Tetramolecular Fluorescence Complementation for Detection of Specific RNAs in Vitro

Stefanie Julia Kellermann,[a] Anna Katharina Rath,[a] and Andrea Rentmeister*[a, b]

Asymmetrically localized RNA has been found in many cell types and organisms and contributes to important developmental processes, such as mating type switching in yeast,[1] body axis formation in Drosophila,[2,3] and directional steering in axons in the mammalian brain.[4] Different mechanisms, including localized transcription, directed degradation, diffusion and anchoring, or active transport, can lead to this asymmetric localization.[4] To better understand the processes leading to asymmetric localization, probes capable of RNA imaging in the complex cellular environment are required. Importantly, these probes should yield a different signal between their bound and unbound forms, or the signal should be substantially enhanced upon binding to the target RNA.

Several strategies to discriminate between bound and unbound probe have been developed to date. In the most straightforward approach, the fluorescence is increased locally by attaching multiple fluorophores to the RNA. This strategy has been used successfully in living cells, but the RNA of interest had to be extended by up to hundreds of nucleotides.[1, 5] This appendage might impair binding or trafficking behavior of the original RNA. In another approach, the FRET effect has been harnessed to provide a distinct signal when two nucleic acid based probes labeled with suitable fluorophores are bound adjacent to each other.[6] Molecular beacons (stem–loop structures modified by a fluorophore and a quencher at the 5'- and 3'-ends that light up upon binding to target RNA) and forced intercalation probes (peptide nucleic acid based probes in which an intercalator dye serves as a base surrogate) represent other widely used approaches.[7] Labeled nucleic acids, however, cannot be produced inside cells, and delivery of nucleic acid-based probes into cells remains a challenge.[7]

Therefore, the sequence specificity and light-up properties realized in molecular beacons combined with the possibility to genetically encode fluorescence represents a valuable alternative. To this end, RNA binding proteins that recognize specific RNA sequences or structures have been fused to split fluorescent reporter proteins. Several proteins, such as fragile X mental retardation protein, the eukaryotic initiation factor 4A, and Pumilio, have been successfully implemented for RNA detection.[6, 11] Pumilio is an RNA binding protein that binds to a stretch of eight nucleotides of single-stranded RNA (ssRNA) in a sequence-specific manner.[9] Crystal structure analysis has revealed the molecular details of the RNA–protein interaction.[9] Therefore, it has become possible to rationally design Pumilio variants with altered specificity.[10] The combination of Pumilio proteins with split green fluorescent protein (GFP) variants was realized by Ozawa[11] and Tilsner[12] and applied to HeLa and epidermis cells. By bridging two RNA binding proteins fused to split-GFP, a trimolecular fluorescence complementation system was established.[8]

Fluorescence complementation of GFP or Venus has been successfully used to probe protein–protein interactions (bimolecular fluorescence complementation) and, to a lesser extent, RNA–protein interactions (trimolecular fluorescence complementation),[11–13] but still suffers from some limitations. Firstly, self-assembly of the two parts of the fluorescent protein has been reported, and this causes significant background signal.[14] This is particularly aggravating because it increases the background fluorescence and thus compromises discrimination between bound and unbound probe. Mutations that reduce this background improved the signal-to-noise ratio.[14] Secondly, split-GFP requires the fusion of both GFP halves to the RNA binding proteins. These GFP halves are very large tags and can cause insolubility or alter the behavior of the resulting fusion proteins. To address the issue of spontaneous protein complementation, Waldo et al. recently engineered an exceptionally well-folded variant of GFP (“superfolder GFP”), which allowed the creation of tripartite split-GFP.[15] In this system, short fragments were removed from GFP, yet still resulting in a GFP detector (GFP1-9) that can be expressed. The short fragments can be fused to proteins of interest to investigate their interactions.[15, 16] Only if two tagged proteins are brought into proximity can the three-body system be complemented by the detector and fluoresce. These three protein fragments show less self-association and less background in the absence of interaction partners. Importantly, the tags that are fused to these proteins are very short and are therefore less likely to affect the protein properties than GFP halves.

