ABRF 2004 ESRG Study

Modified Amino Acids in Edman Sequencing

Members of the Committee

Benjamin J. Madden (Chair) - Mayo Clinic

Daniel C. Brune - Arizona State Univ.

J. Myron Crawford - Yale Univ.

Richard G. Cook - Baylor Col. of Medicine

Nancy D. Denslow - Univ. of Florida

Ryuji Kobayashi - Univ. of Texas M. D. Anderson Cancer Center

John M. Neveu - Harvard Univ.

Laurey Steinke (EB Liaison) Univ. of Nebraska Medical Center

Objectives of the Study

1. Compile data on elution characteristics of modified PTH amino acids with currently used equipment.
2. Test the ability of participating laboratories to correctly identify modified amino acids.
3. Compare initial yields and repetitive yields from laboratories analyzing the sample.

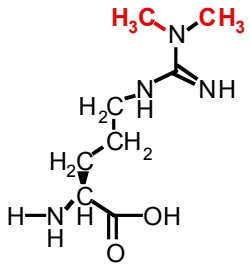
Description of the Sample

The sample was a synthetic polypeptide with the following sequence:

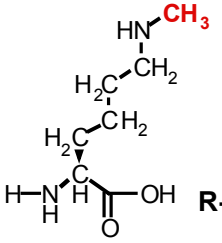
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Tyr-[asymMe$_2$-Arg]-Ala-Trp-[ε-N-Me-Lys]-Pro-[P-Ser]-Ala-[Met-SO]-Tyr-[Cys-S-β-Pam]-Arg-[ε-N-Ac-Lys]-Ser-Trp-[4-OH-Pro]-Tyr-Ala-Lys-Arg

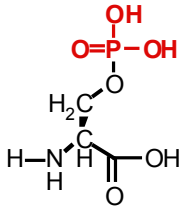
Structures of the modified amino acids are shown below:



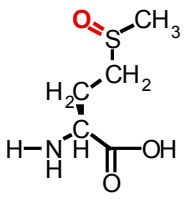
R-2, N,N-Dimethyl arginine



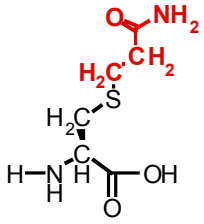
R-5, N-ε-Methyl lysine



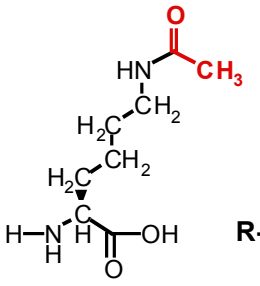
R-7, Phosphoserine



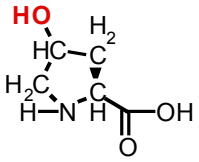
R-9, Methionine sulfoxide



R-11, Cysteine-S-β-propionamide



R-13, N-ε-Acetyl lysine



R-16, 4-Hydroxyproline

Other features of the test peptide:

1. Tyr and Ala residues throughout sequence for repetitive yield calculations.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Tyr-[asymMe$_2$-Arg]-Ala-Trp-[ε-N-Me-Lys]-Pro-[P-Ser]-Ala-[Met-SO]-Tyr-[Cys-S-β-Pam]-Arg-[ε-N-Ac-Lys]-Ser-Trp-[4-OH-Pro]-Tyr-Ala-Lys-Arg

2. Unmodified versions of Arg, Lys, Ser, and Pro for comparison with modified forms:

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Tyr-[asymMe$_2$-Arg]-Ala-Trp-[ε-N-Me-Lys]-Pro-[P-Ser]-Ala-[Met-SO]-Tyr-[Cys-S-β-Pam]-Arg-[ε-N-Ac-Lys]-Ser-Trp-[4-OH-Pro]-Tyr-Ala-Lys-**Arg**

