

ABRF ESRG 2004 Study: Modified Amino Acids in Edman Sequencing

D.C. Brune¹, J.M. Crawford², R.G. Cook³, N.D. Denslow⁴, R. Kobayashi⁵, B.J. Madden⁶, J.M. Neveu⁷, L. Steinke⁸;

¹ Arizona State University, Tempe, AZ, ² Yale University, New Haven CT, ³ Baylor College of Medicine, Houston, TX, ⁴ University of Florida, Gainesville, FL,

⁵ UT MD Anderson Cancer Center, Houston, TX, ⁶ Mayo Clinic, Rochester, MN, ⁷ Harvard University, Cambridge, MA, ⁸ University of Nebraska, Omaha, NE.

Abstract

Edman degradation sequencing relies on comparing HPLC retention times of the PTH amino acids with PTH amino acid standards for amino acid assignments. The elution characteristics of the twenty common amino acids have been well-characterized, which aids in making confident assignments. Modified amino acids may present more of a challenge due to absence of standards and lack of familiarity by the protein sequencer analyst. The ABRF ESRG 2004 sample is the 16th in a series of studies designed to aid laboratories in evaluating their abilities to obtain and interpret amino acid sequence data. Laboratories requesting a sample were sent a tube containing approximately 775 pmoles of a 20 amino acid synthetic peptide composed of several modified amino acids that may be found in proteins or are generated during sample preparation. In addition to filling in an assignment sheet, which included retention times and peak areas, participants were asked to provide specific details about the parameters used for the sequencing run. Modified amino acid elution references were provided and the participants had the option of viewing a list of the modified amino acids that were in the peptide at the ESRG website. The goal of the study consists of two parts: assessment of the ability to correctly assign all the amino acids in the peptide, including the modified amino acids if possible; and the collection and compiling of elution time characteristics based on current methodology of the instruments participating in the study. The resulting compilation of the modified amino acid elution times and running conditions will be accessible at the ABRF ESRG website for future reference.

Materials and Methods

The 2004 ESRG peptide was synthesized using Fmoc chemistry. With the exception of Fmoc-Lys[Me] and Fmoc-Arg[Me]₂, which were coupled manually, all couplings were done on an Applied Biosystems (ABI, Foster City, CA) 433A peptide synthesizer using ABI reagents and standard protected Fmoc amino acids. All protected Fmoc amino acids were from ABI except for the following. Fmoc-Arg[Ph]₂-Wang resin and Fmoc-Hyp[tBu]-OH were obtained from NovaBiochem (San Diego, CA). Fmoc-Arg[Me]₂-OH and Fmoc-Lys[Me][Boc]-OH were obtained from Bachem Bioscience (King of Prussia, PA). Fmoc-Lys[Ac]-OH, Fmoc-Met(O)-OH and Fmoc-Ser[PO(OBzl)-OH]-OH were obtained from AnaSpec (San Jose, CA). After each coupling cycle, unreacted residues were capped with acetic anhydride. The completed peptide was cleaved and deprotected using ethanedithiol: thionisole: trifluoroacetic acid 1: 1: 8 for 3 hrs, precipitated and washed with diethyl ether, and purified on a 10 x 250 mm Vydac C18 reverse phase column and lyophilized. The cysteine alkylation was performed with a one milligram aliquot of lyophilized peptide, solubilized in 450ul 50mM Tris pH 8.1 /50% acetonitrile plus 50ul of 30% acrylamide/ water. After sitting at ambient temperature for 50 minutes, the peptide was purified with multiple injections on a Phenomenex Jupiter C18 2.1mm x 220mm column, eluting with 0.1% TFA/acetonitrile buffers. The peak from each run was pooled and 25ul aliquots were placed in PCR tubes, dried on a Speed Vac spinning concentrator, and stored at -20°C until shipped.

