## Title:

Pseudouridine and N-6 methyladenosine modifications weaken PUF protein/RNA interactions

## Authors:

Pavanapuresan P. Vaidyanathan[1], Ishraq AlSadhan[1], Dawn K. Merriman[2], Hashim Al-Hashimi[2,3], Daniel Herschlag[1,4,5+]

[1]Department of Biochemistry, Stanford University, Stanford, CA 94305, USA

[2]Department of Chemistry, Duke University, Durham, NC 27708, USA.

[3]Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA.

[4]Departments of Chemical Engineering and Chemistry, Stanford University, Stanford, CA 94305, USA

[5]Stanford ChEM-H (Chemistry, Engineering, and Medicine for Human Health), Stanford University, Stanford, CA 94305, USA

[+]Corresponding author: herschla@stanford.edu

## Running Title:

RNA modifications weaken Pumilio binding to RNA

## Keywords:

epitranscriptomics, pseudouridine, N6-methyladenosine, PUMILIO, RNA binding proteins, RNA-protein interactions.

## Abstract

RNA modifications are ubiquitous in biology, with over 100 distinct modifications. While the vast majority were identified and characterized on abundant non-coding RNA such as tRNA and rRNA, the advent of sensitive sequencing-based approaches has led to the discovery of extensive and regulated modification of eukaryotic messenger RNAs as well. The two most abundant mRNA modifications – pseudouridine (Ψ) and N-6 methyladenosine ($m^6A$)– affect diverse cellular processes including mRNA splicing, localization, translation, and decay and modulate RNA structure. Here, we test the hypothesis that RNA modifications directly affect interactions between RNA-binding proteins and target RNA. We show that Ψ and $m^6A$ weaken the binding of the human single-stranded RNA binding protein Pumilio 2 (hPUM2) to its consensus motif, with individual modifications having effects up to ~3 fold and multiple modifications giving larger effects. While there are likely to be some cases where RNA modifications essentially fully ablate protein binding, here we see modest responses that may be more common. Such modest effects could nevertheless profoundly alter the complex landscape of RNA:protein interactions, and the quantitative rather than qualitative nature of these effects underscores the need for quantitative, systems-level accounting of RNA:protein interactions to understand post-transcriptional regulation.

## Introduction

Over a hundred distinct chemical modifications of RNA have been described (Cantara et al 2011; Gilbert et al 2016; Machnicka et al 2013; Schwartz 2016; Karijolich et al 2015). While a majority of these modifications were identified from chemical analysis of abundant non-coding RNA (snoRNA, tRNA and rRNA), novel and sensitive sequencing-based approaches have revealed the widespread prevalence of modified bases such as pseudouridine (ψ, Fig. 1A), N-6 methyladenosine ($m^6A$, Fig 1B), N1-methyladenosine ($m^1A$) and 5-methylcytosine ($m^5C$) in eukaryotic messenger RNAs, collectively termed the epitranscriptome (Carlile et al 2014; Schwartz et al 2014; Lovejoy et al 2014; Meyer and Jaffrey 2014; Liu and Pan 2016 and references therein; Li et al 2016b; Dominissini et al 2016; Khoddami and Cairns 2013; Squires et al 2012).

ψ and $m^6A$ are the most abundant mRNA modifications known, with 0.2 – 0.5% of uridine and ~0.5% of adenosine residues converted to ψ and $m^6A$, respectively (Dominissini et al 2016). Ψs in mRNA are made by ψ synthases that use a guide RNA dependent (Cbf5, Dyskerin) or independent (PUS family proteins) mechanism to recognize and target specific uridines for modification in a manner dependent on nutrient availability and environmental stress conditions (Gilbert et al 2016; Charette and Gray 2000). $m^6A$ modifications are catalyzed by 'writer' methyltransferases (METTL3, METTL14) that target the adenosine within an RRACH (R = G, A; H = A, C, U) consensus motif for modification. Unlike pseudouridines, $m^6A$ modifications are reversible and can be removed by 'eraser' demethylase proteins in a regulated manner (FTO, ALKBH5) (Meyer and Jaffrey 2014; Liu and Pan 2016; Zhao et al 2016).

The abundance and broad distribution of mRNA modifications along with the high degree of conservation of the modifying enzymes suggest that these modifications may have important functional roles in gene expression (Dominissini et al 2016; Hoernes et al 2016; Machnicka et al 2013). Indeed,

pseudouridine (ψ) has been found to cluster in functionally important domains in transfer RNA (tRNA), ribosomal RNA (rRNA) and U2 small nuclear RNA (snoRNA), and shown to affect protein translation and pre-mRNA splicing (Karijolich et al 2015; Charette and Gray 2000; Ge and Yu 2013). ψ was also observed to increase nonsense suppression of mRNA stop codons resulting in altered protein products (Karijolich and Yu 2011). Biochemical and genetic studies have revealed regulatory roles for $m^6A$ modifications in a wide variety of cellular processes including, but not limited to, pre-mRNA splicing, and mRNA transport, translation, and decay (Meyer and Jaffrey 2014; Liu and Pan 2016). $m^6A$ has also been implicated in germline development and fertility, maintenance of the cell cycle and circadian clocks, and in a wide variety of human diseases including cancer (Meyer and Jaffrey 2014; Liu and Pan 2016; Liu and Pan 2015). Despite these advances, we have a limited understanding of how these modifications exert their functions.

