Supplementary Information (SI) for Chemical Science.

This version of the ESI published 22/04/2025 replaces the previous version published 22/04/2025. A typographical error in Figure S9 has been corrected

Supporting Information

Removing redundancy of the NCN codons in vitro for maximal sense codon reassignment

C. A. Jones, C. A. Makovsky, A. K. Haney, A. C. Dutra, C. A. L. McFeely, T. A. Cropp, M. C. T. Hartman

Table of Contents

Figures

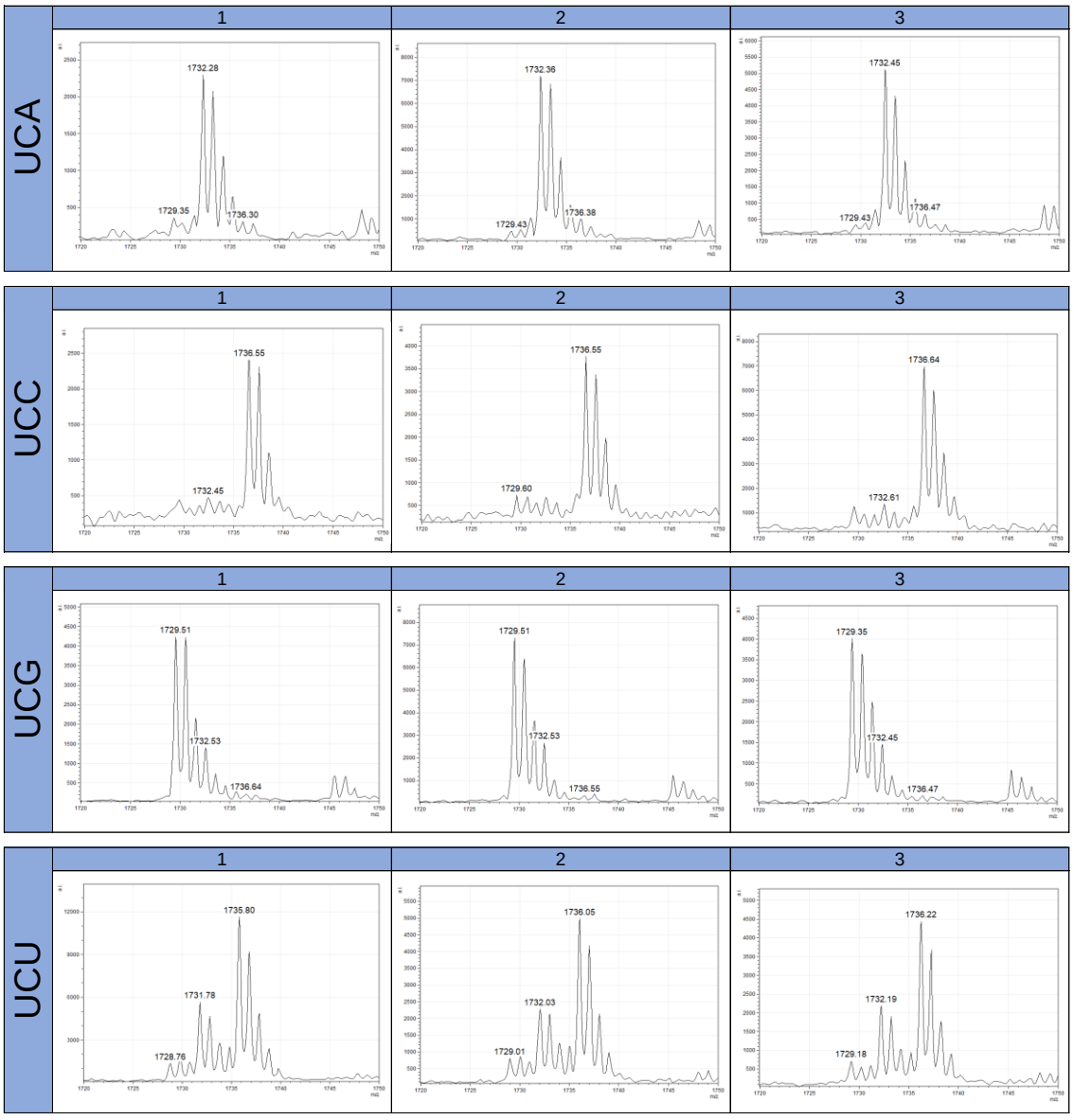
Methods

Tables

Characterization of novel ncAAs

Characterization of reported ncAAs

Supporting Figures:



tRNA	% read average	stdev
$Ser2^{CGA}$	6.30%	3.4
$Ser1^{UGA}$	90%	1.9
$Ser5^{GGA}$	3.50%	1.6

tRNA	% read average	stdev
$Ser2^{CGA}$	10.50%	1.3
$Ser1^{UGA}$	8%	1.8
$Ser5^{GGA}$	81.40%	2

tRNA	% read average	stdev
$Ser2^{CGA}$	87.50%	1.4
$Ser1^{UGA}$	11%	1.6
$Ser5^{GGA}$	1.30%	0.3

tRNA	% read average	stdev
$Ser2^{CGA}$	8.20%	0.9
$Ser1^{UGA}$	29%	1.4
$Ser5^{GGA}$	62.80%	1.1

Figure S1: Serine Isotopic competition data. Mass spectra of resulting peptides from serine isotopic competition assay. Codon reading percentage was calculated as (isotope intensity / total peptide intensity) *100. For peaks less than 3 Da apart, the expected isotopic contribution of the lower mass peak was subtracted from the peak height of the higher isotopic mass peak. In vitro translations occurred in triplicate at 20 µL final volume. Aminoacylated tRNAs were introduced at 5 µM each. mRNAs 1,3,4,17 (Table S3) were added (each mRNA containing a different UCN codon) to a final concentration of 1 µM, and the reactions were incubated at 37 °C for 30 min.

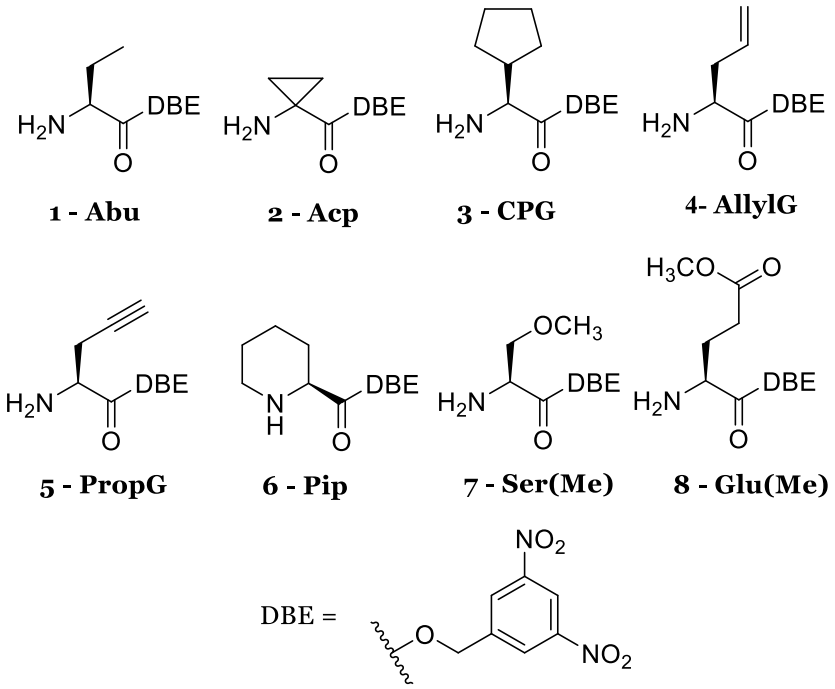
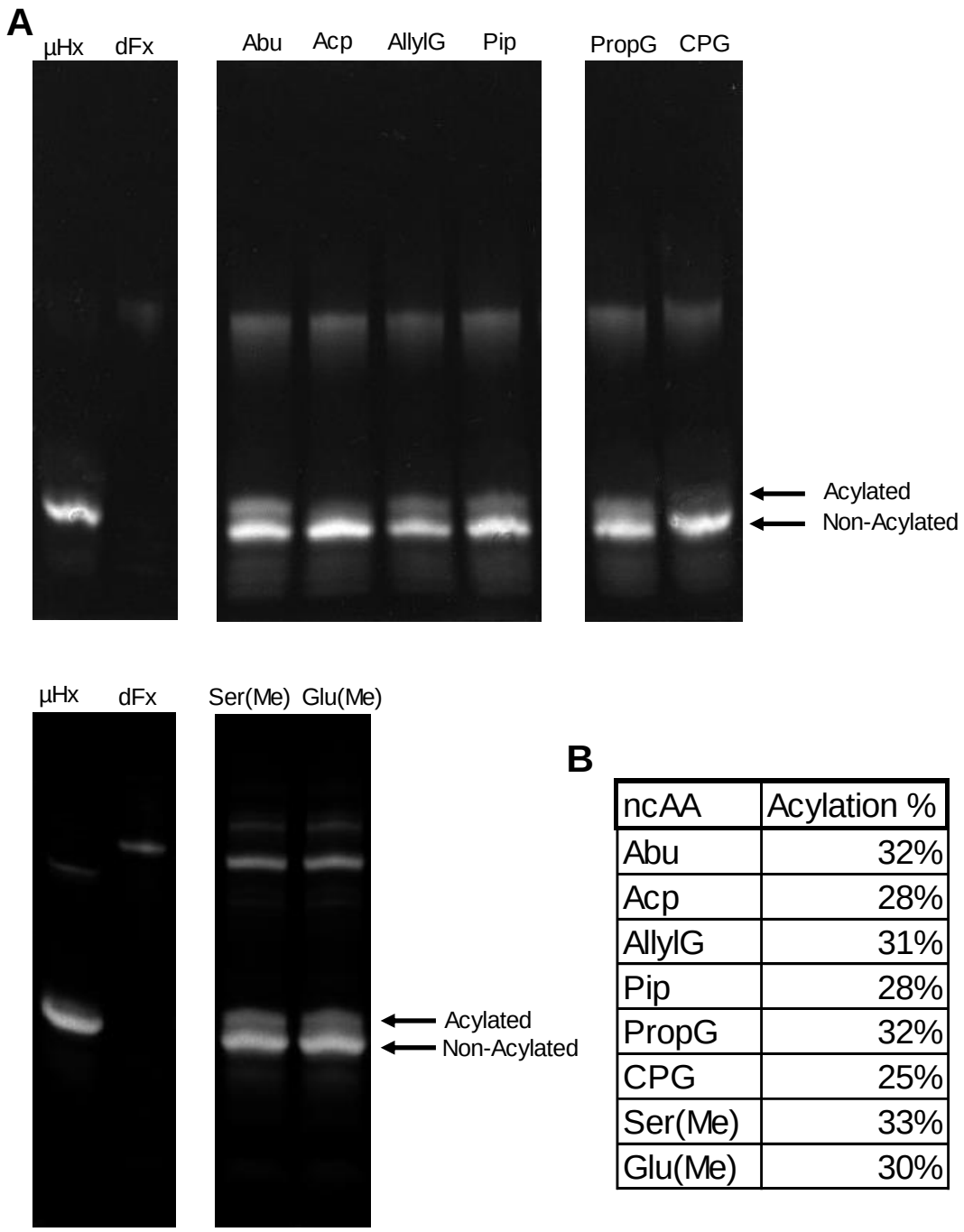


Figure S2: Non-canonical amino acids used in this study. ncAAs contain a dinitrobenzyl ester activating group allowing them to be recognized by the flexizyme for aminoacylation to tRNAs and microhelix RNAs.

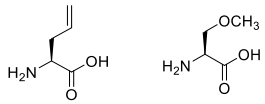


B

ncAA	Acylation %
Abu	32%
Acp	28%
AllylG	31%
Pip	28%
PropG	32%
CPG	25%
Ser(Me)	33%
Glu(Me)	30%

Figure S3: Flexizyme acylations. (A) Acylation of ncAAs **1-6** onto microhelix RNA via flexizymes, taken at 24 hours. Both aminoacylated and non-aminoacylated microhelix RNA is labeled. Gel bands were quantified on BioRad Image Lab 6.0 and intensity comparisons were calculated. (B) Percent acylation of various ncAAs with microhelix RNAs via flexizymes at 24 hours.