3. C-terminal Arg to inhibit washout during sequencing (see above).

4. Trp residues for concentration measurement from A_{280}.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Tyr-[asymMe$_2$-Arg]-Ala-Trp-[ε-N-Me-Lys]-Pro-[P-Ser]-Ala-[Met-SO]-Tyr-[Cys-S-β-Pam]-Arg-[ε-N-Ac-Lys]-Ser-Trp-[4-OH-Pro]-Tyr-Ala-Lys-Arg

Sample Preparation

1. Peptide synthesis was performed on an ABI 433A synthesizer using Fmoc chemistry.
2. Fmoc-Lys(Me) and Fmoc-Arg$(Me)_2$ were coupled manually; all others automatically.
3. Unreacted residues were capped after each coupling cycle with acetic anhydride.
4. The peptide was cleaved from the resin and deprotected by reacting with TFA:ethanedithiol:thioanisole (8:1:1) for 3 hours, followed by precipitating and washing with diethyl ether.
5. The crude peptide was purified on a 10 x 250 mm Vydac C18 column and lyophilized.
6. Cys alkylation was done by adding 50 μl of 30% acrylamide in water to 1 mg peptide in 450 μl of 50 mM Tris, pH 8.1, in 50 % aqueous acetonitrile, and reacting for 50 min at room temperature.
7. The alkylated peptide was purified on a 2.1 x 220 mm Phenomenex Jupiter C18 column.
8. Aliquots of 25 μl were placed in PCR tubes, dried on a Speed Vac, and stored at -20 °C until shipped.
9. The exact quantity of peptide in a sample tube was determined to be 775 picomoles by amino acid analysis.

Characterization of Sample by ESRG

1. Peptide shown to have the predicted mass using MALDI-TOF mass spectrometry.

2. Amino acid analysis was performed and gave expected results.

3. All members of the ESRG sequenced the peptide in their laboratories to confirm the sequence.

Instructions to Participants

1. Dried sample contains 775 picomoles of a 20-amino acid peptide, and can be dissolved in 50 - 100 μl of 0.1 % TFA in 30% acetonitrile in water.

2. Load sample as you would normally, using enough to obtain a complete sequence.

3. A single sequence is expected; provide the assignment, peak area(s) and retention time(s) for the primary amino acid.

4. Some modified amino acids give more than one peak, so provide peak areas and retention times for those peaks too.

Instructions to participants (continued)

5. Use 3-letter codes to identify amino acids. Place tentative calls in parentheses. Use "X" to denote unidentified peaks, and "-" when no peak is observed.

6. An Excel file containing instructions for electronic data submission was sent to participants as an attachment to the e-mailed confirmation of each sample request.

7. For help in identifying modified amino acids, an earlier compilation of relative retention times of some of these by Crankshaw and Grant is posted on the ABRF web site.

8. A list of the specific modified amino acids in the test peptide (but not their positions in the sequence) was available on the web.

Requested Information

1. The amino acid sequence of the peptide.
2. Areas and retention times for peaks on each cycle of sequencing.
3. Picomolar amounts, areas and retention times for standards.
4. Information about sequencer, sample loading, HPLC equipment, gradient, solvents, flow rate, and column.

Sequencer Information

Manufacturer and Model	**16 ABI 49X-HT (avg age of 6.6 +/- 2.3 yrs)** **8 ABI 49X-cLC (avg age of 4.5 +/- 1.7 yrs)** **2 ABI 477A (8 and 11 yrs old, avg age 9.5 yrs)** **1 HP1005A (8 yrs old)** **1 Porton 2090e (14.5 yrs old)**
Overall Age	**2 to11 years Mean 6.6 +/- 2.9 years**
Reagents	**24 used all instrument manufacturer reagents** **1 used all except S4B** **1 only used mfg. R1, S2, R2** **1 said mixed** **1 said no**
TFA Cleavage	**23 liquid and 5 gas phase**
Sample Support	**21 GFF, 4 PVDF, 1 dual column (HP), 1 Porton fiberglass peptide filter** **1 said both GFF and PVDF (probably answered in general)**
Chemistry Cycle	**All ABI users seemed to use the cycles that matched the supp (ie. GFF or PVDF)** **HP user used Version 3.0**
DTT in S2	**8 yes and 20 no**
Other Additives	**27 no and 1 cLC user added TCEP to R4 and R5 (std)**
Sample Solvent	**18 used 0.1% TFA/30% acetonitrile** **2 used 0.1% TFA/20% acetonitrile** **2 used 0.3% TFA/ 30% acetonitrile** **1 used .05% TFA/70% acetonitrile** **1 HP user used 20% acetonitrile** **4 users were unclear**
% Loaded	**Ranged from 1.3 to 100% Mean 23.8% +/- 25.6 %**

