



# Directed evolution of a sequence-specific covalent protein tag for RNA labeling

Rongbing Huang<sup>a</sup> and Alice Y. Ting<sup>a,b,c,d,1</sup>

Contributed by Alice Y. Ting; received October 29, 2024; accepted January 22, 2025; reviewed by Chang C. Liu and David M. Shechner

Efficient methods for conjugating proteins to RNA are needed for RNA delivery, imaging, editing, interactome mapping, and barcoding applications. Noncovalent coupling strategies using viral RNA binding proteins such as MS2/MCP have been widely applied but are limited by tag size, sensitivity, and dissociation over time. We took inspiration from a sequence-specific, covalent protein–DNA conjugation method based on the Rep nickase of a porcine circovirus called “HUH tag”. Though wild-type HUH protein has no detectable activity toward an RNA probe, we engineered an RNA-reactive variant, called “rHUH”, through 7 generations of yeast display–based directed evolution. Our 13.4 kD rHUH has 12 mutations relative to HUH and forms a covalent tyrosine-phosphate ester linkage with a 10-nucleotide RNA recognition sequence (“rRS”) within minutes. We engineered the sensitivity down to 1 nM of target RNA, shifted the metal ion requirement from  $Mn^{2+}$  toward  $Mg^{2+}$ , and demonstrated efficient labeling in mammalian cell lysate. This work paves the way toward a potentially powerful methodology for sequence-specific covalent protein–RNA conjugation in biological systems.

RNA conjugation | directed evolution | protein engineering | RNA technology | RNA–protein interactions

RNAs play a central role in encoding, regulating, and driving biological functions. Increasingly, RNA-based pathways are also being intercepted by therapeutic modalities to treat disease (1). Due to their central importance, technologies to detect, image, manipulate, and engineer RNAs are in great demand. However, many of the most robust molecular tools, especially for deployment in complex biological settings such as living cells, are encoded as proteins, not RNAs. These include fluorescent proteins, CRISPR/Cas enzymes, HaloTag (2), light-switchable proteins such as LOV and CRY/CIBN, endonucleases, viral proteases, and the proximity labeling enzymes APEX (3) and TurboID (4). To use these tools, it is often necessary to establish a molecular link between protein and RNA to bring protein-encoded function to specific RNAs of interest.

Proteins and RNA can be conjugated using chemical methods, but cost, yield, imperfect specificity, and the need for subsequent purification can be prohibitive. Genetically encoded strategies for joining RNA to proteins provide attractive alternatives due to their ease of use, high specificity, and compatibility with living systems. Most common among these are the MS2/MCP system (5, 6), boxB/lambda (7), and PP7/PCP (8). The MS2/MCP system for example consists of a 19-nucleotide MS2 hairpin fused to an RNA of interest, which recruits MCP protein (13.7 kD) fused to a protein of interest. This versatile methodology has been used to image cellular RNAs (6, 8, 9), deliver transcriptional regulators to sgRNAs (10), target proximity labeling enzymes to RNA for discovery of interaction partners (11) and facilitate RNA editing (12). However, a major limitation of these sequence-specific RNA binding proteins is that the linkages are noncovalent. The pairs dissociate over time and cannot withstand extreme conditions such as high temperature, salt, pH, or organic solvent. A genetically encoded but covalent methodology for linking specific proteins and RNA together would offer a distinct paradigm.

To develop such a method, we looked for modular systems that provide covalent RNA–protein conjugation, are fully genetically encoded (no need for chemical derivatization or posttranscriptional/translational modifications), fast, and sequence specific. We did not identify a method that met all these criteria but found that the HUH tag system (13) provides an analogous methodology for covalently coupling single-stranded DNA (ssDNA) to proteins of interest. Derived from a porcine circovirus (PCV2) Rep nickase domain, the HUH tag (herein referred to as dHUH) recognizes a 10-nt ssDNA recognition sequence (herein referred to as dRS for DNA recognition sequence) and forms a covalent adduct via nucleophilic attack of Tyr96 of dHUH onto a specific phosphate backbone linkage of dRS. The outstanding specificity, speed, and sensitivity of this coupling has enabled the HUH tag system to be used for a wide variety of applications, including conjugation of ssDNA barcodes to antibody reagents, tethering

## Significance

To enable the study of RNA in complex biological settings, methods are needed to conjugate RNA to protein-based reporters and enzymes with high sensitivity and specificity. To develop this capability, we started from a natural enzyme that undergoes sequence-specific conjugation to single-stranded DNA and used directed evolution to modify it to react with RNA instead. The resulting “rHUH” protein forms covalent adducts with RNA targets at nanomolar concentrations, within minutes, in a range of settings.

Author affiliations: <sup>a</sup>Department of Genetics, Stanford University, Stanford, CA 94305; <sup>b</sup>Department of Biology, Stanford University, Stanford, CA 94305; <sup>c</sup>Department of Chemistry, Stanford University, Stanford, CA 94305; and <sup>d</sup>Chan Zuckerberg Biohub – San Francisco, San Francisco, CA 94158

Author contributions: R.H. and A.Y.T. designed research; R.H. performed research; R.H. and A.Y.T. analyzed data; and R.H. and A.Y.T. wrote the paper.

Reviewers: C.C.L., University of California Irvine; and D.M.S., University of Washington.

Competing interest statement: A.Y.T. is a scientific advisor to Third Rock Ventures and Nereid Corporation.

Copyright © 2025 the Author(s). Published by PNAS. This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

<sup>1</sup>To whom correspondence may be addressed. Email: [aying@stanford.edu](mailto:aying@stanford.edu).

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2422085122/-/DCSupplemental>.

Published February 26, 2025.

template DNA to Cas9 for genome editing (14), and displaying DNA on nanoparticles (15).

Due to the similarities between ssDNA and RNA, we wondered whether it would be possible to engineer dHUH to recognize and form covalent adducts with short sequence-specific RNA tags (“rRS”) instead of DNA (Fig. 1A). Herein, we describe the engineering of “rHUH” protein for covalent coupling to short (10 nt) RNA motifs (rRS) and the characterization of this conjugation system on yeast and in vitro.