RNA modifications have been proposed to regulate gene expression via multiple molecular mechanisms (Zhao et al 2016). The first and most obvious mechanism is direct recognition of a modification, and indeed the mammalian YTH domain-containing 'reader' proteins specifically bind single-stranded $m^6A$ RNAs and can regulate their function and fate (Meyer and Jaffrey 2014; Wang et al 2014; Liu and Pan 2016; Zhao et al 2016). The $m^6A$ modification position and level in mRNA varies as a function of a complex and dynamic interplay between these 'writers' and 'erasers', as expected for a regulatory phenomenon (Meyer & Jaffrey 2014; Liu & Pan 2016). In contrast, specific ψ reader or eraser proteins have not yet been identified, perhaps reflecting difficulties in identifying the more subtle modification, and in breaking the new ψ C-C bond (Zhao and He 2015). Second, RNA modifications can affect gene expression by altering the stability of RNA structure (Roost et al 2015) and thereby access to RNA binding proteins (RBPs) and the translation or processing machinery (Roost et al 2015; Liu et al 2015). For instance, ψ has been shown to stabilize local secondary structure in RNA via increased base stacking and the ability to coordinate a water molecule through the N1 hydrogen (Davis 1995; Jiang et al 2013; Charette and Gray 2000). Such local structural alterations have been shown to

be important for efficient codon-anticodon recognition (Charette and Gray 2000; Ge and Yu 2013 and references therein) and for spliceosome assembly and function (Wu et al 2016). $m^6A$, on the other hand, destabilizes RNA duplexes (Roost et al 2015; Liu et al 2015). Sites with $m^6A$ tend to be in less structured regions *in vivo* (Spitale et al 2015), suggesting that these modifications could behave as structural 'switches' to regulate the accessibility of the nearby sequences (Liu and Pan 2016).

Here, we address a third likely model of regulation via these and other RNA modifications –altering the interactions between RNA and RNA binding proteins (RBPs). Between 3 - 8% of eukaryotic genomes encode proteins dedicated to binding RNA (Glisovic et al 2008). There is now compelling evidence that many RBPs bind to large sets of cytotopically and functionally related RNAs, presumably to coordinate and control function (Gerber et al 2004; Ule et al 2003; Morris et al 2010). RNA modifications may alter these RBP•RNA pools, by either eliminating certain interacting partners or by subtly altering the balance and identity of bound RNAs and thereby altering gene expression and cellular states.

To test this hypothesis, we determined the effects of ψ and $m^6A$ on the binding of human Pumilio 2 (hPUM2) RNA binding protein to modified RNA. hPUM2 is a member of the highly conserved PUF family of sequence-specific, single-stranded RNA binding proteins (Fig. 1C), binds over 750 unique mRNA targets *in vivo* and plays critical roles in brain and germline development and stem cell maintenance (Morris et al 2008; Galgano et al 2008). To incisively test for direct effects on RNA/protein interactions from these substitutions we systematically probed binding to simple, single-stranded RNAs. We found that ψ and $m^6A$ modifications modestly weaken binding of hPUM2 to target RNA suggesting that there are additional functional roles of RNA modifications and a greater complexity in cellular post-transcriptional regulatory networks.

## Results and Discussion

We set out to ask the general question of whether and to what extent mRNA modifications might alter protein binding, using the RNA binding domain of human Pumilio2 as a model RBP (herein referred to as hPUM2). hPUM2 is composed of 8 PUF repeats each of which recognizes and makes specific interactions with bases in the RNA (Fig. 1C, Lu and Hall 2011). Analysis of crystal structures of human Pumilio proteins in complex with various RNAs (Lu and Hall 2011) suggests that ψ is not expected to directly disrupt hydrogen bond interactions with the protein but that the methyl group of $m^6A$ and the *syn* conformational preferences of the N6 methylamino group (Roost et al 2015) may disrupt interactions. In addition, these modifications could have more subtle effects on stacking, van der Waals interactions, or local sterics and solvation that weaken affinity. To test this hypothesis, we determined the quantitative effects on affinity from introducing Ψ or $m^6A$ residues at each U or A position within the hPUM2 consensus sequence.

### RNA modifications weaken hPUM2 binding to RNA

***Pseudouridine (Ψ) effects on hPUM2 binding.*** We first used a direct gel-shift binding experiment to measure binding affinities of hPUM2 to eight Ψ-containing 10mer RNA oligos (CC**UGUAAAUA**) (Fig. 2A & Fig. S1A-C; Table 1). While weakened binding from these substitutions was observed and direct gel-shift experiments have the apparent advantage of direct observation of the complex of interest at equilibrium, this assay is subject to effects from alteration in equilibrium conditions upon addition of loading buffers and upon placement in the gel well and entry into the gel. These effects are expected to be most pronounced for weaker binders (Batey et al 2001). We therefore also determined affinities by competition, using binding to a radiolabeled unmodified 13mer RNA ($*RU_{13}$) as a common signal (Fig. S1D-E) and the relative ability of each unlabeled modified RNA to reduce binding of $*RU_{13}$. Absolute

values of dissociation constants for the modified RNAs are then readily calculated using the well-established dissociation constant for *$RU_{13}$ (Fig. 2B and S1; also see Methods).

The dissociation constants measured by the two methods are in good agreement with each other, with the exception of the weakest binding, triply modified oligo (3x Ψ), whose binding was 10 fold stronger by competition (Fig. 2C). The weaker apparent affinity observed in the direct assay presumably reflects labeled RNA dissociation that occurs prior to or during gel entry that may be stemmed by higher concentrations of protein nearby to 'rescue' binding. Thus, we adopt the value from the competition assay that compares the binding of each RNA with a common readout, and this assignment was supported by agreement with kinetic measurements (see below).

Our results indicate that changing any of the single uridines to Ψ weakens hPUM2 binding by 2-3 fold, relative to the unmodified control oligo (Fig. 2C and Table 1). Replacing an additional uridine by pseudouridine resulted in a further 2-3 fold decrease in affinity, in agreement with simple energetic additivity (Fig. 2D). Changing all three uridines to Ψs reduced hPUM2 binding ~2-fold more than predicted by an additive model (Fig. 2D), an effect that might represent cumulative error, increased base-stacking from the presence of three Ψs in the free RNA that is disrupted upon binding, or effects within the bound complex.

***N-6 methyladenosine ($m^6A$) effects on hPUM2 binding.*** $m^6A$ modifications have been shown to indirectly facilitate protein binding by destabilizing RNA secondary structure (Roost et al 2015; Liu et al 2015). Here we tested whether these modifications could directly alter the stability of an RNA:protein complex, distinct from indirect structural effects that involve flanking sequences. Using gel shift assays as described above, we measured hPUM2 binding to a series of 11mer RNA oligos in which $m^6A$ replaced 0, 1 or 3 of the adenosines in the PUM2 recognition sequence (CC**UGUAUAUAU**) (Fig. 3A,

3B and S2A). In all cases good agreement was observed between the direct and competition gel-shift methods (Fig. 3C).