HPLC Information

Column Mfg.,Type,Size

All 477A and all 49X-HT used ABI/Brownlee Spheri-5 PTH column 2.1x220 (or 150) mm
7 cLC used ABI/Brownlee Spheri-5 PTH 0.8x250 mm
1 cLC used a Higgins column 1x220 mm
HP used Agilent PTH-column 2.1x250 mm
Porton used Agilent Amino Acid Analysis Col.

Column Age

Ranged from 20 to 7000 cycles
Mean 1012 +/- 1607

Avg. Column Life

ABI 0.8 x 250 mm - 1200-1800 (mean 1130 +/- 484)
ABI 2.1 x 220 mm - 500 - 7500 (mean 1882 +/- 1557)
No info for Higgins and Agilent columns

Solvent A/Additives

For all users of ABI columns (25), it appears all use ABI Solvent A3 equivalent (3.5% THF). All except 1 get it from ABI.
Additives include Premix, acetone, TFA, phosphate buffer, tryptophan, formic acid

Higgins column Š 3.5% THF, Na acetate pH 3.9 not from ABI. Adds heptane sulfonate to A.

HP Š Agilent A, No additives

Porton - 0.01% triethylamine +3.5% THF in 81 mM NaOAc, pH=4.0

Solvent B/Additives

For all users of ABI columns (25) it appears all use ABI Solvent B2 (IPA/acetonitrile) equivalent with 5 making their own. 5 users add DMPTU to B.