Results

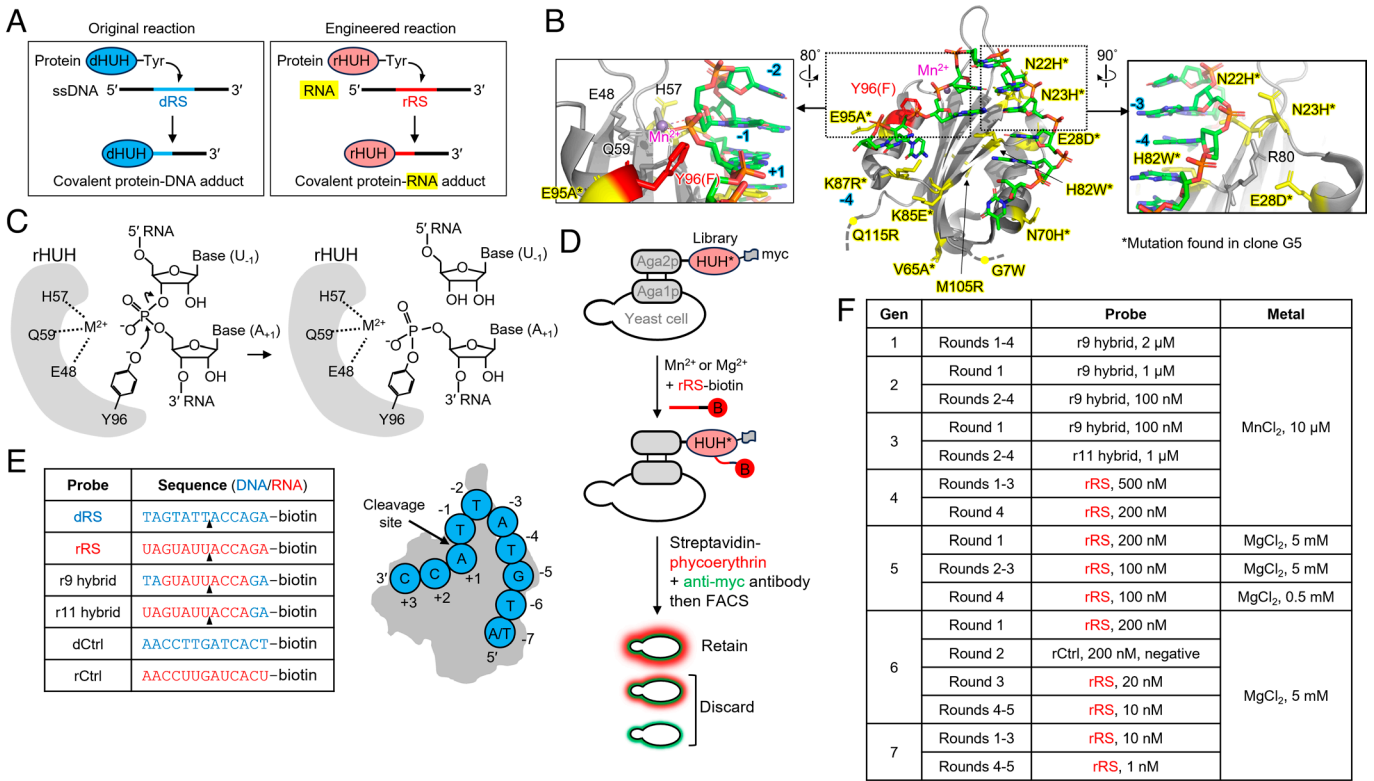
The structure of dHUH in complex with dRS (16) shows extensive contacts between the protein and both phosphate backbone and nucleobases of dRS, which is bent into a U-shape with hydrogen bonding between two base pairs at the hairpin turn (Fig. 1B). The active site of dHUH contains a Mn<sup>2+</sup> ion, coordinated by Glu48, His57, and Gln59 sidechains, which helps to stabilize the pentavalent phosphate transition state generated by nucleophilic attack from Tyr96 (Fig. 1C). After the reaction, the 3' end of dRS is covalently coupled to dHUH via a tyrosine-phosphate ester bond, while the 5' end of dRS (7 nt long) is released.

We established a yeast surface display platform to both characterize and evolve dHUH (Fig. 1D). The 13.3 kD protein was fused to the C-terminal end of the yeast mating protein Aga2p for display on the cell surface. A myc tag was fused to the C-terminal end of

dHUH to facilitate detection by anti-myc antibody. To assess the activity of displayed dHUH, we added 3'-biotin-conjugated DNA probe (biotin-dRS, Fig. 1E) to the cells, washed, and stained with streptavidin-phycoerythrin (PE) conjugate. We observed robust labeling of dHUH by biotin-dRS down to a concentration of 0.1 nM, but no labeling by a scrambled control DNA probe (SI Appendix, Fig. S1A).

We then tested the RNA version of the dRS probe (biotin-rRS). No signal was detected, even at high (1 μM) probe concentrations and long incubation times (2 h). We attempted some rational mutagenesis of dHUH, based on the structure, selecting sidechains that might contact the 2'-OH of bound RNA (SI Appendix, Fig. S1B). None of the 7 mutants we tested showed activity with RNA probe, though some had decreased reactivity toward DNA. We recognized that much more extensive screening would be necessary and proceeded to build a library of dHUH variants for yeast display evolution. To construct the library, we opted for error-prone PCR, rather than more focused mutagenesis of residues in contact with dRS, because long-range interactions could impact both substrate recognition and catalysis. Furthermore, we did not assume that an rRS probe would dock to HUH protein in the same manner as dRS. Our error-prone PCR produced an average of 1 to 2.3 amino acid changes per gene and a library size of ~1.2 × 10<sup>8</sup> (Methods).

We were concerned that the change from dRS to rRS might be too drastic, such that no member of our initial library possessed the ability to recognize a fully RNA substrate. Thus, we performed



**Fig. 1.** Engineering RNA-reactive rHUH protein and directed evolution scheme. (A) Schematic of wild-type (WT) reaction of dHUH protein with ssDNA (the DNA recognition sequence, dRS), and desired conjugation between engineered rHUH protein and RNA (the RNA recognition sequence, rRS). (B) Crystal structure of WT HUH protein in complex with its target ssDNA sequence (dRS), from PDB ID 6WDZ (16). The nucleophilic Tyr96 (mutated to Phe in this structure) is red, the Mn<sup>2+</sup> atom is purple, and the 12 residues we mutated over the course of evolution to rHUH are colored yellow. Zoom on Right shows multiple mutations proximal to -3 and -4 positions of bound dRS. (C) Expected active site chemistry of rHUH, showing nucleophilic attack by deprotonated Tyr96 on RNA backbone, leading to a covalent adduct between rHUH protein and the 3' portion of the target RNA (rRS). Mn<sup>2+</sup>, divalent metal ion such as Mn<sup>2+</sup> or Mg<sup>2+</sup>. (D) Yeast surface display selection scheme. To a library of HUH mutants (HUH\*) is added biotinylated rRS or hybrid sequences [see (E)]. After washing, cells are stained with streptavidin-phycoerythrin (PE) to quantify probe labeling and anti-myc antibody to quantify HUH expression level. FACS sorting is used to enrich cells with high streptavidin-PE/myc intensity ratio. (E) Sequences of ssDNA target (dRS), RNA target (rRS), RNA-DNA hybrids used for early generations of directed evolution, and scrambled control probes. (F) Table of selection conditions used across 7 generations (Gen) of directed evolution. Each generation consisted of 4 to 5 rounds of amplification and sorting, followed by identification of the top clone (named “G1-G7”) and rediversification via error-prone PCR. All labeling times were 2 h at room temperature.

a study of DNA-RNA hybrid substrates, in which single nucleotides or groups of nucleotides in dRS were replaced with RNA (SI Appendix, Fig. S1 C and D). Our results show that several positions of dRS are tolerant to substitution by RNA, while the -3 and -4 positions are quite sensitive (SI Appendix, Fig. S1E). Ultimately, we selected an r9 hybrid probe (Fig. 1E), because we could detect labeling of WT dHUH with high concentrations of this probe (SI Appendix, Fig. S1F).