Replacing a single A with $m^6A$ at position 4 did not affect binding, but the modification at position 6 or 8 decreased binding affinity to RNA by ~2.5 fold each (Fig. 3A and Table 1). The oligo containing three $m^6As$ (3x $m^6A$) bound ~30 fold weaker than the unmodified oligo (Fig. 3A and Table 1), and about 3-4 fold weaker than predicted from an energetic additivity model (Fig. S2B and see below).

**Ψ and $m^6A$ modifications increase the rate of dissociation of the protein-RNA complex**

RNA modifications can weaken binding affinity between protein and RNA by reducing their rate of productive association ($k_{on}$) or by increasing the rate of dissociation of the bound complex ($k_{off}$). Measurement of these kinetic rate constants can provide valuable information about the physical processes underlying the formation or dissociation of the protein-RNA complex. We measured the dissociation rate constants of the 13 unmodified and modified RNA complexes, using a pulse-chase gel shift assay. The protein-RNA complexes were each first allowed to form by mixing a high concentration of protein and trace amounts of the radiolabeled RNA. A large excess of unlabeled 'chase' RNA was then added to the incubation mix to sequester the protein that dissociated from the radiolabeled RNA. Aliquots were removed at regular intervals and monitored for the amount of undissociated radiolabeled complex remaining by loading onto a native gel. As the only variable in a dissociation experiment is the amount of time the complex spends with unlabeled chase RNA before being loaded on a gel, this experiment is not subject to the potential artifacts from direct gel shift equilibrium measurements (E.g., if, say, 50% of the protein dissociates upon entering the gel, it does so for each timepoint and the observed rate constant for loss of protein-bound RNA is unaffected).

We initially carried out these measurements at 25 ˚C, the temperature of the equilibrium binding measurements (Fig. S3A). At this temperature, $k_{off}$ for hPUM2 from unmodified oligos (0x Ψ and 0x

$m^6A$) is ~0.02 $sec^{-1}$ ($t_{1/2}$ ~35 sec). While we were able to measure reliable $k_{off}$ values for oligos with 0 or 1 modifications (8/13 oligos), only limits could be obtained for oligos with multiple base modifications because dissociation was too fast (Table 1). We therefore also measured dissociation rates at 0 ˚C where dissociation is slowed ~2000 fold, allowing all $k_{off}$ values to be quantitatively compared (Table 1 and Fig. S3B). The relative dissociation rates ($k_{rel}$) of the modified oligos at the two temperatures are in excellent agreement (Fig. S3C).

The relative dissociation rates highly correlate with the relative binding affinities between the oligos (Fig. S3D), indicating that the modified bases weaken binding affinity by speeding dissociation of the protein from the RNA and supporting the veracity of the competition $K_D$ measurements (Table 1). Indeed, the only outlier to a linear relationship with slope 1 is the 3x $m^6A$ oligo, which has a 3-fold smaller effect on $k_{off}$ than $K_D$ (Fig. S3D) and thus likely binds to hPUM2 3-fold slower. The 3-fold effect on the binding rate constant matches the 3-fold weaker binding of the 3x $m^6A$ oligo relative to the predicted affinity from an energetic additivity model (Fig. S3D), consistent with formation of a single-stranded structure that is not binding-competent, such as a more stacked state, or less favorable partitioning of a partially bound state such that dissociation is favored over formation of the fully bound complex.

**Physical models for the effects of RNA modifications on protein binding**

NMR and crystallographic studies have revealed that the unique N1 proton of Ψ forms a hydrogen bond with the phosphate backbone in single-stranded RNA (Jiang et al 2013; Davis 1995), and Ψ and $m^6A$ may stabilize base stacking (Davis 1995; Liu and Pan 2016). These effects would weaken protein binding if the stabilizing interactions in free RNA are disrupted in the protein bound complex. Conversely, although not observed in our experiments, modifications could increase protein affinity if there is increased base stacking in bound complexes or enhanced interactions with amino acid side chains.

**RNA modifications: An additional layer of complexity in the regulation of gene expression**

Pseudouridines and N6-methyladenosines are the most abundant mRNA modifications in eukaryotes (ranging from 100 - 1000 Ψ sites (Li et al 2016a) and ~12,000 $m^6A$ sites (Yue et al 2015) in humans). While we used hPUM2 as a convenient *in vitro* test case to study the effects of these modifications, it is likely that Ψ and $m^6A$ disrupt hPUM2:RNA interactions *in vivo* as well.

Human Pumilio proteins, like other PUF family RBPs, recognize a highly conserved core motif, typified by the sequence $UGUA(N)_nAUA$ (Gerber et al 2004; Wickens et al 2002; Galgano et al 2008; Morris et al 2008). Notably, the recognition motif (UGUAR) of the conserved PUS7 Ψ synthase (Carlile et al 2014; Schwartz et al 2014) overlaps with the 5' half of the PUF consensus sequence (Fig. 1C). Prompted by this observation, we used the RMBase database (Sun et al 2015) and identified a number of *in vivo* targets of hPUM2 that appear to harbor these modifications at various positions near or within the binding site (see Supplementary Note S1 for examples).

Modifications near RBP binding sites can play an indirect role by enhancing the accessibility of these sites to RBPs (Spitale et al 2015). For example, we have shown that $m^6A$ can significantly destabilize RNA secondary structure to increase the accessibility of an occluded hPUM2 site (Fig. S4). Conversely, modification-induced structural rearrangements could also lead to decreased accessibility.

Here, we have shown that $m^6A$ and Ψ modifications within hPUM2 binding sites can directly affect the binding affinity of hPUM2, weakening it by 2-3 fold per modification on average. In principle, such effects can occur for many other RNA binding proteins. Indeed, RMBase searches show that RNA modifications appear to be widely prevalent in the binding sites of many different RBPs including hnRNP C, Ago2, Lin28b, TIAL1 among others (Sun et al 2015). While it is possible that there are RBPs

that are affected more profoundly by RNA modifications, modest affinity changes such as those observed herein nevertheless have the potential to substantially alter gene expression.

