Cell Reports, Volume *40*

Supplemental information

N1-methylpseudouridine

found within COVID-19 mRNA vaccines

produces faithful protein products

Kyusik Q. Kim, Bhagyashri D. Burgute, Shin-Cheng Tzeng, Crystal Jing, Courtney Jungers, Junya Zhang, Liewei L. Yan, Richard D. Vierstra, Sergej Djuranovic, Bradley S. Evans, and Hani S. Zaher

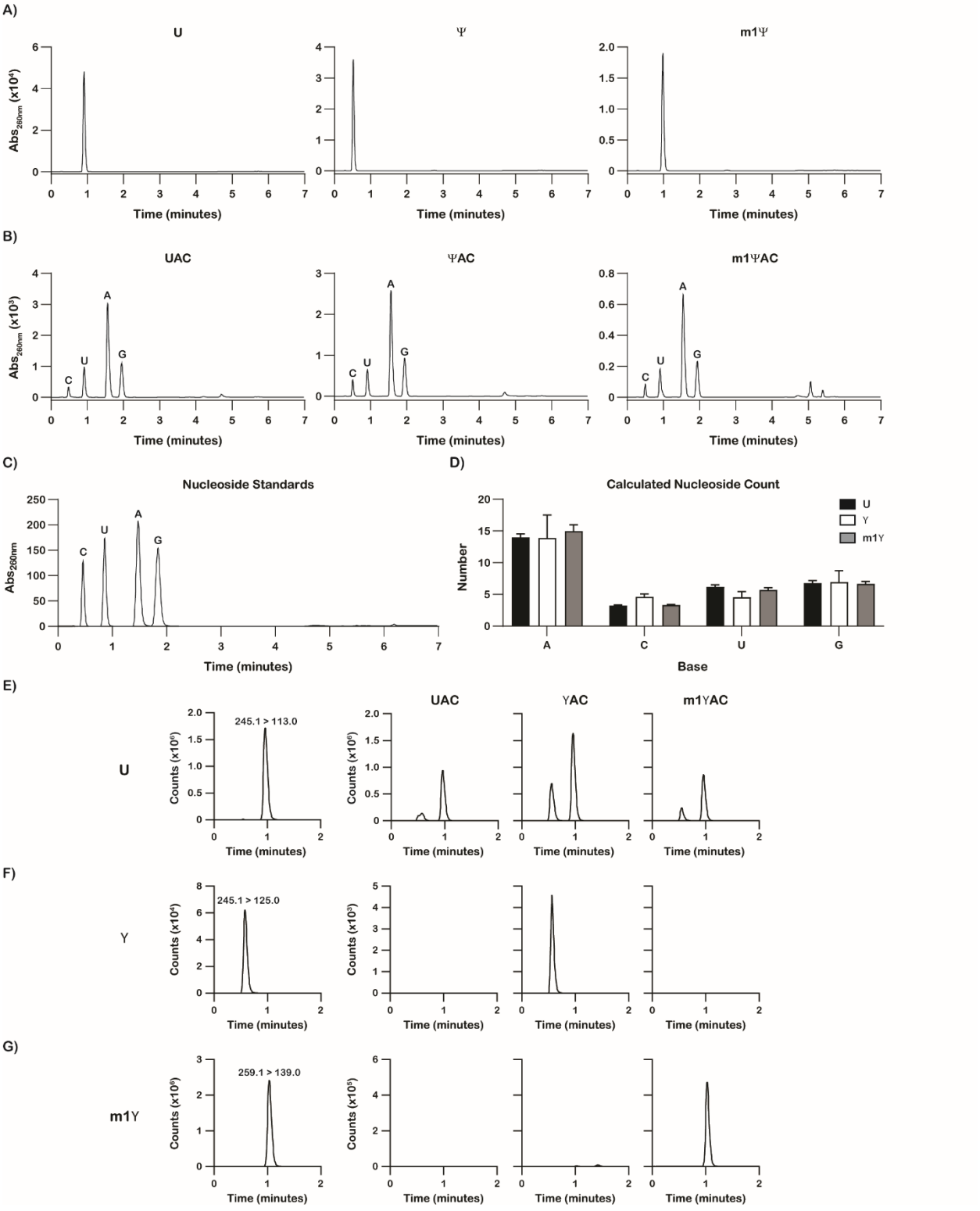
A)
U
Ψ
m1Ψ
Abs260nm (x10⁴)
Time (minutes)
B)
UAC
ΨAC
m1ΨAC
Abs260nm (x10³)
C
U
A
G
Time (minutes)
C)
Nucleoside Standards
Abs260nm
Time (minutes)
D)
Calculated Nucleoside Count
Number
U
Y
m1Y
A
C
U
G
Base
E)
UAC
YAC
m1YAC
U
245.1 > 113.0
Counts (x10⁶)
Time (minutes)
F)
Y
245.1 > 125.0
Counts (x10⁴)
Counts (x10³)
G)
m1Y
259.1 > 139.0
Counts (x10⁶)
Counts (x10⁵)