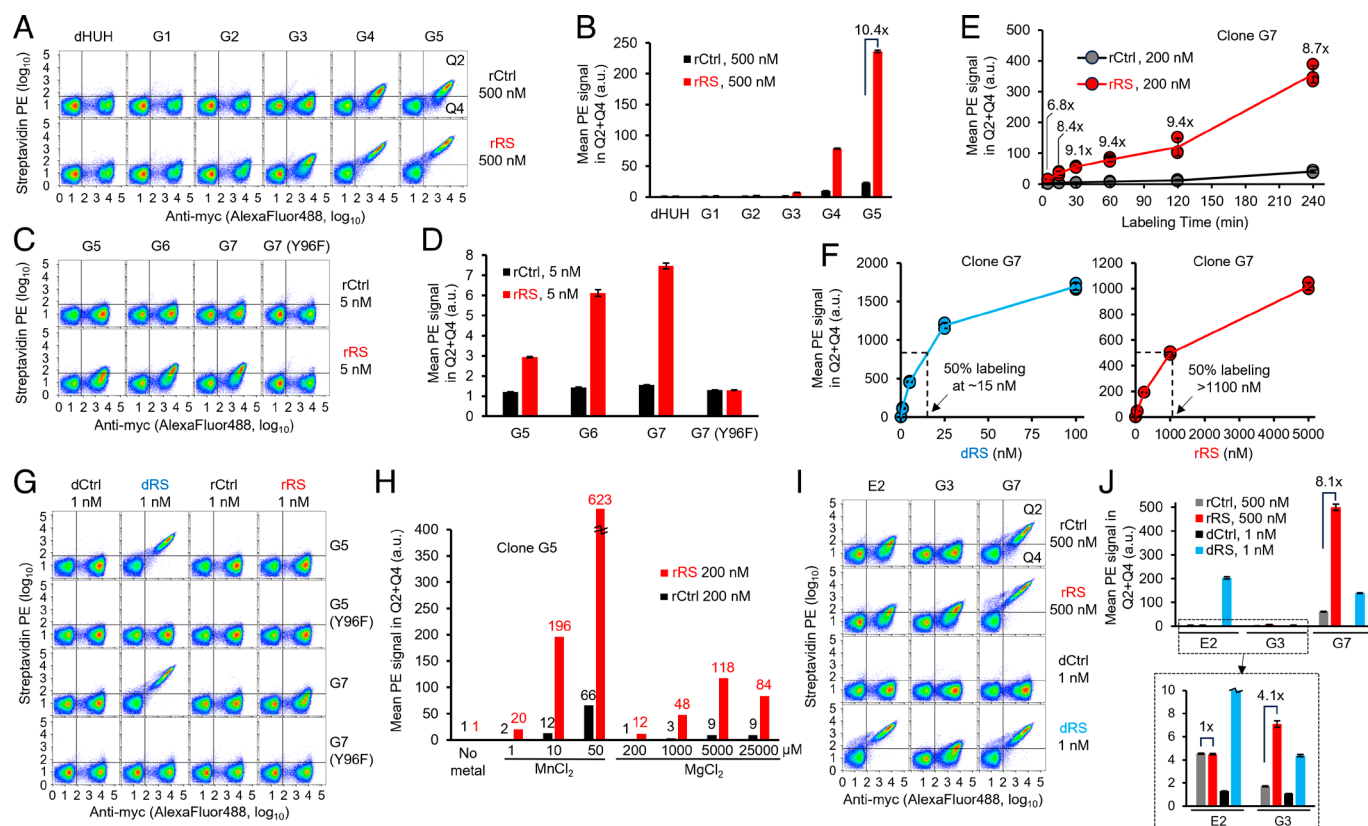
We began evolution by treating the yeast-displayed dHUH library with 2  $\mu$ M of r9 hybrid for 1 h. After washing the cells and staining with streptavidin-PE and anti-myc antibody, we found that ~30% of myc-positive cells showed signal above background. We used FACS to separate this population, amplified the yeast cells, and performed 3 more rounds of selection. Enriched cells were sequenced, and unique clones were compared on the yeast cell surface (SI Appendix, Fig. S2). The best clone, with 3 mutations relative to dHUH, was named G1 and used as the template for the second generation of evolution.

In this manner, we performed 7 total generations of directed evolution, with the winning clone from the previous generation serving as the template for library production for the next generation (Fig. 1F). Generation 2 still utilized the r9 hybrid probe, while Generation 3 used an r11 hybrid with only two nucleotides from DNA. By Generation 4, we had sufficient activity to switch to a pure RNA probe. In Generations 4–7, we progressively decreased the concentration of rRS from 500 nM to 1 nM to select for HUH variants

with higher RNA reactivity. In Generation 5, we also replaced  $\text{MnCl}_2$  with  $\text{MgCl}_2$ , as the latter but not the former is available in cells.

**Results from Directed Evolution.** The mutations found in winning clones from each generation (G1–G7) are shown in SI Appendix, Fig. S2A and Table S3. Other enriched clones are shown in SI Appendix, Table S3 and analyzed in SI Appendix, Fig. S2 B–H. On the yeast surface, we compared the winning clones G1–G7 alongside the original dHUH template for reaction with both dRS and rRS (Fig. 2A–D and SI Appendix, Fig. S3). Fig. 2A and B shows that labeling with rRS is detectable at G3 and jumps dramatically at G4. Reaction with dRS drops somewhat over generations with the most dramatic dip at G3 (SI Appendix, Fig. S3 C and D), perhaps due to the acquisition of H82Y, which may confer strong RNA bias. Notably, G4–G7 also showed significant reactivity toward a scrambled rRS control sequence (“rCtrl”), although the labeling extent was ~10-fold lower than with rRS (Fig. 2B and D).

We performed more careful characterization of clones G5 (which has 9 mutations relative to dHUH) and G7 (which has 3 additional mutations compared to G5; SI Appendix, Fig. S2A). First, we completed a time course experiment, measuring the extent of labeling with 200 nM rRS from 5 min to 4 h. For both G5 and G7, we observed a monotonic increase over this time, while the reaction with scrambled rCtrl rose much less (ninefold difference in labeling extent at 1 h) (Fig. 2E and SI Appendix, Fig. S4 A–C). G7 displayed higher labeling than G5 at all timepoints.



**Fig. 2.** Directed evolution results and characterization of top clones in yeast. (A) FACS comparison of original dHUH template and evolved clones G1–G5, illustrating progress of evolution. Yeast displaying the indicated constructs were labeled with 500 nM rRS for 1 h in the presence of 5 mM  $\text{MgCl}_2$  before streptavidin-PE/anti-myc staining and FACS. rCtrl is a scrambled version of rRS (sequence in Fig. 1E). (B) Quantification of data in (A) and two additional replicates (SI Appendix, Fig. S3A). Labeling of G1–G5 with DNA probes is shown in SI Appendix, Fig. S3 C and D. Mean  $\pm$  SD,  $n = 3$ . (C) FACS comparison of evolved clones G5–G7 on yeast, using 5 nM rRS for 1 h in the presence of 5 mM  $\text{MgCl}_2$ . The Y96F mutation inactivates G7. (D) Quantification of data in (C) and two additional replicates (SI Appendix, Fig. S3B). Mean  $\pm$  SD,  $n = 3$ . (E) Time-dependent labeling of G7 on the yeast surface with 200 nM rRS or rCtrl, in the presence of 5 mM  $\text{MgCl}_2$ . FACS plots in SI Appendix, Fig. S4A. Mean  $\pm$  SD,  $n = 3$ . (F) Labeling of G7 on yeast at various dRS (Left) and rRS (Right) concentrations. Cells were labeled for 5 min in the presence of 50  $\mu$ M  $\text{MnCl}_2$ . FACS plots in SI Appendix, Fig. S5A. Mean  $\pm$  SD,  $n = 3$ . (G) FACS plots showing labeling of G5, G7, and inactive mutants with 1 nM probe for 2 h in the presence of 5 mM  $\text{MgCl}_2$ . (H) Effect of metal ion type and concentration on extent of G5 labeling on yeast, with rRS probe for 2 h. FACS plots in SI Appendix, Fig. S5F. (I) Comparison of G3 and G7 to another evolved RNA-reactive HUH variant, “E2” from ref. 17. Constructs on yeast were labeled with 500 nM rRS or 1 nM dRS for 1 h in the presence of 5 mM  $\text{MgCl}_2$ . (J) Quantification of data in (I) and two additional replicates. Mean  $\pm$  SD,  $n = 3$ .



