HHS Public Access

Author manuscript

Nat Methods. Author manuscript; available in PMC 2021 October 15.

Published in final edited form as:

Nat Methods. 2020 March ; 17(3): 311–318. doi:10.1038/s41592-019-0726-y.

An Optimized Chemical-Genetic Method for Cell-Specific Metabolic Labeling of RNA

Sarah Nainar[1], **Bonnie Cuthbert**[3], **Nathan M. Lim**[1], **Whitney E. England**[1], **Ke Ke**[1], **Kanika Sophal**[2], **Robert Quechol**[2], **David L. Mobley**[1,2], **Celia W. Goulding**[1,3], **Robert C. Spitale**[*,1,2,3]

[(1)]Department of Pharmaceutical Sciences, University of California, Irvine. Irvine, California. 92697

[(2)]Department of Chemistry, University of California, Irvine. Irvine, California. 92697

[(3)]Department of Molecular Biology and Biochemistry, University of California, Irvine. Irvine, California. 92697

Abstract

Tissues and organs are composed of diverse cell types, which poses a major challenge for cell-specific gene expression profiling. Current metabolic labeling methods rely on the inability of mammalian cells to incorporate exogenous pyrimidine analogs, which are then co-opted by ectopically-expressed enzymes. We demonstrate that mammalian cells can incorporate uracil analogs and characterize the enzymatic pathways responsible for high background incorporation. To overcome these limitations, we developed a novel small-molecule/enzyme pair consisting of uridine-cytidine kinase 2 (UCK2) and 2'-azidouridine (2'AzUd). We demonstrate that 2'AzUd is only incorporated in UCK2-expressing cells and characterize selectivity mechanisms using molecular dynamics and X-ray crystallography. Furthermore, this pair can be used to purify and track RNA from specific cellular populations, making it ideal for high-resolution cell-specific RNA labeling. Overall, these results reveal novel aspects of mammalian salvage pathways and serve as a new benchmark for designing, characterizing and evaluating cell-specific biomolecule labeling methodologies.

Introduction

Characterizing gene expression profiles within select cell types is an incredibly challenging task[1]. While cell sorting and ribosome purification methods to profile cell-specific RNA

[*]Correspondence: rspitale@uci.edu.

Author contributions

S.N. and R.C.S. conceived of the study. S.N. performed all metabolic labeling studies and UCK1/2 activity analyses with help from K.S. and K.K. W.E. assisted with data analysis. B.C., R.Q., and C.G. performed X-Ray crystallography analysis and binding studies, with input from R.C.S. and S.N. N.L. and D.L.M. performed molecular dynamics simulations, with input from R.C.S. and S.N. The manuscript was written by S.N. and R.C.S. with input from all authors.

Online content

Supplementary information and images, supplementary notes on X-ray crystallography and molecular dynamics, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://www.nature.com/nmeth/.

expression have been reported, they have significant limitations. The stress of cell sorting can introduce false-positive RNA signatures[2,3], while ribosome purification methods only isolate RNAs associated with ribosomes and enrichment conditions can induce false associations post-lysis[4,5]. These shortcomings underscore the need to develop more stringent methods to track RNA expression in a cell-specific manner.

RNA metabolic labeling utilizes exogenous nucleobase or nucleoside analogs, which are introduced into nascent transcripts through endogenous cellular machinery[6]. These metabolic reporters are endowed with bio-orthogonal functional groups that enable isolation of RNA through selective chemistries[7–10]. The conventional approach for cell-specific labeling relies on expression of a unique, exogenous enzyme in the cell type of interest that converts an inert metabolic intermediate to an active intermediate (Fig. 1)[3]. The gold standard utilizes *Toxoplasma gondii* uracil phosphoribosyl transferase (*Tg*UPRT) to convert the analog 4-thiouracil (4SU) to 4-thiouridine monophosphate, which is further phosphorylated and incorporated into RNA[11]. Notably, the *Tg*UPRT system is selective only if higher-order eukaryotes cannot incorporate uracil or uracil analogs on their own. However, it has recently been demonstrated that the *Drosophila melanogaster* UPRT homolog can incorporate exogenous uracil analogs, limiting the selectivity of *Tg*UPRT labeling[12]. This surprising observation suggests there may be active UPRT homologs in other organisms or alternative endogenous pathways of incorporation that remain unidentified.

In this work, we discovered that mammalian cells are capable of salvaging exogenous uracil analogs via human UPRT (*Hs*UPRT) and more importantly, uridine monophosphate synthase (*Hs*UMPS), a bifunctional enzyme that catalyzes the final two steps of *de novo* pyrimidine biosynthesis. Consequently, this renews the challenge of designing stringent methods for cell-specific metabolic labeling of RNA. We focused on developing a novel nucleoside-enzyme pair through evaluating uridine analogs paired with engineered or wild type (WT) uridine-cytidine kinase 2 (UCK2), a rare kinase isoform with low or no endogenous protein expression. We identified the pairing of 2’-azidouridine (2’AzUd) and WT UCK2, which we characterized through molecular dynamics (MD) simulations and determination of the co-crystal structure of 2’AzUd bound to UCK2. Lastly, we demonstrated the high selectivity of RNA labeling in various cell types and through selective imaging and enrichment.

Endogenous enzymes can incorporate exogenous uracil analogs for metabolic labeling of RNA.

We and others have recently observed that incubation of moderate to high concentrations of uracil analogs such as 5-ethynyluracil (5EU) or 4SU into WT cells results in metabolic incorporation[13–15]. To explore background incorporation, we exposed different mammalian cells to 5EU, with 5-ethynyluridine (5EUd) as a labeling control as it does not rely on uracil salvage (Fig. 2a & b). Figure 2c demonstrates background incorporation in all cell lines from human and mouse tissues, and imaging analysis revealed that 5EU incorporation mirrors 5EUd imaging of RNA synthesis (Supplementary Fig. 1 & 2). Interestingly, we detected 5EU incorporation in primary cortical neurons, which are non-proliferative. We observed a

similar pattern of background incorporation for the analog 4SU (Supplementary Fig. 3)[16]. Notably, our background analysis mimics standard conditions used for *in vivo* studies with *Tg*UPRT/4SU, where concentrations are 1 mM or above, and often administered over the course of several days[17–19].

We identified two candidate enzymes that could utilize uracil or uracil analogs to facilitate labeling: *Hs*UPRT[20] and *Hs*UMPS (Fig. 2d)[21]. UMPS functions through converting orotic acid to orotidine-5'-monophosphate via the N-terminal enzyme orotate phosphoribosyl transferase, but may use uracil or uracil analogs as a substrate[22]. Mammalian tissue body map data demonstrates expression of UPRT and UMPS in many tissue types, which could contribute to background labeling (Fig. 2e)[23].

We overexpressed FLAG-tagged clones of each enzyme in HEK293T cells and observed an increase in the RNA biotinylation signal from both 4SU and 5EU (Fig. 2f–h; Supplementary Fig. 4, 5a). Both enzymes significantly increased 5EU incorporation; however, the increase from *Hs*UMPS was greater (Fig. 2h; Supplementary Fig. 6a). The *Hs*UMPS crystal structure supports this observation, where a spacious pocket near the C5-position may tolerate analogs such as 5EU (Supplementary Fig. 6b). We compared these treatments to 5EU-treated cells expressing *Tg*UPRT (Supplementary Fig. 6c) and observed similar levels of background for both UPRT homologs. Sequence alignments revealed extensive conservation (Supplementary Fig. 7), supporting the observed background labeling. These results suggest that *Hs*UMPS and *Hs*UPRT can facilitate 5EU incorporation into RNA.

To understand the contributions from each endogenous enzyme, we performed siRNA knockdown and assessed the effects on 5EU labeling. Knockdown of *Hs*UMPS revealed a prominent reduction of 5EU incorporation into RNA (Fig. 2i; Supplementary Fig. 8a). Conversely, we observed no significant reduction in labeling signal with *Hs*UPRT knockdown (Fig. 2j & k; Supplementary Fig. 8b), and similar results with 4SU (Supplementary Fig. 5b & c). This demonstrates that background labeling in WT cells is largely due to ubiquitous UMPS enzyme expression (Supplementary Fig. 8c & 9), which reveals a novel function outside of *de novo* metabolism pathways. We anticipate that the contributions from each enzyme will be cell-type-dependent, as their relative expression levels vary across cell types (Fig. 2e). Overall, endogenous UMPS expression in most mammalian tissues may severely limit the stringency of the *Tg*UPRT system.

Identification of the UCK2/2'-azidouridine pair for robust and non-toxic metabolic labeling of RNA.

These striking observations prompted us to search for a novel small molecule-enzyme pair that would enable cell-specific metabolic labeling with higher stringency. Ideally, the pair would intercept pyrimidine triphosphate biosynthesis, as uridine and cytidine are evenly distributed throughout the transcriptome.

The first step in pyrimidine triphosphate production is carried out by uridine-cytidine kinase, which has two isoforms (UCK1 and UCK2; Fig. 3a)[24]. UCK1 is ubiquitously expressed, while UCK2 has detectable mRNA expression only in placenta tissue, and undetectable

or unconfirmed protein expression *in vivo*[25]. Human UCK1 and UCK2 show a sequence similarity of 67%; however, UCK1 has a low tolerance for non-canonical substrates, while the UCK2 active site is more flexible, endowing greater substrate promiscuity [27,28]. We envisioned that UCK2 paired with a uridine or cytidine analog would afford a novel route for cell-specific metabolic labeling.

We began by identifying analogs with no detectable background in WT cells. Consistent with our previous findings, 5-azidomethyluridine (5mAzUd) (Fig. 3b) is not incorporated into cellular RNA[9]. We also found that 2'-azidocytidine (2'AzCd) is incorporated into the RNA of WT cells while 2'-azidouridine (2'AzUd) is not (Fig. 3c; Supplementary Fig. 10). This lack of incorporation was assessed at 1 mM concentrations, which yield high background levels as previously discussed (Fig. 2). As such, we hypothesized that WT or engineered UCK2 could be paired with either 5mAzUd or 2'AzUd to label RNA cell-type specifically.

Inspection of the crystal structure of UCK2[27] bound with cytidine revealed that the 5-position of the nucleobase is crowded by bulky aromatic groups (TYR65/112 and PHE83/114) (Fig. 3d). We hypothesized that mutation of these residues to smaller functional groups would accommodate the azidomethyl- group, enabling analog binding and phosphorylation. We applied the same logic to making UCK2 variants that could accommodate a modification at the 2' position, where we see that ARG166, GLN184, TYR185 and VAL189 could crowd the space below the sugar (Fig. 3e)[27].

We utilized in-cell screening to assay the activities of each UCK2 mutant, wherein we transiently expressed each vector and treated cells with 1 mM analog for 5 h (Supplementary Fig. 11). We observed that WT UCK2 and many of the mutants yielded biotinylation signals after 5mAzUd treatment (Fig. 3f, Supplementary Fig. 12). WT UCK2 outperformed all UCK2 variants tested for 2'AzUd incorporation (Fig. 3g; Supplementary Fig. 12).

Strikingly, we also observed that 5mAzUd treatment resulted in complete cell detachment. We imaged cell morphology over time with 100 μM or 1 mM 5mAzUd treatment and found that cells expressing WT or mutant UCK2 were all affected (Fig. 3h; Supplementary Fig. 13–18). In contrast, we saw no changes to cell morphology in WT or mutant-expressing cells treated with 2'AzUd (Fig. 3h; Supplementary Fig. 19 – 22). To ensure that ectopic expression of UCK2 will not impact cells over time, we created a stable expression HEK293T cell line (Supplementary Fig. 23). We then used Trypan blue staining to determine whether the cell membrane is intact, and an alamarBlue assay to monitor cell proliferation. There were no significant changes with 2'AzUd treatment; however, 5mAzUd treatment consistently resulted in complete cell detachment, which over time led to loss of membrane integrity and reduced proliferation (Fig. 3i & j; Supplementary Fig. 24). We also assessed the effects of 2'AzUd or 5mAzUd labeling on nucleophosmin localization, which leaves the nucleolus when stressed, and saw that neither analog causes this change, in contrast to 4-thiouridine (4SUd) (Supplementary Fig. 25).[29] Furthermore, incorporation of 2'AzUd into RNA did not affect RNA decay rates (Supplementary Fig. 26).

