants of the Pepper aptamer with a reduced number of base pairs in the stem P1 or P3, respectively, and deletion of the protruded junction region J3/2 (M18–M24, Extended Data Figs. 8 and 9a). The activated fluorescence signal was comparable to the wild-type Pepper aptamer until the length of the stem P1 was reduced to one base pair (M18–M20, Extended Data Figs. 8 and 9a,b). The fluorescence intensity was reduced to half on retention of two base pairs in stem P3 (M21, Extended Data Figs. 8 and 9a,b) and became very weak after deletion of the entire stem P3 (M22, Extended Data Figs. 8 and 9a,b). Deletion of the whole bulge junction J3/2 (M23), or deletion of both bulge junction J3/2 and stem P3 (M24), almost abolished the fluorescence (Extended Data Figs. 8 and 9a). In addition, when we expressed these mutants (M21–M24) in mammalian cells, the resulting cell populations and fluorescence intensity showed the same tendency as above (Extended Data Fig. 8 and 9c). Taken together, these results provide a structural and experimental basis to optimize the array of tandem Pepper aptamers in further applications.

Structure of Pepper aptamer bound to HBC analogs. The chemical formula of HBC consists of an acceptor and an electron donor linked by aromatic π conjugation (Fig. 1a). Our previous results showed that the introduction of different electric properties of chemical groupings in the HBC molecule generated a wide range of fluorescence wavelength, from cyan to red[33]. To detect the pattern of these HBC-like fluorophores interacting with Pepper aptamer, we attempted cocrystallization of the Pepper aptamer with HBC analogs, including HBC514 (**2**), HBC508 (**3**), HBC497 (**4**), HBC620 (**5**), HBC525 (**6**) and HBC485 (**7**), and determined the structures of Pepper-HBC514, Pepper-HBC508, Pepper-HBC497, Pepper-HBC620, Pepper-HBC525 and Pepper-HBC485 complexes (Fig. 5 and Extended Data Fig. 10).

As shown in Fig. 5a and Extended Data Fig. 10a, the composition of the binding pocket of HBC514 is similar to HBC, but the introduction of the aliphatic amine group instead of the hydroxy group in the terminus made the ligand lose the hydrogen-bonding interaction with G41. Substitution of the benzene moiety by pyridine (HBC508) or pyrazine (HBC497) generated the same interactions between the Pepper aptamer and the ligand (Fig. 5b,c and Extended Data Fig. 10b,c). The ligand-binding pocket of the Pepper aptamer can also accommodate HBC620, in which the bithiophene with longer length was introduced to improve π conjugation. However, due to the long length of the bithiophene, HBC620 did not form any hydrogen-bonding interaction with the Pepper aptamer (Fig. 5d and Extended Data Fig. 10d). HBC525, in which the electron acceptor was replaced with 2-benzoxazole, retained the ligand interactions between HBC and Pepper aptamer, including the hydrogen-bonding interaction (Fig. 5e and Extended Data Fig. 10e). Another ligand, named HBC485, was substituted with both the aliphatic amine group (introduced in HBC514, Fig. 5a and Extended Data Fig. 10a) and pyrazine (introduced in HBC497, Fig. 5c and Extended Data Fig. 10c), which resulted in a similar interaction with Pepper as for HBC514 (Fig. 5a,f and Extended Data Fig. 10a,f).

Discussion

Over the past two decades, the family of GFP-like FPs has been considerably expanded, and has become one of the most widely used tools in molecular biology, cell biology, biotechnology and medicine, not only for tracing of proteins and cells, but also for monitoring of metabolites, ions and cellular signals in live cells and in vivo[1,34]. Numerous crystal structures of these FPs are available, which share a similar β-can structure, comprising 11 strands of β-sheet with an α-helix inside and short helical segments on the ends of the cylinder. The structures of FPs have greatly facilitated efforts towards improving photophysical properties and spectral spanning, as well as the design of genetically encoded biosensors[1,35,36].

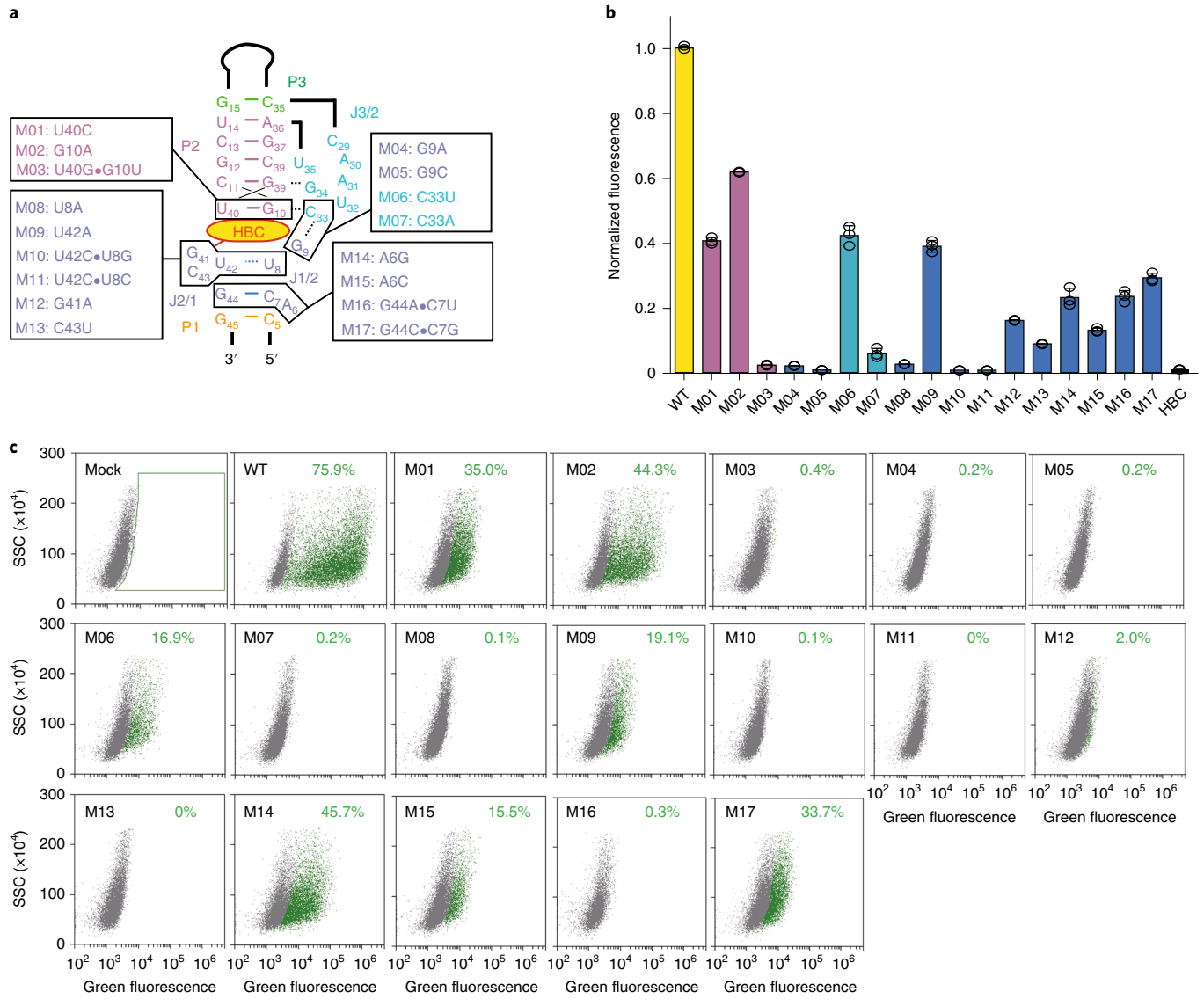


Fig. 4 | In vitro fluorescence assay of the structure-based mutants of Pepper aptamer. a, Structure-based mutants of Pepper aptamer based on the tertiary structure and the binding-pocket composition are shown, along with the schematic secondary structure. **b**, The activated fluorescence of HBC by Pepper mutants from three independently repeated experiments are normalized for comparison with the WT Pepper aptamer. Data represent the mean ± s.d. from three replicates. **c**, Fluorescence-activated cell sorting (FACS) analysis of Pepper fluorescence in mammalian cells. HEK293T cells transfected with plasmid expressing Pepper or Pepper variants were incubated with 1 μM HBC and analyzed using a flow cytometer. HEK293T cells transfected with mock plasmid and labeled with 1 μM HBC were used as controls. The gate (green) was placed on the basis of the control cells to determine the population fraction that exhibited higher fluorescence than the background.

In contrast to this rich set of tools for studies using FPs, the development of FRs, the equivalent of FPs for RNA studies, has turned out to be more difficult than their protein counterparts, as there are no native FR scaffolds that can serve as a starting point for engineering[37]. In 1999, Grate et al. evolved an RNA aptamer to bind malachite green (MG), aiming to specifically inactivate its attached messenger RNA by illuminating the dye to generate reactive oxygen species[38]. It was not until 2003 that Roger Y. Tsien's group raised the concept towards development of fluorescent RNA by using the MG aptamer to specifically bind the MG dye and significantly activate its fluorescence[10]. So far, several fluorescent RNAs based on Hoechst derivatives, cyanine and GFP-derived fluorophores have been developed and some of their tertiary structures have been solved by X-ray crystallography. Most of these FRs contain a G-quadruplex in their fluorophore-binding site. For example, Spinach contains an unusual antiparallel G-quadruplex together with a base triplet and an unpaired G to form the binding pocket[39,40]; Mango folds into mixed parallel and antiparallel G-quadruplexes by using six loops to connect the 12 guanine residues of its three G-quartets[32,41]; Corn forms a homodimer that encapsulates one DFHO (3,5-difluoro-4-hydroxybenzylidene-imidazolinone-2-oxime) at its interprotomer interface, sandwiching it with a G-quadruplex from each protomer[32]. However, previous studies have shown that eukaryotic cells have machinery that specifically inhibits the formation of G-quartet folded structures[42], suggesting that G-quartet-containing FRs may have poor folding propensity in live cells and are not optimal for RNA imaging. DIR2s, second-generation dimethylindole reds, are the only FRs that lack a G-quadruplex motif and adopt a compact, tuning-fork-like architecture consisting of a helix and two short stem loops oriented in parallel to create the ligand-binding pocket[43]. Nevertheless, DIR2s show a high tendency to dimerize, which may interfere with the behavior and function of the tagged RNA.