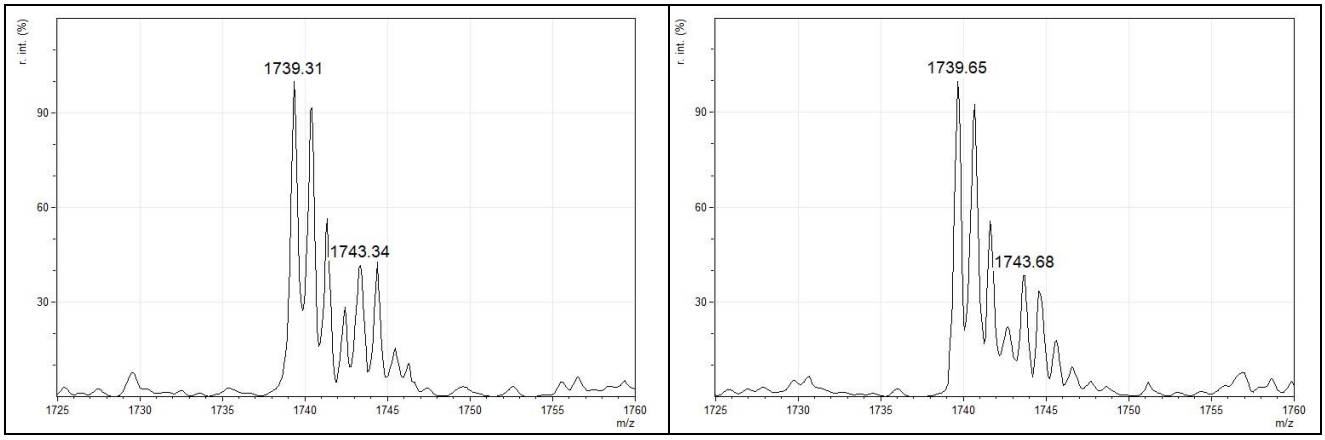
A



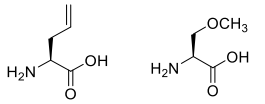
1:1 Ser5^{GGA} : Ser1^{UGA}

tRNA	Avg. Encoding %	stdev
Ser5^{GGA}	75.2	0.93
Ser1^{UGA}	24.8	0.93

Expected Ser5^{CGA} [M+H]$^{+}$: 1739.74
Expected Ser1^{UGA} [M+H]$^{+}$: 1744.02



B



2:1 Ser5^{GGA} : Ser1^{UGA}

tRNA	Avg. Encoding %	stdev
Ser5^{GGA}	89.9	1.34
Ser1^{UGA}	10.1	1.34

Expected Ser5^{CGA} [M+H]$^{+}$: 1739.74
Expected Ser1^{UGA} [M+H]$^{+}$: 1744.02

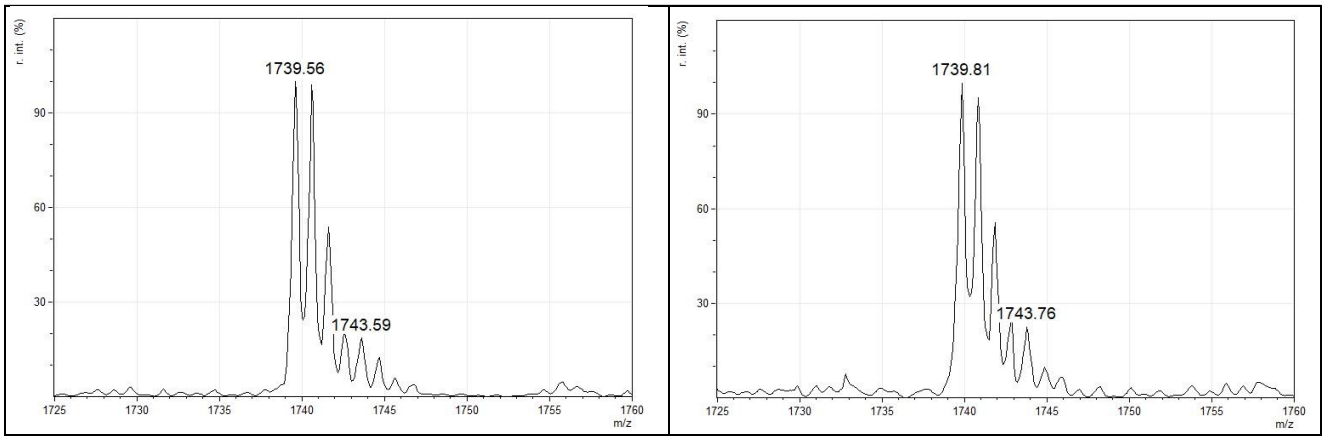


Figure S4: Competition at the UCU codon for translation with different Ser5^{GGA} : Ser1^{UGA} ratios. Both tRNAs competed for incorporation on the UCU codon on mRNA 4. Encoding percentages were calculated by (observed peak intensity) / (sum of both peak intensities) and assumed equivalent ionization efficiencies. To calculate Ser(Me) peak intensity, the expected isotopic contribution of the AllylG peak was subtracted from ser(Me)'s mass peak. (A) Mass spectrometry of ncAA competition containing 15 μM AllylG- Ser5^{GGA} and 15 μM Ser(Me)- Ser1^{UGA}, and their respective encoding percentages. Ser5^{GGA} was observed to encode at a rate of 75.2%, and Ser1^{UGA} was observed to encode at a rate of 24.8%. (B) Mass spectrometry of ncAA competition containing 30 μM AllylG- Ser5^{GGA} and 15 μM Ser(Me)- Ser1^{UGA}, and their respective encoding percentages. This change in ratio from 1:1 to 2:1 increased Ser5^{GGA} selectivity of the UCU codon from 75.2% to 89.9%.

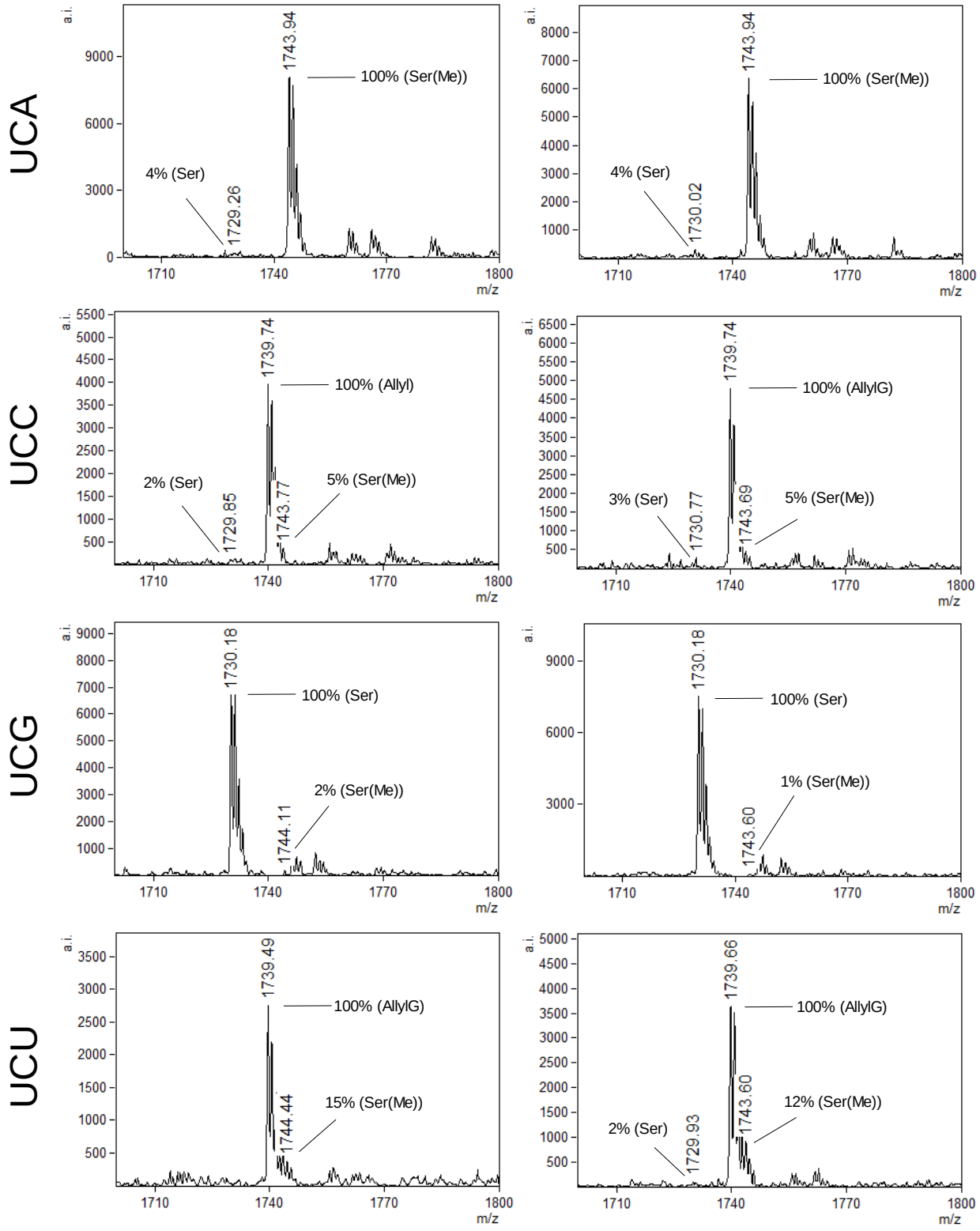
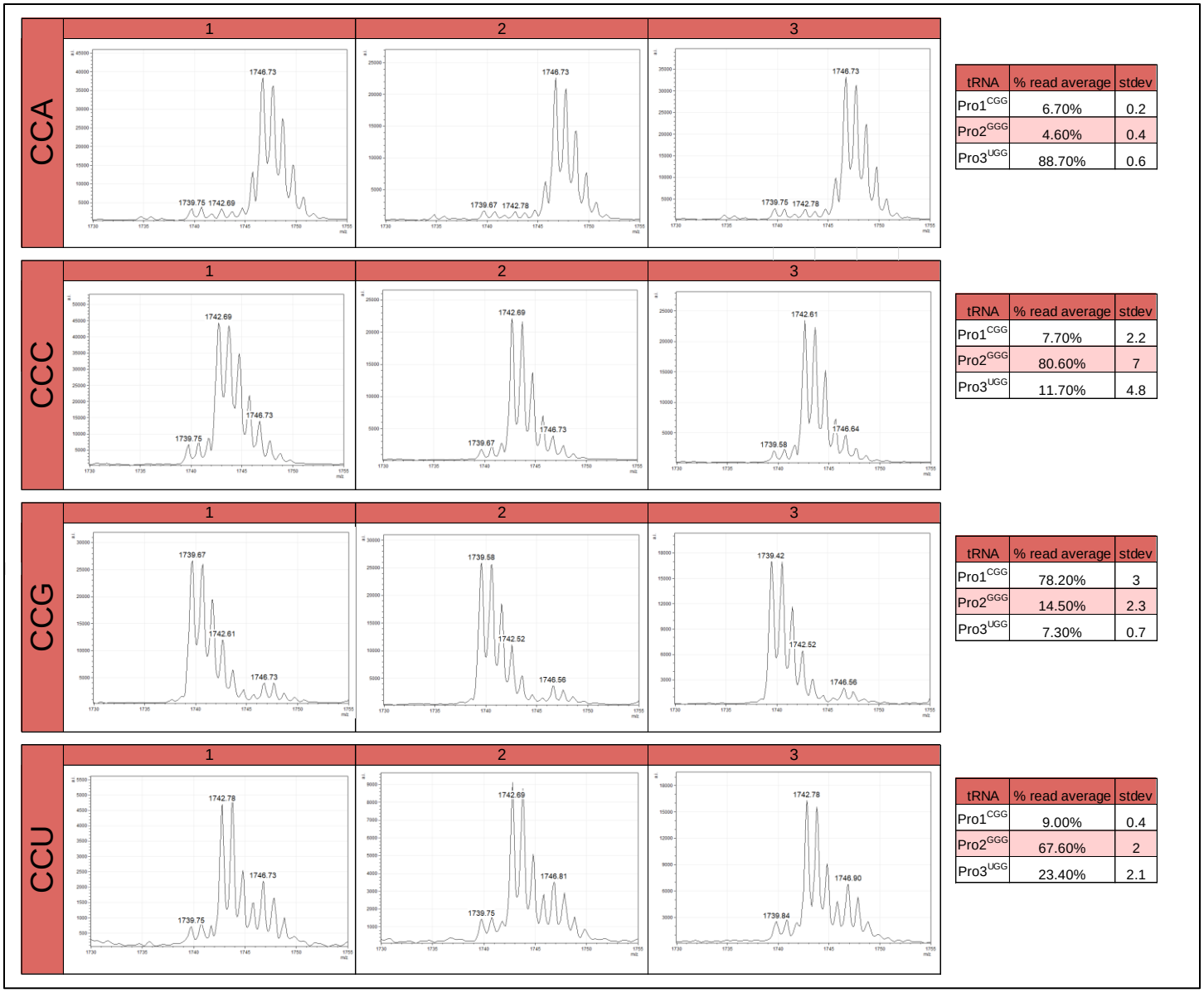


Figure S5: Additional MALDI-MS results for ncAA incorporation for mRNAs 1-4. Assays contained 15 μM Ser(Me)-Ser1^{UGA}, 15μM Ser-Ser2^{CGA}, and 30μM AllylG-Ser5^{GGA}. Translations were incubated for 30 min at 37 °C. To highlight orthogonality, relative peak heights for misincorporation peaks are labeled with the base peak being 100%.



tRNA	% read average	stdev
$Pro1^{CGG}$	6.70%	0.2
$Pro2^{GGG}$	4.60%	0.4
$Pro3^{UGG}$	88.70%	0.6

tRNA	% read average	stdev
$Pro1^{CGG}$	7.70%	2.2
$Pro2^{GGG}$	80.60%	7
$Pro3^{UGG}$	11.70%	4.8

tRNA	% read average	stdev
$Pro1^{CGG}$	78.20%	3
$Pro2^{GGG}$	14.50%	2.3
$Pro3^{UGG}$	7.30%	0.7

tRNA	% read average	stdev
$Pro1^{CGG}$	9.00%	0.4
$Pro2^{GGG}$	67.60%	2
$Pro3^{UGG}$	23.40%	2.1

Figure S6: Proline Isotopic competition data. Mass spectra of resulting peptides from proline isotopic competition assay. Codon reading percentage was calculated as (isotope intensity / total peptide intensity) *100. For peaks less than 3 Da apart, the expected isotopic contribution of the lower mass peak was subtracted from the peak height of the higher isotopic mass peak. In vitro translations occurred in triplicate at 15 µL final volume. Isotopically aminoacylated tRNAs were introduced at 5 µM each. mRNAs **5-8** were added (each mRNA containing a different CCN codon) to a final concentration of 1 µM, and the reactions were incubated at 37 °C for 30 min.