Cellular toxicity is a major concern when using RNA metabolic reporters. However, toxicity is rarely investigated extensively, and only 4SUd has been reported to be toxic[30]. Therefore, we performed viability assays for cells treated with 4SUd, 5EUd, and 2'AzCd, and found that the widely used 4SUd was the most toxic, followed by 5EUd (Supplementary Fig. 27 & 28). Our results demonstrate that in contrast to the other analogs, 2'AzUd is not incorporated into WT cells and the WT UCK2/2'AzUd pair is non-toxic, yielding robust metabolic labeling of RNA.

Comparison of analog binding and enzyme kinetics of UCK1/2 isoforms

To provide evidence for selectivity of 2'AzUd with UCK2, we characterized the dissociation constants of 2'AzUd with purified UCK1 and UCK2 (Supplementary Fig. 29; Table 1). UCK1/2 bind to cytidine and uridine with the highest affinity, and introduction of the 2'-azide lowers affinity for both 2'AzCd and 2'AzUd, where UCK1 affinity for 2'AzUd was the lowest (180.8 ± 17.1 nM). This suggests that UCK1 is more discriminating in ligand binding while UCK2 is more promiscuous[27,31]. However, all affinities were in the nanomolar range.

We determined the Michaelis constants of uridine and cytidine phosphorylation, which agreed with literature values (Table 1; Supplementary Fig. 30)[32,33]. The 2' modification significantly raised these values. There was no measurable activity for 2'AzUd with UCK1, while the K_m of 2'AzUd with UCK2 was 240 mM. Despite agreement in binding affinities, UCK2 shows a striking catalytic preference for 2'AzUd compared to UCK1.

Structural analysis of the 2'AzUd/UCK2 pair

Similar binding affinity preferences but selective catalysis suggests a structural mechanism is involved. Comparing the structures of UCK1 and UCK2 (PDB: 2JEO, 1UEJ), one helix of the lid domain differs[27] (Supplementary Fig. 31). In UCK2, the helix is longer and more structured, purportedly for induced fit binding[27].

To understand the interactions of UCK2 with 2'-azido analogs, we performed docking and MD simulations using the UCK2•cytidine complex (PDB: 1UEJ) as our starting model (Supplementary Note 1). We docked 2'AzUd into the UCK2 active site and found two binding modes. The first is canonical (2'$AzUd_{canc}$; Fig. 4a), mimicking the sugar orientation in the reported structure, whereby the ribose ring oxygen is pointed outward. The second, novel binding mode illustrates the sugar has flipped around the N-glycosidic bond, orienting the oxygen inward (2'$AzUd_{flip}$; Fig. 4b).

We validated our MD predictions by obtaining crystal structures of 2'AzUd-complexed UCK2 with and without ATP (Supplementary Note 2). Figure 4c illustrates electron density of three subunits modeled with pre-catalytic 2'AzUd (UCK2•2'AzUd) (Supplementary Fig. 32). We also obtained a post-catalytic structure of UCK2, wherein a phosphorylation reaction yielded 2'-azidouridine mono-phosphate (UCK2•2'AzUMP) (Fig. 4d; Supplementary Fig. 33). Superimposition of pre- and post-catalytic structures revealed transitions in the lid domain and a loop directly opposite. In the UCK2•2'AzUMP structure,

key residues shift toward the binding pocket and make contacts with the substrate, indicating induced fit binding to promote catalysis (Fig. 4e; Supplementary Note 2).

Structures of pre-catalytic UCK2•2'AzUd and 2'AzUd$_{flip}$ similarly overlay (RMSD 2.88 Å) and have common hydrogen bonds between the structure and MD simulations, while post-catalytic UCK2•2'AzUMP overlays (RMSD 2.44 Å) and has common hydrogen bonds with 2'AzUd$_{canc}$ (Fig. 4f & g, Supplementary Fig. 34 & 35). These observations strongly suggest that the novel flipped binding mode resembles the pre-catalytic state, and the canonical binding mode represents the post-catalytic state (Supplementary Movie 1). UCK2, unlike UCK1, is likely able to achieve a catalytically-active conformation with 2'AzUd using an induced fit mechanism, facilitated by the flipped binding mode.

Demonstration of stringent cell-specific metabolic labeling

Our observations suggest that the UCK2/2'AzUd pair would be sufficiently stringent for cell-specific metabolic labeling, if UCK1 is similarly refractory to 2'AzUd phosphorylation in cells and UCK2 enzyme is not expressed in the targeted cell type. Supplementary Fig. 36 demonstrates that immortalized cell lines (293T and HeLa) possess higher RNA expression of UCK1/2 than any tissues[34], therefore these cell lines can serve as benchmarks for potential off-target labeling. We assessed 2'AzUd labeling in the presence of overexpressed UCK1 or UCK2, confirming expression through Western blot (Fig. 5a–c; Supplementary Fig. 37a & b). 2'AzUd was only incorporated into cellular RNA when UCK2 was overexpressed, consistent with the *in vitro* data. We confirmed overexpression through direct detection of ectopic UCK2 enzyme; in contrast, endogenous UCK2 protein was undetectable in WT cells (Fig. 5c), despite higher reported mRNA expression (Supplementary Fig. 36). These results mirror reports of undetectable or unconfirmed UCK2 protein expression *in vivo*[25].

We tested for spurious incorporation of 2'AzUd in other immortalized cell lines and found no 2'AzUd background, in contrast to uracil analogs (Fig. 5d; Fig. 2). This was true for human and mouse lines, demonstrating that the observed selectivity may be applied across mammalian cells. This also provides evidence that limited UCK2 expression will not produce appreciable background labeling in primary cell and tissue experiments.

Next, we determined the concentration and time needed for 2'AzUd labeling. We observed signal from metabolic incorporation within 15 minutes (Fig. 5e; Supplementary Fig. 37c), which is ideal for tracking transient expression. LC-MS/MS analysis at 12 h post-labeling demonstrates a ~1% substitution rate, similar to other RNA labeling analogs (Supplementary Fig. 38).[35,36] We further observed labeling at concentrations as low as 0.01 mM (Fig. 5f; Supplementary Fig. 37d). Despite the relatively slow kinetics of UCK2 with 2'AzUd, exogenous enzyme expression and multiple-turnover capability facilitate detectable incorporation at very low concentrations.

We further demonstrated labeling specificity by imaging nascent RNA in 2'AzUd-treated cells expressing UCK2-TurboGFP. The strongest signal originated from the nucleolus, the

site of ribosomal RNA synthesis (Fig. 5g). We also see labeling in all cells treated with 2'AzCd, regardless of UCK2 expression.

Lastly, we envisioned utilizing this system to enrich RNAs from a complex cellular mixture. We transfected cells with UCK2-TurboGFP or mCherry, followed by re-plating as a heterogeneous population. After 2'AzUd labeling, we enriched and quantified TurboGFP and mCherry cDNA using qPCR. TurboGFP cDNA was enriched at 5 h compared to the untreated 0 h control, while mCherry was not (Supplementary Fig. 39). Finally, we co-cultured UCK2-expressing human cells with mouse cells (Fig. 5h). Post-labeling and enrichment, human RNA was enriched and mouse RNA was depleted (Fig. 5i). Read density across genes further demonstrated the specificity of enrichment (Fig. 5j). We also evaluated U-richness enrichment bias and observed minimal but statistically significant differences between RNAs at different enrichment levels (Fig. 5k). This is similar to a recent report that robust covalent biotinylation can remove biases.[37] These experiments confirm the specificity of the UCK2/2'AzUd system and establish methods for further RNA expression profiling of unique cell types.

Discussion

An underappreciated facet of developing cell-specific metabolic methods for biomolecular analysis is a comprehensive understanding of the pathways responsible for installing the bio-orthogonal reporter molecules. Herein, we have demonstrated that there are additional considerations when using *Tg*UPRT and uracil analogs. We made the seminal discovery that mammalian cells can incorporate exogenous uracil analogs into RNA, having specifically identified that *Hs*UMPS and *Hs*UPRT enzymes contribute to metabolic flux of these analogs into RNA. This surprising result expands our understanding of mammalian uracil and uracil analog incorporation and cautions against indiscriminate use of uracil analogs in cells with high background activity.

Because endogenous expression of enzymes that enable uracil analog incorporation is pervasive, we focused on intercepting uridine-cytidine mono-phosphorylation. We discovered that pairing UCK2 with 2'AzUd can be a highly-stringent, non-toxic approach for cell-specific metabolic labeling of RNA. Analysis of other non-specific metabolic probes (4SUd, 5EUd and 2'AzCd) revealed a pattern of toxicity wherein base modifications had the most pronounced effects on viability compared to sugar modification.

Biochemical analysis to determine the affinity of 2'AzUd and 2'AzCd with UCK1/2 revealed similar patterns of binding affinity. However, kinetic analysis revealed only UCK2 could phosphorylate 2'AzUd. We performed MD simulations of 2'AzUd in the binding pocket of UCK2, from which we predicted a novel, flipped binding mode. We then confirmed our models through crystal structures of UCK2 complexed with 2'AzUd. Structural analysis revealed necessary interactions for binding and binding modes pre- and post-catalysis, which will guide future analog design and UCK2 engineering. To our knowledge, this is the first comprehensive biochemical and structural analysis of a metabolic enzyme with a bio-orthogonal metabolic reporter. Lastly, we have demonstrated stringent cell-specificity for the UCK2/2'AzUd pair in a series of cellular experiments.

Overall, we have developed and thoroughly characterized a novel approach for cell-specific RNA tagging. We have uncovered a novel aspect of mammalian cell metabolism: the ability to process and incorporate uracil analogs into RNA. This inspired us to identify a new approach for high-resolution RNA labeling. We anticipate that these advancements will serve as a new paradigm for developing such methods and will supply the field with a stringent system for cell-specific RNA tracking and analysis. Finally, we note that 2'AzUd and UCK2 expression plasmids are commercially available (Methods), allowing this system to be easily adopted throughout the scientific community.

Methods

Vectors for transient transfection:

pCMV6-UCK2-GFP (Origene #RG202406)

pCMV6-UCK2-Myc-DDK (Origene #RC202406)

pCMV6-UCK1-Myc-DDK (Origine #RC220876)

UCK1-GFP (UCK1 ORF cloned into pCMV6-GFP vector through restrictions sites Xho1 and SacII)

pCMV6-UPRT-Myc-DDK (Origene #RC205030)

pLJC2-UPMS-3xFLAG (a gift from David Sabatini, Addgene plasmid # 87975) http://n2t.net/addgene:87975 ; RRID:Addgene_87975)

pcDNA3.3-mini-INT-HA-TgUPRT (a gift from Mike Cleary Lab)

pMSCV-IRES-mCherry (a gift from Ellen Rothenberg, Addgene plasmid # 80139) http://n2t.net/addgene:80139 ; RRID:Addgene_80139)

Vector for stable line generation:

Lentivirus EF1a-UCK2-P2A-EGFP; PGK-Puro

List of qPCR primers used:

Target	Forward	Reverse
GAPDH	TGTCAAGCTCATTTCCTGGTAT	CTCTCTTCCTCTTGTGCTCTT
UPRT	ATGGTGACCACTCCAACAGG	TTGCCTCACCGCTTCTCATT
UMPS	CGTCTTCTGAGTCAGGTTGC	ATAAGGCACTCCACACACGG
TurboGFP	CTACCACTTCGGCACCTACC	GTACTTCTCGATGCGGGTGT
mCherry	GATAACATGGCCATCATCAAGGA	CGTGGCCGTTCACGGAG
MYC	AATGAAAAGGCCCCCAAGGTAGTTATCC	GTCGTTTCCGCAACAAGTCCT CTTC
EEF1A	GTCAAGGAAGTCAGCGCCTA	CTTGAACCACGGCATGTTGG
MZT1	TTAGCCATGGCGAGTAGCAG	AGCAGAACGTCCATGGTCTC

All primers were empirically determined to have efficiencies between 95 – 103%.