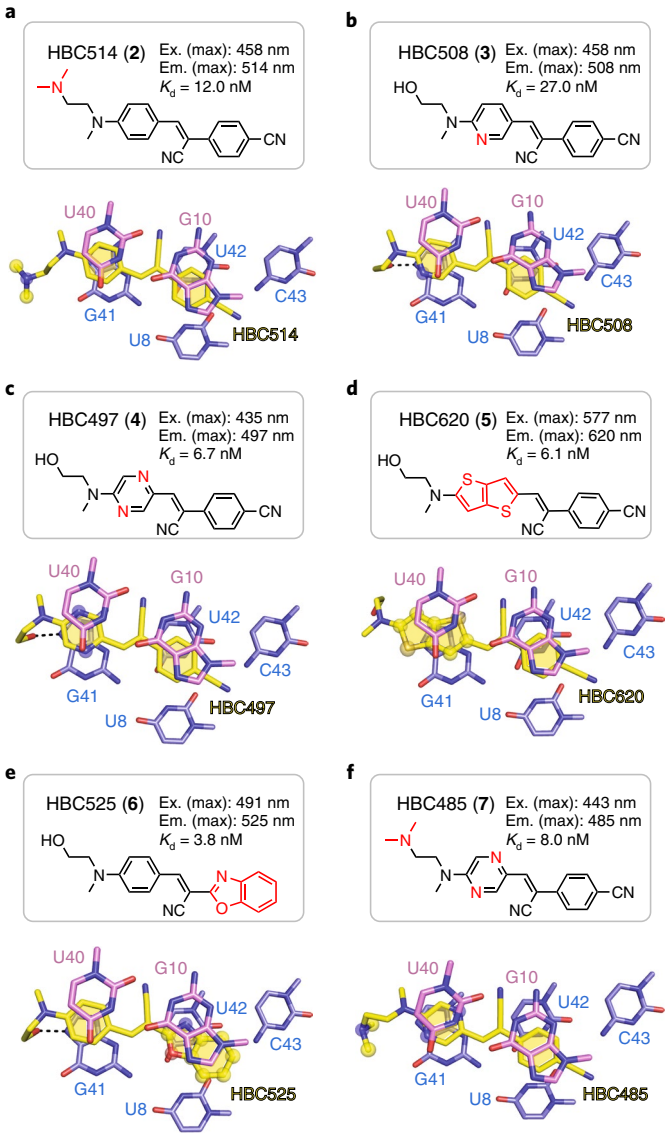


Fig. 5 | The binding-pocket composition of HBC-like fluorophores. a–f, The chemical structures, spectral properties and composition of the binding pockets of HBC514 (**a**), HBC508 (**b**), HBC497 (**c**), HBC620 (**d**), HBC525 (**e**) and HBC485 (**f**). All the modification sites of HBC-like fluorophores are labeled in red in the chemical structures and shown with transparent balls in the binding-pocket composition. Ex. (max) represents the excitation maximum wavelength. Em. (max) represents the emission maximum wavelength. K_d represents the binding dissociation constant between Pepper and each fluorophore molecule.

The Pepper series FRs developed recently fulfill almost all of the favorability criteria for ideal FRs; namely, they should be monomeric, stable, bright and multicolored[33]. In this study, we have determined the structures of Pepper in complex with HBC and its analogs. The Pepper aptamer folds into a monomeric non-G-quadruplex tuning-fork-like architecture composed of a helix and one protruded junction region, which is consistent with previous biochemical studies, which showed that Pepper is independent of potassium ions and probably lacks a G-quartet in its folded topology[33]. Further structure-based analysis and the impact of divalent cations revealed that Pepper is dependent on Mg^{2+} cations. As shown in Fig. 2b, all the residues except U35 in J3/2 are coordinated to M1 (Mg^{2+}) and M2 (Mg^{2+}), which together help to define the tertiary conformation of J3/2. M3 (Mg^{2+}) located in the vicinity of the HBC-binding pocket forms direct coordination interactions with residues lining the binding pocket, and is therefore involved in the stabilization of the binding pocket. M4 (Mg^{2+}) is positioned between J2/1 and stem P1 and clamps them together by coordination. These metal-mediated interactions were also observed in the stabilization of previously reported RNA tertiary structures and involved in their ligand-binding pocket formation[44–46]. The near-planar HBC fluorophore molecule intercalates in the middle of the structure and is sandwiched between one non-G-quadruplex base quadruple and one noncanonical G·U base pair, which are additionally stapled by the terminal coaxial stems as part of the long colinear helix (Fig. 1c,d). The structures of Pepper aptamer in complex with each HBC analogs indicate that the ligand-recognition pocket is capable of tolerating specific modifications on HBC (Fig. 5), which may provide opportunities for further optimization on the fluorophore ligands. Our structure-based fluorescence assays on mutants of the monomeric non-G-quadruplex Pepper aptamer confirm the observations emerging from the tertiary structure (Fig. 4 and Extended Data Figs. 7–9). Nevertheless, no variants showing significantly improved properties were found. It appears that Pepper aptamer is already in its optimal form, which is not surprising as we had already performed extensive truncation and site-directed mutation studies on the D11 aptamer, the precursor of Pepper, in our original report[33]. Alternatively, the photophysical properties of Pepper fluorescent RNAs may be improved by modification of HBC ligands. In the future, computer-aided rational design of HBC analogs, which may have higher affinity or broader spectral properties (for example, infrared emission), may result in Pepper-HBC complexes with improved photophysical properties.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41589-021-00884-6.

Received: 18 April 2021; Accepted: 20 August 2021;
Published online: 01 November 2021

References

1. Shaner, N. C., Steinbach, P. A. & Tsien, R. Y. A guide to choosing fluorescent proteins. *Nat. Methods* **2**, 905–909 (2005).
2. Gall, J. G. & Pardue, M. L. Formation and detection of RNA-DNA hybrid molecules in cytological preparations. *Proc. Natl Acad. Sci. USA* **63**, 378–383 (1969).
3. Bertrand, E. et al. Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* **2**, 437–445 (1998).
4. Daigle, N. & Ellenberg, J. λ_N-GFP: an RNA reporter system for live-cell imaging. *Nat. Methods* **4**, 633–636 (2007).
5. Nelles, D. A. et al. Programmable RNA tracking in live cells with CRISPR/Cas9. *Cell* **165**, 488–496 (2016).
6. Batra, R. et al. Elimination of toxic microsatellite repeat expansion RNA by RNA-targeting Cas9. *Cell* **170**, 899–912.e10 (2017).
7. Yang, L. Z. et al. Dynamic imaging of RNA in living cells by CRISPR-Cas13 systems. *Mol. Cell* **76**, 981–997.e7 (2019).
8. You, M. & Jaffrey, S. R. Structure and mechanism of RNA mimics of green fluorescent protein. *Annu. Rev. Biophys.* **44**, 187–206 (2015).
9. Neubacher, S. & Hennig, S. RNA structure and cellular applications of fluorescent light-up aptamers. *Angew. Chem. Int. Ed. Engl.* **58**, 1266–1279 (2019).
10. Babendure, J. R., Adams, S. R. & Tsien, R. Y. Aptamers switch on fluorescence of triphenylmethane dyes. *J. Am. Chem. Soc.* **125**, 14716–14717 (2003).
11. Sando, S., Narita, A., Hayami, M. & Aoyama, Y. Transcription monitoring using fused RNA with a dye-binding light-up aptamer as a tag: a blue fluorescent RNA. *Chem. Commun.* **33**, 3858–3860 (2008).
12. Murata, A., Sato, S., Kawazoe, Y. & Uesugi, M. Small-molecule fluorescent probes for specific RNA targets. *Chem. Commun.* **47**, 4712–4714 (2011).