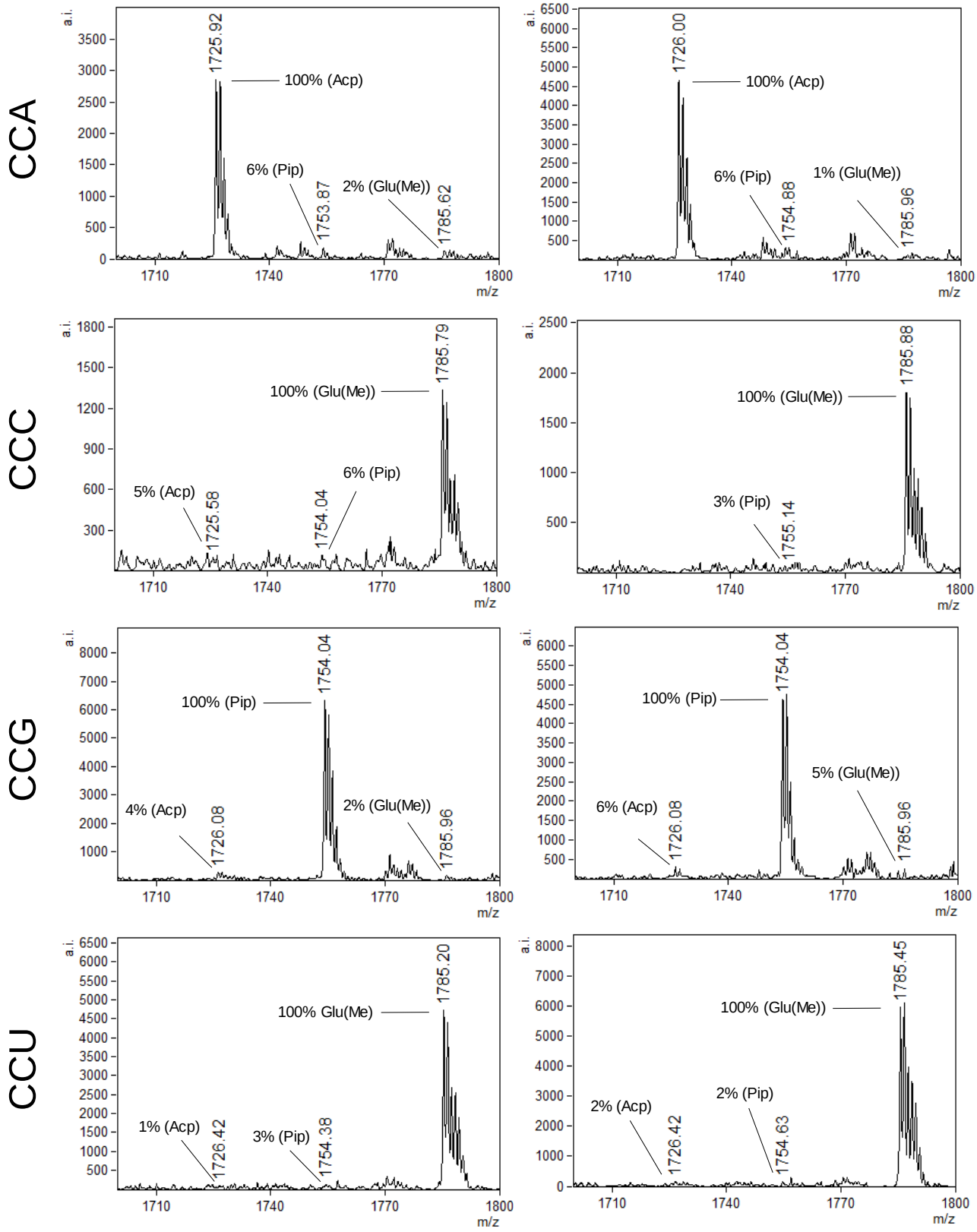


Figure S7: Additional MALDI-MS results for ncAA incorporation for mRNAs 5-8. Assays contain 15μM each of Pip-Pro1^{UGA}, Glu(Me)-Pro2^{CGA}, and Acp-Pro3^{GGA}. Translations were incubated for 30 min at 37 °C. To highlight orthogonality, relative peak heights for misincorporation peaks are labeled with the base peak being 100%.

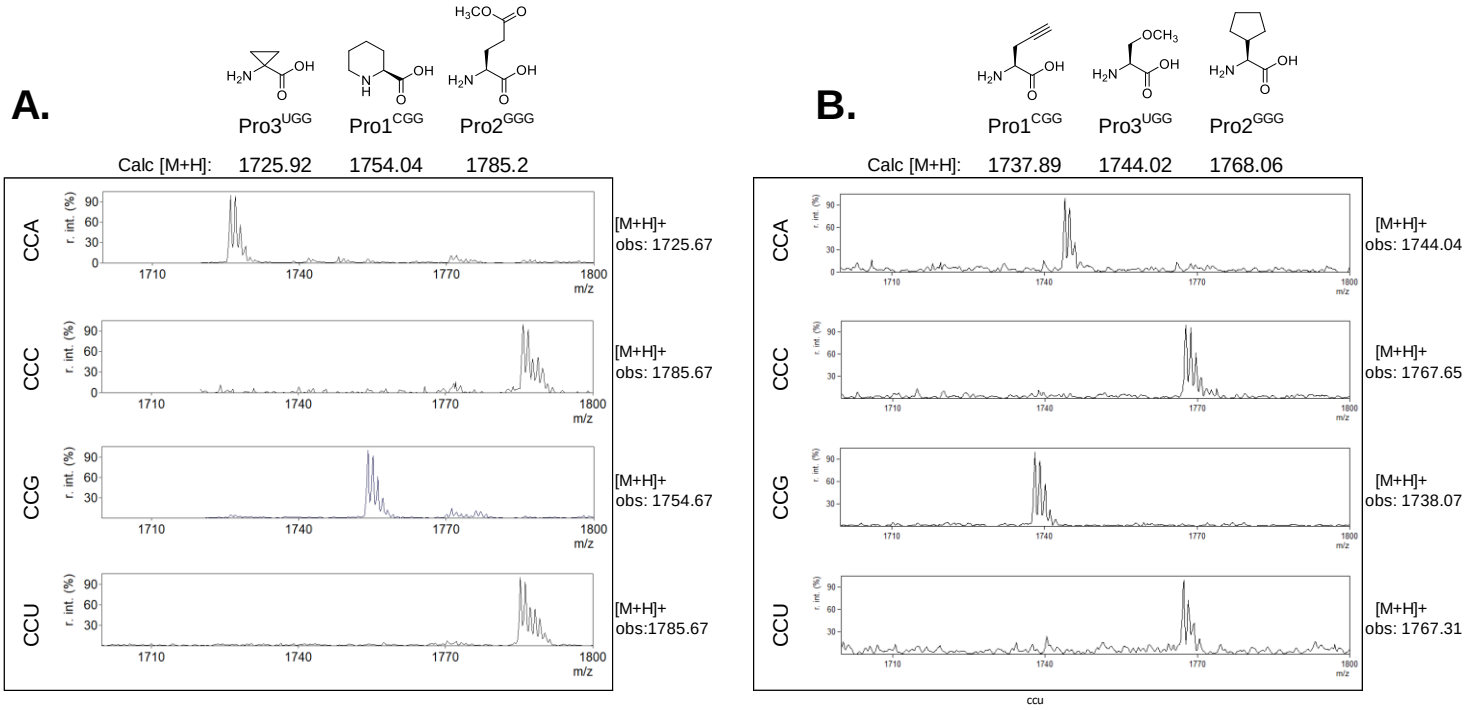
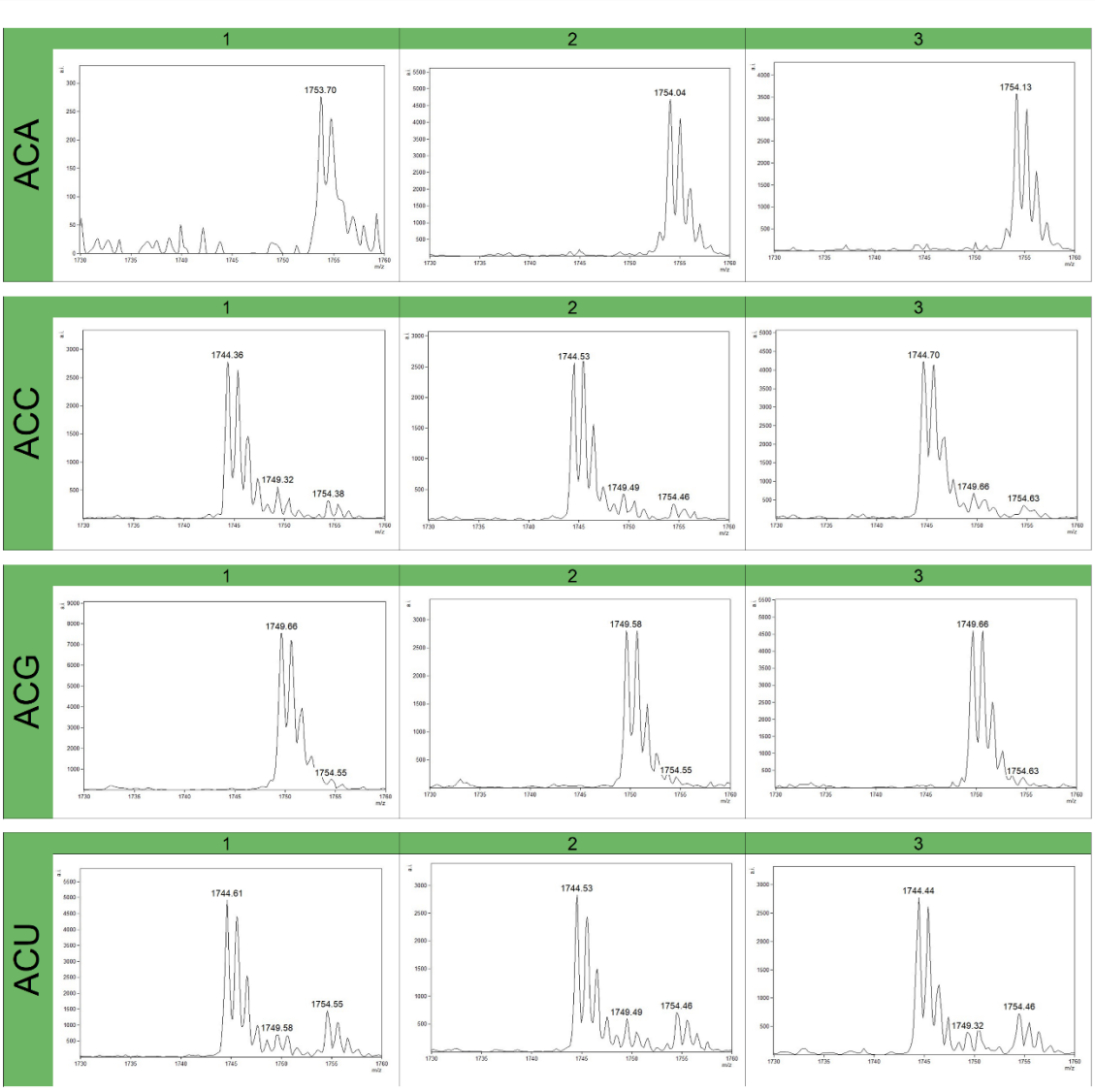


Figure S8: ncAA competition showing effects of ncAA substrate on codon selectivity. A. Codon competition MS data using ncAA-tRNAs Acp-Pro3, Pip-Pro1, and Glu(Me)-Pro2. B. Codon competition MS data using ncAA-tRNAs PropG-Pro1, Ser(Me)-Pro3, and CPG-Pro2. Translations were carried out for 30 min at 37 ºC.



tRNA	% read average	stdev
$Thr1^{GGU}$	4.00%	1.1
$Thr2^{CGU}$	3.00%	1.2
$Thr4^{UGU}$	93.00%	2.3

tRNA	% read average	stdev
$Thr1^{GGU}$	81.60%	3.5
$Thr2^{CGU}$	11.20%	2.6
$Thr4^{UGU}$	7.20%	0.9

tRNA	% read average	stdev
$Thr1^{GGU}$	0.60%	0.3
$Thr2^{CGU}$	91.30%	1.8
$Thr4^{UGU}$	8.20%	2.1

tRNA	% read average	stdev
$Thr1^{GGU}$	70.90%	0.9
$Thr2^{CGU}$	10.10%	2.6
$Thr4^{UGU}$	18.90%	1.8

Figure S9: Threonine Isotopic competition data. Mass spectra of resulting peptides from threonine isotopic competition assay. Codon reading percentage was calculated as (isotope intensity / total peptide intensity) *100. For peaks less than 3 Da apart, the expected isotopic contribution of the lower mass peak was subtracted from the peak height of the higher isotopic mass peak. In vitro translations occurred in triplicate at 15 µL final volume. Isotopically aminoacylated tRNAs are introduced at 5 µM each. mRNAs **9-12** are added (each mRNA containing a different ACN codon) to a final concentration of 1 µM, and the reactions were incubated at 37 °C for 30 min.