Site-directed mutagenesis of UCK2-GFP vector.

Amino acid substitutions in the open reading frame of UCK2 were made using the Q5 Mutagenesis kit (New England Biolabs). Back to back primers containing mutations were designed for the targeted region, allowing for exponential application of the plasmid. Mutants with more than one substitution were made from template containing one previously made mutation. Following kinase, ligation and Dpn1 digestion of the amplification reaction, the mutant plasmids were transformed into chemically competent DH5alpha *E. coli* and selected via ampicillin resistance. Following plasmid isolation and sequence validation, 1–2 μg of each mutant was transfected into HEK293T cells and analyzed for kinase activity via streptavidin-HRP dot blot, described below.

Mutation	Forward	Reverse
Y65A	GGATAGCTTCgccCGTGTCCTTACCTC	TGGCTCAGGATGACCACCTGCTTC
Y65S	GGATAGCTTCagcCGTGTCCTTACCTC	TGGCTCAGGATGACCACCTGCTTC
Y65D	GGATAGCTTCgacCGTGTCCTTACCTC	TGGCTCAGGATGACCACCTGCTTC
F83A	CCAGTTCAACgccGACCACCCGGATGCC	CCCTTCAGGGCTTTGGCCTTCT
Y112A	GATCCCCGTGgccGACTTTGTCTC	TGGACTGTTTTCCCTTCAGTGAT
Y112S	GATCCCCGTGagcGACTTTGTCTC	TGGACTGTTTTCCCTTCAGTGAT
F114A	CGTGTATGACgccGTCTCCCATTCC	GGGATCTGGACTGTTTTCCCTTC
F114S	CGTGTATGACagcGTCTCCCATTC	GGGATCTGGACTGTTTTCCCTTC
F83A; F114S	CCAGTTCAACgccGACCACCCGGATGCC	CCCTTCAGGGCTTTGGCCTTCT
F83A; Y112S	CCAGTTCAACgccGACCACCCGGATGCC	CCCTTCAGGGCTTTGGCCTTCT
Y112S; F114A	cgccGTCTCCCATTCCCGGAAG	tcgctCACGGGGATCTGGACTGT
Y112S; F114S	cagcGTCTCCCATTCCCGGAAG	tcgctCACGGGGATCTGGACTGT
Y65A; Y112S; F114S	GGATAGCTTCgccCGTGTCCTTACCTC	TGGCTCAGGATGACCACCTGCTTC
R166G	GCTCTCACGCggcGTATTAAGGGAC	CGGGTGTCCGCATCTGTATCCAC
R166A	GCTCTCACGCgccGTATTAAGGGACATCAGC	CGGGTGTCCGCATCTGTATCCAC
R166H	GCTCTCACGCcacGTATTAAGGGACATCAGC	CGGGTGTCCGCATCTGTATCCAC
R166Y	GCTCTCACGCtacGTATTAAGGGACATCAGC	CGGGTGTCCGCATCTGTATCCAC
Q184G	GATTTTATCTggcTACATTACGTTCGTCAAGCCTGC	TGCTCAAGATCCCTGCCTCTCTC
Q184A	GATTTTATCTgccTACATTACGTTCGTCAAGCCTGC	TGCTCAAGATCCCTGCCTCTCTC
Q184H	GATTTTATCTcacTACATTACGTTCGTCAAGC	TGCTCAAGATCCCTGCCTCTCTC
Q184Y	GATTTTATCTtacTACATTACGTTCGTCAAGCCTG	TGCTCAAGATCCCTGCCTCTCTC
Y185A	TTTATCTCAGgccATTACGTTCGTCAAG	ATCTGCTCAAGATCCCTGCCT
V189G	CATTACGTTCggcAAGCCTGCCT	TACTGAGATAAAATCTGCTCAAGATC
V189A	CATTACGTTCgccAAGCCTGCCT	TACTGAGATAAAATCTGCTCAAGATC
V189H	CATTACGTTCcacAAGCCTGCCTTTG	TACTGAGATAAAATCTGCTCAAG
V189Y	CATTACGTTCtacAAGCCTGCCTTTG	TACTGAGATAAAATCTGCTCAAG

Cell lines and culture conditions.

HEK293T, HeLa, NIH-3T3, (ATCC) were cultured in DMEM containing glucose and L-glutamine (Corning) supplemented with 10% FBS, 1% penicillin and streptomycin. K562 (ATCC) cells were cultured in RPMI with L-glutamine (Life Technologies) supplemented with 10% FBS, 1% penicillin and streptomycin. Primary cortical neurons (PCN) were isolated as previously described[38] and maintained in Neural Basal media supplemented with 1× B27 supplements, 1% penicillin and streptomycin and 1% Glutamax (Thermo Fisher Scientific). All cell lines were grown at 37 °C, in an atmosphere of 5% CO_2.

Cell transfection and stable line generation.

Cells were transiently transfected with plasmid 1–2 μg of plasmid using jetPRIME DNA transfection reagent (Polyplus Transfection) and were incubated for 24 – 48 h to allow for expression. A UCK2 stable expression cell line was generated in HEK293T cells following transduction with lentivirus encapsulated EF1a-UCK2-P2A-EGFP-PGK-Puro vector. Cells were puromycin selected to obtain UCK2 positive cells, which were then clonally isolated through limiting dilution.

Nucleoside analog treatment in cultured cells.

5-ethynyl uridine (Click Chemistry Tools), 4-thiouracil (Sigma), 4-thiouridine (Cayman Chemical). 5-ethynyluracil was synthesized as previously described[39] 2’azidocytidine and 2’aziduridine were purchased from Carbosynth. Analogs were added at various concentrations and times to complete culture medium from 200 mM DMSO stocks (final concentration of <1% DMSO).

Alamar Blue Assay.

HEK293T cells were seeded at densities of 10,000 cells into a 96-well plate and grown for 1 day in supplemented DMEM. Cells were then incubated for various lengths of time with each analog at titrated concentrations or DMSO, with 0.5% DMSO final. Following incubation, 10 μL of alamarBlue reagent (Thermo Fisher Scientific) was added to each well and incubated for 4 hours to assess the cells’ ability to irreversibly reduce resazurin to resorufin, resulting in a color change from blue to red. Absorbance readings were taken at 570 nm and 600 nm as a reference, and were used to calculate cell proliferation relative to the DMSO-treated control.

Trypan Blue Assay.

HEK293T cells were seeded at densities of 1.5 – 2.5×10^5 into 6 well plates and grown for 18 – 24 hours. WT or UCK2-expressing cells were then incubated for 12, 24 and 48 h with analogs 4SUd, 5EUd, 2’AzCd, and 2ÁzUd at 1 mM final concentration. Following incubation, cells were trypsinized, removed from the plate and spun down. The supernatant was aspirated and the cells were re-suspended in 1 mL of culture medium, after which 10 μL was aliquoted for Trypan staining. To each aliquot, 10 μL of 0.4% Trypan Blue solution (Gibco) was added and 10 μL of this solution was applied to a hemocytometer for assessment of cell viability. Total and dead cell counts was recorded for each of the four corners of the grid.

Imaging of NPM1 localization after nucleoside treatment.

Coverslips in 6-well tissue culture plates were treated with 1× poly-D lysine (Corning) solution for 3 h at 37 °C. HEK293T cells were seeded at densities of 2.5×10^5 and grown to ~50% confluency on glass cover slips. Cells with or without ectopic expression of UCK2 were treated with nucleosides to final concentrations of 500 μM and incubated for 5 hours. After labeling, cells were washed twice with DPBS and fixed for 10 min at room temperature with 3.7% paraformaldehyde. Cells were quenched with 50 mM glycine, DPBS for 5 min, and then washed twice more. Cells were permeabilized in 0.1% Triton-X DPBS for 30 min, and then washed ×2 with DPBS. HEK293T cells were blocked with 10% goat serum/DPSB for 2 h. Cells were than incubated with 1:500 anti-NPM1 (anti-B23; Sigma Aldrich) in 1% BSA/DPBS overnight at 4 °C, followed by washing ×3 with DPBS. Cells were then incubated with 1:1000 anti-rabbit Alexa 647 (Fisher Scientific) in 1% BSA/DPBS for 1 h in the dark at room temperature. Cells were stained with a solution of 1:3000 Hoechst 333242 for 10 min (Thermo Fisher Scientific). Coverslips were briefly washed and then mounted using VectaShield mounting medium (Vector Labs). Slides were imaged via fluorescence confocal microscopy using a 63× oil immersion objective on a Leica 700 Carl Zeiss microscope.

RNA decay measurements.

WT or HEK293T cells ectopically expressing UCK2 were seeded at densities of 2.5×10^5 and grown to ~70% confluency, followed by nucleoside treatment as previously described. Following incubation, the media was decanted, and the cells were washed and chased with clean media containing 20 μM of transcriptional inhibitor Actinomycin D (Sigma Aldrich). Total RNA was extracted from cells by adding 1 ml of Trizol reagent at 0, 1, 2, 4 and 6 h following Actinomycin D treatment. Genomic DNA contamination was diminished by adding TURBO DNase (Thermo Fisher) digestion at 37 °C for 30 min. 2 μg of total RNA was used as template for cDNA synthesis. Reverse transcription was carried out using random hexamers by PrimeScript RT reagent Kit (Takara) according to the manufacturer's protocol. cDNA was diluted 1:10 with nuclease-free water to be used as template for quantitative PCR. The qPCR reaction was performed with the primers specific for the gene of interest according to the manufacturer's protocol: 5 μL of TB Green Advantage qPCR Premix (Takara), 2 μL of forward and reverse primers (10 μM each), and 3 μL of cDNA template. Samples were run in biological duplicates and technical triplicates. Analysis was performed by normalizing the mean Cq values for each time point of each gene to the 0 h time point. After which, the relative abundance of each mRNA was calculated using the 2^(-deltaCq). Values were plotted in Prism Graphpad, followed by analysis through a non-linear regression curve fitting (one-phase decay).[40]

RNA isolation and biotinylation via azide-alkyne click chemistry.

After labeling with azide or alkyne modified nucleosides/nucleobase, total cellular RNA was harvested using Trizol Reagent (Thermo Fisher Scientific) following the manufacturer's instructions, with two final washes with 70% ethanol, followed by resuspension in 30 – 50 μL of nuclease free water. CuAAC or SPAAC reactions were prepared using 5–10 μg of total RNA in solutions containing either: 500 μM $CuSO_4$, 2.5 mM THPTA (Sigma Aldrich),

2.5 mM sodium ascorbate, and 1 mM alkyne or azide-biotin (Sigma) with a final DMSO concentration of 10% or 10 mM Tris-HCL pH 7.4, 10 mM EDTA, 1 mM DBCO-biotin (Sigma Aldrich). CuAAC reactions were incubated with at room temperature for 30 min at 400 RPM and SPAAC reactions were incubated at 37 °C for 1 h. Biotinylation reactions were purified using Zymo RNA clean and concentrator-5 spin column according to the manufacturer instructions, with RNA eluted in 10 μL of nuclease free water.

4SU and 4SUd RNA biotinylation.

Cells were treated as previously described above, and RNA was Trizol extracted. Biotinylation was carried out in a reaction containing or 100 μM MTSEA-biotin (TRC chemicals) or 1 mg/mL iodoacetyl biotin, 1 mM EDTA, 20 mM HEPES pH 7.4 and 5–10 μg of total RNA. The biotinylation reaction was allowed to proceed for 1 h at room temperature in the dark, after which it was column purified using Zymo clean & concentrator-5.

HRP-streptavidin Dot blotting.