We next titrated the concentration of rRS (and for comparison, dRS probe as well). Fig. 2F and *SI Appendix, Fig. S5 A–E* show that both G5 and G7 labeling with rRS increased steadily and did not show signs of saturation, even at the highest practically achievable concentration of 5  $\mu$ M RNA probe. By contrast, labeling with dRS could be easily saturated at  $\sim$ 50 nM probe. Thus, G7 and G5 are still labeled by DNA probe with far greater sensitivity than by RNA probe. In Fig. 2G, G5 and G7 are compared side by side at very low probe concentration – 1 nM dRS or rRS. Whereas labeling of G5 is difficult to detect at this concentration, G7 shows clear signal over background. Negative controls with scrambled rCtrl or the inactivating point mutation Y96F in G7 fail to show labeling.

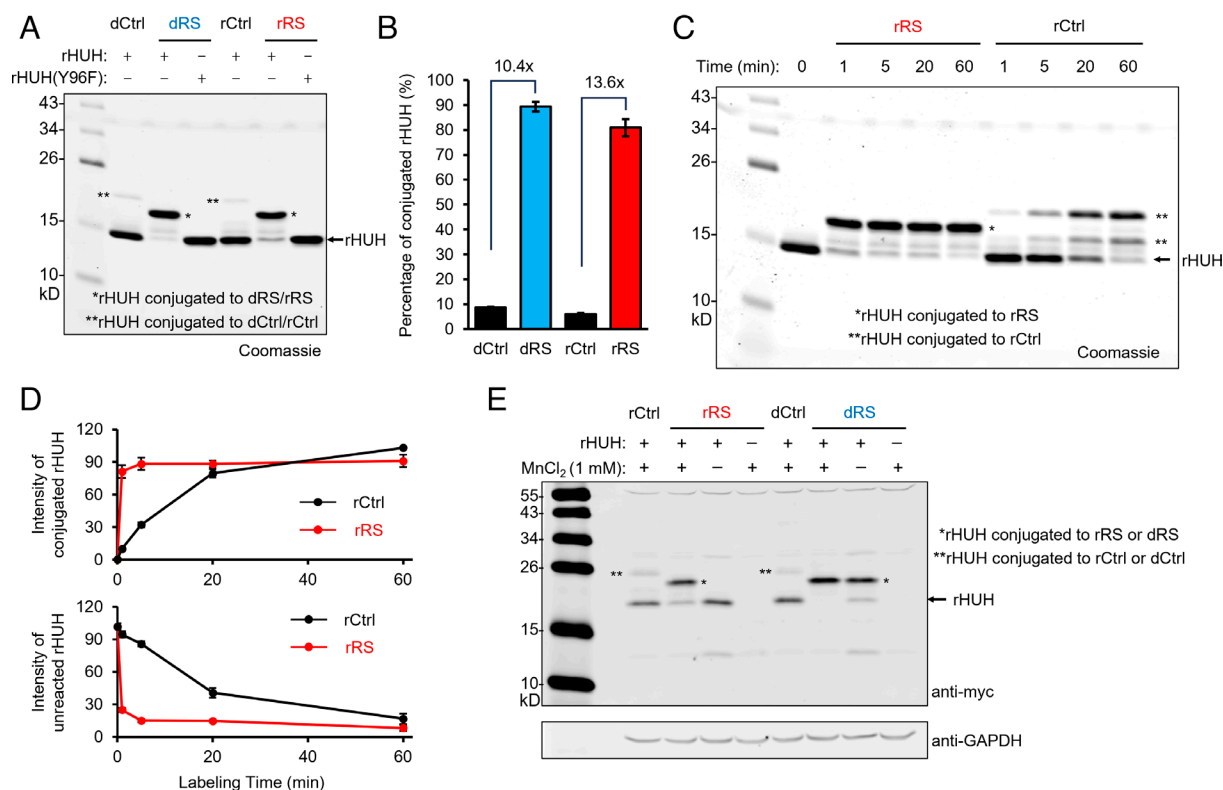
Because the HUH reaction mechanism requires a divalent metal ion ( $Mn^{2+}$  in the crystal structure), we tested the metal ion dependence (Fig. 2H and *SI Appendix, Fig. S5F*). Like WT dHUH, G5 preferred  $Mn^{2+}$  over  $Mg^{2+}$ , but labeling was substantial at 1 to 5 mM of  $MgCl_2$ , which is in the range of physiological  $Mg^{2+}$  concentrations (18, 19).

Mutagenesis of the active site Tyr96 in G5 and G7 to Phe abolished labeling by rRS (Fig. 2D and *SI Appendix, Fig. S4C*), suggesting that these constructs use a similar labeling mechanism to dHUH. To examine the necessity of all 9 mutations in G5 relative to WT dHUH, we reverted each one to the WT amino acid. *SI Appendix, Fig. S6* shows that every reversion decreased labeling with rRS, with the mutations H22N and W82H having the largest effect. His82 in the WT dHUH structure forms a  $\pi$ – $\pi$  stacking interaction with adenosine in the +1 position of dRS. His82 was mutated to Tyr in G3, and further mutated to Trp in G4, both of which may increase  $\pi$ – $\pi$  stabilization.

Altogether, our experiments indicate that our directed evolution produced rHUH candidates with vastly increased reactivity toward RNA probes. Each generation of evolution produced steep improvements in reactivity, with G5 being  $\sim$ 25-fold more active than G3, and G7 being  $\sim$ 2.6-fold more active than G5. Ultimately, G7 could be labeled with as little as 1 nM rRS probe, in the presence of 5 mM  $Mg^{2+}$  rather than  $Mn^{2+}$ , and with a specificity ratio of  $\sim$ 12. We selected G7 as our final rHUH and refer to it as rHUH henceforth.

**In Vitro Characterization of rHUH.** We purified rHUH protein by overexpression in *Escherichia coli* and nickel affinity chromatography (16). Because rHUH has multiple surface-exposed histidines, it does not require a His<sub>6</sub> tag to bind to nickel. To evaluate the reactivity of rHUH toward rRS, we used a gel shift assay, detecting the rHUH protein by Coomassie staining. Upon covalent reaction with rRS, the molecular weight of rHUH shifts from  $\sim$ 15 kD to  $\sim$ 19 kD. Fig. 3A and B and *SI Appendix, Fig. S7A* show the reaction of rHUH with both rRS and dRS, after just 1 min incubation in the presence of 1 mM  $MnCl_2$ . Mutation of the active site tyrosine (Y96F) abolishes labeling with both probes. Consistent with our observations on the yeast surface, rHUH also reacts to a small extent with scrambled control – both rCtrl and dCtrl, with a specificity ratio of 13.6 for RNA and 10.4 for DNA.

We also used purified rHUH protein to perform a time course experiment. The protein was combined with rRS or rCtrl in the presence of 1 mM  $MnCl_2$  for 1 min to 1 h and then analyzed by SDS-PAGE analysis and Coomassie staining (Fig. 3C and D and *SI Appendix, Fig. S7B*). Under these conditions, rHUH conjugation to rRS was complete in 1 min, whereas labeling with rCtrl required an hour to reach a similar labeling intensity.



**Fig. 3.** In vitro characterization of purified rHUH (G7). (A) Gel shift assay showing reaction of purified rHUH with DNA and RNA probes. 5  $\mu$ M rHUH was combined with 10  $\mu$ M probe for 1 min in the presence of 1 mM  $MnCl_2$  and then analyzed by SDS-PAGE and Coomassie staining. Uncropped gel and additional replicates in *SI Appendix, Fig. S7A*. (B) Quantification of data in (A). Mean  $\pm$  SD,  $n = 3$ . (C) Time-dependent labeling of purified rHUH with rRS. 5  $\mu$ M rHUH was combined with 10  $\mu$ M probe in the presence of 1 mM  $MnCl_2$ . Uncropped gel and additional replicates shown in *SI Appendix, Fig. S7B*. (D) Quantification of data in (C). Mean  $\pm$  SD,  $n = 3$ . (E) rHUH labeling in mammalian cell lysate. HEK 239 T cells expressing rHUH were lysed and incubated with 10  $\mu$ M DNA and RNA probes as in (A), for 5 min in the presence of 1 mM  $MnCl_2$ . rHUH protein was then detected by anti-myc western blotting.