These effects of RNA modifications may thus, be comparable to the small –typically <2 fold– effects on mRNA translation and abundance typically observed to arise from microRNA regulation (Bartel 2009). In this regard, RNA modifications may function analogously to microRNAs, as 'rheostats' to fine tune gene expression (Hendrickson et al 2008; Hendrickson et al 2009; Baek et al 2008; Bartel 2009). Our results therefore, reveal an additional layer of complexity in RNA:protein interaction networks and underscore the need for genome-wide quantitative data and modeling to unravel their regulatory logic.

## Materials and Methods

### Protein expression and purification

The RNA-binding domain (residues 706-1059) of *H. sapiens* PUM2 (a kind gift from Andre Gerber, University of Surrey, UK) was cloned into a custom pET28a-based expression vector in frame with an N-terminal His-tag and SNAP tag. The construct was transformed into *E. coli* protein expression strain BL21 (DE3). Protein expression was induced at an OD600 of between 0.6 - 0.8 with 1 mM IPTG at 18 ˚C for ~20 h. Cells were harvested by centrifugation at 4500g for 20 min. Cell pellets were resuspended in Buffer A (20 mM HEPES-Na, pH 7.4, 500 mM potassium acetate (KOAc), 5 % glycerol, 0.2% Tween-20, 10 mM Imidazole, 2 mM DTT, 1 mM PMSF and Complete Mini protease inhibitor cocktail (Roche)) and lysed four times using an Emulsiflex (Avestin). The lysate was clarified by centrifugation at 20,000 g for 20 mins, nucleic acids were then precipitated with dropwise addition of Polyethylene Imine (Sigma) to a final concentration of 0.21% with constant stirring at 4˚C and pelleted by centrifugation at 20,000 g for 20 min. The supernatant was loaded on a Nickel-chelating HisTrap HP column (GE) and washed extensively over a shallow 10 - 25 mM imidazole gradient. Protein was eluted over a linear 25 - 500 mM gradient of imidazole. Peak protein fractions were pooled and desalted into Buffer B (20 mM HEPES-Na, pH 7.4, 50 mM KOAc, 5% glycerol, 0.1% Tween-20, 2 mM DTT) using a desalting column. The His-tag was removed by incubation with TEV protease for 13 – 16 h at 4 ˚C, and the protein solution was loaded again on the HisTrap HP column. The flow-through containing cleaved protein was collected, desalted into Buffer B, and loaded on a Heparin column and washed extensively to remove any bound RNA. Protein was eluted over a linear gradient of potassium acetate from 50 to 1000 mM. Fractions were pooled and desalted into Buffer C (20 mM HEPES-Na, pH 7.4, 100 mM KOAc, 5 % glycerol, 0.1 % Tween-20 and 2 mM DTT), concentrated and diluted two-fold with Buffer C containing 80% glycerol for final storage at -20 ˚C.

**RNA oligos used for binding experiments**

Pseudouridine-containing 10mer RNA oligos (CCUGUAAAUA) were synthesized by the Protein and Nucleic Acid (PAN) facility at Stanford University on an ABI 394 DNA synthesizer using standard RNA protocols and commercially available pseudouridine (Glen Research). After deprotection, RNA oligos were purified using Glen-Pak RNA cartridges (Glen Research) following manufacturer's recommendations.

N6-methyladenosine 11mer oligos (CCUGUAUAUAU**)** were synthesized with a MerMade 6 DNA/RNA synthesizer (Bioautomation) utilizing standard phosphoramidite chemistry and commercially available N6-methyladenosine phosphoramidite (Glen Research). The short $m^6A$ oligos contained an additional 3' uridine in order to simplify their synthesis ($m^6A$ is available as phosphoramidite but not 3'-coupled to a synthesis column). Standard deprotection protocols for base and 2'-hydroxyl deprotection were used, with subsequent purification using Glen-Pak RNA purifications cartridges following manufacturer's recommendations (www.glenresearch.com). Oligos were then ethanol precipitated and dissolved in TE buffer (10 mM Tris-Na, pH 8, 1 mM Na-EDTA). All synthesized RNAs were observed to migrate as a single band on a native polyacrylamide gel. In addition, we observed >95% purity for three $m^6A$ hairpin oligos (Fig. S5), which were tested by reverse-phase HPLC.

RNA oligonucleotides (25 pmol) were radiolabeled using [$\gamma$-$^{32}$P] ATP (Perkin Elmer) and T4 Polynucleotide Kinase (NEB) for 30 min at 37 ˚C and purified from non-denaturing 20% acrylamide gels run in 0.5X TBE (50 mM Tris-Na, 41.5 mM Boric Acid, 0.5 mM Na-EDTA).

**Electrophoretic Mobility Shift Assays (EMSA)**

For equilibrium measurements, serially diluted protein amounts were incubated with trace amounts (<40 pM) of radiolabeled RNA in 20 µl reaction volumes for ≥30 min at 25 ˚C in buffer containing 20 mM

HEPES-Na, pH 7.4, 100 mM KOAc, 0.1 % Tween-20, 5 % glycerol, 0.1 mg/ml BSA, 2 mM $MgCl_2$ and 2 mM DTT. Equilibrium conditions were verified by varying incubation time (2 – 510 min) and the amount of radiolabeled RNA (3-fold), as well as by measuring the dissociation kinetics of the protein-RNA complex (see below). An aliquot of each reaction (7.5 µl) was mixed with 5 µl of ice-cold quench solution containing 2.5 µM unlabeled version of the same RNA, 6.25% Ficoll 400 and 0.075% Bromophenol Blue. A specified aliquot of that mixture (7.5 µl) was immediately loaded on a pre-equilibrated, native 20% acrylamide gel in 0.5X TBE. Gels were run at 750 V for ~1 hr before drying and exposing to phosphorimager screens. Screens were scanned on a Molecular Dynamics Typhoon instrument (GE Healthcare). Bound and free RNA were quantified using TotalLab Quant and fit using the hyperbolic equation (fraction bound=[P]/([P] + $K_D$)) in Kaleidagraph (Synergy Software).