Equal amounts of column-purified, biotinylated RNA were spotted onto 2× SSC equilibrated Hybond-N+ membrane (GE Healthcare), allowing spots to dry in between. The membrane was crosslinked using the auto crosslink function at 254 nm on a UV Stratalinker (Stratagene). The membrane was then blocked in a 10% SDS blocking solution followed by incubation with 1:10,000 dilution of high sensitivity streptavidin-HRP (Thermo Fisher Scientific) in blocking solution. The membrane was washed ×2 in a 1:10 solution of blocking solution and ×2 in Tris-saline buffer. It was then incubated in SuperSignal West Pico Chemiluminescent Substrate (Thermo Fisher Scientific) according to manufacturer's instructions and imaged on a ChemiDoc MP imaging system (Bio-Rad) under the chemiluminescence channel. Following imaging, the blot was stained in a 0.04% methylene blue, 0.3 M sodium acetate solution to visualize all RNA as a loading control. The blot was imaged again using the colorimetric channel.

siRNA knockdown of nucleoside metabolism enzymes in mammalian cells.

siRNA pools targeting human uridine monophosphate synthase (UMPS) and uracil phosphoribosyltransferase (UPRT) (SMARTpool: ON-Targetplus siRNA, Dharmacon) were re-suspended to working concentrations of 20 μM. Cells were transfected with 25 nM of siRNA using INTERFERin transfection reagent, (PolyPlus Jet Fusion) and allowed to incubate for 72 h. Cells were then treated with analog in fresh media as previously described, and were subsequently collected for extraction and dot blotting or RT-qPCR, or lysate was collected for Western blotting.

RT-qPCR analysis of knockdown.

Cells treated with siRNA were isolated and total RNA was extracted. Following RNA precipitation, 1 – 2 μg of RNA was subjected to Dnase Turbo (Thermo Fisher Scientific) treatment according to the manufacturer's instructions, followed by Zymo column purification. RNA was then subjected to reverse transcription using the SuperScript IV Reverse Transcriptase kit (Thermo Fisher Scientific) with oligo(dT) priming. Following first strand synthesis, the reactions were inactivated by heating to 85 °C for 10 min, after

which cDNA was directly subjected to qPCR analysis. qPCR was carried out on a CFX Connect Real-Time PCR Detection System using the iQ Supermix (Bio-Rad). Relative gene expression of UPRT or UMPS to the housekeeping gene GAPDH was determined through the delta-delta Cq expression method.

Western blotting detection of protein ectopic expression and knockdown.

Cells were seeded at a density of 2.5×10^5 into 6 well plates and were grown to 60–70% confluency. Cells were transiently transfected with plasmid using jetPRIME transfection reagent (Polyplus transfection) followed by incubation for 24 – 72 h. After allowing for protein to express, the media was decanted and 300 μL of lysis buffer (150 mM NaCl, 5 mM EDTA pH 8.0, 50 mM Tris, pH 8.0, 1.0% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) spiked with DNase Turbo and RNaseOut (Thermo Fisher Scientific) was added to the cells. Cells were scraped and transferred to tubes, which were incubated on ice for 15 min to allow for lysis to occur. Lysates were spun down at 14,000 RPM, 4 °C for 25 min, and the soluble fraction was collected for protein quantification using the Qubit Protein Assay kit and Qubit fluorimeter (Thermo Fisher Scientific). Equal amounts of protein (10–20 μg) lysate were mixed with 2× Laemelli sample buffer (Bio-Rad) containing β-mercaptoethanol, and boiled at 95 °C for 5 min. Samples were cooled to room temperature, loaded onto a pre-cast 4–20% TGX gel (Bio-Rad) and resolved. Proteins were transferred onto nitrocellulose in a midi-transfer stack using the Trans-Blot Turbo transfer apparatus (Bio-Rad). The membrane was blocked in 5% milk, TBST for 1 h followed by hybridization with primary antibody at room temperature for 1 h. UCK2 over-expression was first detected through 1:500 of anti-UCK2 human (Abcam, Ab167683). UCK2, UMPS, UCK1 and UPRT were detected through 1:1000 dilution of anti-FLAG (Cell Signaling technologies #14793). UMPS was detected at 1:500 dilution of anti-UMPS (Abcam, Ab155763), and UPRT was detected at 1:500 dilution of anti-UPRT (Abcam, Ab111701). The membrane was washed 3 times for 5 min using TBST, followed by incubation with secondary antibody, either 1:2000 anti-mouse HRP or anti-rabbit HRP (Cell Signaling Technologies, #7076, #7074). The membrane was then incubated with SuperSignal West Pico Chemiluminescent Substrate (Thermo Fisher Scientific) according to the manufacturer's instructions and imaged on a ChemiDoc MP imaging system (Bio-Rad) under the chemiluminescence channel. The membrane was subsequently washed and stained with PageBlue Protein Staining solution (Thermo Fisher Scientific) for protein loading controls. Following de-staining, the membrane was imaged under the colorimetric channel.

LC-MS/MS analysis of 2'azidouridine incorporation into cellular RNA.

Wildtype or UCK2-expressing HEK293T cells were treated with 1 mM analog for 12 h as previously described. After RNA extraction with TRIzol (Thermo Fischer Scientific) and purification with RNA clean & concentrator-5 columns (Zymo), 1 μg total RNA samples containing each analog were digested to nucleosides using a Nucleoside Digestion Mix (New England Biolabs). A calibration curve of standard solutions containing Adenosine, Uridine, Guanosine, Uridine, and the nucleoside analogs 2'azidouridine (2'AzUd) and 5-iodocytidine as an internal standard were injected on a Waters Micromass Quattro Premier XE LC-MS system equipped with a triple quadrupole (QQQ) mass detector operating in multiple reaction monitoring (MRM) and positive (+ESI) electrospray ionization modes.

Liquid chromatography was carried out on a Waters Acuity UPLC equipped with a C18 column (4.6 × 150 mm, 5 μm) with a gradient mobile phase consisting of water (+0.2% acetic acid) : acetonitrile. Digested RNA sample reactions were diluted 1:5 and 10 μL was injected. Mass spectrometry data acquisition was recorded in MRM mode and analyzed on Waters MassLynx 4.1 software. The nucleosides 2'AzUd and uridine were identified according to the precursor and product ions found through QuanOpt parameters, with mass transitions of 269.96 to 113.13 and 245.03 to 113.02, respectively. Quantification of nucleoside concentration was determined through a calibration curve of pure standards run simultaneously with the samples. The percent incorporation of 2'AzUd was calculated based on the ratio of 2'AzUd to Uridine.

UCK1 and UCK2 expression and purification.

A plasmid encoding an N-terminally His-tagged UCK1 was purchased from addgene (plasmid number 42374). UCK1 was expressed in *E. coli* BL21 (DE3) Gold cells (Agilent). Cells were grown aerobically at 37 °C in LB medium containing 50 μg/mL ampicillin. UCK2 expression was induced by the addition of 1 mM isopropyl-β-D-thiogalactosidase at an OD_{600} ≈ 0.6–0.8 and expressed overnight at 18 ºC. Cells were collected by centrifugation at 5,500 x g for 10 min and then resuspended in resuspension buffer (50 mM Tris/HCl, pH 7.5, 0.5 M NaCl and 1 mM BME). Cells were lysed in resuspension buffer containing 10 mg/mL lysozyme and 1 mM phenylmethylsulfonyl fluoride (PMSF), and centrifuged for an hour at 4 ºC at 14000 rpm to remove debris. Clarified lysate was loaded onto a column of Ni-NTA beads, Ni-NTA resin was washed with resuspension buffer supplemented with 20 mM imidazole. UCK1 was eluted with stepwise increases in imidazole concentration (50 mM - 2 M) in resuspension buffer. Eluents containing UCK1 were combined, dialyzed overnight into resuspension buffer and concentrated to a ~0.5 mg/mL (~18 μM) using a 30-kDa centrifugal concentrator (Centricon; Millipore).

A plasmid encoding an N-terminal GST fusion UCK2 (residues 1–250) construct as described by Suzuki *et al.* was generated by Genscript in the vector pGEX-6p (Genscript) (Suzuki et al. 2004; Suzuki et al. 2003). UCK2-GST was expressed in BL21 (DE3) Gold cells and lysed as described for UCK1, however UCK2-GST was solubilized in 0.1 M sodium phosphate pH 7.4 and 0.5 M NaCl. Clarified lysate was loaded onto a GSTrap FF column (GE Healthcare), washed with PBS buffer and eluted with a gradient of 50 mM Tris/HCl pH 8, 200 mM NaCl and 10 mM reduced glutathione. Eluted protein was dialyzed into 50 mM Tris/HCl pH 7, 150 mM NaCl, 1 mM EDTA and 1 mM DTT before the GST-tag was removed by digestion with Pierce HRV 3C Protease (Thermo Scientific) at 4 ºC overnight. Remaining GST-tagged UCK2 was separated using glutathione sepharose 4B resin (GE Healthcare,). The untagged UCK2 protein was further purified by gel filtration on a Superdex 200 column (GE Healthcare) in 20 mM Tris/HCl pH 7.5, 150 mM NaCl.

UCK1 and UCK2 fluorescence-based nucleoside binding experiments.

The affinities of UCK1 and UCK2 for uridine, cytidine, and their 2'-azido analogs were determined using fluorescence microscopy. Titrations of nucleoside into UCK1/UCK2 were monitored by fluorescence, observing the change of tyrosine fluorescence intensity with increasing concentration of ligand. UCK1 and UCK2 (100 nM) and the titratable ligand (at

1.25, 2.5 and 5 μM) were prepared in 20 mM Tris/HCl pH 7.5, 150 mM NaCl. Ligands were titrated into protein solutions in 5–500 nM increments with 3 minute incubations between measurements. The tyrosine fluorescence emission was acquired at 303 nm through excitation at 274 nm using a Hitachi F-4500 fluorescence spectrophotometer at 25 °C. The fluorescence emission at 303 nm when plotted against increasing concentrations of ligand provides the dissociation constant (K_d) at half maximal binding. The K_d was determined by the non-linear least squares regression equation for bimolecular model binding: $Y = \frac{(Fmax - Fmin)x}{Kd + x} + Fmax$, where x is the ligand concentration, F_{max} refers to the maximal observed fluorescence and F_{min} to the minimal observed fluorescence. Experiments were performed in triplicate.

NADH-coupled activity assay.

Purified enzyme was subjected to kinetic analysis using an ATPase-coupled assay that monitors the reduction of NADH to NAD+ via decrease in absorbance at 340 nm.[32,41,42] Each reaction contained 20 mM HEPES (pH 7.2), 100 mM KCl, 2 mM $MgCl_2$, 50–300 μM cytidine or uridine for UCK2, 50–8000 μM cytidine or uridine for UCK1, 100 μM – 10 mM 2'AzCd or 2'AzUd for UCK2, 1–40 mM 2'AzCd or 2'AzUd, 1 mM ATP, 250 μM NADH, 250 μM phosphoenolpyruvate, 0.5 U lactate dehydrogenase from rabbit muscle (Sigma Aldrich), 0.5 U pyruvate kinase from rabbit muscle (Sigma Aldrich), and 100 nM of purified UCK1 or UCK2 enzyme. Continuous measurements were taken at 340 nm in 150 μL fill volumes in 96-well UV-transparent plates (Corning), using a Spark Multimode microplate reader (Tecan). Kinetic constants were obtained through a nonlinear fit of the data using the Michaelis-Menten equation with Prism 8 GraphPad software.

RNA fluorescence imaging via CuAAC.

Coverslips in 6-well tissue culture plates were treated with 1× poly-D lysine (Corning) solution for 3 h at 37 °C. HEK293T cells were seeded at densities of 2.5×10^5 and grown to ~50% confluency on glass cover slips. Cells with or without ectopic expression of UCK2 were treated with nucleosides to final concentrations of 1 mM and incubated for 5 hours. After labeling, cells were washed twice with DPBS and fixed for 10 min at room temperature with 3.7% paraformaldehyde. Cells were quenched with 50 mM glycine, DPBS for 5 min, and then washed twice more. Cells were permeabilized in 0.1% Triton-X DPBS for 15 min, and then washed ×2 with DPBS. Cells were then treated with solutions containing 500 μM $CuSO_4$, 2.5 mM THPTA (Sigma), 2.5 mM sodium ascorbate, and 15 μM Alexa-Fluor 568 (Click Chemistry Tools) and incubated at room temperature for 1 h in the dark. Cells were washed ×2 with 0.1% Triton-X DPSB for 2 min, and then ×5 with DPBS for 5 min each on an orbital shaker. Cells were stained with a solution of 1:3000 Hoechst 333242 for 10 min (Thermo Fisher). Coverslips were briefly washed and then mounted using VectaShield mounting medium (Vector Labs). Slides were imaged via fluorescence confocal microscopy using a 63× oil immersion objective on a Leica 700 Carl Zeiss microscope.