13. Arora, A., Sunbul, M. & Jäschke, A. Dual-colour imaging of RNAs using quencher- and fluorophore-binding aptamers. *Nucleic Acids Res.* **43**, e144 (2015).
14. Sunbul, M. & Jäschke, A. SRB-2: a promiscuous rainbow aptamer for live-cell RNA imaging. *Nucleic Acids Res.* **46**, e110 (2018).
15. Braselmann, E. et al. A multicolor riboswitch-based platform for imaging of RNA in live mammalian cells. *Nat. Chem. Biol.* **14**, 964–971 (2018).
16. Filonov, G. S., Moon, J. D., Svensen, N. & Jaffrey, S. R. Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J. Am. Chem. Soc.* **136**, 16299–16308 (2014).
17. Strack, R. L., Disney, M. D. & Jaffrey, S. R. A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA. *Nat. Methods* **10**, 1219–1224 (2013).
18. Paige, J. S., Wu, K. Y. & Jaffrey, S. R. RNA mimics of green fluorescent protein. *Science* **333**, 642–646 (2011).
19. Song, W. et al. Imaging RNA polymerase III transcription using a photostable RNA–fluorophore complex. *Nat. Chem. Biol.* **13**, 1187–1194 (2017).
20. Bouhedda, F. et al. A dimerization-based fluorogenic dye-aptamer module for RNA imaging in live cells. *Nat. Chem. Biol.* **16**, 69–76 (2020).
21. Dolgosheina, E. V. et al. RNA Mango aptamer-fluorophore: a bright, high-affinity complex for RNA labeling and tracking. *ACS Chem. Biol.* **9**, 2412–2420 (2014).
22. Constantin, T. P. et al. Synthesis of new fluorogenic cyanine dyes and incorporation into RNA fluoromodules. *Org. Lett.* **10**, 1561–1564 (2008).
23. Tan, X. et al. Fluoromodules consisting of a promiscuous RNA aptamer and red or blue fluorogenic cyanine dyes: selection, characterization, and bioimaging. *J. Am. Chem. Soc.* **139**, 9001–9009 (2017).
24. Holeman, L. A., Robinson, S. L., Szostak, J. W. & Wilson, C. Isolation and characterization of fluorophore-binding RNA aptamers. *Fold. Des.* **3**, 423–431 (1998).
25. Sunbul, M. & Jäschke, A. Contact-mediated quenching for RNA imaging in bacteria with a fluorophore-binding aptamer. *Angew. Chem. Int. Ed. Engl.* **52**, 13401–13404 (2013).
26. Song, W., Strack, R. L., Svensen, N. & Jaffrey, S. R. Plug-and-play fluorophores extend the spectral properties of Spinach. *J. Am. Chem. Soc.* **136**, 1198–1201 (2014).
27. Dey, S. K. & Jaffrey, S. R. RIBOTACs: small molecules target RNA for degradation. *Cell Chem. Biol.* **26**, 1047–1049 (2019).
28. Li, X. et al. Imaging intracellular *S*-adenosyl methionine dynamics in live mammalian cells with a genetically encoded red fluorescent RNA-based sensor. *J. Am. Chem. Soc.* **142**, 14117–14124 (2020).
29. Steinmetzger, C., Palanisamy, N., Gore, K. R. & Höbartner, C. A multicolor large Stokes shift fuorogen-activating RNA aptamer with cationic chromophores. *Chemistry* **25**, 1931–1935 (2019).
30. Wang, P. et al. Photochemical properties of Spinach and its use in selective imaging. *Chem. Sci.* **4**, 2865–2873 (2013).
31. Han, K. Y., Leslie, B. J., Fei, J., Zhang, J. & Ha, T. Understanding the photophysics of the spinach-DFHBI RNA aptamer-fluorogen complex to improve live-cell RNA imaging. *J. Am. Chem. Soc.* **135**, 19033–19038 (2013).
32. Warner, K. D. et al. A homodimer interface without base pairs in an RNA mimic of red fluorescent protein. *Nat. Chem. Biol.* **13**, 1195–1201 (2017).
33. Chen, X. et al. Visualizing RNA dynamics in live cells with bright and stable fluorescent RNAs. *Nat. Biotechnol.* **37**, 1287–1293 (2019).
34. Shcherbakova, D. M., Baloban, M. & Verkhusha, V. V. Near-infrared fluorescent proteins engineered from bacterial phytochromes. *Curr. Opin. Chem. Biol.* **27**, 52–63 (2015).
35. Zhang, Z., Cheng, X., Zhao, Y. & Yang, Y. Lighting up live-cell and in vivo central carbon metabolism with genetically encoded fluorescent sensors. *Annu. Rev. Anal. Chem.* **13**, 293–314 (2020).
36. Specht, E. A., Braselmann, E. & Palmer, A. E. A critical and comparative review of fluorescent tools for live-cell imaging. *Annu. Rev. Physiol.* **79**, 93–117 (2017).
37. Bouhedda, F., Autour, A. & Ryckelynck, M. Light-up RNA aptamers and their cognate fluorogens: from their development to their applications. *Int. J. Mol. Sci.* **19**, 44 (2017).
38. Grate, D. & Wilson, C. Laser-mediated site-specific inactivation of RNA transcripts. *Proc. Natl Acad. Sci. USA* **96**, 6131–6136 (1999).
39. Huang, H. et al. A G-quadruplex-containing RNA activates fluorescence in a GFP-like fluorophore. *Nat. Chem. Biol.* **10**, 686–691 (2014).
40. Warner, K. D. et al. Structural basis for activity of highly efficient RNA mimics of green fluorescent protein. *Nat. Struct. Mol. Biol.* **21**, 658–663 (2014).
41. Trachman, R. J. 3rd et al. Structural basis for high-affinity fluorophore binding and activation by RNA Mango. *Nat. Chem. Biol.* **13**, 807–813 (2017).
42. Guo, J. U. & Bartel, D. P. RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science* **353**, aaf5371 (2016).
43. Shelke, S. A. et al. Structural basis for activation of fluorogenic dyes by an RNA aptamer lacking a G-quadruplex motif. *Nat. Commun.* **9**, 4542 (2018).
44. Chen, H. et al. Structural distinctions between NAD^+ riboswitch domains 1 and 2 determine differential folding and ligand binding. *Nucleic Acids Res.* **48**, 12394–12406 (2020).
45. Trausch, J. J., Marcano-Velázquez, J. G., Matyjasik, M. M. & Batey, R. T. Metal ion-mediated nucleobase recognition by the ZTP riboswitch. *Chem. Biol.* **22**, 829–837 (2015).
46. Ren, A., Rajashankar, KanagalaghattaR. & Patel, DinshawJ. Global RNA fold and molecular recognition for a *pfl* riboswitch bound to ZMP, a master regulator of one-carbon metabolism. *Structure* **23**, 1375–1381 (2015).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2021

Methods

RNA preparation. To facilitate the crystallization of the Pepper-HBC complex, we tried the Pepper constructs with different length of stems P1 and P3, and introduced stable GNRA and UNCG tetraloops, or the U1A protein-binding loop, to close stem P1 or P3. The target RNA sequence followed by HDV ribozyme was inserted into pUT7 plasmid, which was amplified in *Escherichia coli* and linearized by *Hin*d III restriction endonuclease to generate the DNA template for transcription. In vitro transcription was performed at 37 °C using T7 RNA polymerase for 4 h, which was followed by purification with urea-denatured polyacrylamide gel electrophoresis (urea–PAGE). The target RNA product was visualized with a UV lamp and excised from gel, followed by electroelution using the Elutrap electroelution system (GE Healthcare) at 4 °C in 0.5× Tris–acetate–EDTA (TAE) buffer. The eluted RNA sample was further precipitated by the isopropanol method, washed with 70% ethanol and finally dissolved in diethyl pyrocarbonate (DEPC)-treated double-distilled water after lyophilization.

Crystallization. The fluorophore ligands HBC and its analogs were dissolved at a high concentration of 5 mM in DMSO as stock solutions. Then, 0.05 mM RNA was mixed with 0.25 mM HBC or HBC analog in a final solution containing 40 mM HEPES, pH 7.4, 125 mM KCl, 5 mM $MgCl_2$, 5% DMSO. The mixture of RNA and ligand was annealed at 65 °C for 5 min and cooled on ice for 1 hour to form the RNA-fluorophore complex. Then the RNA-fluorophore complex sample was concentrated tenfold to a final concentration of 0.5 mM RNA and 2.5 mM ligand for crystallization. The crystallization was performed at 16 °C by mixing the RNA-ligand complex with the reservoir well solution at a ratio of 1:1 using the sitting-drop diffusion method. The well-diffracted crystals of Pepper-HBC and Pepper-HBC analogs grew from the similar conditions, containing 0.1 M NH_4Cl, 0.005 M $MgCl_2$, 0.025 M HEPES pH 7.0, 1.25 M 1,6-hexanediol, 35–40% MPD after one week. The crystals were flash-frozen in liquid nitrogen before diffraction data collection. For anomalous diffraction data collection, crystals of the Pepper-HBC complex were transferred into the crystallization solution containing 0.1 M NH_4Cl, 0.005 M $MgCl_2$, 0.025 M HEPES pH 7.0, 1.25 M 1,6-hexanediol, 40% MPD supplemented with 5 mM $Ir(NH_3)_6Cl_3$ or 50 mM $MnCl_2$, or 50 mM CsCl at 4 °C for 12 h to prepare the $Ir(NH_3)_6^{3+}$-, Mn^{2+}- and Cs^+-soaked crystals, respectively.

X-ray diffraction data collection and structure determination. X-ray diffraction data for crystals of the complexes were collected at beamline BL-19U1 at Shanghai Synchrotron Radiation Facility (SSRF). The collected data were processed using HKL3000 (HKL Research)[47]. The phase problem of the Pepper-HBC complex was determined with the SAD technique using the diffracted anomalous signal collected from $Ir(NH_3)_6^{3+}$-soaked crystals in the Phenix suite[48]. The structure model was further built and refined by Coot[49] and Phenix programs[48]. The structures of Pepper-HBC analog complexes were solved by the molecular replacement method with the Pepper-HBC structure as the initial model, and further refined by Coot[49] and Phenix programs[48]. The statistics of crystal diffraction data and structure refinement are listed in Supplementary Tables 1 and 2.

In vitro characterization of Pepper and mutants. To confirm the observation in the tertiary structure of the Pepper-HBC complex, we made structure-based mutations of the Pepper aptamer. The nucleotide replacement mutants of the Pepper aptamer were prepared by in vitro transcription, while the deoxynucleotide replacement mutants (A6dA, G41dG and U42dU) of the Pepper aptamer were synthesized by Shanghai GenePharma.

To evaluate the capability of the Pepper mutants to activate the fluorophore ligand HBC, RNA was prepared in the buffer containing 40 mM HEPES pH 7.4, 125 mM KCl, 5 mM $MgCl_2$ and annealed at 65 °C for 5 min, followed by incubation on ice for 30 min. Before the fluorescence measurement experiment, 5 μM of the fluorophore ligand HBC was mixed with 1 μM annealed RNA and incubated at room temperature for another 30 min.

To explore the impact of the types and concentration of metal ions on HBC fluorescence activation by the Pepper aptamer, the trace amount of Mg^{2+} was removed from the RNA sample by dialysis against a solution containing 0.5 mM EDTA and 6 M urea overnight. Then, the treated RNA was further precipitated by isopropanol, washed with 70% ethanol and finally dissolved in DEPC-H_2O after lyophilization. To determine the influence of divalent metal ions, RNA was annealed at 65 °C for 5 min in 40 mM HEPES pH 7.4, 125 mM KCl and incubated on ice for 30 min. Then, titration of 50 nM refolded RNA in the reaction buffer containing 40 mM HEPES pH 7.4, 125 mM KCl supplemented with 0 mM to 10 mM individual divalent metal ions, such as Mg^{2+}, Ca^{2+}, Mn^{2+} or Ba^{2+}, with increasing concentrations of HBC ranging from 0 μM to 2 μM was performed at room temperature. To test the impact of the monovalent anion K^+ on the fluorophore activation of HBC, RNA was annealed at 65 °C for 5 min in a solution containing 40 mM HEPES pH 7.4, 5 mM $MgCl_2$ and followed by incubation on ice for 30 min. Then, titration of 50 nM RNA with HBC (concentrations ranging from 0 μM to 2 μM) was carried out at room temperature in the reaction buffer containing 40 mM HEPES pH 7.4, 5 mM $MgCl_2$ supplemented with 0 mM to 50 mM KCl.

The resulting data points were fitted to a curve on the basis of the Hill equation: $y = B_{max}x_h/(K_{dh} + x_h)$, where B_{max} is the maximum specific binding parameter in the same units as Y, h is the Hill slope, K_d is the concentration needed to achieve a half-maximum binding at equilibrium. The fluorescence signal was measured using a Synergy Neo2 Multi-Mode Microplate Reader (BioTek) with an excitation wavelength of 485 nm (±20 nm) and an emission wavelength of 530 nm (±20 nm) for ligand HBC. All fluorescence experiments were repeated three times independently.