To estimate the binding affinity between rHUH and rRS, we used the inactive mutant (Y96F) and an electrophoretic mobility shift assay with fluorophore-conjugated rRS probe. *SI Appendix, Fig. S8 A and B* shows a  $K_d$  of  $\sim 1.66 \mu\text{M}$ , suggesting that rHUH has considerably lower affinity for rRS than WT dHUH has for dRS.

**Sequence Specificity of rHUH.** We attempted to improve the specificity ratio of rHUH beyond  $\sim 12$  through incorporation of negative selections in Generations 6 and 7, but these efforts were ultimately not successful. *SI Appendix, Fig. S9* shows the alternative paths explored in Generations 6 and 7 that switch between positive and negative selections. Each time we implemented a negative selection, enriching low streptavidin-PE cells after treatment with 200 to 500 nM rCtrl, both the signal (rRS labeling activity) and background (rCtrl labeling activity) of the resulting library decreased. Both were then restored following rounds of positive selection. At the end of 5 alternating rounds, none of the enriched clones from this path had higher activity than the starting template (G5). We attempted the same strategy in Generation 7 (*SI Appendix, Fig. S9C*) and found that all our enriched clones were identical in sequence to the template (G6). Ultimately the winners of each generation (G6 and G7) derived from the selection paths that largely did not incorporate negative rounds.

Despite the modest specificity ratio, we moved toward rHUH testing in mammalian HEK 293 T cells. We prepared cell lysate and incubated it with 10  $\mu\text{M}$  of rRS or dRS for 5 min. Gel shift analysis in Fig. 3E shows extensive reaction with both probes. Interestingly, rRS reaction requires supplementary addition of 1 mM  $\text{MnCl}_2$  while reaction with dRS does not. Here again, some nonspecific reactions with rCtrl and dCtrl were detected, with specificity ratios similar to those previously observed for purified rHUH (6.3 for RNA and 14.4 for DNA here). We also repeated the experiment with rRS coexpressed in HEK293T cells but found that rHUH was labeled to a similar degree by rRS, rCtrl, and endogenous GAPDH mRNA. These results suggest that a major impediment to the use of rHUH inside cells is the inadequate sequence specificity.

In light of these data, we performed a more careful study of rHUH sequence specificity using in vitro transcribed (IVT) RNAs. As shown in *SI Appendix, Fig. S10 A and B*, the IVT RNA probes contain a target region (rRS or variant), linker or structural sequences, and a constant annealing region on the 3' end to facilitate detection. We labeled yeast displaying rHUH with 10 nM of various IVT RNAs in the presence of  $\text{MnCl}_2$  for 1 h at room temperature. Bound probes were then detected with biotinylated antisense DNA followed by streptavidin-PE and FACS analysis.

We first compared IVT rRS with three scrambled control probes (rCtrl1-3) and found that all three controls were labeled 2 to 2.5-fold less than rRS (*SI Appendix, Fig. S10 C and D*). The relatively low specificity ratio for these IVT RNAs compared to synthetic probes might arise from nonspecific interactions between rHUH and the annealing region of the IVT probes. We also mutated rRS by 1 or 2 nucleotides to generate probes SM1-10 and DM1-6 (*SI Appendix, Fig. S10B*). We found that most single mutants (SM) and all double mutants (DM) of rRS showed 15 to 70% decreased labeling compared to rRS (*SI Appendix, Fig. S10 E and F*). Two mutants of rRS (SM1 and SM7) showed enhanced labeling with rHUH.

To improve the labeling of rRS by rHUH in the context of longer RNAs, we designed a stem-loop rRS to stabilize and protect the rRS recognition sequence from secondary interactions (*SI Appendix, Fig. S10G*). We found that rRS-SL gave 3.6-fold higher labeling than linear rRS probe and a slightly higher speci-

ficity ratio (2.5 instead of 2.1; *SI Appendix, Fig. S10 G and H*). Stem-loop rRS may be beneficial when using rHUH to label longer RNAs.

We performed a BLAST search of sequences similar to rRS present in the human transcriptome. Over 250 sequences differ by only one nucleotide, and two transcripts have a perfect match to rRS. We made an IVT probe from one of these mRNAs, GPR26, that contains the rRS match and 31 flanking nucleotides. This probe labeled rHUH efficiently on the yeast cell surface, while its corresponding control (with the rRS match region replaced by rCtrl1) gave 2.6-fold less labeling. In its native cellular context, full-length GPR26 mRNA may not be as sterically accessible or reactive as our truncated IVT GPR26-based probe. However, these data do highlight the considerable challenge of translating rHUH methodology to the live cell context.

**Comparison of rHUH to E2 from ref. 17.** Recently, another RNA-reactive HUH variant was reported in a preprint by Smiley et al. (17). To evaluate their clone, E2, we performed a side-by-side comparison on yeast to our G3 and G7 (rHUH) constructs (Fig. 2 I and J). E2 has 3 mutations relative to dHUH, one of which is also present in our final rHUH (H82W). We labeled yeast cells with 500 nM RNA probe (rRS) for 1 h and observed a small degree of labeling of E2, similar to the labeling extent of our G3. The labeling of rHUH (G7) was far higher, about 100-fold more under the same conditions. In addition, E2 showed similar reactivity toward rRS and rCtrl (specificity ratio  $\sim 1$ ), suggesting a loss of sequence specificity, while G3 and G7 preferred rRS to rCtrl by 4.1- and 8.1-fold, respectively.

E2 has two mutations (E84Q and K87Y) not present in rHUH. We wondered whether adding these mutations to rHUH might improve its properties. Thus we prepared "G7E2" with two additional E2-derived mutations compared to G7 (rHUH). On the yeast surface, we found that G7E2 was  $\sim 20\%$  less active than G7 toward rRS probe, and similarly specific ( $\sim 10$ -fold) (*SI Appendix, Fig. S11 A and B*).

Differences in library design and selection conditions may account for the  $\sim 100$ -fold greater activity of our rHUH compared to E2. Smiley et al. used saturation mutagenesis of specific positions around the dHUH active site to generate a library, while we used error-prone PCR seven separate times (for Generations 1–7) to diversify across the entire dHUH gene. We performed a total of 30 rounds of selection via FACS, whereas Smiley et al. performed 6 rounds in total, four of which were via MACS, which has a lower dynamic range than FACS. Finally, we lowered RNA probe concentration progressively (from 500 nM to 1 nM) and switched from  $\text{Mn}^{2+}$  to  $\text{Mg}^{2+}$  in Generation 5, while Smiley et al. fixed both the RNA probe concentration (at 50 nM) and metal ion concentration (to 500  $\mu\text{M}$   $\text{Mn}^{2+}$ ). The creation of rHUH likely required these more stringent selection conditions and repeated library diversification.

## Discussion

In this study, we have engineered rHUH, a covalent and sequence-specific protein tag for RNA, through 7 generations of yeast surface display directed evolution. rHUH forms a covalent adduct with a 10 nucleotide RNA recognition sequence called rRS through  $\text{Mn}^{2+}$  or  $\text{Mg}^{2+}$ -dependent formation of a tyrosine-phosphate ester bond. We showed that rHUH can couple to non-structured RNAs on the yeast surface and in vitro, with a sensitivity down to 1 nM RNA probe, and a specificity factor of  $\sim 6$  to 12. This methodology could potentially be used to link both short and long RNA transcripts to a wide range of protein-based probes

including fluorescent proteins, CRISPR-based editors, and proximity labeling enzymes.

rHUH enhances the arsenal of methods available to link RNA to protein. Among covalent strategies, chemical conjugation is dominant (20), but recent studies have begun to exploit enzymes for the task. For example, the RNA-TAG method (21) uses a tRNA guanine transglycosylase to attach a benzylguanine moiety onto an RNA of interest. After purification of the conjugate, the RNA is reacted with a SNAP tag-fused protein for 4 h (21). A simpler methodology uses an engineered variant of uridine-54 tRNA methyltransferase, TrmA, to form a covalent adduct with an RNA hairpin motif (22), but the reaction appears to be slow and the generalizability has not been explored.