For competition experiments, serially diluted, unlabeled modified RNAs were incubated at 25 ˚C for ≥30 min with hPUM2 and radiolabeled RNA (*UCU**UGUAUAUA**UA; $RU_{13}$; ≤0.05 nM) in 20 µl reaction volumes at a concentration of protein (0.5 nM) at which 50-70% of the p*RNA was bound. A 30 min incubation was determined to be sufficient for equilibration based on the controls in the direct binding experiment and measurements of the dissociation kinetics. Reactions were quenched in ice-cold quench solution containing 1 µM unlabeled $RU_{13}$ RNA and immediately loaded on a pre-equilibrated, native 20% acrylamide gel in 0.5X TBE. Gels were run, dried and quantified as above. Competition data were fit to a model for one-site specific inhibition in Kaleidagraph (Synergy Software).

Dissociation rate constants were monitored by gel-shift pulse chase experiments as follows. A saturating amount of hPUM2 protein (typically 10 – 50 nM for 25 ˚C experiments and 100 – 500 nM for 0 ˚C experiments) was incubated with radiolabeled RNA for ≥15 min. A large excess of unlabeled RNA (10-20 times the protein concentration) was then added to the incubation mix. Aliquots were removed at

regular intervals and mixed with ice-cold gel loading buffer (as above without unlabeled RNA); 7.5 µl of this mix was loaded on a pre-equilibrated, native 20% acrylamide gel in 0.5X TBE; and gels were run, dried and quantified as above. Dissociation rate constants were extracted from a fit to a model for single exponential decay.

The active fraction for hPUM2 was determined by titration experiments. Briefly, varying protein amounts were incubated with a fixed, saturating concentration (26 nM) of $RU_{13}$ RNA and reactions were incubated at 25 ˚C for ≥30 min. An aliquot of each reaction (7.5 µl) was mixed with 5 µl ice-cold quench solution containing 2.5 µM unlabeled $RU_{13}$ and electrophoresed as above. The active fraction of the hPUM2 preparation used in these experiments was ~55% and the dissociation constants reported here have been corrected appropriately by the amount of active protein.

**UV Melting Experiments**

Absorbance versus temperature profiles for the structured $m^6A$-containing RNA oligos were collected on a Shimadzu 1800 UV-vis spectrophotometer using an eight-chamber quartz microcuvette. RNA oligos were diluted two-fold in buffer containing 40 mM sodium cacodylate, pH 7, 2 M NaCl and 1 mM Na-EDTA. Melt profiles were collected at 260 nm in duplicate for at least three different concentrations of each RNA with a constant heating rate of 1 ˚C/min and subsequently analyzed using LIFFT program (https://github.com/DasLab/LIFFT).

## Acknowledgements

We thank Dr. Andre Gerber, University of Surrey, UK for the kind gift of the human Pumilio 2 expression plasmids. We are also grateful to the Protein and Nucleic Acid Facility at Stanford for the synthesis of the ψ -containing oligos. This work was supported by a grant from the National Institutes of Health (NIH) to D.H. and H.A-H. (P01 GM066275). We thank members of the Herschlag and Al-Hashimi laboratories for helpful discussions.

## References

Baek D, Villén J, Shin C, Camargo FD, Gygi SP, Bartel DP. 2008. The impact of microRNAs on protein output. *Nature*. 455(7209): 64-71

Bartel DP.2009. MicroRNAs: target recognition and regulatory functions. *Cell* 136: 215-233.

Batey RT, Sagar MB, Doudna JA. 2001. Structural and energetic analysis of RNA recognition by a universally conserved protein from the signal recognition particle. *J Mol Biol* 307: 229-24

Cantara WA, Crain PF, Rozenski J, McCloskey JA, Harris KA, Zhang X, Vendeix FA, Fabris D, Agris PF. 2011. The RNA modification database, RNAMDB: 2011 update. *Nucleic Acids Res* 39: D195-D201.

Carlile TM, Rojas-Duran MF, Zinshteyn B, Shin H, Bartoli KM, Gilbert WV. 2014. Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature*. 515: 143-146

Charette M, Gray MW. 2000. Pseudouridine in RNA: what, where, how, and why. *IUBMB life* 49: 341-351.

Davis DR.1995. Stabilization of RNA stacking by pseudouridine. *Nucleic Acids Res* 23: 5020-5026.

Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S, Peer E, Kol N, Ben-Haim MS, Dai Q, Di Segni A, Salmon-Divon M, Clark WC. 2016. The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. *Nature* 530: 441-446.

Galgano A, Forrer M, Jaskiewicz L, Kanitz A, Zavolan M, Gerber AP. 2008. Comparative analysis of mRNA targets for human PUF-family proteins suggests extensive interaction with the miRNA regulatory system. *PLoS One* 3: e3164.

Ge J, Yu YT. 2013. RNA pseudouridylation: new insights into an old modification. *Trends Biochem Sci* 38: 210-218.

Gerber AP, Herschlag D, Brown PO. 2004. Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast. *PLoS Biol* 2: E79.

Gilbert WV, Bell TA, Schaening C. 2016. Messenger RNA modifications: Form, distribution, and function. *Science* 352: 1408-1412.

Glisovic T, Bachorik JL, Yong J, Dreyfuss G. 2008. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett* 582(14):1977-86

Hendrickson DG, Hogan DJ, Herschlag D, Ferrell JE, Brown PO. 2008. Systematic identification of mRNAs recruited to argonaute 2 by specific microRNAs and corresponding changes in transcript abundance. *PLoS One*. 3(5): e2126.

Hendrickson DG, Hogan DJ, McCullough HL, Myers JW, Herschlag D, Ferrell JE, Brown PO. 2009. Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA. *PLoS Biol*. 7(11): e1000238

Hoernes TP, Hüttenhofer A, Erlacher MD. 2016. mRNA modifications: Dynamic regulators of gene expression? *RNA Biology* 13: 760-765.

Jiang J, Aduri R, Chow CS, SantaLucia J. 2013. Structure modulation of helix 69 from Escherichia coli 23S ribosomal RNA by pseudouridylations. *Nucleic Acids Res* 42: 3971-3981.

Karijolich J, Yi C, Yu YT. 2015. Transcriptome-wide dynamics of RNA pseudouridylation. *Nat Rev Mol Cell Biol* 16: 581-585.

Karijolich J, Yu YT. 2011. Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature* 474: 395-398.