DNA cloning, cell culture and transfection. For mammalian expression of the Pepper aptamer and its mutants, template DNA encoding ribozyme-flanked[50] Pepper was synthesized by Shanghai Generay Biotech. The template DNA was amplified and inserted into the original single guide RNA (sgRNA) expression plasmid using the Hieff Clone One Step Cloning Kit (Yeasen), which was linearized by PCR amplification to remove the native sgRNA scaffold[51]. Reverse PCR was performed using the primers containing different mutations in the Pepper sequence. The linearized fragments were self-ligated using the Hieff Clone One Step Cloning Kit (Yeasen). The plasmids were verified via sequencing by Shanghai Jie Li Biotech. HEK293T cells were cultured in DMEM (high glucose) supplemented with 10% FBS. The cell line was authenticated and mycoplasma-negative. The cell line was cultured at 37 °C in a humidified atmosphere of 95% air and 5% CO_2 and split every 2 days or at confluence. Unless indicated, transient transfection was performed using the Hieff Trans liposomal transfection reagent (Yeasen, catalog no. 40802ES02) according to the manufacturer's protocol. In detail, cells were seeded at 70–90% confluence before transfection. For a 96-well plate, 0.1 μg of DNA and 0.25 μl of Hieff Trans liposomal transfection reagent were diluted in 25 μl of Opti-MEM (Gibco). The DNA–lipid complex was added to the cells after a 20-min incubation at room temperature.

Flow cytometry. HEK293T cells were transfected with the plasmid expressing Pepper, or Pepper variant, according to the procedure described above. After 36 hours, the cells were digested and resuspended in the 4% FBS/1× PBS solution containing 1 μM HBC and 20 mM $MgSO_4$. The fluorescence was analyzed using a CytoFLEX-S flow cytometer (Beckman Coulter) with an excitation of 488/8 nm and an emission of 525/40 nm. The gate was placed on the basis of the control cells labeled with 1 μM HBC to determine the population fraction that exhibited higher fluorescence than the background. Processing and analysis of data were performed using the Cytexpert program (Beckman Coulter).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Atomic coordinates and structure factors for Pepper aptamer in complex with HBC and the analogs have been deposited at the Protein Data bank (www.rcsb.org) under accession numbers 7EOH, 7EOK, 7EOL, 7EOM, 7EON, 7EOO and 7EOP. The crystal structures of $Ir(NH_3)_6^{3+}$-soaked, Mn^{2+}-soaked and Cs^+-soaked crystals have been deposited at the Protein Data bank under accession numbers 7EOG, 7EOI and 7EOJ, respectively. Source data are provided with this paper.

References

47. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the integration of data reduction and structure solution–from diffraction images to an initial model in minutes. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 859–866 (2006).
48. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
49. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
50. Litke, J. L. & Jaffrey, S. R. Highly efficient expression of circular RNA aptamers in cells using autocatalytic transcripts. *Nat. Biotechnol.* **37**, 667–675 (2019).
51. Shao, S. et al. Long-term dual-color tracking of genomic loci by modified sgRNAs of the CRISPR/Cas9 system. *Nucleic Acids Res.* **44**, e86 (2016).

Acknowledgements

We thank the staff members of the Large-scale Protein Preparation System, BL-17B, BL-18U1 and BL-19U1 beamlines at the National Facility for Protein Science in Shanghai (NFPS), Zhangjiang Laboratory, China for providing technical support and assistance in data collection and analysis. We thank the staff of the BL-17U1 beamline at SSRF for their assistance in X-ray data collection. We thank the core facility of the Life Sciences Institute (LSI), Zhejiang University, for technical assistance. This work was supported by the National Natural Science Foundation of China (32022039, 31870810, 91940302 and 91640104 to A.R., 91857202 and 21937004 to Y.Y.), the National Key Research and Development Program of China (2017YFA050400 and 2019YFA0904800 to Y.Y., 2019YFA0110500 to L.Z.), the outstanding youth fund of

Zhejiang Province (LR19C050003 to A.R.), the Shanghai Science and Technology Commission (18JC1411900 to Y.Y.) and the new faculty start-up funds from Zhejiang University (to A.R.), Shanghai Municipal Education Commission-Frontier Research Base of Optogenetic Techniques for Cell Metabolism (2021 Sci & Tech 03-28 to Y.Y. and X.C.) and the Fundamental Research Funds for the Central Universities (to Y.Y. and X.C.).

Author contributions

K.H. undertook all of the crystallization screening and optimization, the diffraction data collection, the structure determination and the in vitro characterization of Pepper aptamer with the assistance of C.L. and Q.S. under the supervision of A.R. X.C. performed the live-cell characterization of Pepper aptamer with the assistance of H.L. under the supervision of Y.Y. L.Z. synthesized HBC and the analogs. The structures were analyzed by A.R. and K.H. The paper was written jointly by A.R. and Y.Y. with input from the remaining authors.

Competing interests

Y.Y., L.Z. and X.C. are named inventors of patent applications nos. 201910352348X, 201910348701.7, PCT/CN2020/087311, PCT/CN2020/087415. The remaining authors declare no competing interests.

Additional information

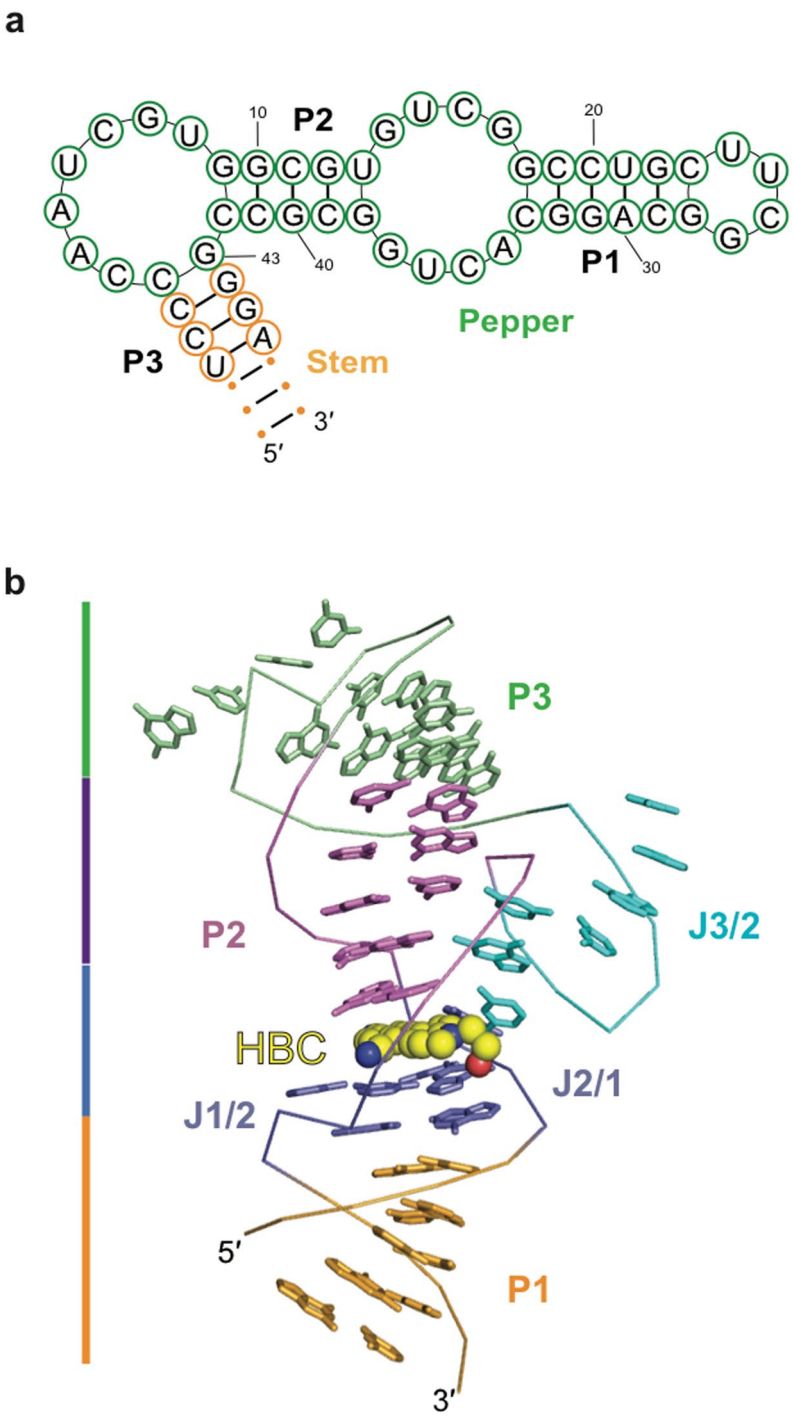
Extended data is available for this paper at https://doi.org/10.1038/s41589-021-00884-6.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41589-021-00884-6.