Here we report the engineering of a tetramolecular fluorescence complementation (TetFC) system that allows sequence-specific RNA detection based on protein complementation. Our system is based on two variants of the Pumilio homology domain of Homo sapiens Pumilio, a three-way split-GFP, and the RNA of interest (Figure 1). We used the wild type Pumilio 1 homology domain from Homo sapiens (Pum-WT), which binds to the nanos response element (box A), and a previously described variant ("Pum-Var1" or "MUT6-2/7-2") that binds to the

University of Hamburg, Department of Chemistry
Institute of Biochemistry and Molecular Biology
Martin Luther King Platz 6, 20146 Hamburg (Germany)
E-mail: rentmeister@uni-hamburg.de

[2] Dr. A. Rentmeister
The Hamburg Centre for Ultrafast Imaging
Luruper Chaussee 149, 22761 Hamburg (Germany)

These authors contributed equally to this work.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201200734.
For protein complementation, we designed RNA sequences containing box A and box B in tandem (box AB), connected by a five-nucleotide linker (Figure 1). Furthermore, we constructed controls to test the specificity of the complementation system: box AC (box A and nonbinding box C) should allow binding of S10-Pum-WT but not of Pum-Var1-S11. It thus provides a stringent test for unspecific complementation. Mixing all four components of our tetramolecular system (S10-Pum-WT, Pum-Var1-S11, GFP1-9, RNA) resulted in a fluorescence signal, whereas without RNA or protein components there was no signal (Figure 3). The fluorescence signal started to develop rapidly and could be used to reliably discriminate between target RNA and controls after approximately 10 min (Figure 3B). The fluorescence of reconstituted GFP was stable and could be analyzed by PAGE (Figure 3B). Importantly, replacing box AB with control box AC did not result in a detectable signal in the polyacrylamide (PAA) gel, thus suggesting that binding of both S10-Pum-WT and Pum-Var1-S11 is required for signal emission.

We next analyzed time and concentration dependency of the fluorescence signal. Using equimolar amounts of Pumilio fusion proteins and RNA (as well

![Figure 1](https://www.chembiochem.org/content/16/20/15463/fig/1)

**Figure 1.** Tetramolecular fluorescence complementation (TetFC) for sequence-specific RNA detection. A) Binding of two Pumilio proteins to the target RNA at designated sites forms a complex that can recruit GFP1-9, thereby leading to fluorescence. S10-Pum-WT (red) binds specifically to box A, and Pum-Var1-S11 (black) binds to box B. S10 and S11 are strands derived from GFP. In TetFC, binding of GFP1-9 can only occur when tags S10 and S11 are positioned at an appropriate distance. B) Modules of the RNA constructs used in this study. C) Possible scenarios of Pumilio binding to RNA constructs and subsequent protein complementation. Semicircles represent Pumilio constructs. Reconstituted GFP is shown as a green barrel. Lines represent S10 and S11 tags.

![Figure 2](https://www.chembiochem.org/content/16/20/15463/fig/2)

**Figure 2.** Binding of tagged Pumilio proteins to different RNA constructs. A) Representative gel-shift experiments of indicated RNA constructs with S10-Pum-WT and Pum-Var1-S11 proteins. B) Evaluation of the box A and box B gel-shift experiments shown in A). Disassociation constants for box A with S10-Pum-WT (4 nM, black trace) and Pum-Var1-S11 (253 nM, black), as well as box B with S10-Pum-WT (74 nM, red) and Pum-Var1-S11 (18 nM, red) were determined from at least two independent experiments. Arrows indicate RNA-protein complexes. RNA concentrations are based on cpm; details can be found in the Experimental Section.

It should be mentioned that lower $K_d$ values for the untagged constructs have been reported in the literature. In these calculations, the protein concentrations were based on the percentage of functional protein in each preparation, which could account for the difference. From our data we conclude that the attachment of tags can influence the binding of Pumilio to ssRNA constructs. The attachment of the short tags used here, however, reduced the affinity only slightly and did not affect the specificities.