The long-term goal of rHUH engineering is protein–RNA conjugation inside living cells, where the ability to visualize, analyze, and manipulate specific RNA species would be very highly enabling. The major challenges to translating rHUH methodology to the cell interior are sensitivity, specificity, and metal ion availability. While rRNAs and tRNAs are highly abundant [ $\sim 2$  to  $25\ \mu\text{M}$  (23)], many mRNA species are present at just 1 to 100 nM or even less inside cells (23). An effective rHUH tag would need to react rapidly with target RNA species in this concentration regime, while avoiding off-target reactions with thousands of higher-abundance RNAs. The current specificity ratio of our rHUH (ratio of reaction with rRS versus scrambled control) is only 6 to 12, far less than what is required for specific tagging in living cells. Indeed, our preliminary test of rHUH–rRS labeling in HEK 293 T cells showed a lack of specific conjugation. Finally, WT dHUH requires 10 to 500  $\mu\text{M}$  of  $\text{Mn}^{2+}$ , which is scarce inside cells; protein-bound and unbound pools of  $\text{Mn}^{2+}$  total only  $\sim 1\ \mu\text{M}$  (24). We used directed evolution to shift the divalent metal ion requirement toward  $\text{Mg}^{2+}$ , which is present at 0.5 to 1 mM in free form (19), but further engineering is likely necessary for high activity in the cellular milieu.

Our yeast display directed evolution platform was able to produce RNA labeling activity from a starting template with no detectable reactivity toward RNA on the yeast surface. We observed remarkable improvements in activity toward RNA over 7 generations of directed evolution, culminating in our best clone, G7, which is  $\sim 780$ -fold more active than G1, our first-generation winner. Nevertheless, we found that our selection platform was highly limited when it came to engineering rHUH sequence specificity. We attempted many rounds of negative selection but only managed to decrease *both* specific and nonspecific reactivity together. From generation 3 onward, only the activity improved, with little to no gain in specificity ratio. A different selection scheme, that simultaneously (rather than alternately) selects for high rRS labeling and minimal off-target labeling might yield improved results. Another goal of future work is to optimize the RNA target sequence. Our study of rRS single mutants revealed two constructs with higher labeling than rRS (*SI Appendix, Fig. S10 E and F*). By mixing and matching mutations, we obtained a second-generation “rRS2” with threefold higher labeling of rHUH on yeast than the original rRS [which mirrors exactly the original dRS used by HUH tag (16) (*SI Appendix, Fig. S12*)]. Further optimization of the rRS is likely to improve both the specificity and sensitivity of this methodology. Overall, our study lays the groundwork for a versatile covalent protein tag for investigating the biology of RNAs.

## Methods

**Cloning.** The double-digested vectors and the PCR amplified inserts with 20 to 50 homologous overhangs on both ends were ligated by Gibson assembly. Ligated plasmid products were introduced by heat shock transformation into competent XL1-Blue bacteria.

**Yeast Cell Culture.** For yeast display, *S. cerevisiae* strain EBY100 cultured in yeast extract peptone dextrose (YPD) complete medium was transformed with the yeast-display plasmid pCTCON2 (25) using the Frozen E-Z Yeast Transformation II kit (Zymo Research) according to manufacturer protocols. Transformed cells were selected on synthetic dextrose plus casein amino acid (SDCAA) plates and propagated in SDCAA medium at  $30^\circ\text{C}$ . Protein expression was induced by inoculating saturated yeast culture into SD/RCAA (SDCAA medium with 90% of dextrose replaced with raffinose) and incubated at room temperature for 4 h; then, 2% of galactose was added to the culture and the incubation was continued at room temperature with shaking overnight.

**Generation of rHUH Libraries for Yeast Display.** Libraries of rHUH mutants were generated by error-prone PCR according to published protocols (26). 100 ng of the template in vector pCTCON2 was amplified for 20 cycles with 0.4  $\mu\text{M}$  forward and reverse primers:

F: 5'-CTAGTGGTGGAGGAGGCTCTGTGGAGGCGGTAGCGAGGCGGAGGGTCGGTAGC-3'

R: 5'-TATCAGATCTCGAGCTATTACAAGTCCTCTCAGAAATAAGCTTTTGTCGGATCC-3' and 1x ThermoPol Reaction Buffer (NEB), 10 units of Taq DNA

polymerase (NEB), 5, 10, or 20  $\mu\text{M}$  of 8-oxo-2'-deoxyguanosin-5'-triphosphate (8-oxo-dGTP), 1, 2, or 4  $\mu\text{M}$  of 2'-deoxy-P-nucleoside-5'-triphosphate (dNTP) for the low, medium, and high mutagenesis libraries, respectively, 200  $\mu\text{M}$  of dNTP in a 100  $\mu\text{l}$  total volume. 300 ng of the gel-purified PCR products were reamplified for another 30 cycles under normal PCR conditions using following the primer pair:

F: 5'-CAAGGCTCGCAGGCTAGTGTGTGAGGAGGCTCTGTGTG-3'

R: 5'-CTACACTGTTGTATCAGATCTCGAGCTATTACAAGTC-3'

The amplified PCR products were gel purified and combined with BamHI–NheI linearized pCTCON2 vector (4  $\mu\text{g}$  insert/1  $\mu\text{g}$  vector) and electroporated into electrocompetent *S. cerevisiae* EBY100 (26). The electroporated cultures were rescued in 2 mL YPD medium for 1 h at  $30^\circ\text{C}$  with no shaking. The rescued cell suspension was transferred to 100 mL of SDCAA medium supplemented with 50 units/mL penicillin and 50  $\mu\text{g}/\text{mL}$  streptomycin and grown for 2 d at  $30^\circ\text{C}$ . The library size was estimated by 100- to 10,000-fold diluting a small fraction of the rescued cell and counting number of the colonies grown by the diluted fraction on SDCAA plates. The low, medium, and high mutagenesis libraries were each  $\sim 4 \times 10^7$  in size which made the combined library  $\sim 1.2 \times 10^8$  in size. The grown libraries were analyzed by sequencing and resulted in 1, 1.6, and 2.3 amino acids mutated per gene, for the low, medium, and high mutagenesis libraries, respectively.

**Biotinylated dRS and rRS Probes.** All biotinylated probes were ordered from IDT with 3' Biotin-TEG modifications. All probes were received as dried pellets and dissolved to 1 mM stocks with  $\text{H}_2\text{O}$  upon arrival. A fraction of 1 mM stocks was further diluted with  $\text{H}_2\text{O}$  into 100  $\mu\text{M}$  secondary stocks. The secondary stocks were aliquoted and frozen under  $-80^\circ\text{C}$ . The 1 mM stocks were frozen under  $-80^\circ\text{C}$  in the original tube.