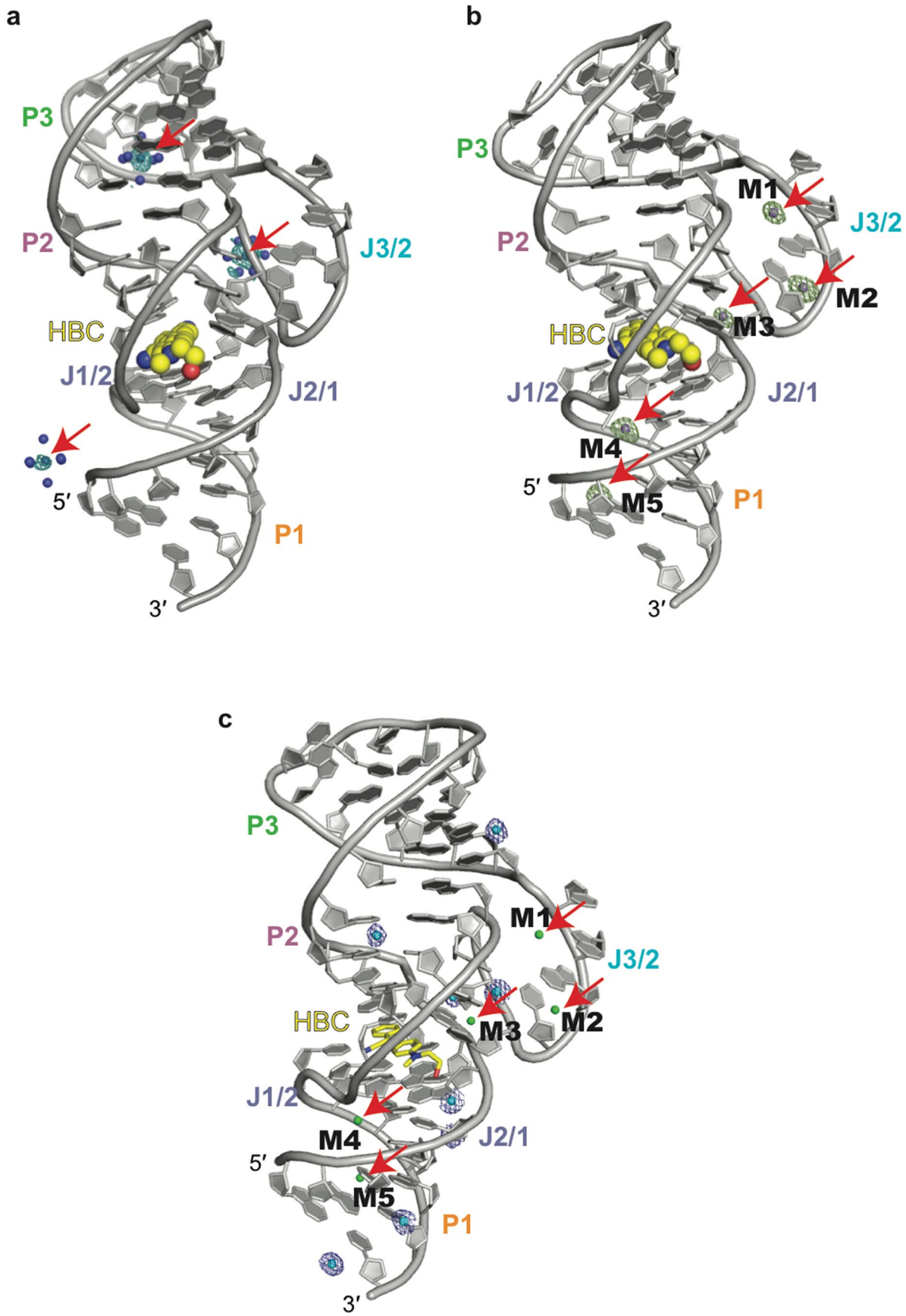
Correspondence and requests for materials should be addressed to Linyong Zhu, Yi Yang or Aiming Ren.

Peer review information *Nature Chemical Biology* thanks Charles Dann and other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

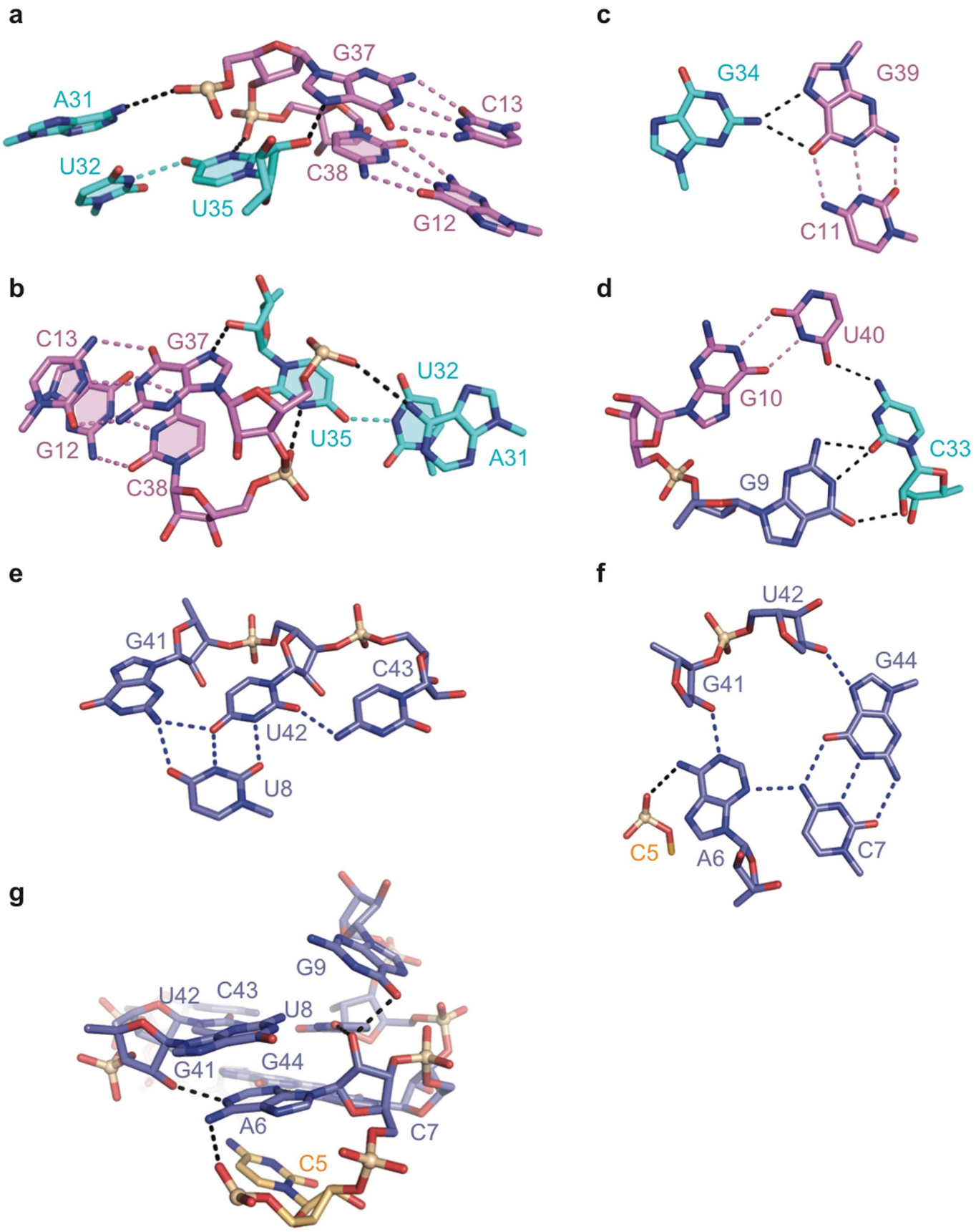


Extended Data Fig. 1 | The sequence and structure of Pepper aptamer. a, The sequence and predicted secondary structure of the original selected Pepper aptamer. **b**, The tertiary structure of Pepper aptamer is shown in ribbon representation with the bases of the residues shown in stick and the bound HBC shown in sphere representation. Stems P1, P2, P3 and the zipped junction J1/2 and J2/1 stacks continuously and form the long helix.



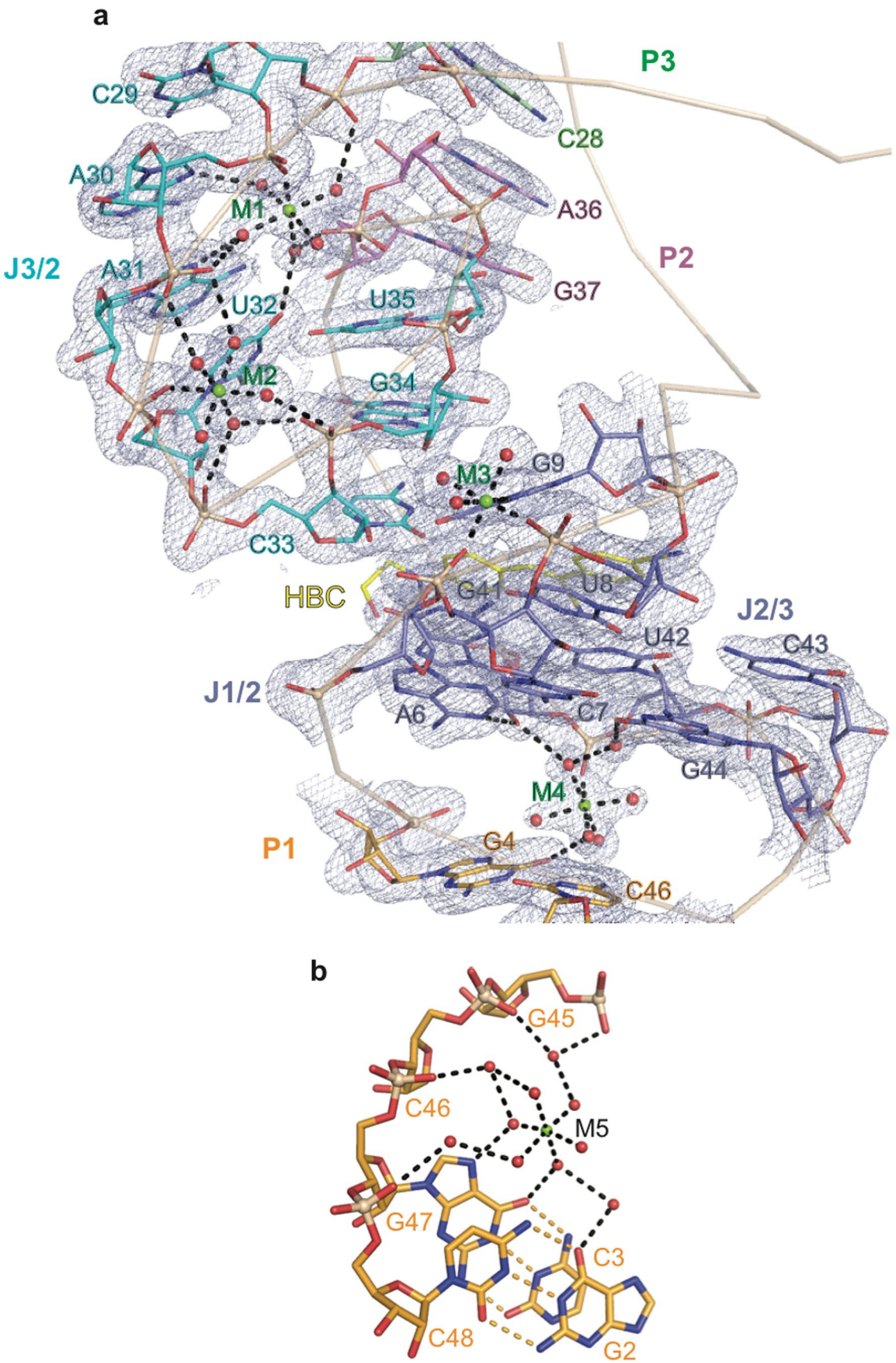
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Anomalous difference electron density map for $Ir(NH_3)_6^{3+}$-soaked, Mn^{2+}-soaked and Cs^{+}-soaked Pepper aptamer crystals. a, Anomalous electron density map contoured at level 3.0 σ for $Ir(NH_3)_6^{3+}$ sites (labelled with red arrow) in each asymmetric unit of Pepper aptamer were used to determine the phase problem of the structure. **b**, Anomalous electron density map contoured at level 3.0 σ for Mn^{2+} sites (shown as purple balls, labelled with red arrow) in each asymmetric unit of Pepper aptamer. **c**, Anomalous electron density map contoured at level 3.0 σ for Cs^{+} sites (shown as cyan balls) in each asymmetric unit of Pepper aptamer. The bound Mg^{2+} sites are shown as green ball and labelled with red arrow.