Figure S1: LC-MS/MS validation of synthetic model RNAs. A) Diode array detector chromatograms of peaks detected at 260 nm for uridine, pseudouridine, and N1-methylpseudouridine standards. **B)** Diode array detector chromatograms of peaks detected at 260 nm for the UAC, YAC, and m1YAC model mRNAs. The identity of each peak is labeled, as confirmed by comparison to canonical nucleoside standards. **C)** Diode array detector chromatogram of peaks detected at 260 nm for the four canonical ribonucleosides. An equimolar amount of cytidine, uridine, adenosine, and guanosine were run as standards. **D)** Plot showing the calculated number of each nucleoside in each model mRNA. Plotted are the average of the three calculated counts with error bars representing the standard deviation around the mean. Pseudouridine co-eluted with cytidine, which is reflected in the loss of a U and a gain of a C for the ΨAC mRNA. **E)** On the left is a plot of the mass spectrometry counts for the uridine standard; the mass transition used to detect the ion product of the fragmentation reaction is denoted above the peak. On the right are plots of the mass spectrometry counts for the three model mRNAs at the given mass transition. **F)** On the left is a plot of the mass spectrometry counts for a pseudouridine standard; the mass transition used to detect the ion product of the fragmentation reaction is denoted above the peak. On the right are plots of the mass spectrometry counts for the three model mRNAs at the given mass transition. **G)** On the left is a plot of the mass spectrometry counts for the N1-methylpseudouridine standard; the mass transition used to detect the ion product of the fragmentation reaction is denoted above the peak. On the right are plots of the mass spectrometry counts for the three model mRNAs at the given mass transition. Related to Figure 1.