Co-culture enrichment of biotinylated transcripts.

HEK293T cells were separately transfected with mCherry or UCK2-GFP plasmid for 24 hours, following by 1:1 re-plating together to generate a mixed culture. Cell were allowed to grow and express for another 16 h, followed by treatment with 1 mM 2'AzUd for 0 or 5 h. Following RNA extraction and SPAAC biotinylation, RNA was DNase treated and 1 μg was subjected to reverse transcription with oligo(dT) priming. Following cDNA synthesis, 1:10 of the reaction was set aside and diluted 1:100 to use as an input control. The remainder of the reaction was subjected to streptavidin bead enrichment using MyOne Streptavidin C1 beads (Thermo Fisher Scientific) with 10 μL of beads per 1 μg of input RNA. Prior to hybridization, beads were washed ×4 in binding buffer (1 M NaCl, 100 mM Tris-HCl pH 7.0, 10 mM EDTA, 0.2% (v/v) Tween-20) before blocking in RNase-free glycogen (Life Technologies) for 1 h at room temperature. After blocking, beads were washed twice more and re-suspended in the original volume using binding buffer. cDNA reactions were added to beads and the reactions were incubated for 20 min at room temperature with rotation to ensure that the beads were mixing. Following hybridization, beads were sequestered on magnetic rack and the supernatant was removed. Beads bound by cDNA:RNA hybrids were transferred to new tubes and were washed ×4 with high salt wash buffer (4 M NaCl, 100 mM Tris-HCl pH 7.0, 10 mM EDTA, 0.2% (v/v) Tween-20) followed by ×2 washes with nuclease free water. Beads were then spun down and re-suspended in 50 μL of cDNA elution buffer (10 U RNase H, 1× RNase H buffer (New England Biolabs), 4 μg RNase A/T1 mix (Thermo Fisher Scientific), 12.5 mM D-biotin) and incubated at 37 °C for 30 min at 750 RPM. Following incubation 1 μL of DMSO was added and samples were heated to 95 °C for 8 min. The eluted cDNA was purified using the Zymo DNA Clean and Concentrator-5 kit with cDNA eluted in 20 μL of nuclease free water.

qPCR analysis of enriched turboGFP from co-culture.

Quantification of input and enriched samples was assessed by qPCR analysis, which was carried out as previously described. Briefly, 1 μL of input or enriched cDNA from 0 h and 5 h 2'AzUd treatments were profiled for mCherry and turboGFP expression, after which the fold enrichment of each gene was calculated against the 0 h treatment, negative control: $2^{[(Cq\text{-}INPUT_{5\ hr} - Cq\text{-}ENRICH_{5\ hr}) - (Cq\text{-}INPUT_{0\ hr} - Cq\text{-}ENRICH_{0\ hr})]}$. Enrichment analysis was performed in biological duplicates and technical triplicates.[43]

Human-mouse co-culture enrichment and RNA sequencing.

HEK293T expressing UCK2-GFP and wildtype NIH3T3 cells were grown in 10 cm co-cultures, followed by labeling with 2'AzUd as previously described. Following treatment, the media was decanted and the cells were washed briefly ×2 with DPBS before treatment with Trizol reagent according to the manufacturer's instructions. After precipitation, the RNA pellets were washed ×2 with 70% ethanol followed by resuspension in 40 μL of nuclease free water. The RNA was treated with DNase TURBO (Thermo Fisher Scientific) for 30 min at 37 °C, followed by column cleanup using Zymo IIICG columns (Zymo Research) and resuspension to concentrations of 1–1.5 μg/μL. RNA was then subjected to biotinylation with DBCO biotin (Sigma Aldrich) as previously described in 150 μL reaction volumes. Enrichment was carried out using 50 – 100 μg of biotinylated RNA.

MyOne Streptavidin C1 beads (Thermo Fisher Scientific) were washed ×4 in bead binding buffer (1 M NaCl, 100 mM Tris-HCl pH 7.0, 10 mM EDTA, 0.2% (v/v) Tween-20) before resuspension in their original volume with bead binding buffer. Biotinylated RNA was added to the beads and incubated for 20 min at room temperature with end-over-end rotation to ensure that the beads were mixing. Following hybridization, beads were sequestered on the magnetic rack and the supernatant was removed. Beads bound with RNA were transferred to new tubes and were washed ×4 with 42 °C high salt wash buffer (4 M NaCl, 100 mM Tris-HCl pH 7.0, 10 mM EDTA, 0.2% (v/v) Tween-20) followed by ×3 washes with nuclease free water. Beads were then spun down and re-suspended in 100 μL of 95% formamide, 10 mM EDTA elution buffer, and were incubated at 90 °C for 5 min at 1000 RPM. Tubes were then cooled for 1 min on ice, before addition of 500 uL of Trizol and extraction as previously described. Purified, biotin-enriched RNA along with input biotinylated RNA were carried through to library preparation using the Tru-seq Stranded total RNA library Prep Human/Mouse/Rat kit (Ilumina) with ribosome depletion. Superscript IV (Thermo Fisher Scientific) was used for first strand synthesis. Sequencing was carried out on an Illumina HiSeq 4000 with SR100. Sequencing read quality was evaluated using FastQC[44]. Reads were mapped to a combined human-mouse genome index built from GRCh38 (GENCODE annotations v30) and GRCm38 (GENCODE annotations vM21) reference genomes[45] using STAR[46]. Reads mapping equally well to multiple positions were discarded. Gene counts were also quantified using STAR[46]. Differential expression analysis of human and mouse genes was performed using DESeq2[47]. Per-position depth of coverage was quantified using samtools[48] and visualized in the UCSC Genome Browser [49].

Reporting Summary.

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

Statistics and Reproducibility.

Figure 2c, 5EU labeling HeLa dot blot experiment was independently repeated twice, HEK293T and NIH3T3 were independently repeated three times, all with similar results. K562 and PCN experiments were performed once, but with two biotinylation and blot replicates having similar results.

Figure 2h. Analysis was performed from a sample size (n) of 3 independent experiments each for *Hs*UPRT and *Hs*UMPS, with ** $p = 0.0015$ and ** $p = 0.0086$ (95% CI), respectively, as determined by a two-tailed Student's *t*-test. Error bars represent standard deviation.

Figure 2k. Bar chart showing the mean integrated chemiluminescence signal of siRNA knockdown dot blots, normalized to the scramble (middle) signal. Analysis was performed from a sample size (n) of 4 independent experiments each for *Hs*UPRT and *Hs*UMPS, with ns $p = 0.8891$ and **** $p = 1.61E-08$ (95% CI), as determined by a two-tailed Student's *t*-test. Error bars represent standard deviation.

Figure 3h, representative images of cell toxicity mediated through 5mAzUd in the absence or presence of UCK2 or UCK2 variant. This experiment was performed once as a 5 –

24 h time titration at 0, 0.1 and 1 mM concentrations for the WT; Y112A; F114S; F83A/F114S; and Y112S/F114S variants (those demonstrating the highest labeling). Cell imaging experiment for 2'AzUd in the absence or presence of UCK2 was performed twice, with similar results. Representative imaging of R166A was performed once as a 5 – 72 h time titration at 0, 0.1 and 1 mM concentrations.

Figure 3j, Alamar analysis shows the mean percent cell proliferation of independent cell cultures with sample sizes (n) of 4 for UCK2+2'AzUd 12 h, UCK2+2'AzUd 48 h, 5mAzUd 5 h, 5mAzUd 24 h, UCK2+5mAzUd 12 h, and UCK2+5mAzUd 24 h. Sample sizes (n) of 5 for 2'AzUd 5 h, UCK2+2'AzUd 24 h, 5mAzUd 48 h, UCK2+5mAzUd 5 h, and UCK2+5mAzUd 48 h. Sample sizes (n) of and 6 for 2'AzUd 12 h, 2'AzUd 24 h, 2'AzUd 48 h, 2'AzUd 72, UCK2+2'AzUd 5 h, UCK2+2'AzUd 72 h, 5mAzUd 12 h, 5mAzUd 72 h, and UCK2+5mAzUd 72 h. Error bars represent standard deviation. Data represents independent biological replicates (cell-culture replicates) and error bars represent standard deviation.

Figure 5c, Anti-FLAG UCK1 western blot was independently repeated twice, with similar results. Anti-FLAG UCK2 western blot was independently repeated twice, with similar results. Anti-UCK2 western blot was independently repeated four times, with similar results.

Figure 5k. Plot of U-richness in enriched RNAs. Significant comparisons are denoted with a horizontal bar. The sample size is n=3 with ** $p = <0.01$ and *** $p = <0.001$, as determined by Type II ANOVA and Tukey's test.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank members of the Spitale lab for their careful reading and critique of the manuscript. The Spitale lab is supported by startup funds from the University of California, Irvine, and NIH 1DP2GM119164 (RCS) and 5R21MH113062 (RCS). RCS is a Pew Biomedical Scholar. SN is supported as a Vertex Fellow and NSF BEST-IGERT (DGE-1144901). DLM appreciates financial support from the National Institutes of Health (1R01GM108889–01). This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1321846 (N.L.). UC Cancer Research Coordinating Committee Grant, CTR-18–522186 (C.W.G.) and R.Q was supported by MARC (GM-69337) and MBRS+MSD (GM-055246), and B.C by a UCI Chancellor's ADVANCE Postdoctoral Fellowship. We thank the Advanced Light Source at Berkeley National Laboratories and the Stanford Synchrotron Radiation Lightsource for their invaluable help in data collection. This work was also made possible, in part, through access to the Genomics High Throughput Facility Shared Resource of the Cancer Center Support Grant (P30CA-062203) at the University of California, Irvine and NIH shared instrumentation grants 1S10RR025496–01, 1S10OD010794–01, and 1S10OD021718–01.

Data availability.

All structures and PDBs have been deposited in the RCSB and will be released upon acceptance of the manuscript. RNA sequencing datasets have been deposited in GEO under the accession number GSE136638

References (Main Text)

1. Jung J & Jung H. Methods to analyze cell type-specific gene expression profiles from heterogeneous cell populations. Anim Cells Syst 20, 113–117, doi:10.1080/19768354.2016.1191544 (2016).