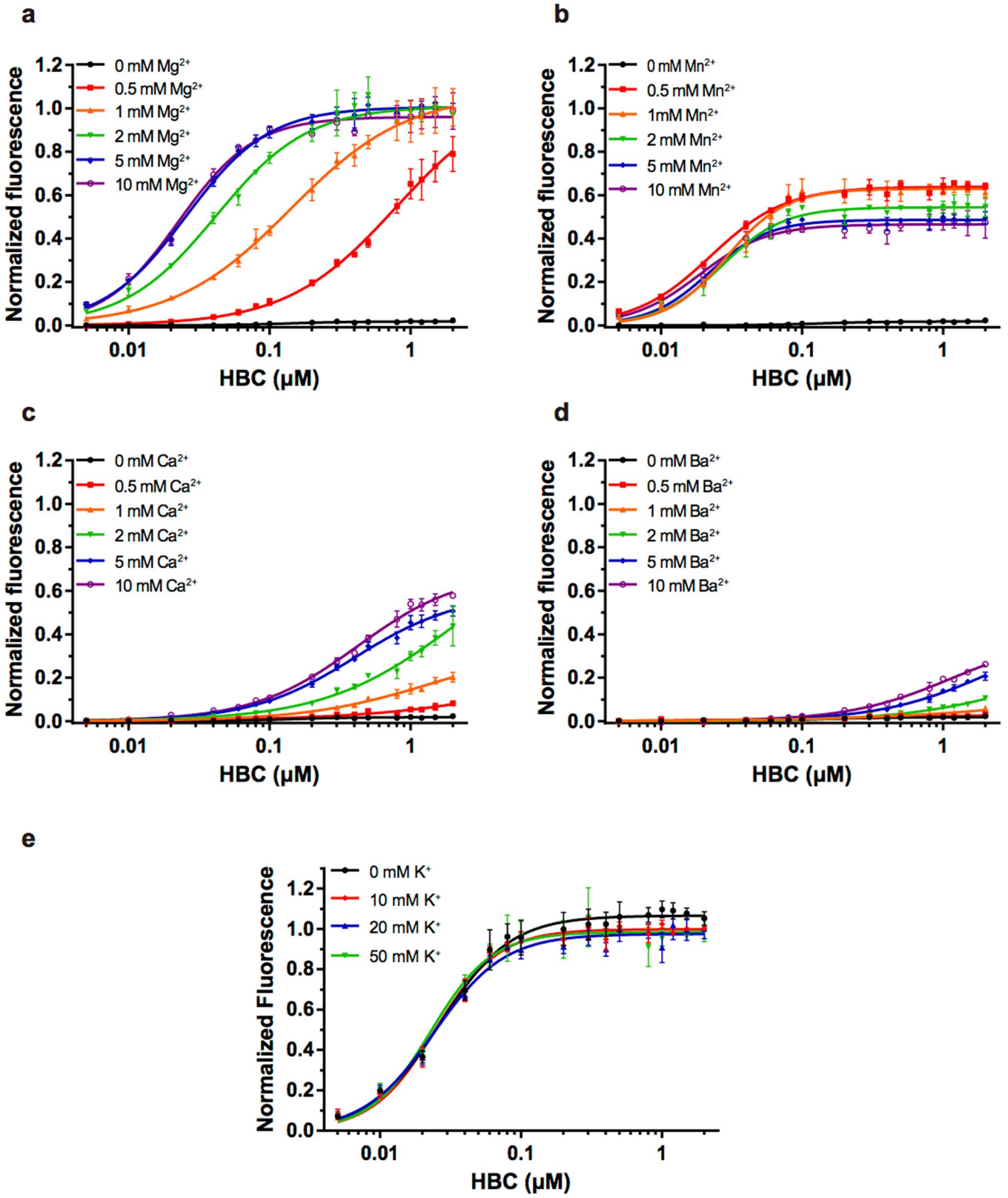


Extended Data Fig. 3 | See next page for caption.

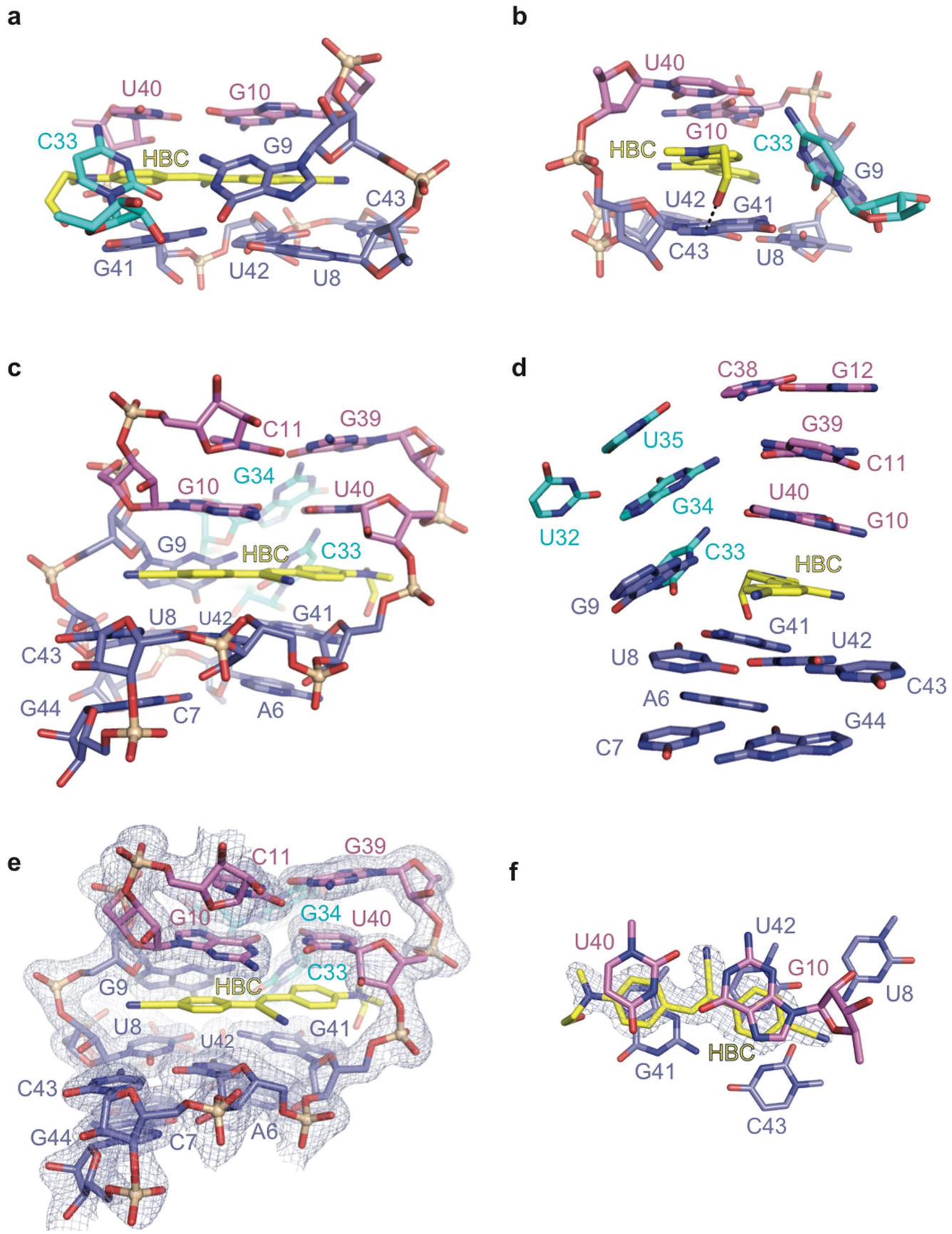
Extended Data Fig. 3 | Structural details of Pepper aptamer. a, b, Interaction between junction J3/2 and stem P2. A31 (J3/2) forms one hydrogen bond with the phosphate of G37 (stem P2). N3 of U32 forms one hydrogen bond with O4 of U35 within J3/2. Then, U35 forms two additional hydrogen bonds with the base of G37 and the phosphate of C38 from stem P2. **c**, A base triple is formed by G34 (J3/2) and the base pair G39-C22 from stem P2. **d**, C33 (J3/2) forms one hydrogen bond with U40 from the terminal of stem P2 and three additional hydrogen bonds with the Watson-Crick edge of G9 from J1/2, hence a base quadruple G10-U40-C33-G9 is formed in the bottom of J3/2. **e**, Three consecutive nucleotides G41, U42 and C43 from J2/1 are positioned in the same plane and form hydrogen-bonding interactions with the Watson-Crick edge of U8 from J1/2. **f**, G44 (J2/1) forms a canonical Watson-Crick base pair with C7 (J1/2). The base of A6 forms extensive interaction with the phosphate of C5, the base of C7 and the sugar of G41. Additionally, the sugar of U42 forms one hydrogen bond with the base of G44. It is notable that A6, G41 and U42 adopt 2′-*endo* sugar pucker conformation in the structure. **g**, 2′-OH of A6 pointes upwards and forms hydrogen bond interaction with the base of U8 and G9. The base of A6 forms two hydrogen bonds with the sugar of G41 and the phosphate of C5.



Extended Data Fig. 4 | The identified divalent metals and the involved interaction in Pepper structure. a, The composite omit electron density map (contoured at level 1.0 σ) for four metal cations M1, M2, M3 and M4 (shown as green balls) and the coordinated residues in Pepper aptamer. **b**, M5 located in the major groove of stem P1 and forms indirect coordination with the phosphates of G45, G46, G47, G48 and the base of G47, G2 from stem P1.