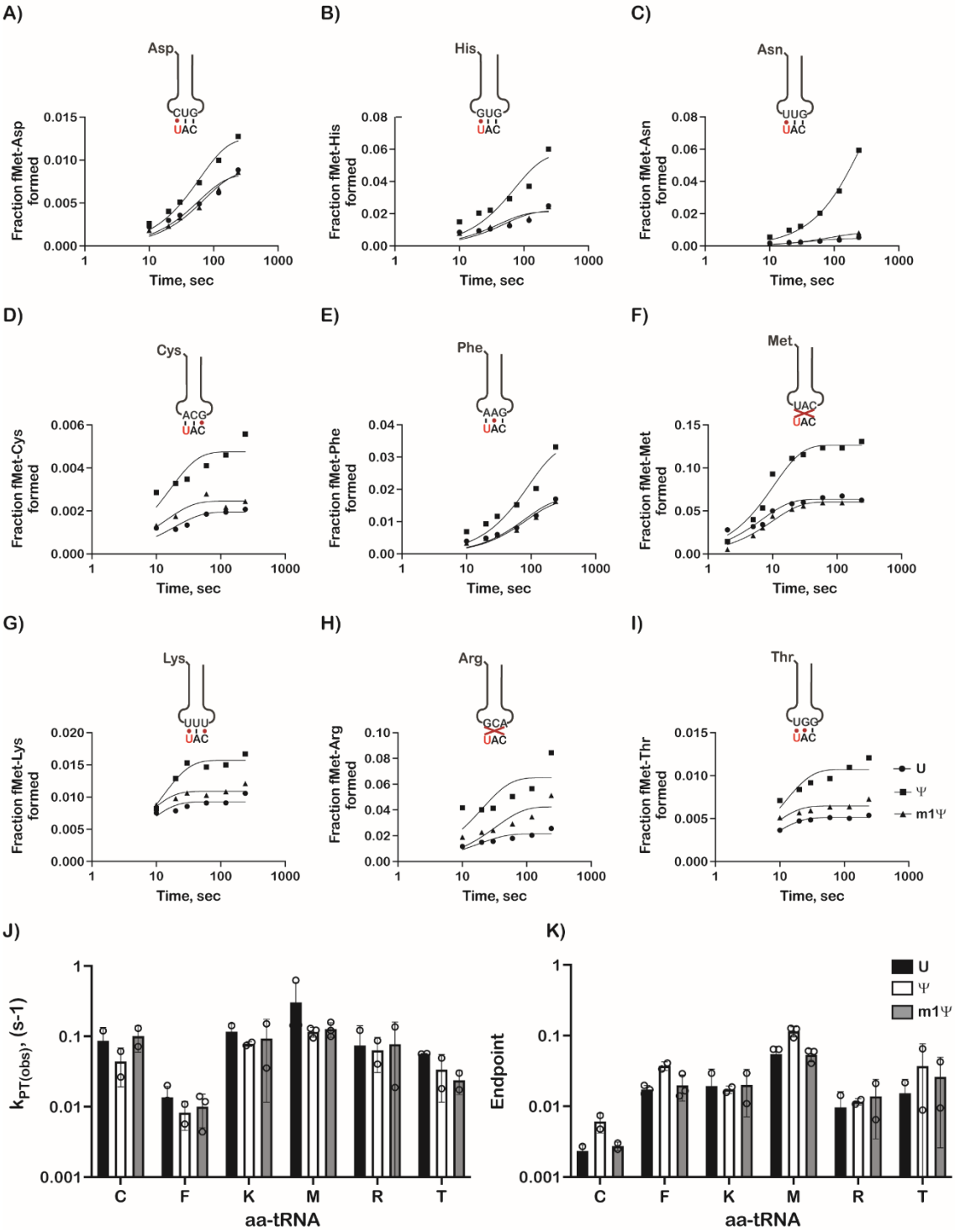

Figure S2: Pseudouridine increases misincorporation of near and non-cognates while N1-methylpseudouridine does not. A-I) Representative time courses of the indicated near and non-

cognate tRNA ternary complexes and UAC, ΨAC, or Met-m1ΨAG initiation complexes. The codon (UAC; modification in red), near cognate tRNA, and the mismatch (red dot) are indicated. All reactions were conducted at least in duplicates. **J-K)** Bar graph showing the measured observed rates of peptide-bond formation and reaction end points, respectively, in the presence of 1 μM initiation complex and 2.5 μM denoted ternary complex. Plotted are the average values with error bars representing the standard deviation around the mean. Related to Figure 3.