2. Feng H, Zhang X & Zhang C. mRIN for direct assessment of genome-wide and gene-specific mRNA integrity from large-scale RNA-sequencing data. Nat Commun 6, 7816, doi:10.1038/ncomms8816 (2015). [PubMed: 26234653]
3. Handley A, Schauer T, Ladurner AG & Margulies CE Designing Cell-Type-Specific Genome-wide Experiments. Mol. Cell 58, 621–631, doi:10.1016/j.molcel.2015.04.024 (2015). [PubMed: 26000847]
4. Riley KJ, Yario TA & Steitz JA Association of Argonaute proteins and microRNAs can occur after cell lysis. Rna 18, 1581–1585, doi:10.1261/rna.034934.112 (2012). [PubMed: 22836356]
5. Mili S & Steitz JA Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. Rna 10, 1692–1694, doi:10.1261/rna.7151404 (2004). [PubMed: 15388877]
6. Rabani M et al. Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. Nat Biotechnol 29, 436–442, doi:10.1038/nbt.1861 (2011). [PubMed: 21516085]
7. Zheng Y & Beal PA Synthesis and evaluation of an alkyne-modified ATP analog for enzymatic incorporation into RNA. Bioorg Med Chem Lett 26, 1799–1802, doi:10.1016/j.bmcl.2016.02.038 (2016). [PubMed: 26927424]
8. Fauster K et al. 2 ‘-Azido RNA, a Versatile Tool for Chemical Biology: Synthesis, X-ray Structure, siRNA Applications, Click Labeling. Acs Chem Biol 7, 581–589, doi:10.1021/cb200510k (2012). [PubMed: 22273279]
9. Nainar S et al. Metabolic Incorporation of Azide Functionality into Cellular RNA. ChemBioChem 17, 2149–2152, doi:10.1002/cbic.201600300 (2016). [PubMed: 27595557]
10. Jao CY & Salic A. Exploring RNA transcription and turnover in vivo by using click chemistry. P Natl Acad Sci USA 105, 15779–15784, doi:10.1073/pnas.0808480105 (2008).
11. Cleary MD, Meiering CD, Jan E, Guymon R & Boothroyd JC Biosynthetic labeling of RNA with uracil phosphoribosyltransferase allows cell-specific microarray analysis of mRNA synthesis and decay. Nat Biotechnol 23, 232–237, doi:10.1038/nbt1061 (2005). [PubMed: 15685165]
12. Ghosh AC, Shimell M, Leof ER, Haley MJ & O’Connor MB UPRT, a suicide-gene therapy candidate in higher eukaryotes, is required for Drosophila larval growth and normal adult lifespan. Sci Rep 5, 13176, doi:10.1038/srep13176 (2015). [PubMed: 26271729]
13. Nguyen K et al. Cell-Selective Bio-orthogonal Metabolic Labeling of RNA. J Am Chem Soc 139, 2148–2151, doi:10.1021/jacs.6b11401 (2017). [PubMed: 28139910]
14. Hida N et al. EC-tagging allows cell type-specific RNA analysis. Nucleic Acids Res 45, e138, doi:10.1093/nar/gkx551 (2017). [PubMed: 28641402]
15. Zajaczkowski EL et al. Bio-orthogonal Metabolic Labeling of Nascent RNA in Neurons Improves the Sensitivity of Transcriptome-Wide Profiling. ACS Chem Neurosci 9, 1858–1865, doi:10.1021/acschemneuro.8b00197 (2018). [PubMed: 29874042]
16. Miller MR, Robinson KJ, Cleary MD & Doe CQ TU-tagging: cell type-specific RNA isolation from intact complex tissues. Nat Methods 6, 439–441, doi:10.1038/nmeth.1329 (2009). [PubMed: 19430475]
17. Tomorsky J, DeBlander L, Kentros CG, Doe CQ & Niell CM TU-Tagging: A Method for Identifying Layer-Enriched Neuronal Genes in Developing Mouse Visual Cortex. eNeuro 4, doi:10.1523/ENEURO.0181-17.2017 (2017).
18. Chatzi C, Zhang Y, Shen R, Westbrook GL & Goodman RH Transcriptional Profiling of Newly Generated Dentate Granule Cells Using TU Tagging Reveals Pattern Shifts in Gene Expression during Circuit Integration. eNeuro 3, doi:10.1523/ENEURO.0024-16.2016 (2016).
19. van Velthoven CTJ, de Morree A, Egner IM, Brett JO & Rando TA Transcriptional Profiling of Quiescent Muscle Stem Cells In Vivo. Cell Rep 21, 1994–2004, doi:10.1016/j.celrep.2017.10.037 (2017). [PubMed: 29141228]
20. Li J et al. Identification and characterization of human uracil phosphoribosyltransferase (UPRTase). J Hum Genet 52, 415–422, doi:10.1007/s10038-007-0129-2 (2007). [PubMed: 17384901]

21. Jones ME Pyrimidine Nucleotide Biosynthesis in Animals - Genes, Enzymes, and Regulation of Ump Biosynthesis. Annu Rev Biochem 49, 253–279, doi:DOI 10.1146/annurev.bi.49.070180.001345 (1980). [PubMed: 6105839]

22. Suchi M et al. Molecular cloning of the human UMP synthase gene and characterization of point mutations in two hereditary orotic aciduria families. Am J Hum Genet 60, 525–539 (1997). [PubMed: 9042911]

23. Petryszak R et al. Expression Atlas update--an integrated database of gene and protein expression in humans, animals and plants. Nucleic Acids Res. 44, D746–752, doi:10.1093/nar/gkv1045 (2016). [PubMed: 26481351]

24. Huang M & Graves LM De novo synthesis of pyrimidine nucleotides; emerging interfaces with signal transduction pathways. Cell. Mol. Life Sci 60, 321–336 (2003). [PubMed: 12678497]

25. Uhlen M et al. Tissue-based map of the human proteome. Science 347, doi:UNSP 1260419 10.1126/science.1260419 (2015).

26. Malami I & Abdul AB Involvement of the uridine cytidine kinase 2 enzyme in cancer cell death: A molecular crosstalk between the enzyme and cellular apoptosis induction. Biomed Pharmacother 109, 1506–1510, doi:10.1016/j.biopha.2018.10.200 (2019). [PubMed: 30551402]

27. Suzuki NN, Koizumi K, Fukushima M, Matsuda A & Inagaki F. Structural basis for the specificity, catalysis, and regulation of human uridine-cytidine kinase. Structure 12, 751–764, doi:DOI 10.1016/j.str.2004.02.038 (2004). [PubMed: 15130468]

28. Van Rompay AR, Norda A, Linden K, Johansson M & Karlsson A. Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases. Mol Pharmacol 59, 1181–1186 (2001). [PubMed: 11306702]

29. Burger K et al. 4-thiouridine inhibits rRNA synthesis and causes a nucleolar stress response. RNA Biol 10, 1623–1630, doi:10.4161/rna.26214 (2013). [PubMed: 24025460]

30. Burger K et al. 4-thiouridine inhibits rRNA synthesis and causes a nucleolar stress response. Rna Biology 10, 1623–1630, doi:10.4161/rna.26214 (2013). [PubMed: 24025460]

31. Van Rompay AR, Norda A, Linden K, Johansson M & Karlsson A. Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases. Mol Pharmacol 59, 1181–1186 (2001). [PubMed: 11306702]

32. Tomoike F, Nakagawa N, Kuramitsu S & Masui R. A Single Amino Acid Limits the Substrate Specificity of Thermus thermophilus Uridine-Cytidine Kinase to Cytidine. Biochemistry 50, 4597–4607, doi:10.1021/bi102054n (2011). [PubMed: 21539325]

33. van Kuilenburg ABP & Meinsma R. The pivotal role of uridine-cytidine kinases in pyrimidine metabolism and activation of cytotoxic nucleoside analogues in neuroblastoma. Bba-Mol Basis Dis 1862, 1504–1512, doi:10.1016/j.bbadis.2016.05.012 (2016).

34. Hishiki T, Kawamoto S, Morishita S & Okubo K. BodyMap: a human and mouse gene expression database. Nucleic Acids Res 28, 136–138, doi:DOI 10.1093/nar/28.1.136 (2000). [PubMed: 10592203]

35. Kubota M et al. Expanding the Scope of RNA Metabolic Labeling with Vinyl Nucleosides and Inverse Electron-Demand Diels-Alder Chemistry. Acs Chem Biol 14, 1698–1707, doi:10.1021/acschembio.9b00079 (2019). [PubMed: 31310712]

36. Jao CY & Salic A. Exploring RNA transcription and turnover in vivo by using click chemistry. Proc Natl Acad Sci U S A 105, 15779–15784, doi:10.1073/pnas.0808480105 (2008). [PubMed: 18840688]

37. Duffy EE et al. Tracking Distinct RNA Populations Using Efficient and Reversible Covalent Chemistry. Mol Cell 59, 858–866, doi:10.1016/j.molcel.2015.07.023 (2015). [PubMed: 26340425]

References (Methods)

38. Zajaczkowski EL et al. Bio-orthogonal Metabolic Labeling of Nascent RNA in Neurons Improves the Sensitivity of Transcriptome-Wide Profiling. ACS Chem. Neurosci 9, 1858–1865, doi:10.1021/acschemneuro.8b00197 (2018). [PubMed: 29874042]

39. Nguyen K et al. Cell-Selective Bio-orthogonal Metabolic Labeling of RNA. J. Am. Chem. Soc 139, 2148–2151, doi:10.1021/jacs.6b11401 (2017). [PubMed: 28139910]

40. Ratnadiwakara M & Änkö M-L mRNA Stability Assay Using Transcription Inhibition by Actinomycin D in Mouse Pluripotent Stem Cells. Bio-protocol 8, e3072, doi:10.21769/BioProtoc.3072 (2018). [PubMed: 34532533]
41. Tomoike F et al. Indispensable residue for uridine binding in the uridine-cytidine kinase family. Biochem. Biophys. Rep 11, 93–98, doi:10.1016/j.bbrep.2017.07.002 (2017). [PubMed: 28955773]
42. Agarwal KC, Miech RP & Parks RE in Methods Enzymol. Vol. 51 483–490 (Academic Press, 1978). [PubMed: 211390]
43. Pfaffl MW A new mathematical model for relative quantification in real-time RT–PCR. Nucleic Acids Res. 29, e45-e45, doi:10.1093/nar/29.9.e45 (2001). [PubMed: 11328886]
44. Brown J, Pirrung M & McCue LA FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. Bioinformatics, doi:10.1093/bioinformatics/btx373 (2017).
45. Frankish A et al. GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res 47, D766–D773, doi:10.1093/nar/gky955 (2019). [PubMed: 30357393]
46. Dobin A et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21, doi:10.1093/bioinformatics/bts635 (2013). [PubMed: 23104886]
47. Love MI, Huber W & Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550, doi:10.1186/s13059-014-0550-8 (2014). [PubMed: 25516281]
48. Li H et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079, doi:10.1093/bioinformatics/btp352 (2009). [PubMed: 19505943]
49. Kent WJ et al. The human genome browser at UCSC. Genome Res 12, 996–1006, doi:10.1101/gr.229102 (2002). [PubMed: 12045153]

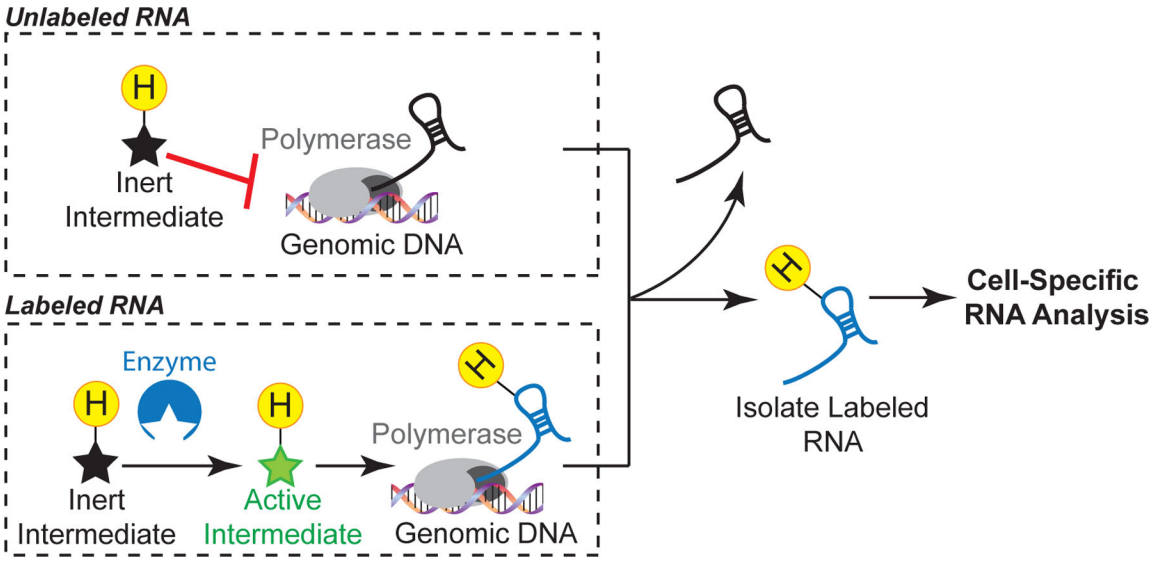

Figure 1: Schematic of cell-specific metabolic labeling of RNA.