Khoddami V, Cairns BR. 2013. Identification of direct targets and modified bases of RNA cytosine methyltransferases. *Nature Biotechnol* 31: 458-464.

Li X, Ma S, Yi C. 2016a. Pseudouridine: the fifth RNA nucleotide with renewed interests. *Curr Opin Chem Biol* 33: 108-116.

Li X, Xiong X, Wang K, Wang L, Shu X, Ma S, Yi C. 2016b. Transcriptome-wide mapping reveals reversible and dynamic N1-methyladenosine methylome. *Nat Chem Biol* 5: 311-316

Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T. 2015. N6-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* 518: 560-564.

Liu N, Pan T. 2016. N6-methyladenosine–encoded epitranscriptomics. *Nat Struct Mol Biol* 23: 98-102.

Liu N, Pan T. 2015. RNA epigenetics. *Transl Res* 165: 28-35.

Lovejoy AF, Riordan DP, Brown PO. 2014. Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in S. cerevisiae. *PLoS One* 9: e110799.

Lu G, Hall TMT. 2011. Alternate modes of cognate RNA recognition by human PUMILIO proteins. *Structure* 19: 361-367.

Machnicka MA, Milanowska K, Osman Oglou O, Purta E, Kurkowska M, Olchowik A, Januszewski W, Kalinowski S, Dunin-Horkawicz S, Rother KM, et al. 2013. MODOMICS: a database of RNA modification pathways--2013 update. *Nucleic Acids Res* 41: D262-D267.

Meyer KD, Jaffrey SR. 2014. The dynamic epitranscriptome: N6-methyladenosine and gene expression control. *Nat Rev Mol Cell Biol* 15: 313-326.

Morris AR, Mukherjee N, Keene JD. 2008. Ribonomic analysis of human Pum1 reveals cis-trans conservation across species despite evolution of diverse mRNA target sets. *Mol Cell Biol* 28: 4093-4103.

Morris AR, Mukherjee N, Keene JD. 2010. Systematic analysis of posttranscriptional gene expression. *Wiley Interdiscip Rev Syst Biol Med* 2: 162-180.

Roost C, Lynch SR, Batista PJ, Qu K, Chang HY, Kool ET. 2015. Structure and thermodynamics of N6-methyladenosine in RNA: a spring-loaded base modification. *J Am Chem Soc* 137: 2107-2115.

Schwartz S. 2016. Cracking the epitranscriptome. *RNA* 22: 169-174.

Schwartz S, Bernstein DA, Mumbach MR, Jovanovic M, Herbst RH, León-Ricardo BX, Engreitz JM, Guttman M, Satija R, Lander ES. 2014. Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* 159: 148-162.

Spitale RC, Flynn RA, Zhang QC, Crisalli P, Lee B, Jung J-W, Kuchelmeister HY, Batista PJ, Torre EA, Kool ET. 2015. Structural imprints in vivo decode RNA regulatory mechanisms. *Nature* 519: 486-490.

Squires JE, Patel HR, Nousch M, Sibbritt T, Humphreys DT, Parker BJ, Suter CM, Preiss T. 2012. Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res* 40: 5023-5033.

Sun W-J, Li J-H, Liu S, Wu J, Zhou H, Qu L-H, Yang J-H. 2015. RMBase: a resource for decoding the landscape of RNA modifications from high-throughput sequencing data. *Nucleic Acids Res* 44: D259-D265.

Ule J, Jensen KB, Ruggiu M, Mele A, Ule A, Darnell RB. 2003. CLIP identifies Nova-regulated RNA networks in the brain. *Science* 302: 1212-1215.

Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, Fu Y, Parisien M, Dai Q, Jia G. 2014. N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* 505: 117-120.

Wickens M, Bernstein DS, Kimble J, Parker R. 2002. A PUF family portrait: 3'UTR regulation as a way of life. *Trends Genet* 18: 150-157.

Wu G, Adachi H, Ge J, Stephenson D, Query CC, Yu YT. 2016. Pseudouridines in U2 snRNA stimulate the ATPase activity of Prp5 during spliceosome assembly. *EMBO J*. 35(6): 654-667.

Yue Y, Liu J, He C. 2015. RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. *Genes Dev* 29: 1343-1355.

Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, Dirks RM, Pierce NA. 2011. NUPACK: analysis and design of nucleic acid systems. *J Comput Chem* 32: 170-173.

Zhao BS, He C. 2015. Pseudouridine in a new era of RNA modifications. *Cell Res* 2: 153-154

Zhao BS, Roundtree IA, He C. 2016. Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol* 10.1038/nrm.2016.132

## Figure Legends:

**Fig 1.** Pseudouridine (Ψ) and N6-methyladenosine ($m^6A$) are the two most abundant mRNA modifications in eukaryotes. *(A)* Ψ, the 5' ribosyl isomer of uridine, is made by pseudouridine synthases via guide-RNA dependent and independent mechanisms. *(B)* N-6 methyl modifications to adenosines are made by $m^6A$ 'writers' and removed by 'erasers'. *(C)* Cartoon representation of a structural model of human Pumilo2 in complex with an 8mer RNA whose sequence is shown below. PDB ID: 3Q0Q (Adapted from Lu and Hall 2011). The Pus7 recognition motif (UGUAR) in the sequence is boxed in blue.

**Fig 2.** Ψ weakens hPUM2 binding affinity to target RNA. Binding isotherms (*A*) and inhibition curves (*B*) measuring hPUM2 affinities to RNA oligos (CCUGUAAAUA) containing 0 (green circles), 1 (blue symbols) and 3 (red circles) Ψ modifications. Representative curves from two independent replicates shown. *(C)* Bar graph comparing hPUM2 $K_D$ values measured directly (grey) or by competition (black). *(D)* Additivity of the effects of individual Ψ modifications. Bar graph showing the observed (black) $K_D$ competition values for RNA oligos containing two or more Ψ modifications compared to values predicted based on energetic additivity from the individual modifications (grey). The predicted $K_D$ ($K_D$ *pred*) of an oligo containing multiple Ψs was calculated as follows. The observed $K_D$s ($K_D$ *obs*) of the singly modified oligos were first normalized by the $K_D$ of the unmodified oligo (0x Ψ) to yield the corresponding $K_D$ *rel* and $K_D$ *pred* for an oligo with Ψs at positions *i* and *j* was calculated as:

$$K_D\ pred\ (i,j)\ =\ K_D\ rel\ (i) \times K_D\ rel\ (j) \times K_D\ obs\ (0x\ \Psi)$$

**Fig 3.** $m^6A$ weakens hPUM2 binding affinity to target RNA. Binding isotherms (*A*) and inhibition curves (*B*) measuring hPUM2 affinities to RNA oligos (CCUGUAUAUAU) containing 0 (green circles), 1 (blue symbols) and 3 (red circles) $m^6A$ modifications. Representative curves from two independent replicates

are shown. *(C)* Bar graph comparing hPUM2 $K_D$ values measured directly (grey) or by competition (black).