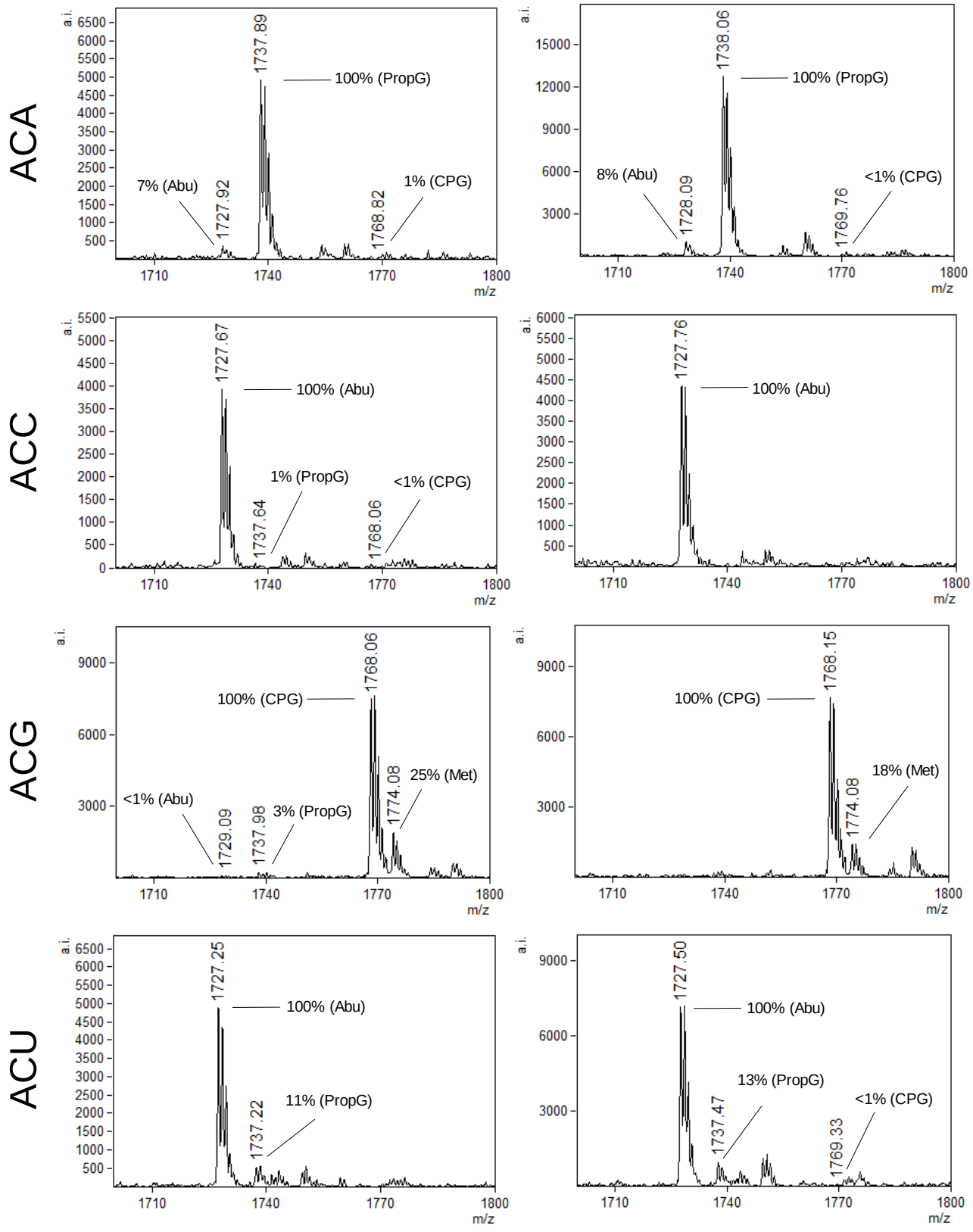
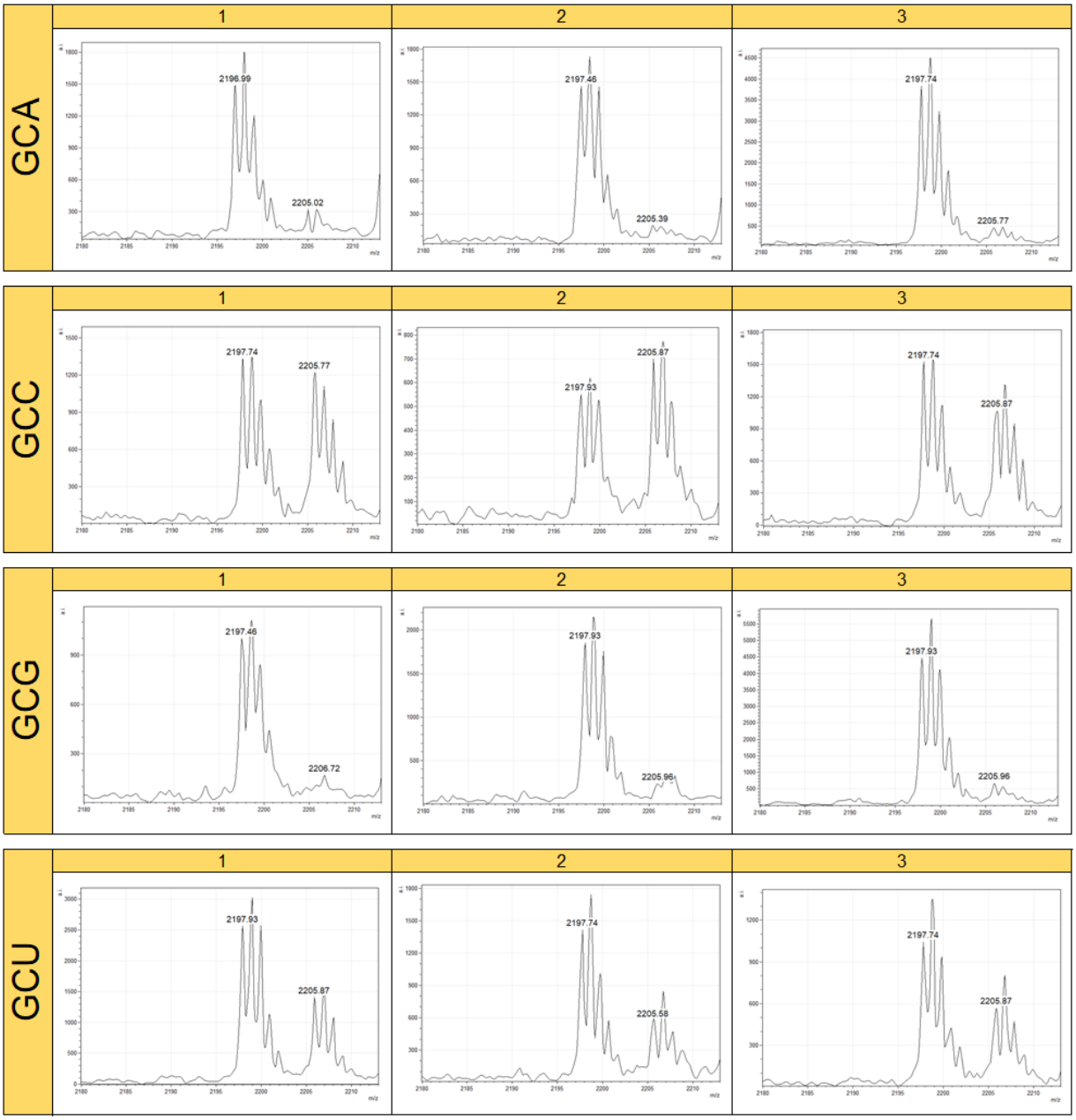


Figure S10: Additional MALDI-MS results for ncAA incorporation for mRNAs 9-12. Assays contain 15μM each of Abu-Thr1^{GGU}, CPG-Thr2^{CGU}, and PropG-Thr4^{UGU}. Translations were incubated for 30 min at 37 °C. To highlight orthogonality, relative peak heights for misincorporation peaks are labeled with the base peak being 100%.

A

	M	6xH	A	M	S	P	Q	P	V	D	S	W	V	stop	stop
mRNA 13:	AUG	CA U/C	GC**A**	AUG	AGC	CCG	CAG	CCG	GUG	GAU	AGC	UGG	GUG	UAG	UAG
mRNA 14:	AUG	CA U/C	GC**C**	OCT	AGC	CCG	CAG	CCG	GUG	GAU	AGC	UGG	GUG	UAG	UAG
mRNA 15:	AUG	CA U/C	GC**G**	DEC	AGC	CCG	CAG	CCG	GUG	GAU	AGC	UGG	GUG	UAG	UAG
mRNA 16:	AUG	CA U/C	GC**U**	FEB	AGC	CCG	CAG	CCG	GUG	GAU	AGC	UGG	GUG	UAG	UAG

B



tRNA	% read average	stdev
$Ala1^{UGC}$	90.20%	2.1
$Ala2^{GGC}$	9.80%	2.1

tRNA	% read average	stdev
$Ala1^{UGC}$	51.70%	7.8
$Ala2^{GGC}$	48.30%	7.8

tRNA	% read average	stdev
$Ala1^{UGC}$	90.70%	2.8
$Ala2^{GGC}$	9.30%	2.8

tRNA	% read average	stdev
$Ala1^{UGC}$	67.20%	3.7
$Ala2^{GGC}$	32.80%	3.7

Figure S11: Alanine Isotopic competition data. (A) mRNAs used to test the alanine codons. (B) Mass spectra of resulting peptides from alanine isotopic competition assay. Codon reading percentage was calculated as (isotope intensity / total peptide intensity) *100. For peaks less than 3 Da apart, the expected isotopic contribution of the lower mass peak was subtracted from the peak height of the higher isotopic mass peak. In vitro translations occurred in triplicate at 15 µL final volume. Isotopically aminoacylated tRNAs are introduced at 5 µM each. mRNAs **13-16** (Table **S2,S3**) are added (each mRNA containing a different GCN codon) to a final concentration of 1 µM, and the reactions were incubated at 37 °C for 30 min.

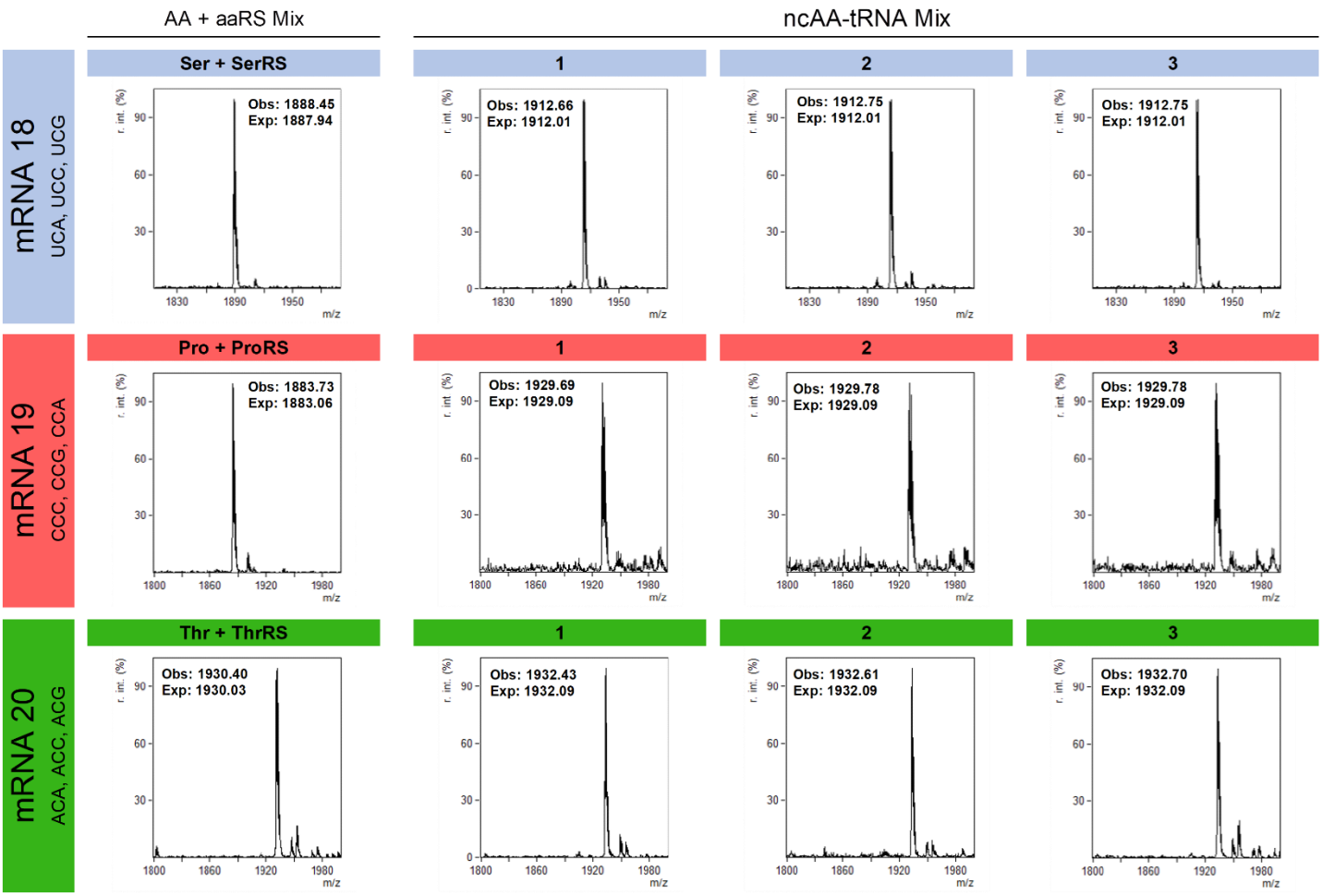


Figure S12: All MALDI-MS results for the triple codon mRNAs. The top data (blue) shows the serine codons, the middle, proline (red), and bottom, threonine (green) codons. Positive controls containing the native AARS/AA pair are shown at the left. Assays contained the same respective ncAA-tRNA pairs and concentrations as shown in Figure 5. All translations were carried out at 30 µL volumes and were translated at 37 °C for 60 minutes.