**General Method for Labeling Yeast-Displayed rHUH Variants.** All reagents used in rHUH labeling are in RNase-free format or as clean as possible. Yeast cells induced overnight were washed twice with PBST buffer (PBS, pH 7.4, supplemented with 0.2% Tween-20) and resuspended in PBST supplemented with 10  $\mu\text{M}$   $\text{MnCl}_2$  or 5 mM  $\text{MgCl}_2$  (or indicated otherwise) and RNase inhibitor cocktail (RiboLock, Invitrogen). The indicated amount of probe was added to the cells and incubated at room temperature with gentle shaking. After labeling, the yeast cells were washed once with PBST supplemented with 5 mM EDTA, followed by incubation with anti-myc chicken IgY antibody (Exalpha Biological, 1:400) and then with streptavidin-PE (Jackson ImmunoResearch, 1:400) and Goat anti-chicken IgY antibody Alexa Fluor 488 (AF488) conjugate (Invitrogen, 1:400). The labeled yeast cells were analyzed by the ZE5 FACS cell analyzer (Bio-Rad).

**General Methods for Yeast Display-Based Directed Evolution.** For the first round of selection for each generation, the electroporated yeast libraries were grown for 2 d to reach saturation. Then,  $\sim 1.5 \times 10^9$  of yeast cells from the saturated culture were inoculated into 165 mL of SD/RCAA medium and incubated at room temperature for 4 h; then, 2% of galactose was added to the culture and the incubation was continued at room temperature with shaking overnight.  $\sim 6 \times 10^8$  of yeast cells from the overnight induced culture were labeled following the method described in section “General methods for yeast displayed rHUH labeling”. The labeled yeast cells were resuspended in 10 mL of ice-cold PBS and sorted on a BD FACS Aria II cell sorter (BD Biosciences) with appropriate lasers and emission



filters for PE and AF488. The yeast cells were processed at a rate of ~20,000 cells per second and the top 0.5 to 1% of cells with highest PE and AF488 signals were collected in SDCAA medium containing 1% penicillin-streptomycin and incubated at 30 °C for 3 d. Limited by the sorting speed and FACS sorter availability,  $\sim 3 \times 10^8$  of yeast cells were processed by the sorter, and  $\sim 2 \times 10^6$  of yeast cells were collected.

For the second round,  $\sim 2 \times 10^7$  of yeast cells were labeled and sorted following similar procedures. For the third and later rounds,  $\sim 4 \times 10^6$  of yeast cells were used. For positive selections, the top 0.5 to 1% of cells with highest PE and AF488 signals were collected. For negative selections, the top 1% of cells with highest AF488 signal and lowest PE signal were collected.

After the last round of selection, the collected yeast cells were grown to saturation, and 1 mL of saturated culture was removed for DNA extraction using the ZymoPrep yeast Plasmid Miniprep II (Zymo Research) kit according to manufacturer protocols (using 10  $\mu$ L zymolyase). The extracted DNA was transformed into competent XL1-Blue bacteria. After the emergence of colonies on the bacteria culture plate, 12 to 24 colonies were picked up and subjected to standard plasmid amplification protocols. The amplified plasmids were analyzed by Sanger sequencing.

**Mammalian Cell Culture and Transfection.** HEK293T cells from ATCC (passage number <25) were cultured as a monolayer in growth media (DMEM, high glucose, Gibco) supplemented with 10% fetal bovine serum (Avantor) at 37 °C under 5% CO<sub>2</sub>. For transient expression, cells were typically transfected at approximately 80% confluency using 1 mg/mL PEI max solution (polyethylenimine HCl max pH 7.3). For inducing expression or rHUH, 100  $\mu$ g/mL of doxycycline was added to the medium.

**Bacterial Expression and Purification of rHUH Protein.** For bacterial expression and purification of rHUH protein, we cloned the rHUH sequence into the pTD68 vector (Addgene 123643) with all N-terminal tags removed and a C-terminal myc tag added. We found that rHUH from G5 onward could bind to Ni-NTA without a His6 tag. Competent BL21 *E. coli* was transformed with the rHUH expression plasmid by heat shock. Cells were then grown in LB medium containing 100 mg/L carbenicillin at 37 °C and 220 r.p.m. until an OD<sub>600</sub> of ~0.6 to 0.8 was reached. Protein expression was induced with 0.05 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG); then, the culture was shifted from 37 °C to room temperature during the induction period. After overnight growth, the bacteria were pelleted by centrifugation at 7,000  $\times g$  for 3 min at room temperature, the supernatant was discarded, and the pellet was stored at -80 °C.

The frozen pellet was thawed at room temperature and transferred to ice immediately after thawing. 10 mL of lysis buffer (B-PER Bacterial Protein Extraction Reagent, Thermo Scientific, supplemented with 1% protease inhibitor cocktail and 1 mM PMSF) per 500 mL of culture was added to the pellet. The pellet was resuspended and kept on ice for at least 10 min. Then, the lysate was sonicated using a Misonix sonicator (1-s on, 1-s off, for a total of 60-s on, ice chill for 2 min. Repeat 4 times). The sonicated lysate was clarified by centrifugation for 10 min at 18,000  $\times g$  at 4 °C, and the supernatant was clarified by the same method again, and then, the supernatant was carefully transferred to a 10-mL conical with 2 mL Ni-NTA agarose bead slurry (Invitrogen) per 10 mL lysate and incubated at 4 °C for 30 min with gentle rotation. Then, the slurry was placed in a gravity column and washed twice with 10 CV of washing buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 30 mM imidazole). The protein was eluted with 5 CV of elution buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 300 mM imidazole). The purity was analyzed by Coomassie Blue staining.

The eluted protein was further purified by using a HiLoad 16/600 Superdex 75 pg column (Cytiva) with FPLC buffer (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM EDTA, and 1 mM DTT). The collected fractions were concentrated using Amicon Ultra-15 Centrifugal Filter Units, 3,000-kDa cutoff.

**EMSA Assay.** Purified rHUH(Y96F) protein was diluted into binding buffer (50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM MnCl<sub>2</sub>, and 1 U/ $\mu$ L Ribolock RNase inhibitor cocktail) to final concentrations ranging from 50 nM to 5000 nM. The diluted proteins were incubated with 10 nM rRS-F488 (UAGUUAUACCAGA-Alexa

Fluor 488) for 2 h at room temperature and mixed with 1/6 volume of 50% glycerol and resolved on 6% PAGE TEB gels (Invitrogen) at room temperature (using 0.5x TBE buffer, 100 V running for 40 min). Gels were imaged using a Typhoon imager. The unbound rRS-F488 intensity was quantified using ImageJ (Wayne Rasband, NIH). Line regression was performed in GraphPad Prism using nonlinear fit with one-site binding hyperbola.  $K_d$  values are calculated by GraphPad Prism.

**IVT and labeling of IVT probes on yeast.** The DNA template for IVT is a dsDNA containing a T7 promoter (TAATACGACTCACTATAGGG) and corresponding DNA sequence of the listed probe. IVT was performed with the MEGascript T7 Transcription Kit (Thermo Fisher) following the manufacturer's protocol and the RNAs were purified and concentrated using RNA Clean & Concentrator-5 kit (Zymo Research). Yeast cells induced overnight were washed twice with PBST buffer (PBS, pH 7.4, supplemented with 0.2% Tween-20) and resuspended in PBST supplemented with 10  $\mu$ M MnCl<sub>2</sub> and RNase inhibitor cocktail (RiboLock, Invitrogen). The yeast cells were incubated with 10 nM of IVT probe for 1 h at room temperature and the reaction was quenched by changing the buffer to PBST supplemented with 5 mM EDTA. Then, the yeast cells were incubated with high salt PBS (PBS supplemented with 300 mM NaCl) containing 100 nM antisense biotin-DNA oligo (AAGTATTACCAGAACTTAACATCATGGTGT-biotin) for 1 h, followed by incubation with anti-myc chicken IgY antibody (Exalpha Biological, 1:400) and then with streptavidin-PE (Jackson ImmunoResearch, 1:400) and Goat anti-chicken IgY antibody AlexaFluor 488 (AF488) conjugate (Invitrogen, 1:400). The labeled yeast cells were analyzed by the ZE5 FACS cell analyzer (Bio-Rad).