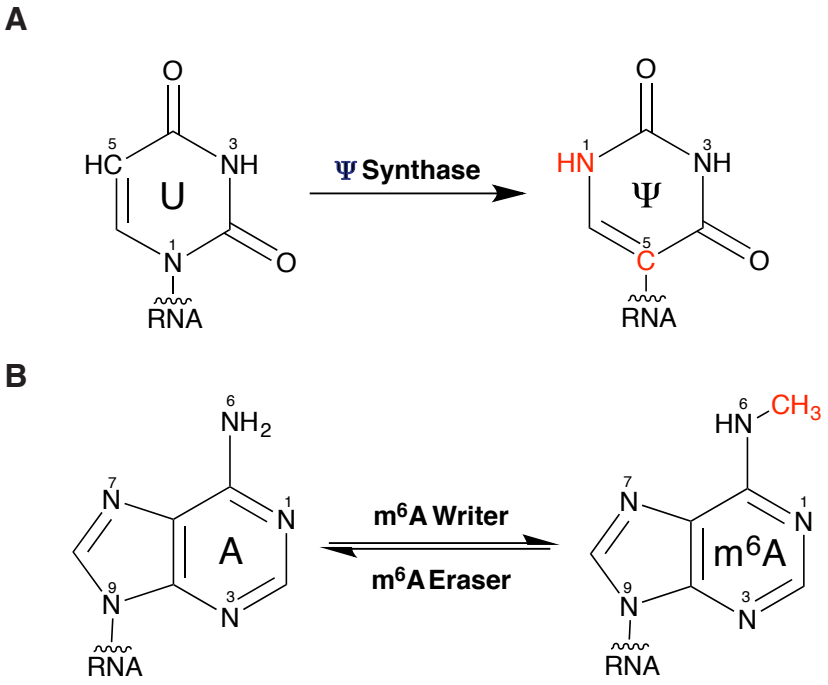


Fig 1

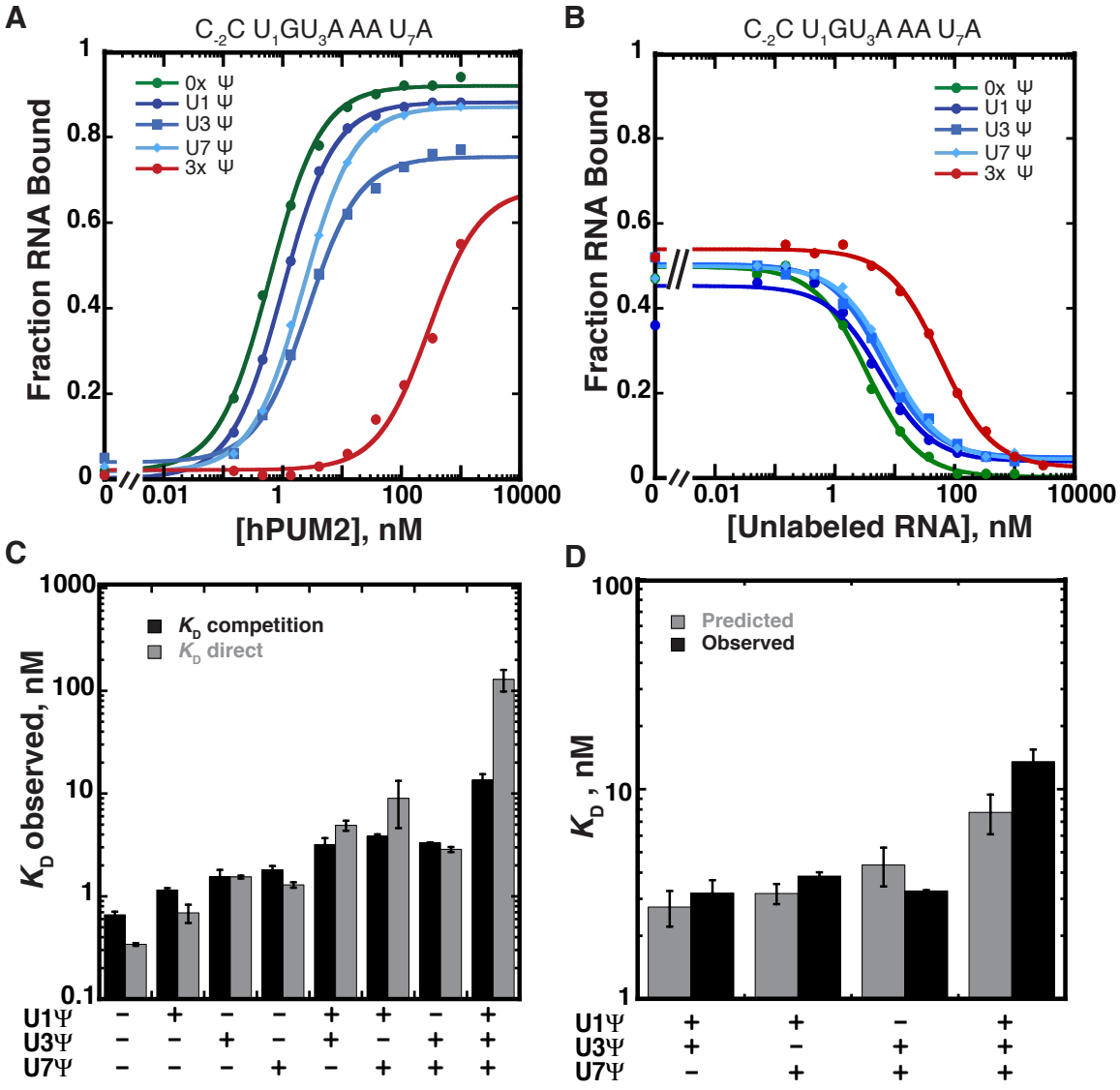


Fig 2

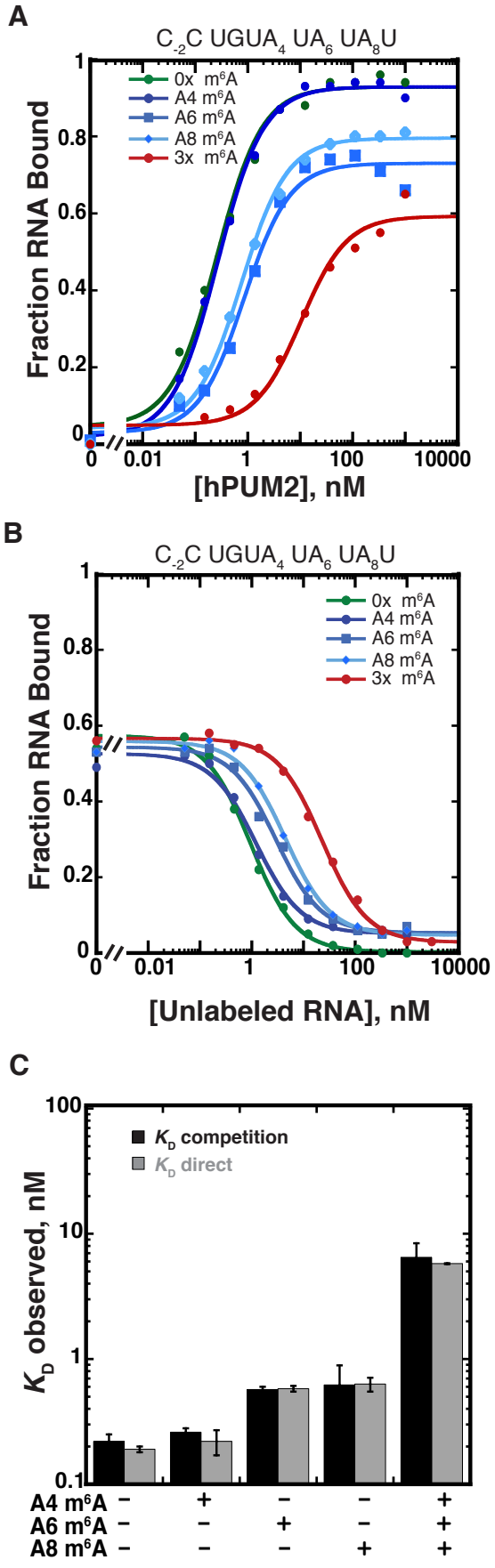


Fig 3

**Table 1: Thermodynamic and Kinetic Measurements of hPUM2 Binding of Modified RNAs**

| RNA Oligo | 25 ˚C | | | | 0 ˚C |
|---|---|---|---|---|---|
| | $K_D$ direct (nM) | $K_D$ comp[a] (nM) | $k_{off}$ (x$10^{-2}$ $s^{-1}$) | $k_{on\ (calc)}$ (x$10^{7}$ $M^{-1}$ $s^{-1}$)[b] | $k_{off}$ (x$10^{-5}$ $s^{-1}$) |
| UCUUGUAUAUAUA ($RU_{13}$) | 0.14 (0.11, 0.18)[b] | 0.34 (0.15, 0.66)[b] | - | - | - |
| | | | | | |
| CCUGUAAAUA | 0.34 ± 0.01 | 0.66 ± 0.05 | 2.3 ± 0.3 | 3.5 | 1.2 |
| CCψGUAAAUA | 0.70 ± 0.1 | 1.15 ± 0.05 | 3.8 | 3.3 | 2.8 |
| CCUGψAAAUA | 1.54 ± 0.05 | 1.5 ± 0.3 | 6.2 | 4.1 | 5.2 |
| CCUGUAAAψA | 1.3 ± 0.1 | 1.8 ± 0.2 | 5.3 | 2.9 | 2.7 |
| CCψGψAAAUA | 4.9 ± 0.6 | 3.2 ± 0.5 | >13.6 | >4.3 | 15.6 |
| CCψGUAAAψA | 9 ± 4 | 3.9 ± 0.2 | >14.2 | >3.6 | 10.8 |