Primer name		Sequence (5'-3')
Flexizyme + Minihelix		
dFx REV	**CM_3_024**	ACCTAACGCCATGTACCCTTTCGGGGATGCGGAAATCTTTCGATCCTATAGTGAGTCGTATTACGCC
Microhelix REV	**CM_3_020**	TGGCGGCTCTGCGAACAGAGCCTATAGTGAGTCGTATTACGCC
T7 FWD	**CM_3_023**	GGCGTAATACGACTCACTATAG
mRNA primers		
Oligo-UniFWD	CAM_015	GGCGTAATACGACTCACTATAGGGTTAACTTTACCGAAGGAGGAAAGA
Oligo-UniREV	CAM_016	CTACTATTTGTCATCGTCGTCTTTATAATC
HFSW- UCA	CAM019	CCGAAGGAGGAAAGAATGCATTTTTCATGGGATTATAAAGACGACGATG
HFSW- UCC	CJ_1_005	CCGAAGGAGGAAAGAATGCACTTTTCCTGGGATTATAAAGACGACGATG
HFSW - UCG	CAM021	CCGAAGGAGGAAAGAATGCATTTTTCGTGGGATTATAAAGACGACGATG
HFSW - UCU	CAM018	CCGAAGGAGGAAAGAATGCATTTTTCTTGGGATTATAAAGACGACGATG
HFPW - CCA	CAM024	CCGAAGGAGGAAAGAATGCATTTTCCATGGGATTATAAAGACGACGATG
HFPW - CCC	CAM023	CCGAAGGAGGAAAGAATGCATTTTCCCTGGGATTATAAAGACGACGATG
HFPW - CCG	CAM025	CCGAAGGAGGAAAGAATGCATTTTCCGTGGGATTATAAAGACGACGATG
HFPW - CCU	CJ_1_006	CCGAAGGAGGAAAGAATGCACTTCCCTTGGGATTATAAAGACGACGATG
HFTW - ACA	MH_2_022	CCGAAGGAGGAAAGAATGCATTTTACATGGGATTATAAAGACGACGATG
HFTW - ACC	MH_2_023	CCGAAGGAGGAAAGAATGCATTTTACCTGGGATTATAAAGACGACGATG
HFTW - ACG	MH_2_024	CCGAAGGAGGAAAGAATGCATTTTACGTGGGATTATAAAGACGACGATG
HFTW - ACU	MH_2_025	CCGAAGGAGGAAAGAATGCATTTTACTTGGGATTATAAAGACGACGATG
Triple-Ser	CAM031	CCGAAGGAGGAAAGAATGAAATCATATTCCTATTCGGATTATAAAGACGACGATG
Triple-Pro	CAM032	CCGAAGGAGGAAAGAATGAAACCCAAACCGTATCCAGATTATAAAGACGACGATG
Triple-Thr	CJ.1.016	CCGAAGGAGGAAAGAATGAAAACATATACCTATACGGATTATAAAGACGACGATG
H6-Ala - FWD	EO03	GGCGTAATACGACTCACTATAGGGTTAACTTTACCGAAGGAGGAAAGA
H6-Ala - GCA	GK83.GCA	CCGAAGGAGGAAAGAATGCATCATCACCATCACCATGCAATGAGCCCGCAG
H6-Ala - GCC	GK85.GCC	CCGAAGGAGGAAAGAATGCATCATCACCATCACCATGCCATGAGCCCGCAG
H6-Ala - GCG	GK87.GCG	CCGAAGGAGGAAAGAATGCATCATCACCATCACCATGCGATGAGCCCGCAG
H6-Ala - GCU	GK89.GCU	CCGAAGGAGGAAAGAATGCATCATCACCATCACCATGCTATGAGCCCGCAG
H6-Ala - REV	GK.UR	CTACTACACCCAGCTATCCACCGGCTGCGGGCTCAT
t7 tRNAs		
Ser1 FWD	CAM_2_08	GGCGTAATACGACTCACTATAGGAAGTGTGGCCGAGCGGTTGAAGG
Ser2 FWD	CAM_2_11	GGCGTAATACGACTCACTATAGGAGAGATGCCGGAGCGGCTGAACGGACCG
Ser5 FWD	CAM_2_17	GGCGTAATACGACTCACTATAGGTGAGGTGTCCGAGTGGTTGAAGGAGCACGCCTG
Ser1 REV	CAM_2_10	TGGCGGAAGCGCAGAGATTCGAACTCTGGAACCC
Ser2 REV	CAM_2_13	TGGCGGAGAGAGGGGGATTTGAACCCCCGGTAGAGT

Ser5 REV	CAM_2_46	TGGCGGTGAGGGGGGGATTCGAACCCCCGATACGTTG
Ser1 MID	CAM_2_09	CCGAGCGGTTGAAGGCACCGGTCTTGAAAACCGGCGACCCGAAAGGGTTCCAGAGTTCGAATCT
Ser2 MID	CAM_2_12	CGGCTGAACGGACCGGTCTCGAAAACCGGAGTAGGGGCAACTCTACCGGGGGTTCAAATC
Ser5 MID	CAM_2_18	GTGGTTGAAGGAGCACGCCTGGAAAGTGTGTATACGGCAACGTATCGGGGGTTCGAATCC
ProUniFWD	CAM_2_050	GAAATTAATACGACTCACTATAGAGACAACCAGGAGTCTATAAAAT
Pro1 REV	CAM_2_058	TmGGTCGGTGATAGAGGATTCGAACCTCCGACCCCTTCGTCCCGAAC
Pro2 REV	CAM_2_061	TmGGTCGGCACGAGAGGATTTGAACCTCCGACCCCCGACACCCC
Pro3 REV	CAM_2_064	TmGGTCGGCGAGAGAGGATTCGAACCTCCGACCCACTGGTCCCAAACC
Pro1 MID1	CAM_2_056	GACAACCAGGAGTCTATAAAATCACCGCTGAAGAGACTGGACGAAACCAATAGGTCCGGT
Pro1 MID2	CAM_2_057	GAAACCAATAGGTCCGGTGATTGGCGCAGCCTGGTAGCGCACTTCGTTCGGGACGAAGGG
Pro2 MID1	CAM_2_059	GACAACCAGGAGTCTATAAAATTGCCGCTGAAGAGACTGGACGAAACCAATAGGTCCGG
Pro2 MID2	CAM_2_060	GAAACCAATAGGTCCGGCACGTAGCGCAGCCTGGTAGCGCACCGTCATGGGGTGTCGGGG
Pro3 MID1	CAM_2_062	GACAACCAGGAGTCTATAAAATCGCCGCTGAAGAGACTGGACGAAACCAATAGGTCCGG
Pro3 MID2	CAM_2_063	GAAACCAATAGGTCCGGCGAGTAGCGCAGCTTGGTAGCGCAACTGGTTTGGGACCAGTGG
Thr1 FWD	CAM_2_029	GGCGTAATACGACTCACTATAGCTGATATGGCTCAGTTGGTAGA
Thr2 FWD	MH.2.010	GGCGTAATACGACTCACTATAGCCGATATAGCTCAGTTGGT
Thr4 FWD	CAM_2_053	GGCGTAATACGACTCACTATAGCCGACTTAGCTCAGTAGGT
Thr1 REV	CAM_2_031	TGGTGCTGATACCCAGAGTCGAA
Thr2 REV	MH.2.012	TmGGTGCCGATAATAGGAGTCGAACCTACGACCTTCGCATT
Thr4 REV	CAM_2_055	TmGGTGCCGACTACCGGAATCGAACTGGTGACCTACTGATTA
Thr1 MID	CAM_2_045	GGCTCAGTTGGTAGAGCGCACCCTTGGTAAGGGTGAGGTCCCCAGTTCGACTCTGGGTATC
Thr2 MID	MH.2.011	CGATATAGCTCAGTTGGTAGAGCAGCGCATTCGTAATGCGAAGGTCGTAG
Thr4 MID	CAM_2_054	CGACTTAGCTCAGTAGGTAGAGCAACTGACTTGTAATCAGTAGGTCACCAGT
Ala1 FWD	CAM_2_038	GGCGTAATACGACTCACTATAGGGGCTATAGCTCAGCTGGGAGAG
Ala1 MID	CAM_2_039	CTCAGCTGGGAGAGCGCCTGCTTTGCACGCAGGAGGTCTGCGGTTCGATCC
Ala1 REV	CAM_2_040	TmGGTGGAGCTATGCGGGATCGAACCGCAG
Ala2 FWD	CAM_2_041	GGCGTAATACGACTCACTATAGGGGCTATAGCTCAGCTGGGA
Ala2 MID	CAM_2_042	GCTATAGCTCAGCTGGGAGAGCGCTTGCATGGCATGCAAGAGGTCAGCGGTTCGATCCCGC
Ala2 REV	CAM_2_043	TmGGTGGAGCTAAGCGGGATCGAACCGC

Table S1: DNA oligos used in this paper. mG stands for 2'-*O*-methyl G.

RNA Sequences:

RNA	RNA Sequence (5'-3')
dFx	GGATCGAAAGATTTCCGCATCCCCGAAAGGGTACATGGCGTTAGGT
uHelix	GGCTCTGTTCGCAGAGCCGCCA
mRNA 1-UCA	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUUUUUCAUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 2-UCC	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCACUUUUCCUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 3-UCG	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUUUUUCGUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 4-UCU	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUUUUUCUUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 5-CCA	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUUUUCCAUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG

mRNA 6-CCC	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUUUUCCCUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 7-CCG	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUUUUCCGUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 8-CCU	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCACUUCCCUUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 9-ACA	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUUUUACAUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 10-ACC	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUUUUACCUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 11-ACG	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUUUUACGUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 12-ACU	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUUUUACUUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 13-GCA	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUCAUCACCAUCACCAUGCAAUGAGCCCGCAGCCGGUGGAUAGCUGGGUGUAGUAG
mRNA 14-GCC	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUCAUCACCAUCACCAUGCCAUGAGCCCGCAGCCGGUGGAUAGCUGGGUGUAGUAG
mRNA 15-GCG	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUCAUCACCAUCACCAUGCGAUGAGCCCGCAGCCGGUGGAUAGCUGGGUGUAGUAG
mRNA 16-GCU	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUCAUCACCAUCACCAUGCUAUGAGCCCGCAGCCGGUGGAUAGCUGGGUGUAGUAG
mRNA 17-UCC	GGGUUAACUUUACCGAAGGAGGAAAGAAUGCAUUUUUCCUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 18-Triple-Ser	GGGUUAACUUUACCGAAGGAGGAAAGAAUGAAAUCAUAUUCCUAUUCGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 19-Triple Pro	GGGUUAACUUUACCGAAGGAGGAAAGAAUGAAACCAAAACCCUAUCCGGAUUAUAAAGACGACGAUGACAAAUAGUAG
mRNA 20-Triple Thr	GGGUUAACUUUACCGAAGGAGGAAAGAAUGAAAACAUAUACCUAUACGGAUUAUAAAGACGACGAUGACAAAUAGUAG
Ser1^{UGA} tRNA	GGAAGUGUGGCCGAGCGGUUGAAGGCACCGGUCUUGAAAACCGGCGACCCGAAAGGGUUCCAGAGUUCGAAUCUCUGCGCUUCCGCCA
Ser2^{CGA} tRNA	GGAGAGAUGCCGGAGCGGCUGAACGGACCGGUCUCGAAAACCGGAGUAGGGGCAACUCUACCGGGGGUUCAAAUCCCCCUCUCUCCGCCA
Ser5^{GGA} tRNA	GGUGAGGUGUCCGAGUGGUUGAAGGAGCACGCCUGGAAAGUGUGUAUACGGCAACGUAUCGGGGGUUCGAAUCCCCCCCUCACCGCCA
Pro1^{CGG} tRNA	CGGTGATTGGCGCAGCCTGGTAGCGCACTTCGTTCGGGACGAAGGGGTCGGAGGTTCGAATCCTCTATCACCGACCA
Pro2^{GGG} tRNA	CGGCACGTAGCGCAGCCTGGTAGCGCACCGTCATGGGGTGTCGGGGGTCGGAGGTTCAAATCCTCTCGTGCCGACCA
Pro3^{UGG} tRNA	CGGCGAGTAGCGCAGCTTGGTAGCGCAACTGGTTTGGGACCAGTGGGTCGGAGGTTCGAATCCTCTCTCGCCGACCA
Thr1^{GGU} tRNA	GCTGATATGGCTCAGTTGGTAGAGCGCACCCTTGGTAAGGGTGAGGTCCCCAGTTCGACTCTGGGTATCAGCACCA
Thr2^{CGU} tRNA	GCCGATATAGCTCAGTTGGTAGAGCAGCGCATTCGTAATGCGAAGGTCGTAGGTTCGACTCCTATTATCGGCACCA
Thr4^{UGU} tRNA	GCTGATATAGCTCAGTTGGTAGAGCGCACCCTTGGTAAGGGTGAGGTCGGCAGTTCGAATCTGCCTATCAGCACCA
Ala1^{UGC} tRNA	GGGGCUAUAGCUCAGCUGGGAGAGCGCUUGCAUGGCAUGCAAGAGGUCAGCGGUUCGAUCCCGCUUAGCUCCACCA
Ala2^{GGC} tRNA	GGGGCUAUAGCUCAGCUGGGAGAGCGCCUGCUUUGCACGCAGGAGGUCUGCGGUUCGAUCCCGCAUAGCUCCACCA

Table S2: RNA sequences used in this paper

Methods

General

The d3 (L-Serine-2,3,3-D_3) and +7 (L-Serine-$^{13}C_3$-2,3,3-D_3-^{15}N) serine isotopes, the +7 (L-Proline-2,3,3,4,4,5,5-D_7) proline isotopes, the +10 (L-Threonine-$^{13}C_4$-2,3,4,4,4-D_5-^{15}N) threonine isotope, and the +8 (L-Alanine-$^{13}C_3$-2,3,3,3-D_3-^{15}N) alanine isotope were purchased from Cambridge Isotope Laboratories. The d3 (L-Proline-2,5,5-D_3) was purchased from Cayman Chem. The +5 (L-Threonine-$^{13}C_4$-^{15}N) threonine isotope was purchased from Sigma-Aldrich. Amino acids were dissolved in concentrations of 10-100 mM depending on solubility and the pH of the stock solution was adjusted to 7.4 with KOH (1 M), monitoring with pH paper.

Equipment

Urea-polyacrylamide gel electrophoresis (PAGE) gels were prepared using the SequaGel system (National diagnostics) and cast using the mini-PROTEAN 3 system (Bio-Rad). Urea-PAGE gels were imaged on a Chemidoc MP imaging system (Bio-Rad). Sample absorbance was measured on a NanoDrop ND1000 SpectroPhotometer. RNA and tRNA concentrations were quantified by Qubit fluorometer. Mass spectrometry experiments were performed on a Voyager-DE Pro BioSpectrometry Workstation (Applies Biosystems) under reflectron positive mode. MALDI-MS figures were created from the raw ASCII file data using MMass software after background subtraction.

Transcription of mRNAs / tRNAs / Flexizymes:

DNA oligos were either purchased from IDT or prepared on a MerMade MM-6 oligonucleotide synthesizer. DNA templates for mRNAs encoding a 5' T7 promoter sequence and a Shine Dalgarno sequence followed by a peptide coding sequence for in-vitro transcription were prepared for in vitro transcription using the polymerase chain reaction (PCR). tRNAs, flexizyme, and minihelix DNA templates contained their respective sequence as opposed to a Shine Dalgarno and encoding sequence. DNA templates used for RNAs are available in table S2. PCR mixtures were prepared to a final volume of 300-500 µL containing 1x Q5 reaction buffer (NEB), 0.2 mM dNTPs (NEB), and 0.02 U/µL Q5 DNA polymerase (NEB). Forward and reverse DNA primers were added at 0.5 µM, and middle primers were added at 0.5 nM. PCR was carried out on a DNA Engine thermocylcer (BioRad), beginning with an initial 1 min denaturing step at 95 ºC, followed by twenty cycles of the following: 98 ºC for 10 seconds, oligo Tm for 30 seconds, 72 ºC for 30 seconds. PCR product was confirmed by running 4 µL on a 3% agarose gel, and were then purified via acidic phenol/chloroform extraction followed by ethanol precipitation.

The pelleted PCR products were resuspended in DI-water and added to a transcription assay containing Tris/Triton (40 mM, pH 7.8), spermidine (2.5 mM), $Mg(OAc)_2$ (25 mM), dithiothreitol (10 mM), ATP (5 mM), GTP (5 mM), CTP (5 mM), UPT (5 mM), GMP (5 mM), Ribosafe (0.2 U/µL), inorganic pyrophosphatase (0.001 mg/mL), and T7 polymerase (0.2 µM). The reactions were incubated at 37 °C overnight, and quenched via incubation with DNAse (37 °C, 15 min). Transcribed RNAs were purified by using the Monarch RNA Cleanup Kit (NEB) or by gel electrophoresis. For gel purification, the target band was excised, UV shadowed, and the sample was extracted from the gel via crush and soak (the excised gel was crushed in a tube, 2-3 mL of KOAc (300 mM, pH 5.2) was added, heated at 80 ºC for 30 min, then frozen at -20 ºC, and allowed to thaw to rt). The aqueous fraction was then collected from the mixture via syringe filtration, ethanol precipitated, and the final pellets were resuspended in water. Sequences of all RNAs used are available in table S2.

Quantification of flexizyme substrate yield[1,2]:

The ability of flexizymes to aminoacylate ncAAs to RNAs was determined through acid urea-PAGE gel electrophoresis. dFx flexizyme (25 µM) and microhelix RNA (25 µM) were added to HEPES-KOH (50 mM, pH 7.5) and were heated to 95 °C for 3 minutes, then slowly cooled to rt. $MgCl_2$ was added to a final concentration of 600 mM and incubated at rt for 5 min. The tubes were then transferred to ice and the ncAA-DNB (25 mM in DMSO) was added to a final concentration of 5 mM. At varying time intervals 800 ng of tRNA was removed and

analyzed on an acid urea-PAGE gel according to the protocols published by Lee[1]. Gels were prepared using the Sequagel system, however replacing the buffer solution with 50 mM NaOAc pH 5.2. Gels were run for 4 hours at 4 ºC with NaOAc (50 mM, pH 5.2) as the running buffer and were subsequently stained with GelRed. Acylation yield was determined by comparison of band intensities of acylated and unacylated microhelix RNA (Figure **S3**, Table **S1**).

Isotopic precharging of tRNAs via aminoacyl synthetases:

Individual tRNAs were mixed with HEPES-KOH (30mM, pH 7.5), and KCl (25mM) and were heated to 95 °C for 3 min then cooled to rt. Next, $MgCl_2$ (15 mM), BME (2 mM), ATP (6 mM), inorganic pyrophosphatase (0.001 mg/mL), dialyzed BSA (0.09 mg/mL), and 200 µM AA and 0.1-1 µM aminoacyl synthetase were added and incubated at 37 °C for 1 hour. The following mixture was acidic phenol:chloroform extracted, then pelleted via ethanol precipitation. The pellets were stored at -80 °C for up to 2-3 days then resuspended in KOAc (1 mM, pH 5.2) for use in in-vitro translation. The tRNAs were mixed in equivolume amounts and the mixture was checked to ensure even amounts in translation using a reductive amination MALDI assay. Any deviations from a 1:1:1 mixture was compensated for by adjusting the volumes added to translation. Pellets frozen in 1 mM KOAc (pH 5.2) at -80 ºC were stored for up to a month before use.

Precharging of tRNAs by flexizymes:

tRNAs were pre-charged by flexizymes using the protocol previously reported by Suga. The final reaction components included: the tRNA of interest (25 µM), dFx flexizyme (25 µM), HEPES-KOH (50 µM), $MgCl_2$ (600 mM), the DNB-ncAA of interest (5 mM final in DMSO, total reaction mixture was 20% DMSO final), and were brought to a final volume with RNAse-free water. To prepare the reaction, the HEPES-KOH, tRNA, dFx flexizyme, and water were combined and heated at 95°C for 3 min, then lowered by 0.2 °C/sec to 25 °C. The $MgCl_2$ was then added and incubated at rt for 5 min, then the tubes were transferred to ice and the DNB-ncAA in DMSO was added. The tubes were then incubated on ice until the ideal reaction time was reached (Table S1). The reactions were quenched with 80 µL of NaOAc (300 mM, pH 5.2) and 200 µL room-temperature ethanol per 20 µL of reaction volume. The tubes were then vortexed and centrifuged at 17,000 x g at rt for 15 min, and the supernatant was removed. The pellets were then resuspended with 100 µL of 70% EtOH containing NaOAc (100 mM, pH 5.2) and centrifuged at 17,000 x g for 5 min at rt, then the supernatant was removed. This step was repeated. 100 µL of 70% ethanol was added to the tubes, and the tubes were centrifuged at 17,000 x g for 3 min. The supernatant was removed, and the pellets were allowed to air-dry for 5-10 min at rt. Dry pellets were stored at -80 °C for up to a week.

Isotopic codon competition and sense codon reassignment:

All experiments were performed with a customized version of the PURE cell-free in vitro translation system. All translations, unless otherwise specified, were performed on a 15 µL scale for 30 min at 37°C. All AAs and aminoacyl synthetases (aaRSs) required to decode the mRNA were added to the translation with the exception of the AA-aaRS pairs being reassigned. The final concentration of components in the translation were as follows: 2.1 mg/mL total E. coli tRNA, HEPES-KOH (50 mM, pH 7.6), KOAc (100 mM), $Mg(OAc)_2$ (6 mM), dithiothreitol (1 mM), creatine phosphate (20 mM), 10-formyltetrahydrofolate (0.1 mM), ATP (1.5 mM), GTP (1.5 mM), IF1 (2.7 µM), IF2 (0.4 µM), IF3 (1.5 µM), RF1 (0.3 µM), RF3 (0.17 µM), RRF (0.5 µM), EF-G (0.52 µM), MTF (0.6 µM), Ef-Tu (10 µM), Ef-TS (8 µM), MAGE ribosomes (1.2 µM), inorganic pyrophosphatase (0.1 µg/mL), creatine kinase (4 µg/mL), nucleoside diphosphate kinase (4 µg/mL), myokinase (3 µg/mL), AAs (100 mM), aaRSs (0.1-1.0 µM), and then the mRNA template (1 µM). For the isotopic codon competitions, the pre-charged AA-tRNA pellets were resuspended in KOAc (1 mM, pH 5.2) and added into the translation to a final concentration of 5 µM each. For sense codon reassignment, the precharged ncAA-tRNAs were resuspended and added to translations to a final concentration of 15 µM each (with the exception of Ser5, which was instead added at 30 µM). Our mRNA templates encoded a C-terminal FLAG-tag and final products were collected by adding 7.5 µL of anti-FLAG M2 magnetic beads (Sigma-Aldrich) and binding on a tumbler at rt for 1 h, before subsequently performing three 200 µL washes with Tris-buffered saline (TBS) buffer, and eluting into 50 µL of

1% aqueous trifluoracetic acid (TFA). The final peptides were de-salted via zip-tipping, eluted with 4-Chloro-α-cyanocinnamic acid (ClCA) matrix (6.2 mg/mL, 7:3 MeCN:0.2% TFA) and analyzed via MALDI-MS. Quantitative data analysis was performed using ASCII data files. Peak heights were used after background subtraction to determine the relative fraction of each isotopomeric peptide. For peaks less than 4 Da apart, the expected isotopic contribution of the lower mass peak was subtracted from the peak height of the higher mass peak.

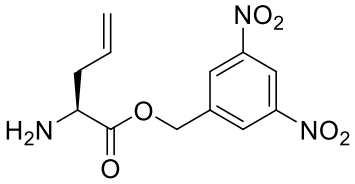
Synthesis of dFx substrates[3]:

1 equivalent (50mg) of the boc-protected non-canonical amino acid, 1.1 equivalents of 3,5-dinitrobenzyl chloride, and 5.6 equivalents of TEA were dissolved in DMF (1-2 mL) and reacted at room temperature overnight. Typically a color change was observed from yellow to dark red. The reaction was then quenched with diethyl ether (~ 2 mL) and then the resulting mixture was added to a separatory funnel with several diethyl ether washes. This solution was washed with 2x 1 M HCl (15 mL), 2x with $NaHCO_3$ (15 mL) and 5x with brine (15 mL). The resulting organic layer was then dried with $MgSO_4$ and filtered, then the remaining solvent was removed via rotary evaporation, resulting in an amber oil.