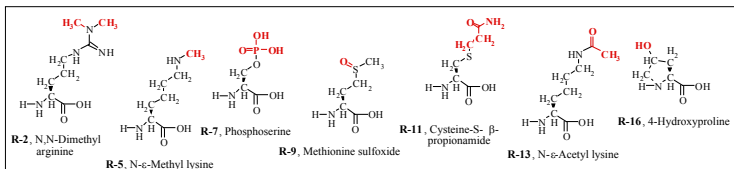


Figure 1. Structures of modified amino acids and sequence of test peptide.

Fac. No.	Inst. Type	% Loaded	Init. Yield	Repet. Yield	PC	TC	PW	TW	X	Tyr	dmR	Ala	Trp	mmK	Pro	pS	Ala	oxM	Tyr	spC	Arg	ack	Ser	Trp	hyP	Tyr				
1	ABI 494 HT	20%	72.1%	88.6%	18	0	0	0	2	Tyr	X	Ala	Trp	X	Pro	pS	Ala	oxM	Tyr	spC	Arg	ack	Ser	Trp	hyP	Tyr				
2	ABI 494 cLC	2%	50.7%	88.1%	18	0	2	0	0	Tyr	mmK	Ala	Trp	dmR	Pro	pS	Ala	oxM	Tyr	spC	Arg	ack	Ser	Trp	hyP	Tyr				
3	ABI 494 HT	50%	79.4%	88.4%	19	1	0	0	0	Tyr	dmR	Ala	Trp	mmK	Pro	pS	Ala	oxM	Tyr	spC	Arg	ack	Ser	Trp	hyP	Tyr				
4	ABI 491 HT	30%	58.1%	89.3%	20	0	0	0	0	Tyr	dmR	Ala	Trp	mmK	Pro	pS	Ala	oxM	Tyr	spC	Arg	ack	Ser	Trp	hyP	Tyr				
5	ABI 49x cLC	N.D.	N.D.	91.2%	8	4	2	4	2	Tyr	(meT)	Ala	Trp	(dmK)	Pro	(Cys)	Ala	Met	Tyr	Gly	Arg	Ala	(Ser)	(Trp)	(meH)	(Tyr)				
6	ABI 494 HT	10%	150.0%	88.2%	17*	0	0	2	1	Tyr	(mmK)	Ala	Trp	(spC)	Pro	X	Ala	Met	Tyr	camC	Arg	ack	Ser	Trp	hyP	Tyr				
7	ABI 49x HT	30%	50.8%	88.2%	20†	0	0	0	0	Tyr	meR	Ala	Trp	mmK	Pro	pS	Ala	Met	Tyr	spC	Arg	ack	Ser	Trp	hyP	Tyr				
8	ABI 494 cLC	3%	134.3%	87.7%	18	0	2	0	0	Tyr	mmK	Ala	Trp	dmR	Pro	X	Ala	oxM	Tyr	spC	Arg	ack	Ser	Trp	hyP	Tyr				
9	ABI 494 HT	20%	40.4%	92.0%	17	0	2	0	1	Tyr	mmK	Ala	Trp	meH	Pro	pS	Ala	Met	Tyr	spC	Arg	ack	Ser	Trp	hyP	Tyr				
10	ABI 492 cLC	1.3%	75.3%	87.4%	18	0	2	0	0	Tyr	dmR	Ala	Trp	mmK	Pro	(oxM)	Ala	Met	Tyr	spC	Arg	ack	(pS)	Trp	hyP	Tyr				
11	ABI 494 HT	20%	53.7%	91.5%	14	3	0	0	3	Tyr	X	Ala	Trp	X	Pro	(pS)	Ala	Met	Tyr	X	Arg	(ack)	Ser	(hyP)	Tyr					
12	ABI 494 HT	20%	62.9%	89.4%	17	3	0	0	0	Tyr	(dmR)	Ala	Trp	(mmK)	Pro	pS	Ala	(MeI)	Tyr	spC	Arg	ack	Ser	Trp	hyP	Tyr				
13	ABI 494 HT	10%	73.6%	91.3%	18	1	1	0	0	Tyr	Tyr	Ala	Trp	mmK	Pro	pS	Ala	Met	Tyr	spC	Arg	ack	Ser	Trp	hyP	Tyr				
14	ABI 477 A	60%	3.5%	95.6%	13	0	1	0	6	Tyr	X	Ala	Trp	X	Pro	Pro	Ala	Met	Tyr	X	Arg	X	Ser	Trp	X	Tyr				
15	ABI 494 HT	40%	13.1%	92.3%	20	0	0	0	0	Tyr	dmR	Ala	Trp	mmK	Pro	pS	Ala	oxM	Tyr	spC	Arg	ack	Ser	Trp	hyP	Tyr				
16	ABI 494 HT	100%	50.2%	89.8%	18	0	2	0	0	Tyr	mmK	Ala	Trp	dmR	Pro	pS	Ala	Met	Tyr	spC	Arg	ack	Ser	Trp	hyP	Tyr				
17	ABI 494 HT	30%	83.4%	88.2%	17	1	0	2	0	Tyr	(mmK)	Ala	Trp	(dmR)	Pro	(pS)	Ala	Met	Tyr	spC	Arg	ack	Ser	Trp	hyP	Tyr				
18	HP G1005A	5%	54.9%	91.2%	11	0	2	0	7	Tyr	Arg	Ala	Trp	X	Pro	X	Ala	Met	Tyr	X	X	Ala	X	Trp	X	Tyr				
19	ABI 494 HT	6.7%	155.7%	87.1%	14	0	2	0	4	Tyr	PECys	Ala	Trp	X	Pro	X	Ala	Met	Tyr	X	Arg	His	Ser	Trp	X	Tyr				
20	ABI 477	100%	N.D.	N.D.	10	0	5	0	5	X	Tyr	Ala	Trp	mmK	Pro	Pro	X	Ala	X	Tyr	Cys-adct	Arg	ack	Ser	Ser	X	X			
					Total	301	13	14	10	22																				