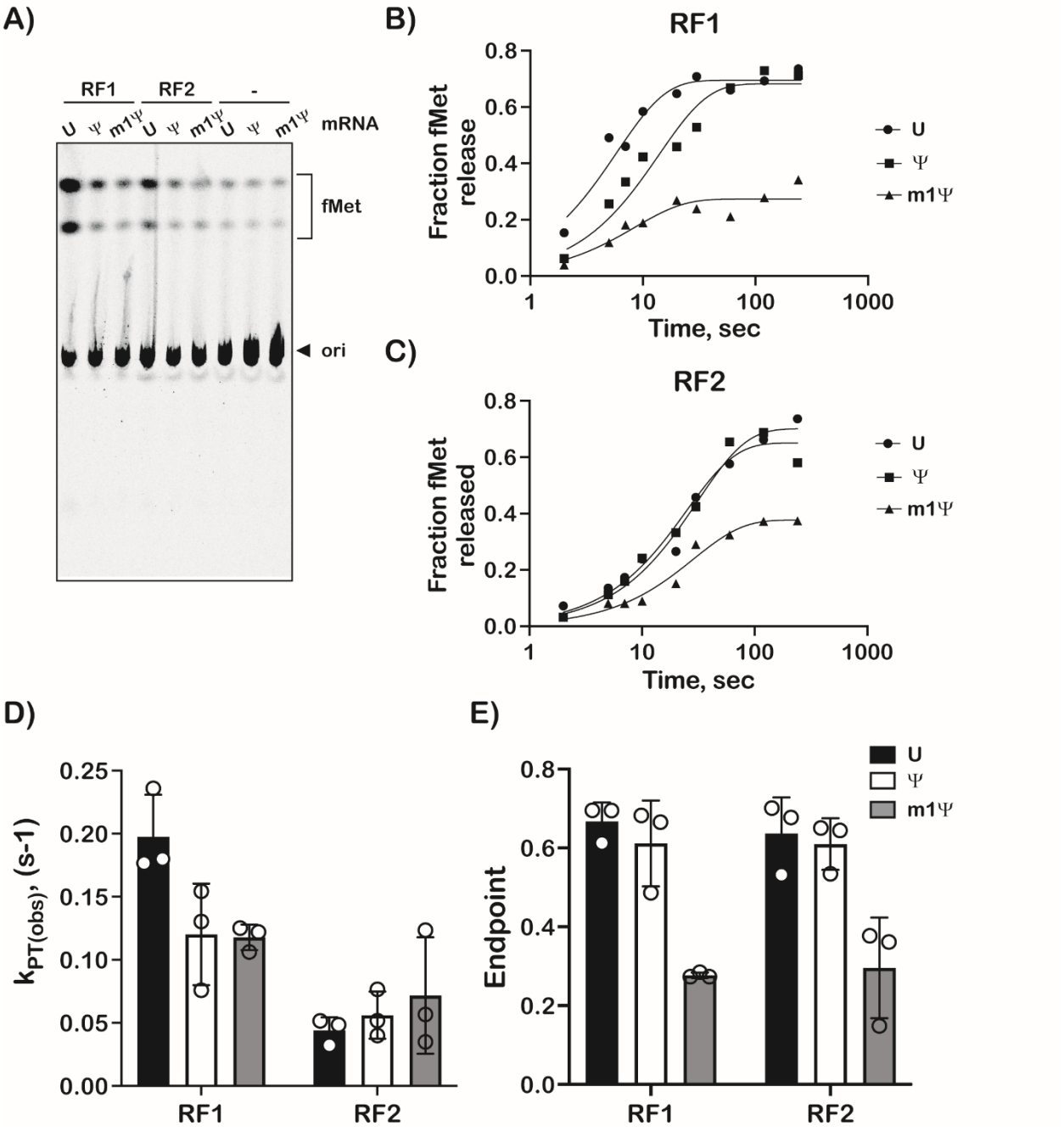

Figure S3: The presence of N1-methylpseudouridine in mRNA increases the accuracy of stop-codon recognition by release factors. A) A representative phosphorimage of an electrophoretic TLC showing peptide release on near-stop codon UAC, ΨAC, and m1ΨAC in the presence and

absence of RF1 and RF2. All reactions were conducted in at least duplicates. **B-C)** Kinetics of f-Met peptide release on an unmodified and modified codon; UAC codon (circles), ΨAC (squares), m1ΨAC (Triangles) in the presence of RF1 **(B)** and RF2 **(C)**. Representative time courses are shown. **D-E)** Bar graph showing the measured observed rates of peptide-bond formation and reaction end points, respectively, in the presence of 1 μM initiation complex and 2.5 μM denoted ternary complex. Plotted are the average values determined from three independent time courses with error bars representing the standard deviation around the mean. Related to Figure 3.