For cell-specific metabolic labeling of RNA, an “inert intermediate” (nucleobase or nucleoside), which cannot be processed by wildtype cells, is introduced. Only cells ectopically expressing an enzyme, which can convert an “inert intermediate” to an “active intermediate,” can yield metabolic labeling of RNA. H = chemical handle for functionalization.

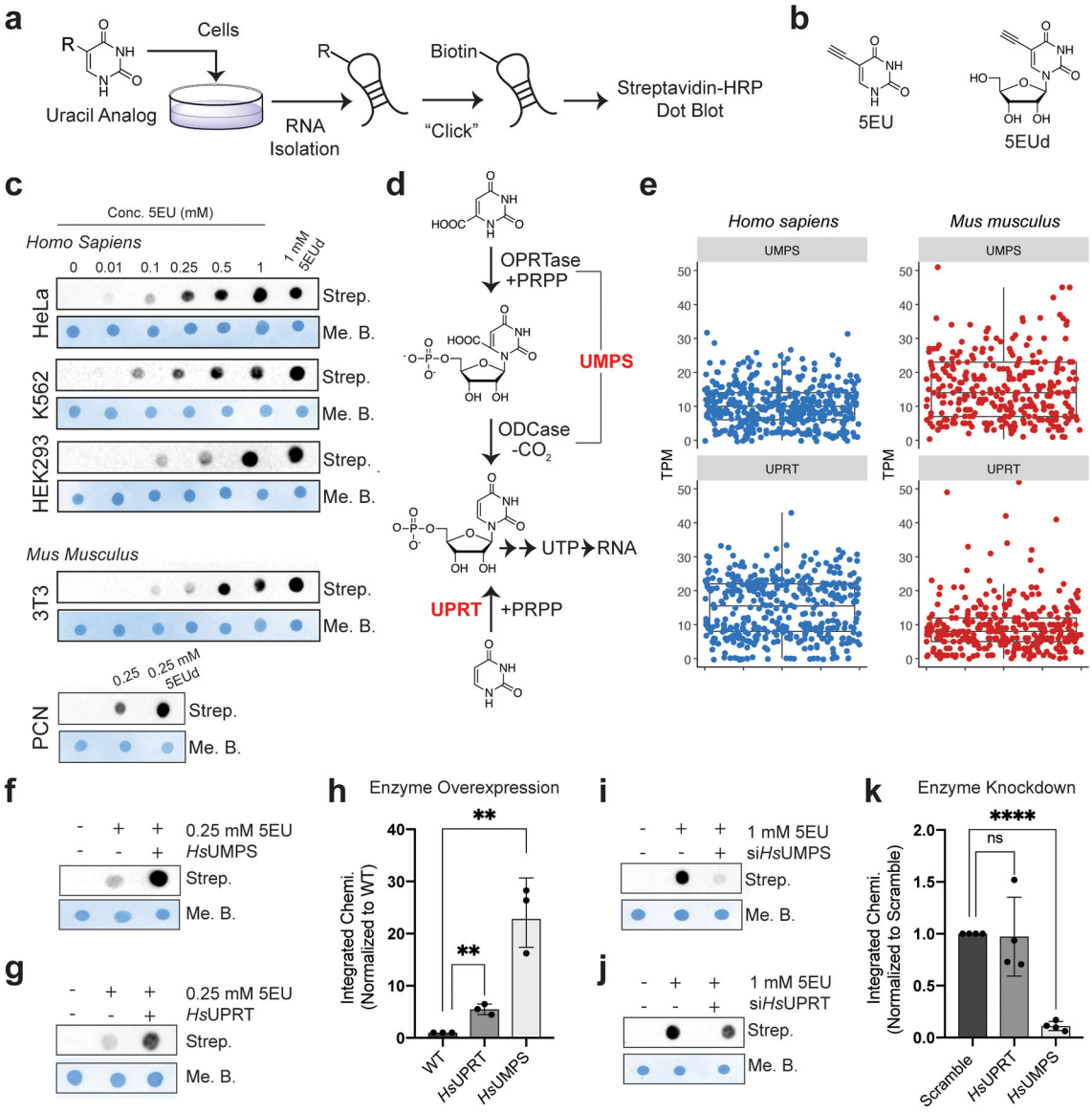

Figure 2: Mammalian cells are capable of salvaging exogenous uracil analogs for eventual incorporation into cellular RNA.

a. Schematic of screening experiments for assaying exogenous uracil analog incorporation into cellular RNA. **b.** Structure of 5-ethynyluracil (for screening; 5EU) and 5-ethynyluridine (analog that is incorporated by wildtype cells; 5EUd) **c.** Time course dot-blot results. Cells were cultured with 5EU for 5 h at indicated concentrations. Total RNA was isolated and biotinylated via azide-alkyne click chemistry and analyzed by streptavidin-HRP dot blot. Strep. = streptavidin-HRP. Me.B. = methylene blue. PCN = primary cortical neurons. **d.** Known enzymatic pathways that could result in salvage of uracil or uracil analogs for subsequent RNA metabolic labeling. **e.** RNA sequencing results from mouse and human body map tissue analysis demonstrate that both UPRT and UMPS are highly expressed in a variety of tissues. **f.** Dot blot demonstrating that in the presence of overexpressed *Hs*UMPS-FLAG construct, there is a strong increase of 5EU incorporation into HEK293T cellular RNA. *Hs = homo sapiens.* This was independently repeated three times, with similar

results. **g.** Dot blot demonstrating that in the presence of overexpressed *Hs*UPRT-FLAG construct, there is a modest increase in incorporation of 5EU into HEK293T cellular RNA. This was independently repeated three times, with similar results. **h.** Bar chart showing the mean integrated chemiluminescence signal of overexpression dot blots, normalized to the WT signal. Analysis was performed from a sample size (n) of 3 independent experiments each for *Hs*UPRT and *Hs*UMPS, with ** $p = 0.0015$ and ** $p = 0.0086$ (95% CI), respectively, as determined by a two-tailed Student's *t*-test. Error bars represent standard deviation. **i.** Dot blot demonstrating a reduction in 5EU signal when treated with siRNA against endogenous *Hs*UMPS. This was independently repeated four times, with similar results. **j.** Dot blot demonstrating no change in 5EU signal when treated with siRNA against endogenous *Hs*UPRT. This was independently repeated four times, with similar results. **k.** Bar chart showing the mean integrated chemiluminescence signal of siRNA knockdown dot blots, normalized to the scramble (middle) signal. Analysis was performed from a sample size (n) of 4 independent experiments each for *Hs*UPRT and *Hs*UMPS, with ns $p = 0.8891$ and **** $p = 1.61E-08$ (95% CI), as determined by a two-tailed Student's *t*-test. Error bars represent standard deviation.

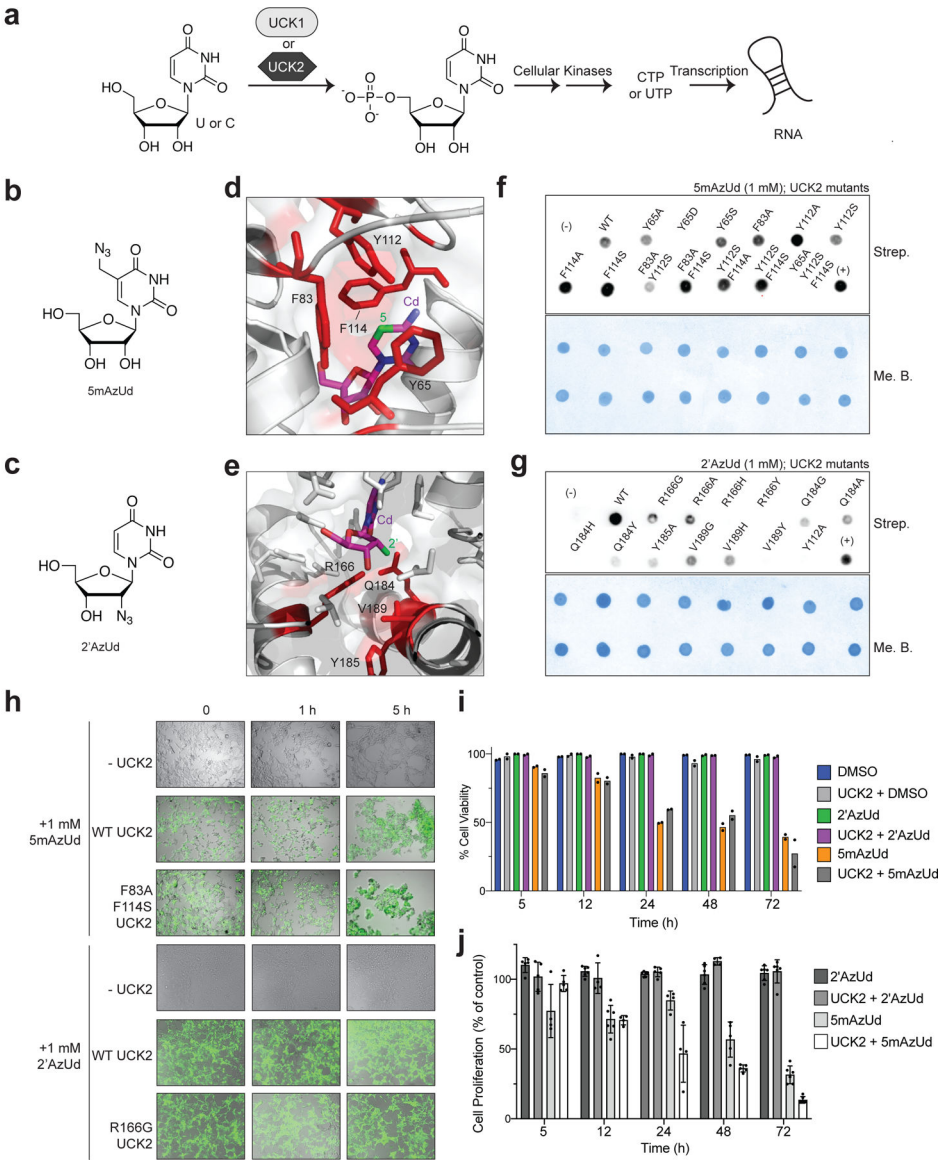

Figure 3: Cellular screening reveals a UCK2/2'AzUd pair amenable for metabolic labeling of RNA.

a. Schematic of pyrimidine salvage pathway through uridine-cytidine kinase and eventual RNA biosynthesis. **b.** Structure of 5mAzUd. **c.** Structure of 2'AzUd. **d.** Crystal structure of the UCK2•cytidine complex (PDB: 1UEJ) with residues selected for mutation to accommodate 5mAzUd highlighted in red (Y65, F83, Y112, F114). C5-position is highlighted in green. **e.** Crystal structure of the UCK2•cytidine complex with residues selected for mutation to accommodate 2'AzUd highlighted in red (R166, Q184, Y185, V189). 2'-position is highlighted in green. **f**. Dot blot results from in-cell screening of UCK2 mutants with 5mAzUd. HEK293T cells WT or mutant UCK2-TurboGFP were treated with 1 mM analog for 5 h. RNA was isolated and biotinylated using azide-alkyne click chemistry, followed by streptavidin-HRP dot blot detection of biotinylated RNA. (−) is the negative control with no UCK2-TurboGFP transfection. (+) is the positive control with 5-ethynyluridine incubation, followed by click chemistry with azide biotin. This was

independently repeated twice, with similar results. **g.** Dot blot results from screening UCK2 mutants with 2'AzUd (as in **2f**). This was independently repeated twice, with similar results. **h.** Representative HEK293T cell imaging highlighting changes to cell morphology, and potential toxicity of 5mAzUd-treated cells (1 mM) expressing WT or UCK2 mutants (top). **i.** Trypan blue assay of HEK293T cell permeability demonstrates that 5mAzUd has an overall greater effect on cell viability than 2'AzUd, relative to DMSO controls. WT or stable-expression UCK2 cells were treated with 1 mM analog for indicated times. These results are from two independent cell culture experiments. **j.** Alamar blue assay demonstrates changes to HEK293T cell proliferation in 5mAzUd-treated WT or UCK2-stable expression cells. Percent proliferation was normalized to DMSO control-treated cells (0.5% final) for each respective time point. Data represents independent biological replicates (cell-culture replicates) and error bars represent standard deviation.