HP Š Agilent B, No additives

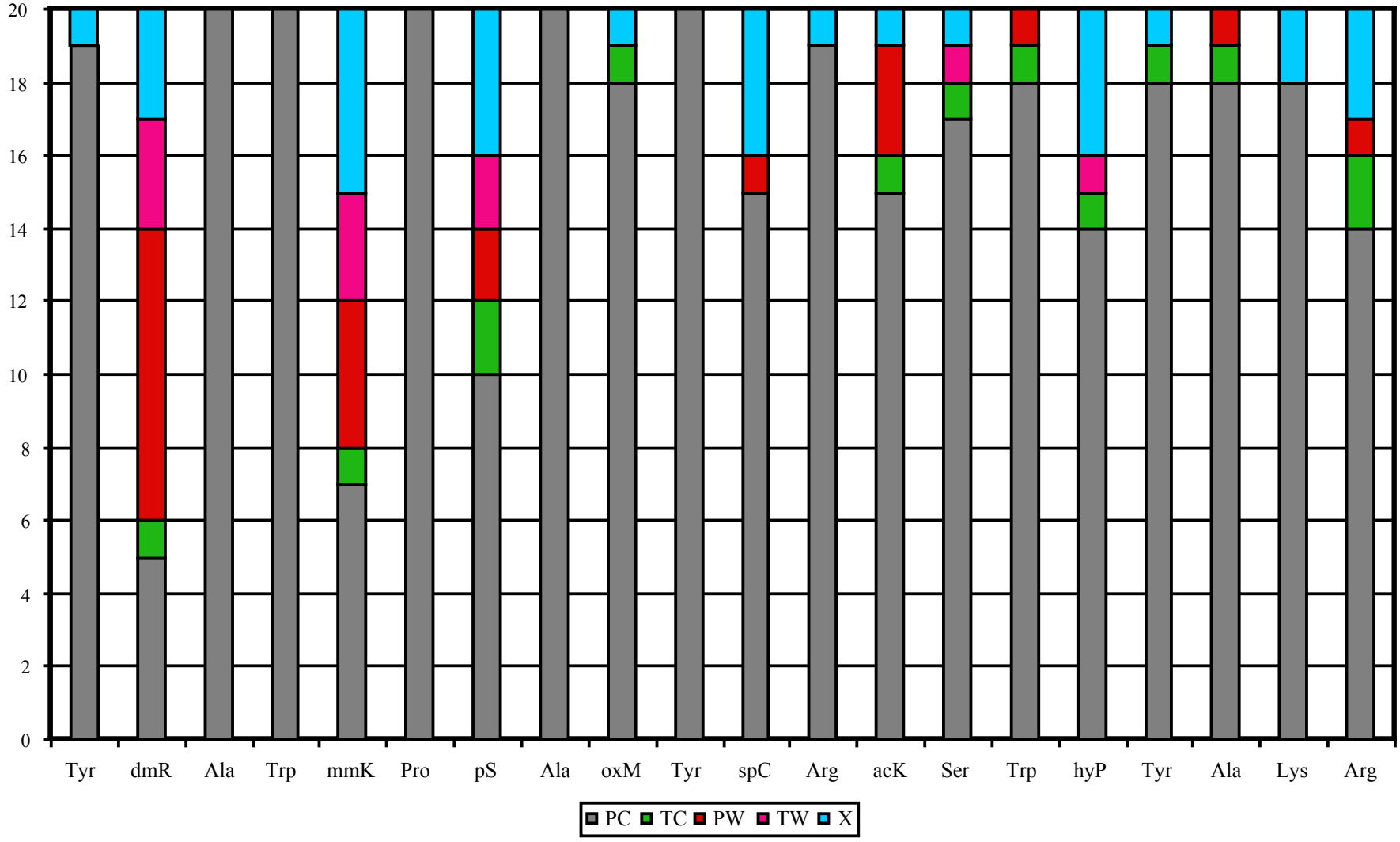
Higgins Š IPA/acetonitrile (homemade). No additives.

Porton Š Acetonitrile. No additives.

Solvent C/Additives

Only the HP user Š acetonitrile, no additives

Graphic Representation of Identification Accuracy



Fac. No.	Inst. Type	% Loaded	PC	TC	PW	TW	X
1	ABI 494 HT	20%	18	0	0	0	2
2	ABI 494 cLC	2%	18	0	2	0	0
3	ABI 494 HT	50%	19	1	0	0	0
4	ABI 491 HT	30%	20	0	0	0	0
5	ABI 49x cLC	N.D.	8	4	2	4	2
6	AB1 494 HT	10%	17□	0	0	2	1
7	ABI 49x HT	30%	20ą	0	0	0	0
8	ABI 494 cLC	3%	18	0	2	0	0
9	ABI 494 HT	20%	17	0	2	0	1
10	ABI 492 cLC	1.3%	18	0	0	2	0
11	ABI 494 HT	20%	14	3	0	0	3
12	ABI 494 HT	20%	17	3	0	0	0
13	ABI 494 HT	10%	18	1	1	0	0
14	ABI 477 A	60%	13	0	1	0	6
15	ABI 494 HT	40%	20	0	0	0	0
16	ABI 494 HT	100%	18	0	2	0	0
17	ABI 494 HT	30%	17	1	0	2	0
18	HP G1005A	5%	11	0	2	0	7
19	ABI 494 HT	6.7%	14	0	2	0	4
20	ABI 477	100%	10	0	5	0	5
		Total	**301**	**13**	**14**	**10**	**22**

□(includes one s-proprionamide-Cys identified as CAM-Cys)