(includes one s-propionamide-Cys identified as CAM-Cys) * oxM by MALDI or confirmed by MALDI
(includes dmR identified as meR)

Table 1. Facility number (Fac. No), Positive Correct (PC) - shown in black on table, Tentative Correct (TC) - in green, Positive Wrong (PW) - in red, Tentative Wrong (TW) - in pink, No call (X) in blue, N.D. could not be determined from information submitted. Amino acids: dmR, N,N-dimethylarginine; mmK, ε-N-methyl lysine; oxM, methionine sulfoxide; ack, ε-N-acetyl lysine; hyP, 4-hydroxyproline (see below for structures). Identification of oxM as Met was counted as correct, because of reduction of that amino acid to Met during sequencing. Company abbreviations are Applied Biosystems (ABI), Hewlett Packard/Agilent (HP). Initial yields were calculated from the formula: I.Y. = ((775 pmol) x (% Loaded) + ((area for Tyr residue) x (picomolar amount of Tyr standard)/(area for Tyr standard)). Some calculated initial yields (shown in red) were greater than 100%, indicating a problem with the reported data. Repetitive yields were calculated from the slopes of Excel trend lines for plots of Log A vs. sequencing cycle, where Log A values were the logs of the areas for the Tyr and Ala residues in the sequence. The repetitive yield is the inverse log of the slope of the trend line.

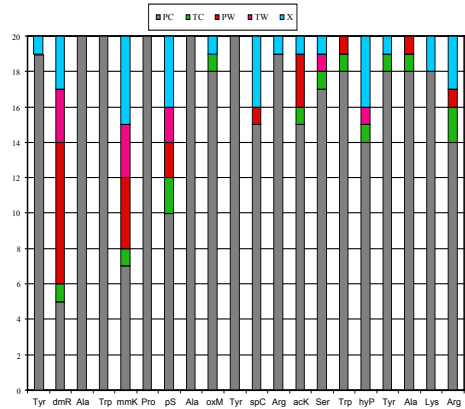


Figure 2. Graphic representation of identification accuracy.