![Image](https://www.chembiochem.org/content/16/20/15463/fig/3)

**Figure 3.** Binding of tagged Pumilio proteins to RNA constructs and subsequent protein complementation. A) Time-dependent fluorescence signal of tetramolecular system (S10-Pum-WT, Pum-Var1-S11, GFP1-9), we found that the fluorescence intensity was indeed concentration dependent. The signal was detectable as early as 10 min after GFP1-9 addition (250 nM Pum Protein and RNA in Figure 3B) then increased linearly and leveled off after 8–10 h, before finally decreasing. The signal was higher and developed faster when higher concentrations of GFP1-9 were used (data not shown). We were able to detect RNA down to 16 nM under the conditions used (Figure 3B, right).
We anticipated that the length of the linker between the RNA motifs might affect the efficiency of complementation. To find the optimal distance we fused box A and box B with linkers of three, five, or seven nucleotides. In fluorescence complementation experiments, the five-nucleotide linker showed the highest fluorescence intensity. Boxes linked by three or seven nucleotides showed only 17 or 31% relative fluorescence, respectively (Figure 3C). This suggests that the distance between the S10 and S11 tags significantly influences either Pumilio binding or the efficiency of protein complementation. Consequently, we used the five-nucleotide linker for all other constructs, including controls.

A major limitation of current protein complementation approaches is background signal resulting from formation of functional fluorescence protein, even though the tested proteins do not interact. We performed quantitative leave-one-out control measurements for each of the four components. In all cases, no signal could be detected if one of the three proteins was absent (Figure 3C). This suggests that the distance between the S10 and S11 tags significantly influences either Pumilio binding or the efficiency of protein complementation. Consequently, we used the five-nucleotide linker for all other constructs, including controls.

A major limitation of current protein complementation approaches is background signal resulting from formation of functional fluorescence protein, even though the tested proteins do not interact. We performed quantitative leave-one-out control measurements for each of the four components. In all cases, no signal could be detected if one of the three proteins was absent (Figure 3C). This suggests that the distance between the S10 and S11 tags significantly influences either Pumilio binding or the efficiency of protein complementation. Consequently, we used the five-nucleotide linker for all other constructs, including controls.

To fully benefit from the nucleic acid recognition properties of two Pumilio proteins, it is imperative to minimize background in the presence of just one of the motifs. We therefore tested the specificity of the RNA recognition by using the constructs listed in Figure 1C. None of the alternative RNA constructs (box AC, box AA, box BA, box BB) showed significant fluorescence (Figure 3D). The low background for box AC, box AA, and box BB suggests that binding of just one of the Pumilio proteins (e.g., S10-Pum-WT to box AC or to box AA) is not sufficient to yield fluorescence complementation. The conventional split-GFP, however, would be expected to cause fluorescence background in these scenarios because of unspecific assembly of split-GFP halves. From our results, we concluded that TetFC has advantages in specificity compared with conventional split-GFP.

Furthermore, for box AA, fluorescence complementation might arise from unspecific binding of Pum-Var1-S11 to box A, as observed to a low extent in EMSA (Figure 2). Nevertheless, the background for box AA was only ~5%, thus indicating that the Pum-Var1-S11 binds with high specificity to the target RNA motif. Importantly, controls box AA and box BB differ from box AB at only two nucleotides, thus highlighting the sequence specificity of this protein-based RNA detection system.

Background signals might also develop when two Pumilio proteins bind to motifs located on different RNA molecules. In TetFC however, addition of discrete box A and box B RNAs only caused a background signal in the order of 2–3% (Figure 3D).

To test if the direction of binding was important, we switched the order of the box A and box B motifs. In this sce-
nario, the Pumilio proteins were expected to bind to the RNA, but with the S10 and S11 tags oriented to the outer ends of the RNA motifs (Figure 1C). For box BA, we observed up to 10% fluorescence intensity—more than for any other control. We cannot rule out that the RNA bends, thereby bringing the S10 and S11 tags into closer proximity than anticipated from the linear drawing. Still, the signal from box BA could be well distinguished from the signal from box AB sequence, even though these constructs differ at only two positions.