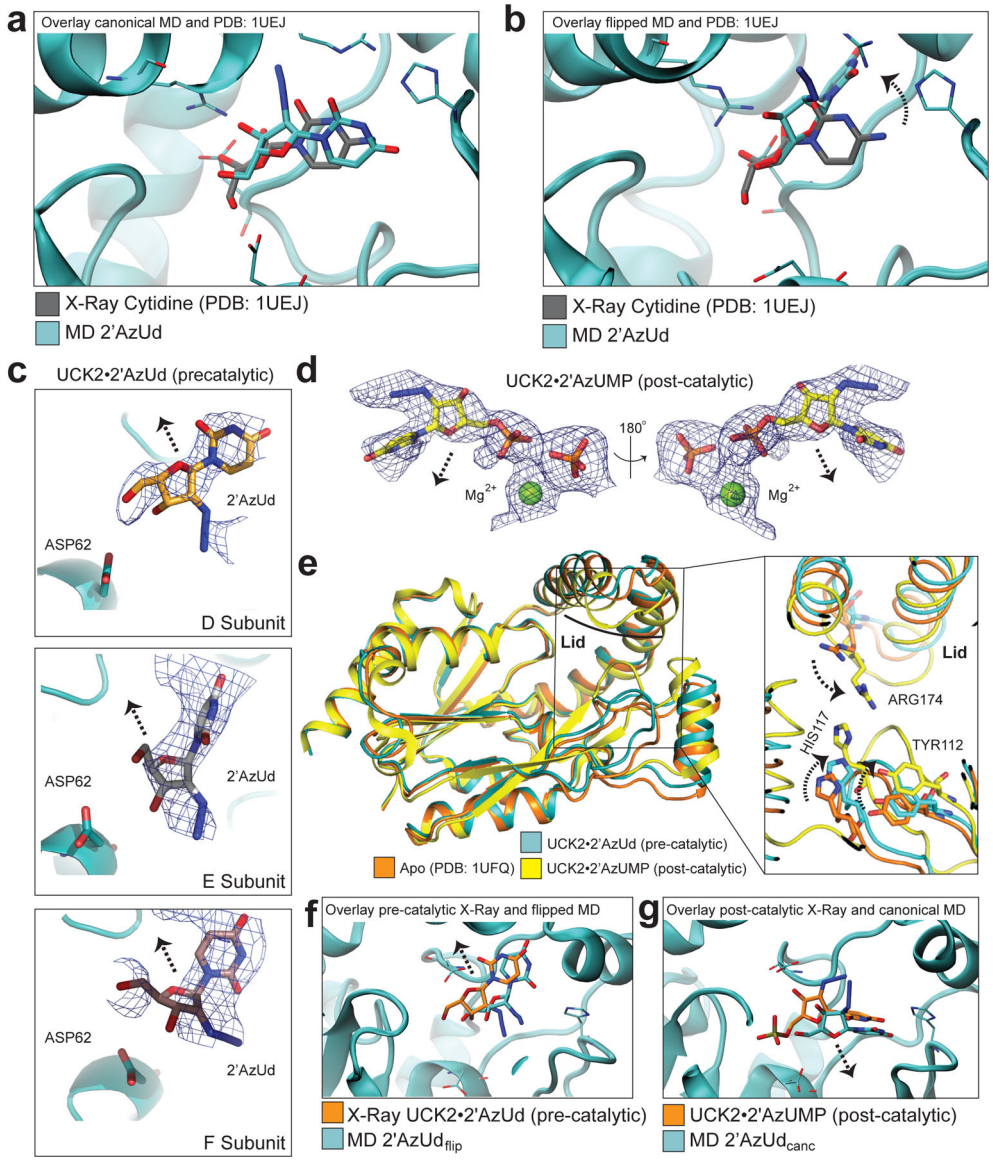

Figure 4: Structural analysis of the UCK2/2'AzUd pair reveals the mechanism for selectivity.

a. Overlay of simulated 2'AzUd in the canonical binding mode ($2'AzUd_{canc}$) and cytidine from PDB: 1UEJ. The ribose oxygen is oriented away from the active site in both the predicted MD conformation and 1UEJ. **b.** Overlay of simulated 2'AzUd in the flipped binding mode ($2'AzUd_{flip}$) and cytidine from PDB: 1UEJ. The ribose of $2'AzUd_{flip}$ in the active site is oriented opposite to that of the canonical binding mode (denoted with a dashed arrow). **c.** Electron density of 2'AzUd in the pre-catalytic states in subunits D, E and F. Each of the orientations for 2'AzUd are from the same view as in a. and b. **d.** Electron density of 2'AzUMP in the post-catalytic state, where the substrate is flipped (facing down and out of the active site) in an orientation opposite that of the pre-catalytic structures. **e.** Superimposition of 2'AzUMP, 2'AzUd and apo-UCK2 structures demonstrating conformational changes in the lid domain and opposing loop containing key

residues in the pre- and post-catalytic states. **f.** Overlay of 2'AzUd$_{flip}$ MD and pre-catalytic 2'AzUd structure. **g.** Overlay of 2'AzUd$_{canc}$ MD and post-catalytic 2AzUMP structure.

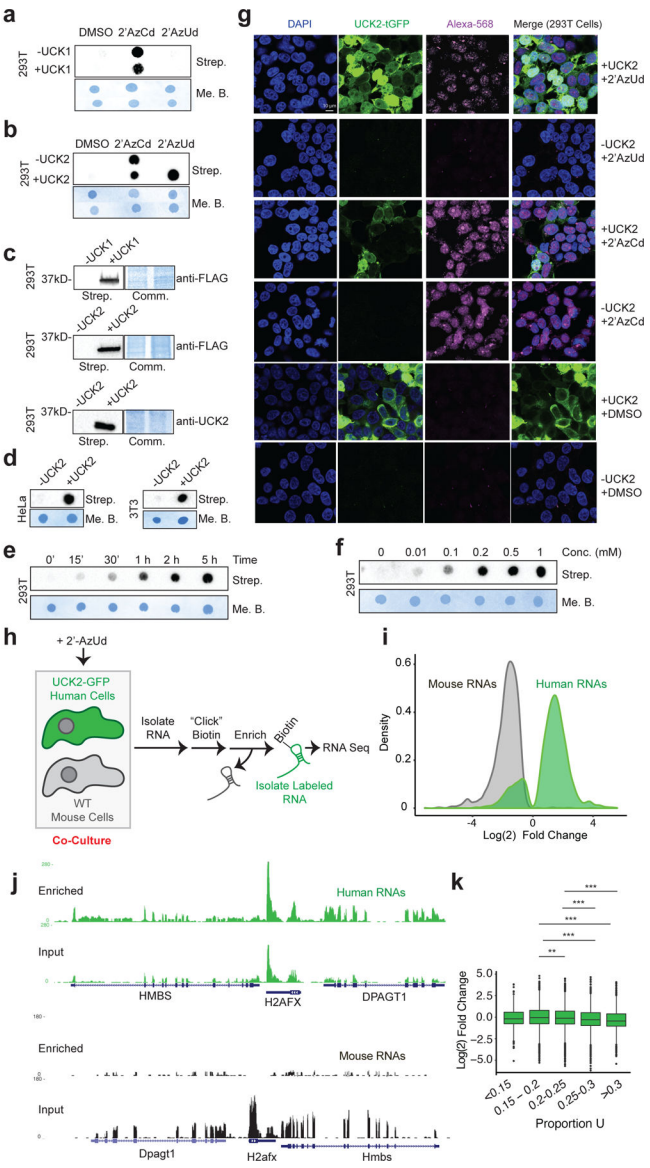

Figure 5: Demonstration of cell-specific metabolic labeling with UCK2/2'AzUd pair.

a. Dot blot demonstrating no incorporation of 2'AzUd in the presence of UCK1-FLAG overexpression. This experiment was independently repeated twice, with similar results. **b.** Dot blot demonstrating incorporation of 2'AzUd only in the presence of UCK2-FLAG overexpression. This experiment was independently repeated twice, with similar results. **c.** Western blots detecting overexpression of UCK1-FLAG and UCK2-FLAG constructs. Overexpression was confirmed with anti-FLAG or additionally UCK2 was also detected via anti-UCK2. **d.** Additional mammalian cells not expressing UCK2 are unable to introduce 2'AzUd into their RNA unless UCK2 is overexpressed. Both HeLa and NIH3T3 experiments were independently repeated twice, with similar results. **e.** Time course analysis of 2'AzUd incorporation into cellular RNA from cells expressing UCK2-FLAG and treated with 1 mM analog. This experiment was independently repeated twice, with similar results. **f.** Concentration titration of 2'AzUd incorporation into cellular RNA in

cells expressing UCK2-FLAG after 5 h. This experiment was independently repeated twice, with similar results. **g.** Cellular imaging demonstrates 2'AzUd is exclusively incorporated into RNA of cells expressing UCK2-TurboGFP (green). HEK293T cells transfected with UCK2-TurboGFP (or mock) were incubated with 1 mM 2'AzUd for 5 h. Cells were fixed, permeabilized and clicked with Alexa-568 alkyne followed by imaging of labeled RNA (red). This experiment was repeated twice, with similar results. **h.** Schematic of co-culture experiment used to test specificity of labeling. **i.** Distribution plot of RNAs enriched and depleted from the co-culture experiment. **j.** Genome browser track of human and mouse loci demonstrating enrichment of human RNAs and depletion of mouse RNAs. **k.** Plot of U-richness in enriched RNAs. Significant comparisons are denoted with a horizontal bar. The sample size is n=3 with ** $p = <0.01$ and *** $p = <0.001$, as determined by Type II ANOVA and Tukey's test.

Nat Methods. Author manuscript; available in PMC 2021 October 15.

Table 1:

Binding and kinetic analysis of UCK1/2 with canonical and 2’-azido analogs.

Values represent the mean of three experiments, with error calculated as standard deviation.

Substrate	Kd (nm)	Kd (nm)	Km (mM)	Km (mM)	Vmax (umol/mg/min)	Vmax (umol/mg/min)	kcat (1/s)	kcat (1/s)	kcat/Km 1/(s*M)	kcat/Km 1/(s*M)
	UCK1	UCK2	UCK1	UCK2	UCK1	UCK2	UCK1	UCK2	UCK1	UCK2
Uridine	56.4 (+/− 11.9)	5.8 (+/− 3.2)	4.06 (+/− 0.82	0.15 (+/− 0.03)	11.56 (+/ −1.42)	78.77 (+/− 6.31)	5.44 (+/− 0.67)	37.40 (+/− 0.46)	1.3 (+/−0.3) X 10^3	2.4 (+/− 0.5) X 10^5
Cytidine	15.9 (+/− 1.1)	3.7 (+/− 0.5)	1.92 (+/− 0.32	0.13 (+/− 0.01)	13.9 (+/− 0.77)	70.61 (+/− 2.49)	6.54 (+/− 0.36)	33.52 (+/− 1.18)	3.4 (+/− 0.06) X 10^3)	2.6 (+/− 0.2) X10^5
2’azidouridine	180.8 (+/− 17.1)	13.4 (+/− 4.8)	n.d.	n.d.	n.d.	2.16 (+/− 0.46)	n.d.	1.03 (+/− 0.22)	n.d.	4.3 (+/− 1.2)
2’azidocytidine	65.7 (+/− 10.6)	17.8 (+/− 6.3)	118 (+/− 62)	205 (+/− 43)	0.23 (+/ −0.10)	3.64 (+/− 0.69)	0.11 (+/− 0.05)	1.73 (+/− 0.33)	0.9 (+/−0.6)	8.4 (+/− 1.8)