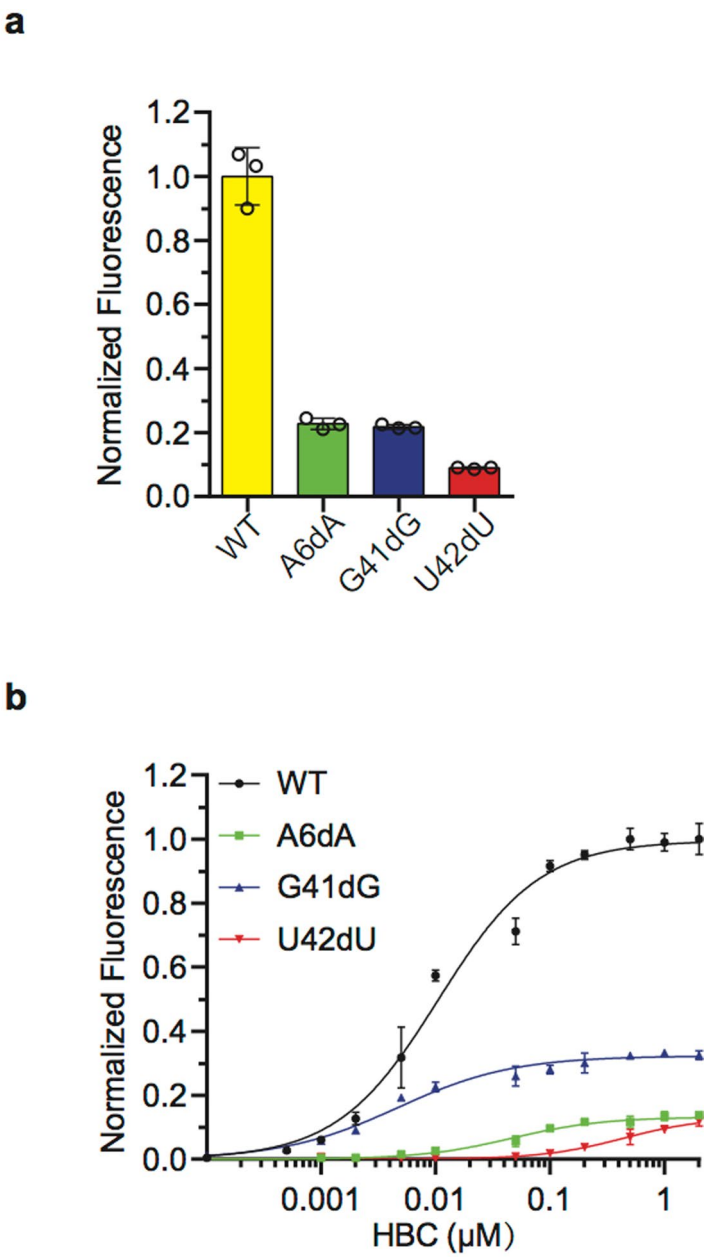


Extended Data Fig. 5 | The metal ions dependence assay of Pepper aptamer. a-d, The fluorescence activation assay of HBC by Pepper aptamer in the presence of Mg^{2+} (**a**), Mn^{2+} (**b**), Ca^{2+} (**c**) and Ba^{2+} (**d**). 50 nM refolded RNA in the reaction buffer containing 40 mM HEPES, pH 7.4, 125 mM KCl supplemented with individual divalent metal ions at different concentration ranging from 0 mM to 10 mM was titrated with increasing concentrations of HBC ranging from 0 μM to 2 μM. **e**, The activated fluorescence of HBC by Pepper aptamer in the presence of K^+ at different concentration from 0 mM to 50 mM. The titration of 50 nM refolded RNA with HBC (concentrations ranging from 0 μM to 2 μM) in the reaction buffer containing 40 mM HEPES, pH 7.4, 5 mM $MgCl_2$ supplemented with 0 mM to 50 mM KCl was carried out. Three independent experiments were carried out with similar results. Data represent the mean ± s.d. from three replicates.

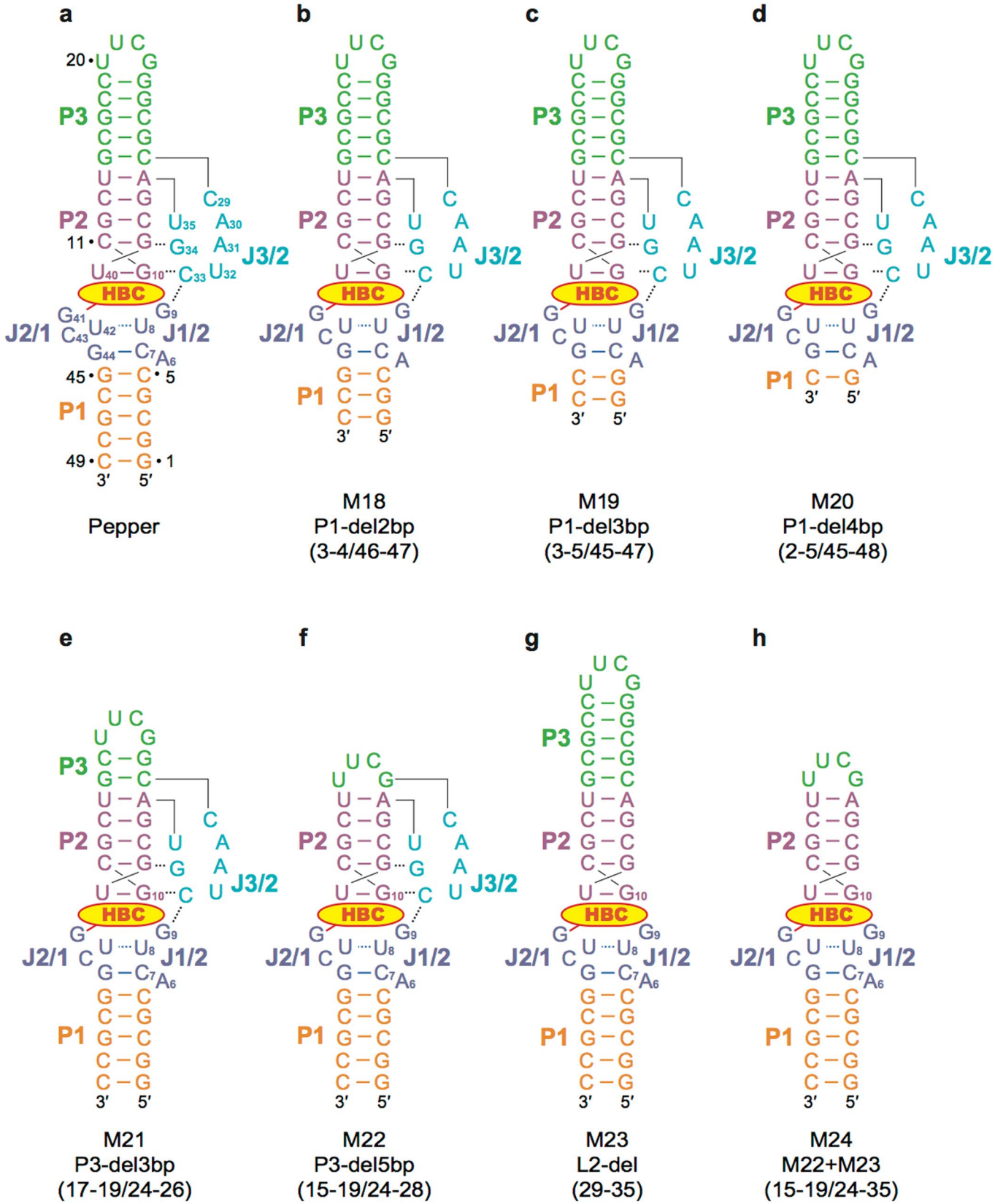


Extended Data Fig. 6 | See next page for caption.

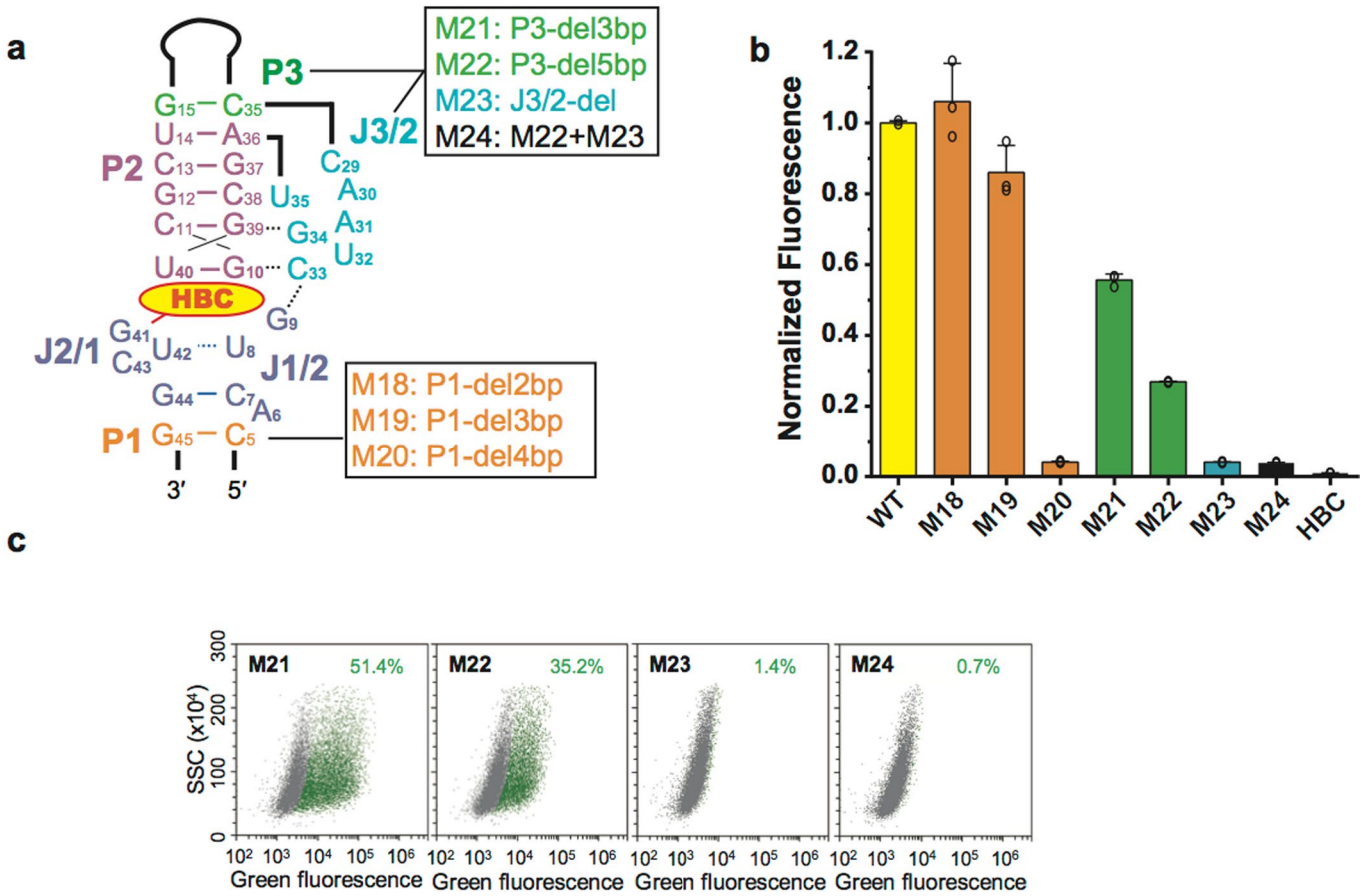
Extended Data Fig. 6 | The composition of HBC binding pocket in Pepper tertiary structure. a, b, Two rotated views of the four-sided binding box shown in Fig. 3c that capsulate ligand HBC in the tertiary structure. **c, d**, The binding box of HBC is further wrapped by the base triples G39-C11-G34 and A6-C7-G44 between stem P2 and the zipped junction J1/2 and J2/1. **e, f**, The composite omit electron density map (contoured at 1.0 σ level) of the residues involved in the formation of HBC-binding pocket (**e**) and the bound HBC (**f**) are shown respectively.