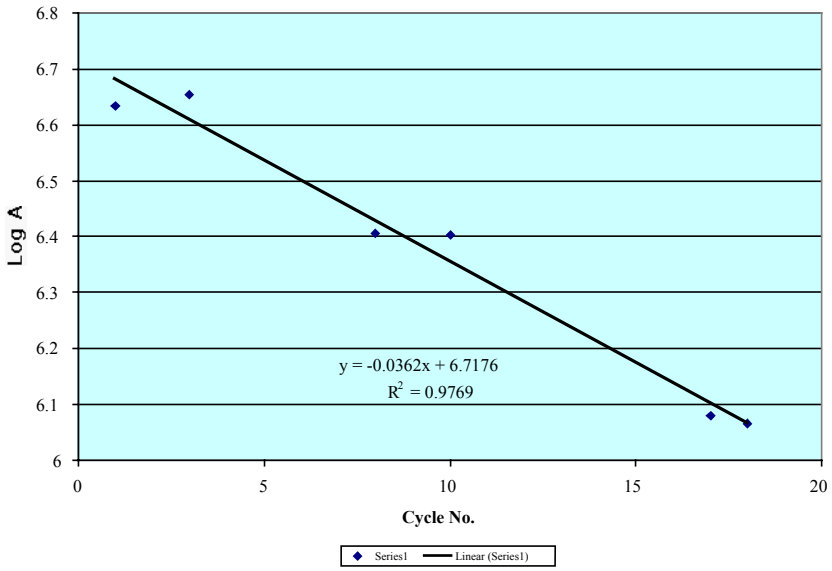
ą(includes dmR identified as meR)

Initial & Repetitive Yields

1. **Initial yields calculated from the equation:**

$$\text{I.Y.} = \frac{775\ \text{pmol} \times (\%\ \text{Loaded}) \times (\text{Area of Tyr Std peak})}{(\text{Area of Tyr 1 peak}) \times (\text{pmol Tyr Std})}$$

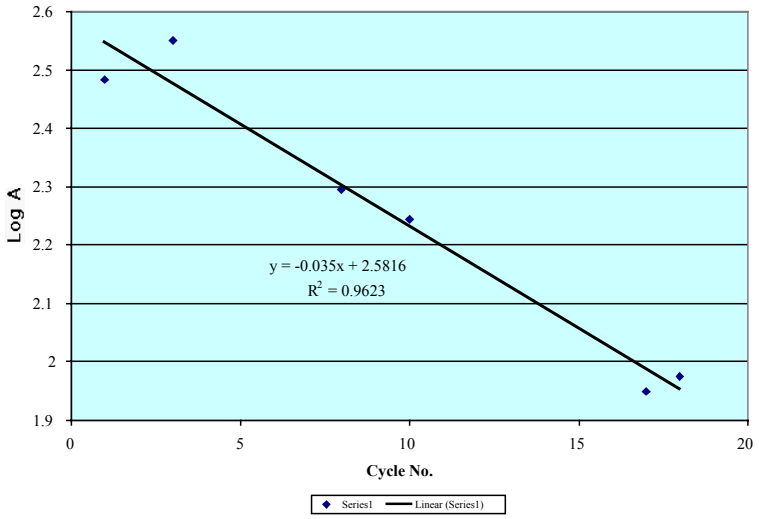
2. **Repetitive yields calculated from slope of trend line through plot of log A as a function of sequencing cycle, where values of A are peak areas of Tyr and Ala residues in the sample sequence. The slope of the trend line is the log of the R.Y.**