The crude boc-protected dinitrobenzyl-ester was then treated with 4M HCl:dioxane (~ 2 mL) at rt for 30 min. The reaction was quenched with diethyl ether (3 mL), and the resulting precipitate was collected and washed with diethyl ether, then dried via rotary evaporation. The precipitate was then dissolved in 1:3 methanol:ethyl acetate, and was recrystallized via slow addition of hexanes, yielding a white crystalline solid. The final product was then collected via gravity filtration.

Characterization of Novel ncAAs:

3,5-dinitrobenzyl (S)-2-aminopent-4-enoate hydrochloride (AllylG):



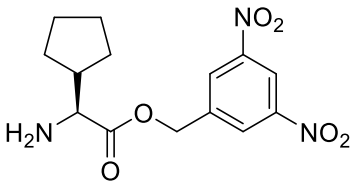
Yield. 26%

^{1}H NMR: (400MHz, D_2O): δ 9.02 (t, J=2 Hz, 1H), 8.64 (s, 2H), 5.61-5.71 (m, 1H), 5.41-5.48 (q, J=13 Hz, 3.96 Hz, 3H), 5.18-5.19 (d, J=4.12, 2H), 5.15 (s, 1H), 4.25-4.28 (q, J=5.5Hz, 1.4, 1H), 2.61-2.75 (m, 2H).

^{13}C NMR: (100 MHz, D_2O): δ 169.20, 148.34, 138.66, 129.78, 128.87, 121.56, 119.17, 66.05, 52.13, 38.71, 34.02

DART-MS: $C_{12}H_{14}N_3O_6^+$ $[M+H]^+$ Expect: 296.0877 m/z; Obs: 296.0866 m/z

3,5-dinitrobenzyl (S)-2-amino-2-cyclopentylacetate hydrochloride (CPG):



Yield: 25%

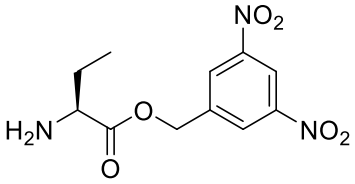
^{1}H NMR: (400MHz, D_2O): δ 9.02 (s, 1H), 8.65 (s, 2H), 5.45-5.46 (q, J=13.1 Hz, 3.4 Hz, 2H), 4.07 (d, J=7.72 Hz, 1H), 2.30 (m, 1H), 1.74 (m, 2H), 1.50 (m, 5H), 1.30 (m, 3H)

^{13}C NMR: (100 MHz, D_2O): δ 169.83, 148.43, 138.79, 128.94, 119.24, 66.00, 56.62, 40.68, 28.61, 28.51, 24.65, 24.46

DART-MS: $C_{14}H_{18}N_3O_6^+$ $[M+H]^+$ Expect: 324.1190 m/z; Obs: 324.1166 m/z

Characterization of previously reported ncAAs

L-2-Aminobutyric acid-3,5-dinitrobenzyl ester hydrochloride (Abu):



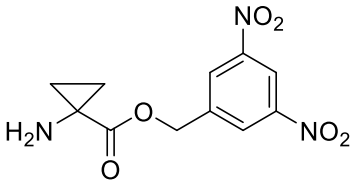
Yield. 47%

NMR spectra matched literature values.[4]

^{1}H NMR: (400 MHz, DMSO-d6): δ 8.827-8.832 (d, J=2.1 Hz, 1H), 8.75 (d, J=2.1 Hz, 2H), 8.627 (S, 3H), 5.52 (s, 2H), 4.12-4.15 (t, J=6.1, 1H), 1.86-1.93 (m, 2H), 0.94-0.98 (t, J = 7.5, 3H)

DART-MS: $C_{11}H_{14}N_3O_6^+$ $[M+H]^+$ Expect: 284.0877 m/z; Obs: 284.0866 m/z

3,5-dinitrobenzyl 1-aminocyclopropane-1-carboxylate hydrochloride (Acp):



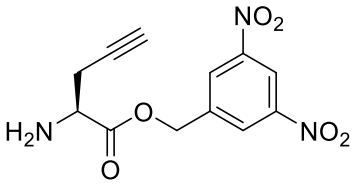
Yield 50%

NMR spectra matched literature values. [5]

^{1}H NMR: (400 MHz, DMSO): δ 9.09 (s, 3H), 8.82 (s, 1H), 8.76 (s, 2H), 5.47 (s, 2H), 1.49-1.52 (m, 5H)

DART-MS: $C_{11}H_{12}N_3O_6^+$ $[M+H]^+$ Expect: 282.0721 m/z; Obs: 282.0723 m/z

3,5-dinitrobenzyl (S)-2amino-4-butyne-1-carboxylate hydrochloride (PropG):



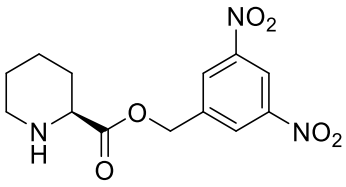
Yield 38%

NMR spectra matched literature values. [5]

^{1}H NMR: (400MHz, DMSO-d6): δ 8.83 (t, J=2.1Hz, 1H), 8.87 (d, J=2.1Hz, 2H), 5.51-5.59 (q, J=13.4Hz, 6.8Hz, 2H), 4.42 (t, J=5.4Hz, 1H), 3.10 (t, J=2.6Hz, 1H), 2.83-2.97 (m, 2H)

DART-MS: $C_{12}H_{12}N_3O_6^+$ [M+H]$^+$ Expect: 294.0721 m/z; Obs: 294.0678/z

3,5-dinitrobenzyl (S)-piperdine-2-carboxylate hydrochloride (Pip):



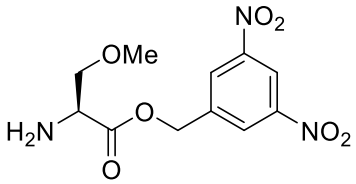
Yield 45%

NMR spectra matched literature values.[4]

^{1}H NMR: (400MHz, D_2O): δ 9.01 (s,1H), 8.62 (s, 2H), 5.44 (s, 3H), 4.09-4.13 (dd, J=3.5,8.2, 1H), 3.41-3.44 (d, J=12.7Hz, 1H), 2.97-3.04 (m, 1H), 2.27-2.32 (dd, J=3.5,10.5, 1H), 1.82 (m, 3H), 1.71 (m, 1H), 1.56(m, 3H)

DART-MS: $C_{13}H_{16}N_3O_6^+$ [M+H]$^+$ Expect: 310.1034 m/z; Obs: 310.1037 m/z

3,5-dinitrobenzyl O-methyl-L-serinate hydrochloride (Ser(Me)):



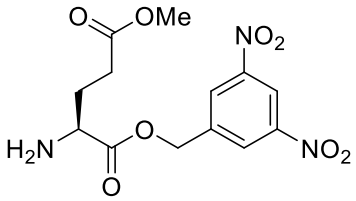
Yield 74%

NMR spectra matched literature values [3]

^{1}H NMR: (400MHz, DMSO): δ 8.81-8.82 (t, J=2.1, 1H), 8.71 (s, 5H), 5.50-5.62 (q, J= 13.8, 20.4, 2H), 4.53 (s, 1H), 3.8-3.9 (dd, J=3.9, 6.7, 1H), 3.32 (s, 4H).

DART-MS: $C_{11}H_{14}N_3O_7^+$ [M+H]$^+$ Expect: 300.0826 m/z; Obs: 300.0798 m/z

(S)-1-(3,5-dinitrobenzyl) 5-methyl 2-aminopentanedioate hydrochloride (Glu(Me)):