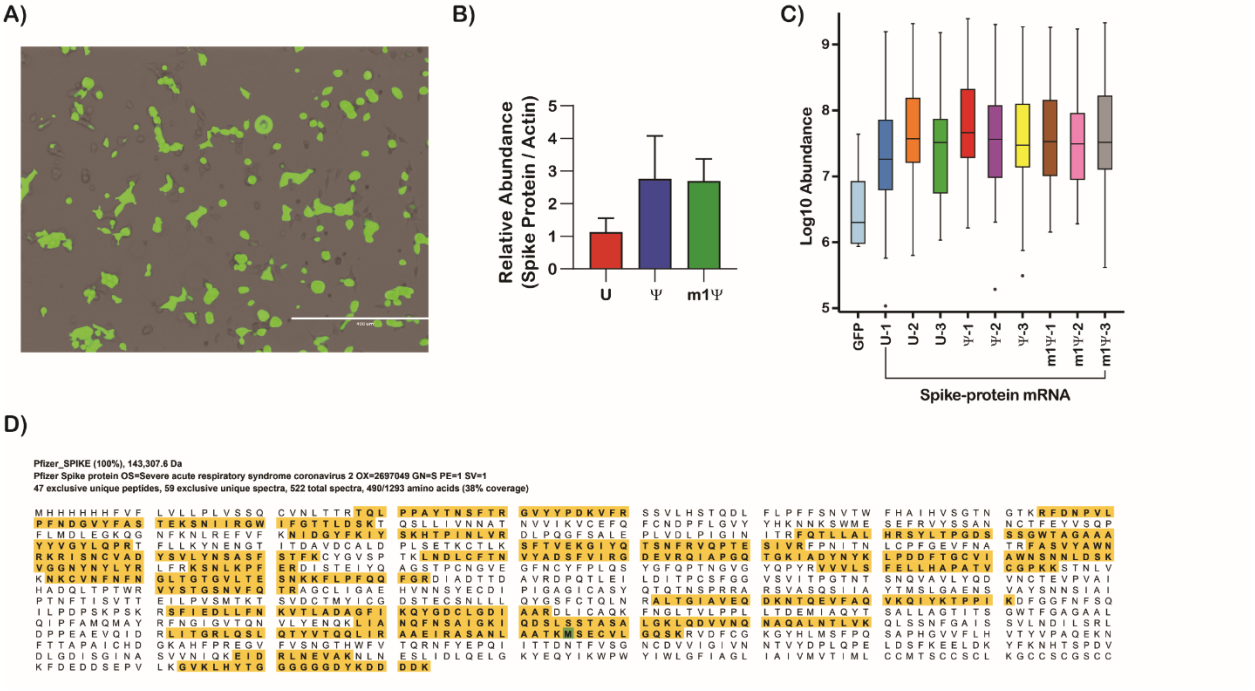

Figure S4: Amino acid substitution products from HEK 293 cells can be detected by mass spectrometry A) Microscope image of HEK 293 cells transfected with EGFP. A white scale bar is included for reference and represents a distance of 400 μm. **B)** Relative abundances of spike protein translated from mRNAs containing the denoted modification, as quantified from the immunoblot. **C)** Normalized abundance of peptides from each sample matching the wild-type spike protein sequence, as determined by label-free quantitation from LC-MS/MS analysis. **D)** Coverage of the spike protein resulting from on-bead digestion and MS/MS analysis. Amino-acids matched to a MS/MS spectrum are highlighted in yellow while amino-acids highlighted in green denote a post-translational modification (ie. Oxidation). Related to Figure 4.