| CCUGψAAAψA | 2.9 ± 0.2 | 3.30 ± 0.05 | >8.9 | >2.7 | 7.7 |
| CCψGψAAAψA | 129 ± 31 (791 ± 9[d]) | 13.5 ± 1.9 | >7.9 | >0.6 | 23.7 |
| | | | | | |
| CCUGUAUAUAU | 0.14 ± 0.0 | 0.22 ± 0.03 | 1.9 ± 0.3 | 8.6 | 1.1 |
| CCUGUAUAUAU | 0.19 ± 0.05 | 0.26 ± 0.02 | 2.5 | 9.6 | 1.5 |
| CCUGUAUAUAU | 0.40 ± 0.11 | 0.57 ± 0.03 | 6.2 | 10.9 | 3.2 |
| CCUGUAUAUAU | 0.42 ± 0.0 | 0.62 ± 0.27 | 5.0 | 8.1 | 2.6 |
| CCUGUAUAUAU | 5.3 ± 0.6 | 6.5 ± 1.9 | - | - | 13.9 |

[a] $K_D$ comp derived by competition against labeled $RU_{13}$ RNA

[b] Mean and (95% confidence intervals) are shown

[c] $k_{on}$ values were calculated from the equation $k_{on} = \frac{k_{off}}{K_D\ comp}$

[d] $K_D$ calculated by forcing the fit through a max Fraction RNA Bound of 1.


# Pseudouridine and N-6 methyladenosine modifications weaken PUF protein/RNA interactions

Pavanapuresan P Vaidyanathan, Ishraq AlSadhan, Dawn K Merriman, et al.

*RNA* published online January 30, 2017