To be useful for RNA detection in vivo, TetFC has to work in complex environments, including the presence of high amounts of other RNA species or cellular components. We therefore tested whether fluorescence complementation occurs in the context of total RNA isolated from Escherichia coli cells. We could indeed specifically detect box AB RNA down to 16 nM in 10 μg of RNA isolated from E. coli cells (Figure 4). Control RNA box AC, however, did not yield a significant signal under these conditions (Figure 4). Importantly, TetFC worked in E. coli cell lysate, and showed much higher fluorescence with box AB than with box AC (Figure 4C). These results suggest that the GFP1-9 is available (not trapped by other biomolecules or assemblies).

In conclusion, our study demonstrates that tetramolecular fluorescence complementation is a powerful means to detect ssRNA with sequence specificity. In the presence of target RNA, the two Pumilio proteins bind to their target RNA with high affinity and become immobilized. The formation of this protein–RNA complex reduces the tetramolecular system—which has little likelihood of spontaneous reassembly—to a regular two-body interaction. The detector GFP1-9 can then bind, and reconstituted GFP starts to develop fluorescence immediately and can be used to discriminate between target RNA and controls after approximately 10 min.

In the absence of target RNA, the three proteins cannot assemble spontaneously and thus background fluorescence is only 1.4% (70-fold increase in signal upon target binding). If one of the three protein components is omitted, the background signal is negligible (0.01–0.02%). This rise in fluorescence lies in the same range as molecular beacons, for which a 25-fold increase has been reported, and a 200-fold increase should be possible.[7a]

TetFC allows discrimination between closely related RNAs, a valuable feature, which can be used to detect single nucleotide polymorphisms (SNPs) or distinguish between microRNAs.[7b] We showed that within a strand of 16 nucleotides, two mutations can be detected specifically. As the binding properties of Pumilio proteins can be altered by rational design[10] and directed evolution,[11] we anticipate that our system can be applied to monitor a wide range of target RNAs.

Importantly, TetFC makes full use of the 16-nucleotide recognition sequence accessible by the combination of two Pumilio proteins. Even when a precomplex between RNA and one of the Pumilio proteins forms, no significant unspecific complementation is observed. This improvement in specificity should prove valuable for intracellular applications. Already, the complementation has worked reliably in a complex background of cellular RNA as well as in lysate from E. coli, thus suggesting that detection in vivo should also be possible. In contrast to molecular beacons, all components of the TetFC system can be produced by the cellular machinery. Therefore, TetFC has potential to assess levels of endogenous RNAs in living cells.

**Experimental Section**

Full experimental details are provided in the Supporting Information.

**Acknowledgements**

A.R. gratefully acknowledges funding by the Emmy Noether programme of the DFG and by the Fonds der Chemischen Industrie. This work has been supported by the excellence cluster “The Hamburg Centre for Ultrafast Imaging—Structure, Dynamics and Control of Matter at the Atomic Scale” of the DFG. A.K.R. thanks Pro Exzellenzia for a doctoral scholarship. We thank Prof. Hall for plasmids for Pum-WT (hSPUM1-HD) and Pum-Var1 (MUT6-2/7-2). We would like to thank Prof. Waldo (Los Alamos National Laboratory) for providing us with the plasmid for GFP1-9. We thank Marina Mutas and Julia Sandberg-Meinhardt for excellent technical assistance.
Tetramolecular Fluorescence Complementation for Detection of Specific RNAs in Vitro

Short fuses: RNA can be detected sequence specifically by using two RNA binding proteins fused to short strands derived from GFP and an additional large GFP fragment. The tetramolecular system has a high signal-to-noise ratio, discriminates between closely related RNA, and works in RNA preparations as well as E. coli cell lysate.