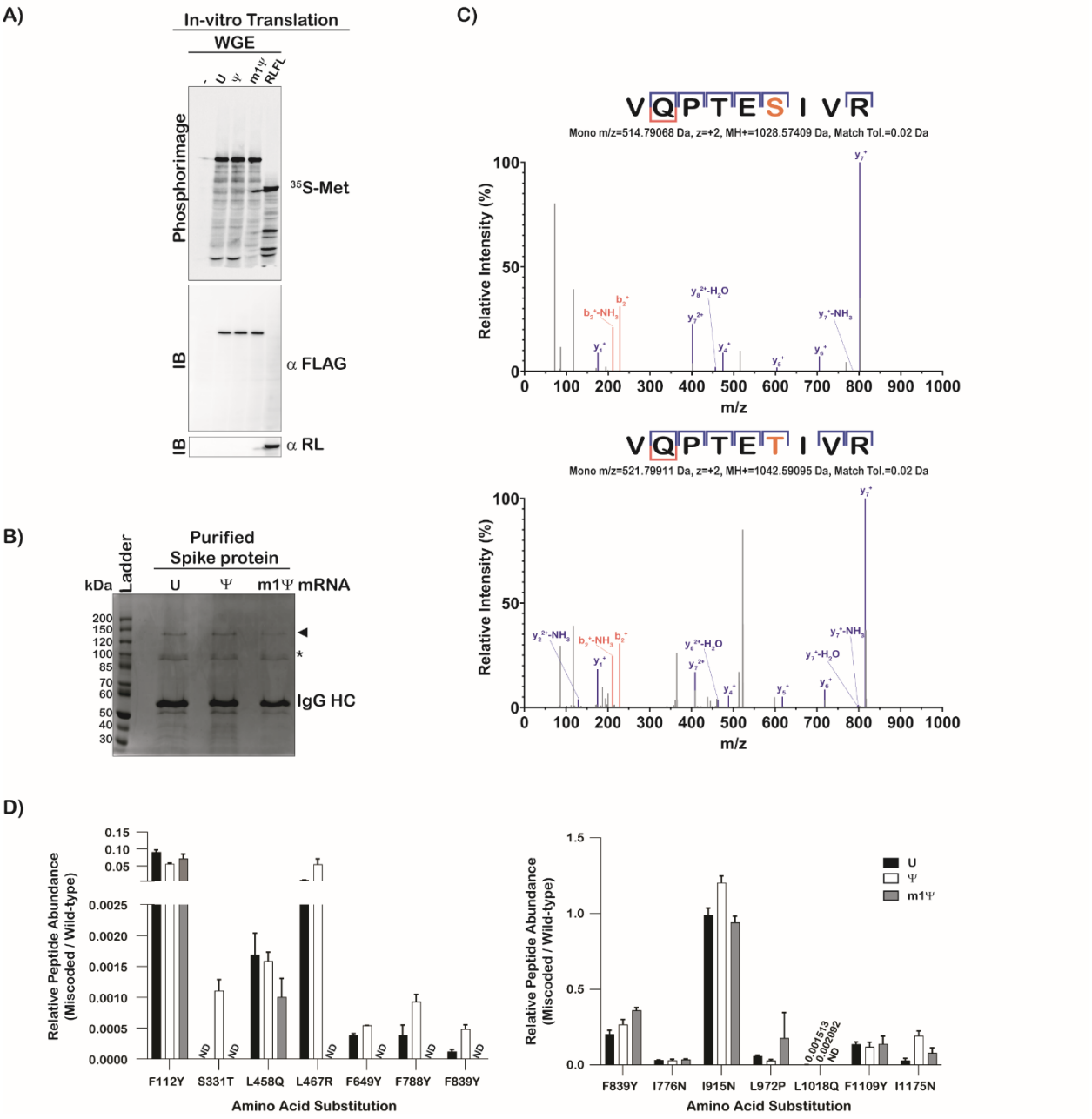

Figure S5: Unlike Ψ-containing SARS-CoV-2 spike protein mRNA, m1Ψ-containing spike protein mRNA does not increase amino-acid misincorporation frequency during translation in wheat-germ extracts. A) At top is a phosphorimage of a PVDF membrane with protein products transferred from the SDS-polyacrylamide gel used to resolve them from the indicated cell-free

translation reactions in the presence of the depicted mRNA constructs. In-vitro protein synthesis was monitored through the addition of [^{35}S]-Methionine to the reactions. At bottom are immunoblotting analysis of the same PVDF membrane with the indicated antibodies. **B)** Image of a Coomassie-stained SDS-polyacrylamide gel used to assess the purity of FLAG-immunoprecipitated spike proteins produced from translation reactions containing the indicated modification. **C)** Fragmentation spectra of a wild-type sequence peptide and its substituent miscoded product. b and y ions are denoted in red and blue, respectively, while the substituted amino acid is denoted in orange. The difference in the m/z for the y_4, y_5, y_6, y_7, and y_8 peaks between the wild-type peptide (top) and miscoded product (bottom) corresponds to a serine to threonine substitution. The nominal mass difference between serine and threonine is 14 Da.