	494 HT (n = 16)		cLC (n = 7)		477A (n=1)		HP G1005A (n = 1)		Porton 2090e (n=1)			
	Av RTnA's	Std Dev of RTnA's	Scale of Value	Av RTnA's of RTnA's	Std Dev of RTnA's	RTnA's	Observed value	RTnA's	Observed value	RTnA's	Observed value	
D	-0.30	0.03	4.29	-0.30	0.01	6.30	-0.30	5.68	-0.47	3.79	-0.31	6.37
N	-0.27	0.01	4.71	-0.28	0.01	6.77	-0.26	6.59	-0.23	7.35	-0.28	6.83
S	-0.22	0.01	5.35	-0.23	0.01	7.56	-0.21	7.59	-0.19	7.96	-0.24	7.55
Q	-0.20	0.01	5.62	-0.21	0.01	7.91	-0.19	8.00	-0.15	8.58	-0.21	7.90
T	-0.18	0.01	5.84	-0.19	0.01	8.18	-0.17	8.37	-0.17	8.25	-0.20	8.24
G	-0.16	0.01	6.11	-0.17	0.01	8.51	-0.15	8.89	-0.13	8.88	-0.17	8.69
E	-0.13	0.01	6.55	-0.14	0.00	9.07	-0.13	9.32	-0.36	5.44	-0.15	9.01
H	-0.05	0.01	7.68	-0.06	0.02	10.36	-0.05	10.93	-0.05	10.16	-0.05	10.48
A	0.00	0.00	8.30	0.00	0.00	11.30	0.00	12.03	0.00	10.86	0.00	11.35
R	0.11	0.02	9.73	0.10	0.02	12.90	0.07	13.48	0.13	12.84	0.12	13.24
Y	0.15	0.01	10.37	0.16	0.01	13.95	0.12	14.82	0.08	12.14	0.15	13.75
P	0.33	0.01	12.69	0.33	0.01	16.74	0.27	17.88	0.23	14.27	0.36	17.09
M	0.38	0.01	13.35	0.38	0.01	17.60	0.32	18.76	0.26	14.85	0.39	17.63
V	0.40	0.01	13.66	0.41	0.01	17.98	0.34	19.26	0.30	15.38	0.40	17.86
W	0.57	0.01	15.92	0.57	0.01	20.74	0.52	23.08	0.41	16.98	0.57	20.59
F	0.62	0.02	16.83	0.62	0.01	21.57	0.59	24.63	0.49	18.27	0.62	21.38
I	0.66	0.02	17.09	0.66	0.01	22.10	0.64	25.67	0.51	18.57	0.65	21.77
K	0.68	0.02	17.35	0.68	0.01	22.46	0.67	26.37	0.44	17.57	0.67	22.43
L	0.70	0.03	17.63	0.70	0.01	22.75	0.70	27.01	0.53	18.90	0.69	22.15
Av App			4.28			6.33		[5.78]		[6.37]		[6.37]
Av Leu			17.62			22.79		[27.01]		[18.90]		[22.45]
Av Ala			8.30			11.30		[12.03]		[10.86]		[11.35]

Table 2. To compensate for variations in the actual elution times of amino acid standards, the retention time (RT) for each standard was normalized to Ala using the following procedure: The retention time of Ala was subtracted from the retention time of each amino acid and this was divided by the total time interval between Asp and Leu (the first and last standards to elute). RTnA = the decimal fraction of this interval between the time when each amino acid eluted and the time when Ala eluted, with negative values indicating amino acids eluting before Ala. Std Dev RTnA is the calculated standard deviation for each RTnA value. Scaled values were determined by multiplying the RT difference between Asp and Leu by RTnA and adding the product to the RT for Ala. In cases where data were available from only one instrument of a particular type, only RTnA values and observed values for RTs are reported. Elution peaks for the HP G1005A that are not in chronological order are underlined. Results in this table (and Table 3) include data from instruments operated by members of the Edman Sequencing Research Group.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Tyr	(asmMe ₂ -Arg)	Ala	Trp	(ε-N-Me-Lys)	Pro	[P-Ser]	Ala	[Met-SO]	Tyr	[Cys-S-Pam]	Arg	(ε-N-Ac-Lys)	Ser	Trp	[4-OH-Pro]	Tyr	Ala	Lys	Arg

Figure 1. Structures of modified amino acids and sequence of test peptide.