**In Vitro Labeling of Purified rHUH or rHUH-Containing Cell Lysate and SDS-PAGE Analysis.** Purified rHUH proteins were diluted in 50 mM Tris-HCl, pH 7.5, 50 mM NaCl, and 1 mM MnCl<sub>2</sub> (or otherwise indicated). DNA/RNA probes were added and mixed by a brief vortex. One-fifth volume of 6  $\times$  SDS Protein Loading buffer supplemented with 5 mM EDTA was added to quench the reaction. And the reaction was further quenched by heat denaturation at 95 °C for 5 min. For HEK293T cells expressing rHUH, the cells were lysed with Nonidet P40 (NP40) lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 1% NP40) on ice for 10 min. The lysate was clarified by centrifugation for 10 min at 20,000  $\times g$  at 4 °C. The supernatant was transferred to a new tube and mixed with 1 mM MnCl<sub>2</sub> and DNA/RNA probes. After the reaction, one fifth volume of 6  $\times$  SDS Protein Loading buffer supplemented with 5 mM EDTA was added to quench the reaction. And the reaction was further quenched by heat denaturation at 95 °C for 5 min. The denatured mixtures were resolved by a 15% Tris-Glycine gel with 10% SDS and Coomassie stained (Abcam, InstantBlue, ab119211). The stained gels were imaged by a LI-COR Odyssey CLx gel imager.

**Western Blot Analysis.** The denatured protein samples were resolved by a 15% Tris-Glycine gel with 10% SDS and then transferred to a nitrocellulose membrane. The membrane was briefly washed by TBST (100 mM Tris-HCl, pH 7.5, NaCl 150 mM, and 0.05% Tween-20) followed by incubation with 5% nonfat milk (Lab Scientific bioKEMIX) in TBST for 1 h. Then, the membrane was incubated with anti-myc chicken IgY antibody (Exalpha Biological, 1:5,000, in TBST) at 4 °C overnight. After three washes with TBST, the membrane was incubated with IRDye 680RD Donkey anti-chicken secondary antibody (LI-COR Biosciences, 1:5,000, in TBST, protected from light) for 1 h. Then, the membrane was washed three times with TBST and imaged by a LI-COR Odyssey CLx gel imager.

**Data, Materials, and Software Availability.** All study data are included in the article and/or [SI Appendix](#).

**ACKNOWLEDGMENTS.** We thank David A. Bushnell for advice on dHUH structure and members of the A.Y.T. laboratory for technical support and advice. This work was supported by the NSF Molecular Foundations for Biotechnology grant number 2330686, the Chan Zuckerberg Biohub – San Francisco, the Stanford Bio-X, and the Bachrach Foundation.

1. Y. Zhu, L. Zhu, X. Wang, H. Jin, RNA-based therapeutics: An overview and prospectus. *Cell Death Dis.* **13**, 644 (2022).
2. M. Metelev *et al.*, Direct measurements of mRNA translation kinetics in living cells. *Nat. Commun.* **13**, 1852 (2022).
3. F. M. Fazal *et al.*, Atlas of subcellular RNA localization revealed by APEX-Seq. *Cell* **178**, 473–490. e426 (2019).

4. W. Qin, K. F. Cho, P. E. Cavanagh, A. Y. Ting, Deciphering molecular interactions by proximity labeling. *Nat. Methods* **18**, 133–143 (2021).
5. E. Tutucci *et al.*, An improved MS2 system for accurate reporting of the mRNA life cycle. *Nat. Methods* **15**, 81–89 (2018).
6. B. Wu, C. Eliscovich, Y. J. Yoon, R. H. Singer, Translation dynamics of single mRNAs in live cells and neurons. *Science* **352**, 1430–1435 (2016).

7. A. Khong, T. Matheny, T. N. Huynh, V. Babl, R. Parker, Limited effects of m(6)A modification on mRNA partitioning into stress granules. *Nat. Commun.* **13**, 3735 (2022).
8. J. M. Halstead *et al.*, An RNA biosensor for imaging the first round of translation from single cells to living animals. *Science* **347**, 1367–1371 (2015).
9. W. Li, A. Maekiniemi, H. Sato, C. Osman, R. H. Singer, An improved imaging system that corrects MS2-induced RNA destabilization. *Nat. Methods* **19**, 1558–1562 (2022).
10. D. M. Shechner, E. Hacisuleyman, S. T. Younger, J. L. Rinn, Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. *Nat. Methods* **12**, 664–670 (2015).
11. S. Han *et al.*, RNA-protein interaction mapping via MS2- or Cas13-based APEX targeting. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 22068–22079 (2020).
12. J. Biswas, R. Rahman, V. Gupta, M. Rosbash, R. H. Singer, MS2-TRIBES evaluates both protein-RNA interactions and nuclear organization of transcription by RNA editing. *iScience* **23**, 101318 (2020).
13. K. N. Lovendahl, A. N. Hayward, W. R. Gordon, Sequence-directed covalent protein-DNA linkages in a single step using HUH-Tags. *J. Am. Chem. Soc.* **139**, 7030–7035 (2017).
14. W. Han *et al.*, Efficient precise integration of large DNA sequences with 3'-overhang dsDNA donors using CRISPR/Cas9. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2221127120 (2023).
15. W. Guo, Y. Mashimo, E. Kobatake, M. Mie, Construction of DNA-displaying nanoparticles by enzymatic conjugation of DNA and elastin-like polypeptides using a replication initiation protein. *Nanotechnology* **31**, 255102 (2020).
16. K. J. Tompkins *et al.*, Molecular underpinnings of ssDNA specificity by Rep HUH-endonucleases and implications for HUH-tag multiplexing and engineering. *Nucleic Acids Res.* **49**, 1046–1064 (2021).
17. A. T. Smiley *et al.*, Sequence-directed covalent protein-RNA linkages in a single step using engineered HUH-Tags. bioRxiv [Preprint] (2024). <https://doi.org/10.1101/2024.08.13.607811> (Accessed 13 August 2024).
18. A. S. Moomaw, M. E. Maguire, The unique nature of mg2+ channels *Physiology (Bethesda)* **23**, 275–285 (2008).
19. H. Zhou, D. E. Clapham, Mammalian MagT1 and TUSC3 are required for cellular magnesium uptake and vertebrate embryonic development. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 15750–15755 (2009).
20. K. Klabenkova, A. Fokina, D. Stetsenko, Chemistry of peptide-oligonucleotide conjugates: A review. *Molecules* **26**, 5420 (2021).
21. E. M. Tota, N. K. Devaraj, RNA-TAG mediated protein-RNA conjugation. *ChemBiochem* **24**, e202300454 (2023).
22. T. S. Smith, M. A. Zoltek, M. D. Simon, Reengineering a tRNA methyltransferase to covalently capture new RNA substrates. *J. Am. Chem. Soc.* **141**, 17460–17465 (2019).
23. A. F. Palazzo, E. S. Lee, Non-coding RNA: What is functional and what is junk? *Front. Genet.* **6**, 2 (2015).
24. S. Kahali *et al.*, A water-soluble, cell-permeable Mn(II) sensor enables visualization of manganese dynamics in live mammalian cells. *Chem. Sci.* **15**, 10753–10769 (2024).
25. G. Chao *et al.*, Isolating and engineering human antibodies using yeast surface display. *Nat. Protoc.* **1**, 755–768 (2006).
26. D. W. Colby *et al.*, Engineering antibody affinity by yeast surface display. *Methods Enzymol.* **388**, 348–356 (2004).