D) Plots showing the relative peptide abundance of each detected wild-type peptide and its substituent miscoded product. Plotted are the means from technical replicates with error bars representing the standard deviation around the mean. The plot on the left displays the results of the first biological replicate mass spectrometry run, while the plot on the right displays the results of the second run. ND denotes the miscoded product was not detected in either technical replicate. Related to Figure 4.

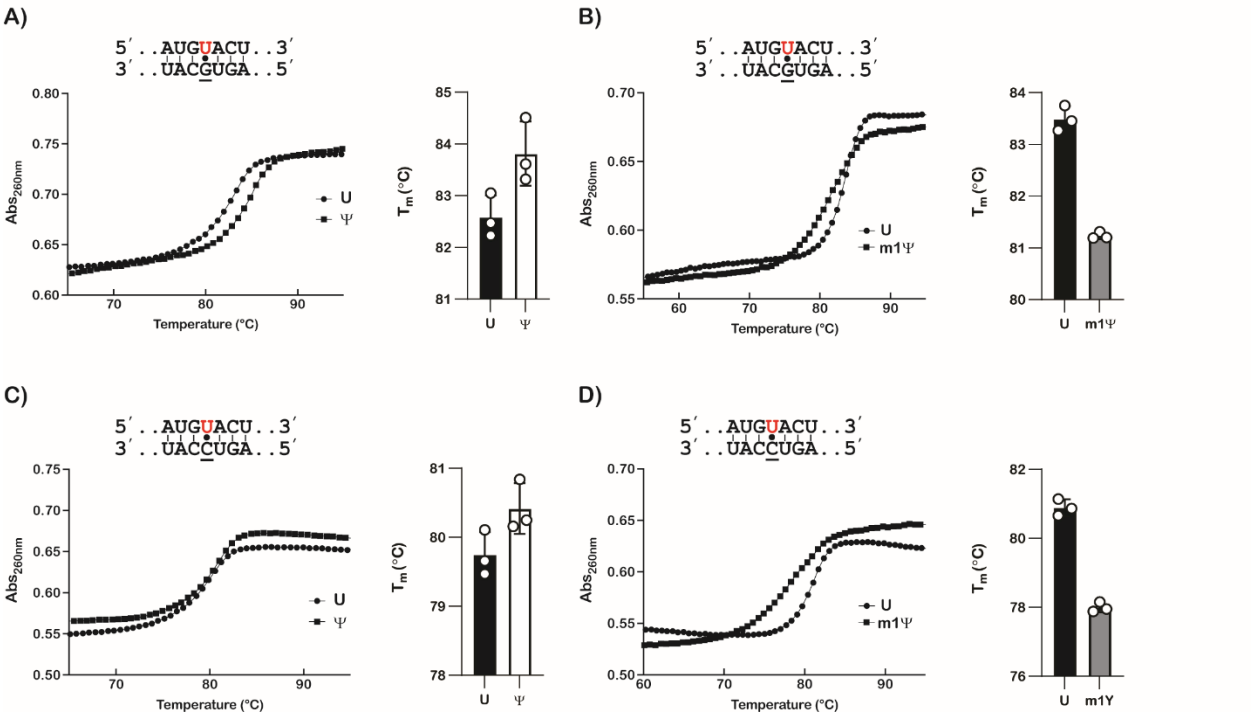

Figure S6: Pseudouridine stabilizes formation of mismatched RNA duplexes whereas N1-methylpseudouridine does not. A-D) Scatterplots showing the change in absorbance at 254 nm as a function of temperature for the indicated duplexes, with an accompanying bar graph showing the determined melting temperature for the same duplexes. Plotted are the means of the calculated melting temperature from three replicate experiments, with the error bars representing the standard deviation around the mean. **A, C)** correspond to duplexes containing Ψ while **B, D)** correspond to duplexes containing m1Ψ. The modified base is denoted in red. Related to Figure 5.