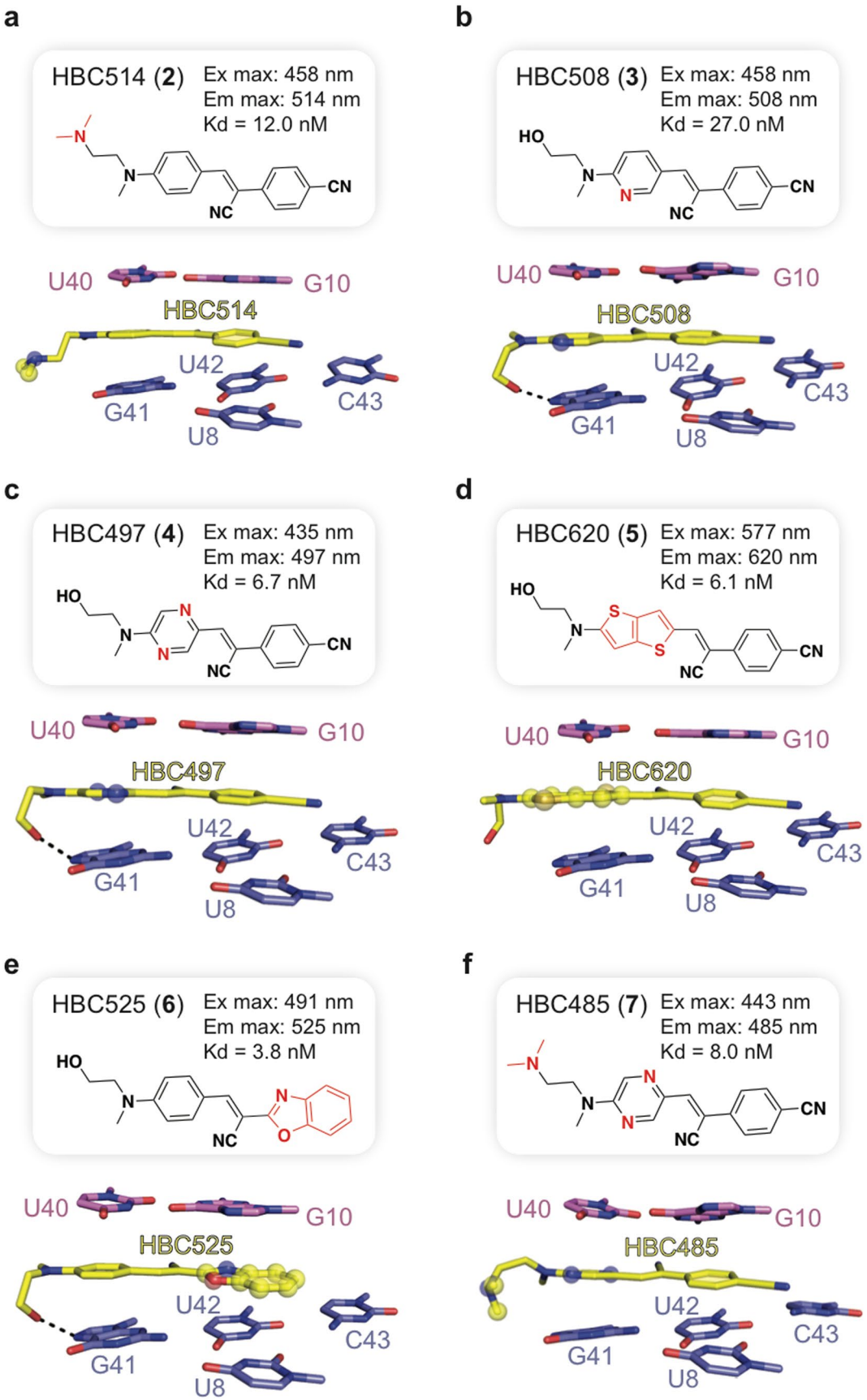
Extended Data Fig. 7 | Effect with deoxynucleotide mutations in Pepper aptamer. a, The activated fluorescence of HBC by Pepper mutants A6dA, G41dG and U42dU from three independently-repeated experiments are normalized for comparison with the wild-type (WT) Pepper aptamer. **b**, The titration of Pepper mutants A6dA, G41dG and U42dU with HBC were carried out in the buffer containing 40 mM HEPES, pH 7.4, 125 mM KCl and 5 mM $MgCl_2$ at room temperature. All the titrations are performed three times independently. Data represent the mean ± s.d. from three replicates.



Extended Data Fig. 8 | The secondary structure of Pepper aptamer and selected mutants M18-24. a, The secondary structure of Pepper aptamer used in crystallization. **b-d**, The secondary structure of Pepper aptamer mutants M18 (**b**), M19 (**c**) and M20 (**d**) with truncated stem P1. **e, f**, The secondary structure of Pepper aptamer mutants M21 (**e**) and M22 (**f**) with truncated stem P3. **g**, The secondary structure of Pepper aptamer mutant M23, in which the bulge junction J3/2 is deleted. **h**, The secondary structure of Pepper aptamer mutant M24, in which the bulge junction J3/2 and stem P3 are deleted.



Extended Data Fig. 9 | *In vitro* and *in vivo* fluorescence assay of selected mutants for M18-M24. a, Structure-based mutants of Pepper aptamer focusing on the truncation of Pepper aptamer. **b**, The *in vitro* fluorescence activation of HBC by Pepper mutants are normalized for comparison with the wild-type (WT) Pepper aptamer. All the tests are performed three times independently. Data represent the mean ± s.d. from three replicates. **c**, FACS analysis of Pepper fluorescence in mammalian cells. HEK293T cells transfected with plasmid expressing Pepper or Pepper variant were incubated with 1 μM HBC and analyzed using a flow cytometer. HEK293T cells transfected with mock plasmid and labeled with 1 μM HBC were used as the controls. Gate (green) was placed based on the control cells to determine the population fraction that exhibited higher fluorescence than the background.



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | The binding pockets composition of HBC-like fluorophores. The chemical structures, the spectral properties and the composition of the binding pockets of HBC514 (**a**), HBC508 (**b**), HBC497 (**c**), HBC620 (**d**), HBC525 (**e**) and HBC485 (**f**) in rotated view as shown in Fig. 5. All the modification sites of HBC-like fluorophores are labelled in red in the chemical structure and shown with transparent ball in the binding pocket composition. The excitation (Ex.) and emission (Em.) wavelengths, and the binding dissociation constant of each HBC-like fluorophore were shown with the compound as well.

nature research

Corresponding author(s): Aiming Ren

Last updated by author(s): Apr 22, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Descriptions of all the data collection are provided in the method section. The X-ray diffraction data were collected at the National Facility for Protein Science in Shanghai (NFPS), Zhangjiang Lab, China. A CytoFLEX-S flow cytometer (Beckman Coulter) was used for FACS analysis.
Data analysis	Descriptions of all the data analysis and the used software are provided in the method section and the supplementary materials. Cytexpert program (Beckman Coulter) was used to process flow cytometry data.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The template plasmid of pepper transcription will be provided upon request. Details of the transcription, purification, crystallization, fluorescence assay experiments have been provided in the method section and supplementary materials. All requests should be addressed to Aiming Ren. Atomic coordinates and structure factors for Pepper aptamer in complex with HBC and the analogs have been deposited at the Protein Data bank (www.rcsb.org) under accession number 7EOH, 7EOK, 7EOL, 7EOM, 7EON, 7EOO and 7EOP, respectively. The crystal structures of Ir(NH3)63+-soaked, Mn2+-soaked and Cs+-soaked crystals have been deposited at the Protein Data bank under accession number 7EOG, 7EOI and 7EOJ, respectively. Source data are provided with this paper.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample sizes chosen for each experiment were indicated in the method section or figure legends.
Data exclusions	No data was removed from our replicated fluorescence assay experiments. Some diffraction outliers of the crystals were automatically excluded by the established program according to the criteria.
Replication	All the attempts to replicate the experiments were successful.
Randomization	Not relevant to the study
Blinding	The authors were not aware of the results before performing the experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T(GNHu44) cells were purchased from Cell Bank of Chinese Academy.
Authentication	No cell lines were authenticated
Mycoplasma contamination	All cell lines have been tested negative for mycoplasma contamination by PCR methods
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	To measure the fluorescence of Pepper and its varaints in living cells, the transfected cells were digested and resuspended in the 4% FBS/1× PBS solution containing 1 μM HBC and 20 mM MgSO4.
Instrument	A CytoFLEX-S flow cytometer (Beckman Coulter) was used to acquire data.
Software	Data was collected using Cytexpert program (Beckman Coulter).
Cell population abundance	All reported populations were greater than 20,000 cells.
Gating strategy	Gate was placed based on the control cells labeled with 1 μM HBC to determine the population fraction that exhibited higher fluorescence than the background.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.