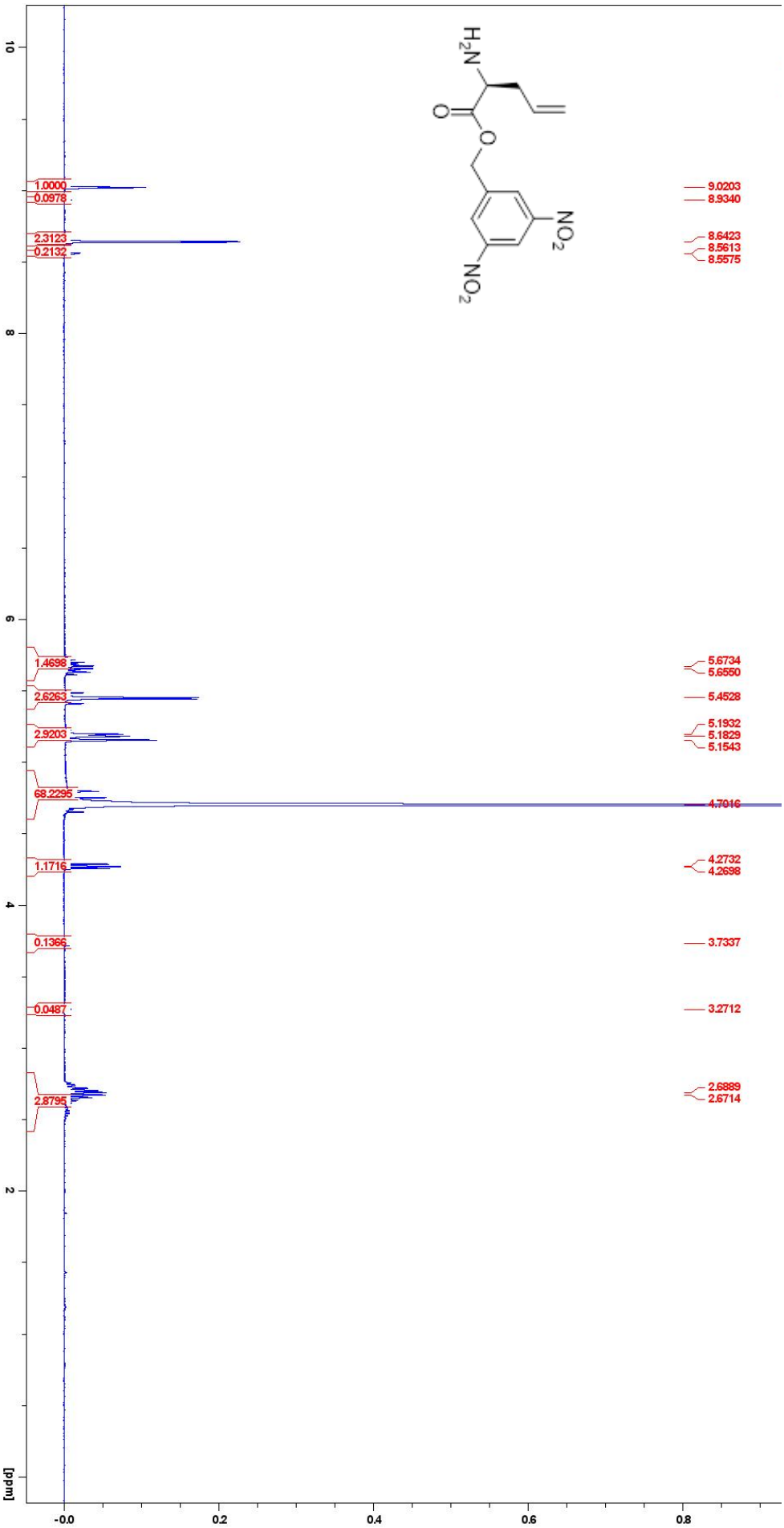


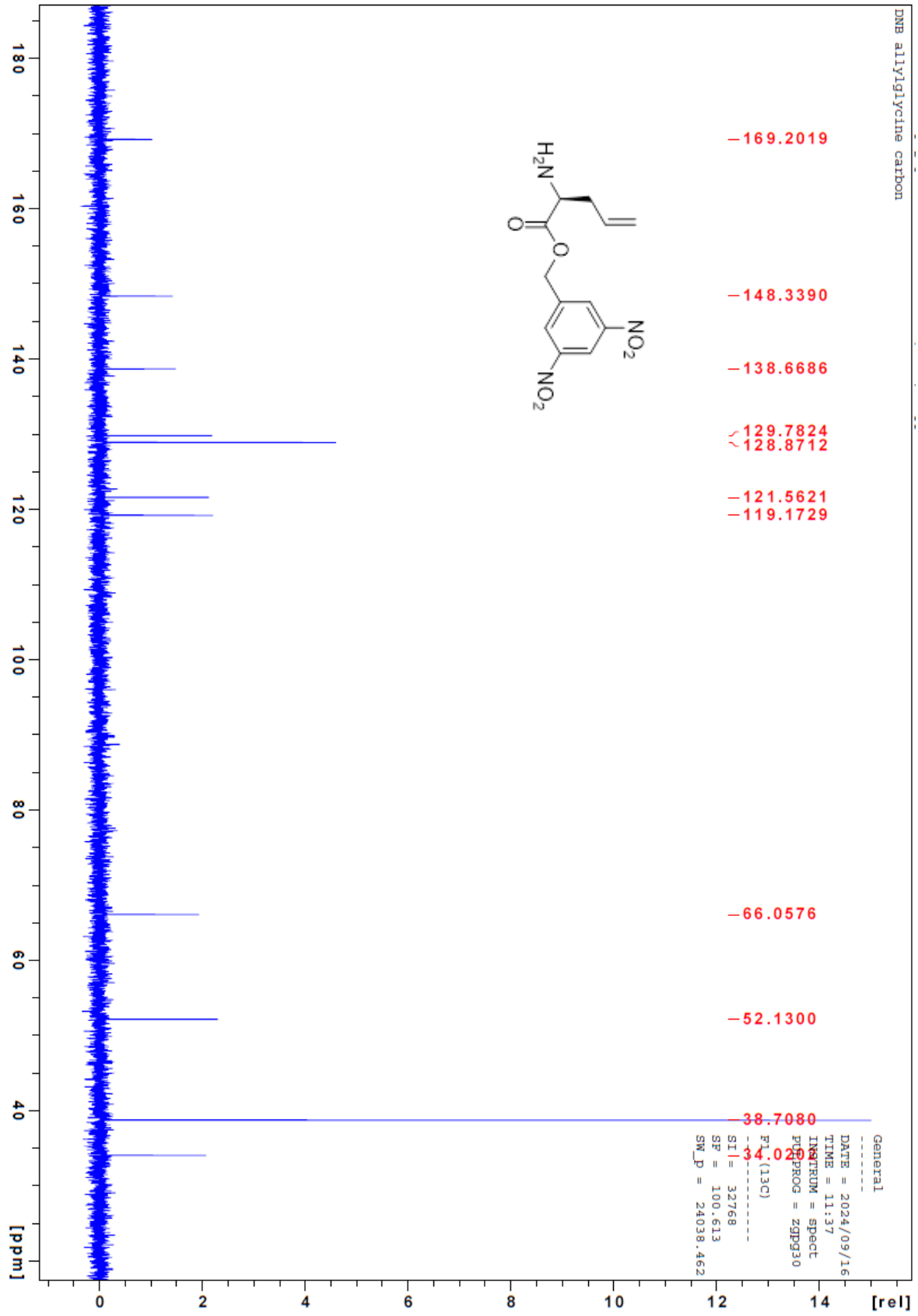
Yield 75%

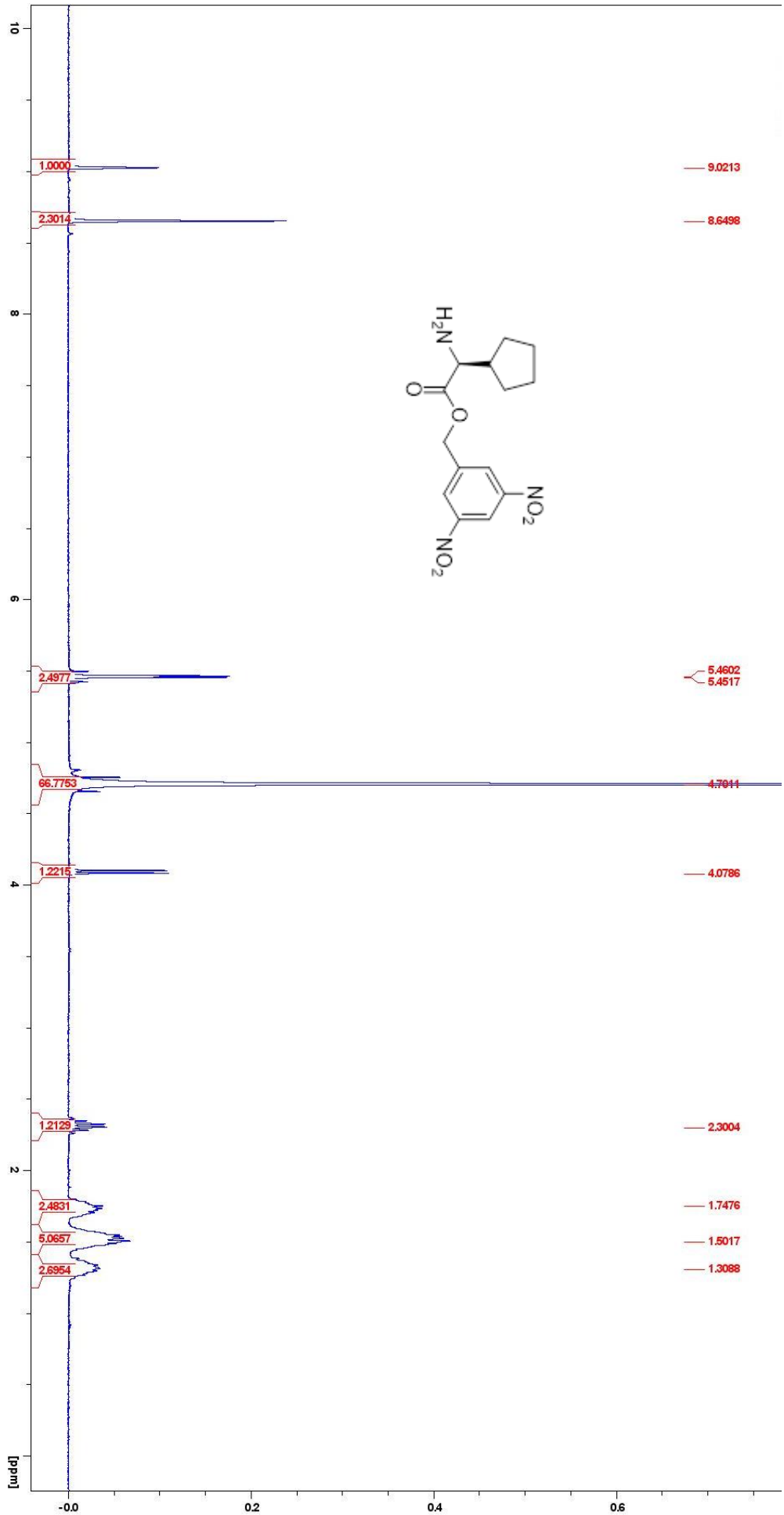
NMR spectra matched literature values [3]

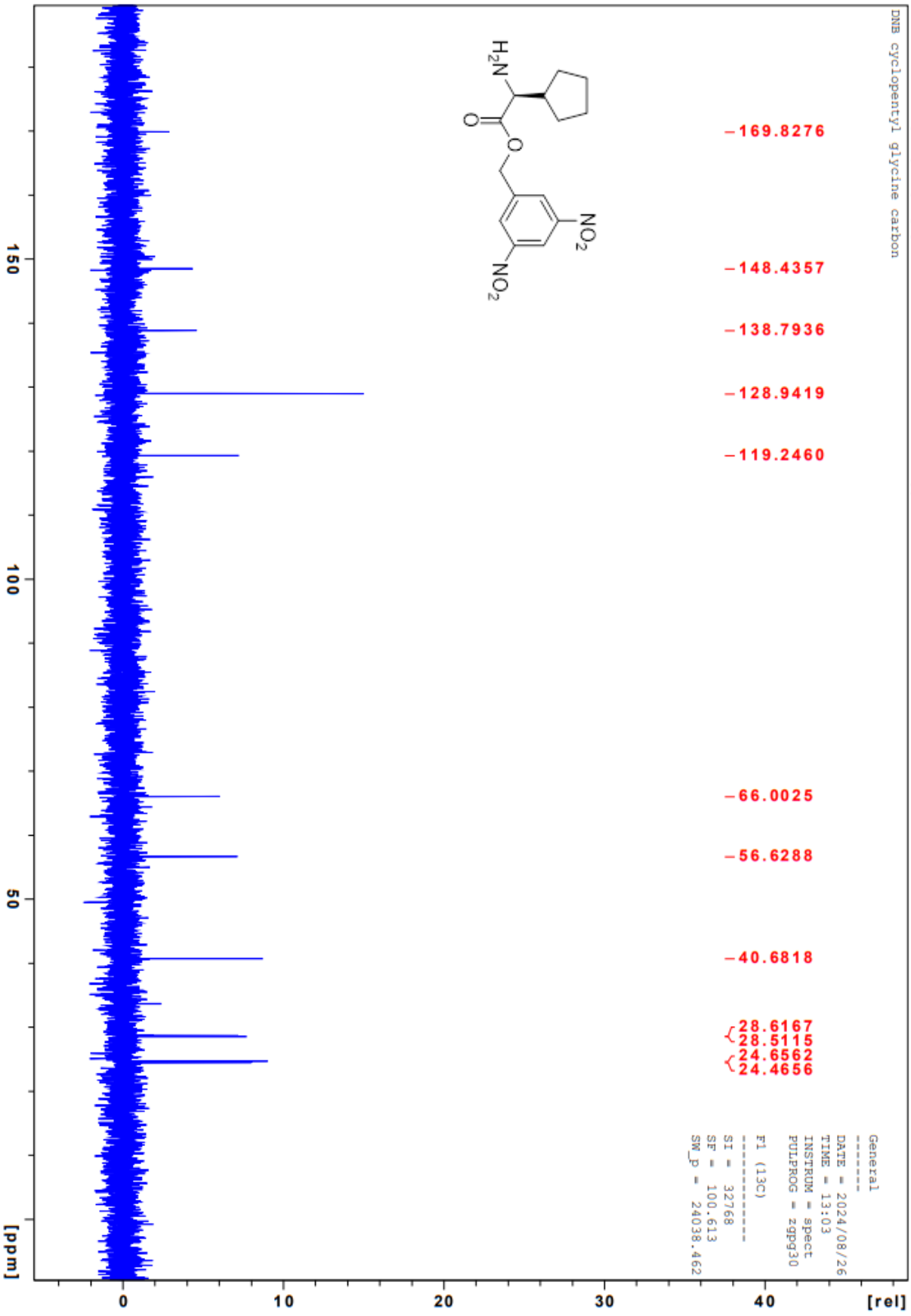
^{1}H NMR: (400MHz, D_2O): δ 8.88 (s, 1H), 8.61 (s, 2H), 5.37-5.46 (q, J = 13.2 Hz, 8 Hz, 2H), 4.16-4.20 (t, 1H), 3.57 (s ,3H), 2.42-2.52 (m, 2H), 2.12-2.23 (m, 2H)

DART-MS: $C_{13}H_{16}N_3O_8^+$ $[M+H]^+$ Expect: 342.0932 m/z; Obs: 342.0910 m/z


1.0000
0.0978
2.3123
0.2132
1.4698
2.6263
2.9203
68.2295
1.1716
0.1366
0.0487
2.8795
9.0203
8.9340
8.6423
8.5613
8.5575
5.6734
5.6550
5.4528
5.1932
5.1829
5.1543
4.7016
4.2732
4.2698
3.7337
3.2712
2.6889
2.6714
[ppm]


DNB allylglycine carbon
169.2019
148.3390
138.6686
129.7824
128.8712
121.5621
119.1729
66.0576
52.1300
38.7080
34.0200
[ppm]
[rel]
General
DATE = 2024/09/16
TIME = 11:37
INSTRUM = spect
PULPROG = zgpg30
F1 (13C)
SI = 32768
SF = 100.613
SW_p = 24038.462


10
8
6
4
2
[ppm]
-0.0
0.2
0.4
0.6
1.0000
2.3014
2.4977
66.7753
1.2215
1.2129
2.4831
5.0657
2.6954
9.0213
8.6498
5.4602
5.4517
4.7011
4.0786
2.3004
1.7476
1.5017
1.3088
H2N
O
O
NO2
NO2


DNB cyclopentyl glycine carbon
−169.8276
−148.4357
−138.7936
−128.9419
−119.2460
−66.0025
−56.6288
−40.6818
28.6167
28.5115
24.6562
24.4656
General
DATE = 2024/08/26
TIME = 13:03
INSTRUM = spect
PULPROG = zgpg30
F1 (13C)
SI = 32768
SF = 100.613
SW_p = 24038.462
[ppm]
[rel]

References

(1) Lee, J.; Schwieter, K. E.; Watkins, A. M.; Kim, D. S.; Yu, H.; Schwarz, K. J.; Lim, J.; Coronado, J.; Byrom, M.; Anslyn, E. V.; Ellington, A. D.; Moore, J. S.; Jewett, M. C. Expanding the Limits of the Second Genetic Code with Ribozymes. *Nat Commun* **2019**, *10*, 5097.

(2) Goto, Y.; Katoh, T.; Suga, H. Flexizymes for Genetic Code Reprogramming. *Nat Protoc* **2011**, *6*, 779–790.

(3) Passioura, T.; Liu, W.; Dunkelmann, D.; Higuchi, T.; Suga, H. Display Selection of Exotic Macrocyclic Peptides Expressed under a Radically Reprogrammed 23 Amino Acid Genetic Code. *J Am Chem Soc* **2018**, *140*, 11551–11555.

(4) Zhang, X.-D.; Wang, Y.-S.; Xiang, H.; Bai, L.-W.; Cheng, P.; Li, K.; Huang, R.; Wang, X.; Lei, X. Nucleoside Modification-Based Flexizymes with Versatile Activity for TRNA Aminoacylation. *Chem Commun* **2024**, *60*, 1607–1610.

(5) Chan, A. I.; Sawant, M. S.; Burdick, D. J.; Tom, J.; Song, A.; Cunningham, C. N. Evaluating Translational Efficiency of Noncanonical Amino Acids to Inform the Design of Druglike Peptide Libraries. *ACS Chem Biol* **2023**, *18*, 81–90.