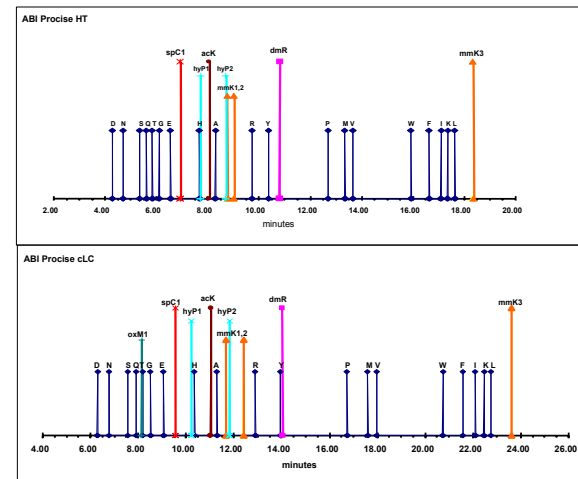


Figure 3. Time lines for elution of standard and modified amino acids on the Procise HT and Procise cLC, the two main types of sequencers used in this study. The data for these figures are taken from Tables 2 and 3.

cycle	Tyr	Procise HTs (n = 16)				Procise cLCs (n = 8)				ABI 477A (n = 1)				HP G1005A (n = 1)				Porton 2090e (n = 1)			
		scaled Values	Av SRTnA's	Std Dev SRTnA's	Relative Yield	Scaled Values	Av SRTnA's	Std Dev SRTnA's	Relative Yield	Std Dev	R.T.	SRTnA	Relative Yield	Std Dev	R.T.	SRTnA	Relative Yield	Std Dev	R.T.	SRTnA	Relative Yield
1	Tyr	10.39	0.16	0.01	104%	10%	13.95	0.16	0.00	103%	19%	14.46	0.11	91%	12.11	0.08	98%	13.76	0.15	96%	
2	dmR	10.81	0.19	0.04	98%	20%	14.04	0.17	0.04	90%	19%	14.77	0.13	242%	12.81	0.13	74%	15.77	0.27	92%	
3	Ala	8.34	0.00	0.01	123%	10%	11.30	0.00	0.00	121%	6%	11.86	-0.01	107%	10.77	-0.01	104%	11.34	0.00	106%	
4	Trp	15.95	0.57	0.01	39%	16%	20.76	0.58	0.01	37%	16%	23	0.51	95%	16.88	0.40	72%	20.58	0.57	103%	
5	mmK1	8.77	0.04	0.02	8%	1%	11.68	0.02	0.00	10%	23	-	-	0%	-	-	0%	-	-	0%	
5	mmK2	9.05	0.06	0.02	7%	3%	12.40	0.07	0.00	3%	-	-	-	0%	-	-	0%	-	-	0%	
5	mmK3	18.37	0.75	0.04	82%	22%	23.62	0.75	0.02	69%	29%	-	-	0%	17.35	0.43	85%	23.07	0.73	101%	
6	Pro	12.74	0.33	0.01	65%	14%	16.77	0.33	0.01	57%	12%	17.94	0.28	116%	14.18	0.22	61%	17.03	0.35	120%	
7	pS1	5.43	-0.22	0.02	5%	2%	7.49	-0.23	0.00	4%	1%	-	-	0%	-	-	0%	-	-	0%	
7	pS2	9.79	0.11	0.04	8%	2%	13.63	0.14	0.00	4%	1%	-	-	0%	-	-	0%	12.33	0.06	71%	
7	pS3	10.62	0.17	0.04	3%	1%	14.76	0.21	0.00	2%	-	-	-	0%	-	-	0%	-	-	0%	
7	pS4	12.41	0.31	0.30	3%	1%	15.05	0.23	0.00	4%	-	-	-	0%	-	-	0%	-	-	0%	
8	Ala	8.34	0.00	0.01	78%	13%	11.33	0.00	0.00	76%	13%	12.01	0.00	102%	10.76	-0.01	102%	11.3	0.00	104%	
9	oxM1																				