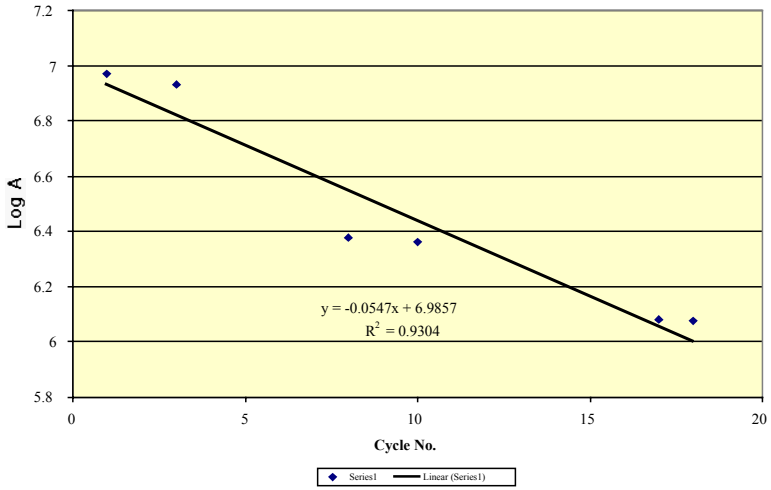
Curious R.Y. Observation

Facilities fell into 2 groups with respect to R.Y.

Facilities 9, 11, 12, 13, 14, 15, 18, 19 (linear)



Facilities 1, 2, 3, 4, 5, 6, 7, 8, 10, 16, 17 (lower R.Y. at beginning)



Correlation to Low Initial Yield

Lower repetitive yields at beginning of sequence generally correlated with lower overall repetitive yields (<90 %). Therefore, these lower repetitive yields were not due to sample washout, in which case repetitive yields would have dropped off later in the sequence.

No correlation was seen between low repetitive yields and type of sample support (GFF vs. PVDF), amount of sample loaded, sequencing chemistry (liquid TFA delivery vs. gas), or type of sequencer (although the HP sequencer, n=1, had a good beginning R.Y.). Pro was present in first half of sequence, but HyPro was present in the latter half.

Tabulated I.Y. & R.Y. Data

Fac. No.	Inst. Type	% Loaded	Init. Yield	Repet. Yield	PC	TC	PW	TW	X
1	ABI 494 HT	20%	72.1%	88.6%	18	0	0	0	2
2	ABI 494 cLC	2%	50.7%	88.1%	18	0	2	0	0
3	ABI 494 HT	50%	79.4%	88.4%	19	1	0	0	0
4	ABI 491 HT	30%	58.1%	89.3%	20	0	0	0	0
5	ABI 49x cLC	N.D.	N.D.	91.2%	8	4	2	4	2
6	AB1 494 HT	10%	150.0%	88.2%	17□	0	0	2	1
7	ABI 49x HT	30%	50.8%	88.2%	20ą	0	0	0	0
8	ABI 494 cLC	3%	134.3%	87.7%	18	0	2	0	0
9	ABI 494 HT	20%	40.4%	92.0%	17	0	2	0	1
10	ABI 492 cLC	1.3%	75.3%	87.4%	18	0	0	2	0
11	ABI 494 HT	20%	53.7%	91.5%	14	3	0	0	3
12	ABI 494 HT	20%	62.9%	89.4%	17	3	0	0	0
13	ABI 494 HT	10%	73.6%	91.3%	18	1	1	0	0
14	ABI 477 A	60%	3.5%	95.6%	13	0	1	0	6
15	ABI 494 HT	40%	13.1%	92.3%	20	0	0	0	0
16	ABI 494 HT	100%	50.2%	89.8%	18	0	2	0	0
17	ABI 494 HT	30%	83.4%	88.2%	17	1	0	2	0
18	HP G1005A	5%	54.9%	91.2%	11	0	2	0	7
19	ABI 494 HT	6.7%	155.7%	87.1%	14	0	2	0	4
20	ABI 477	100%	N.D.	N.D.	10	0	5	0	5

Retention Time (R.T.) Analysis

Retention times were normalized to Ala using the following formula:

$$RTnA = \frac{RTx - RTA}{RTL - RTD}$$

where RTx = R.T. of amino acid x, RTA = R.T. of Ala, RTL = R.T. of Leu, and RTD = R.T. of Asp.
Values of RTnA were determined for all of the standard amino acids and averaged for each type of sequencer.

Sample RTnA (SRTnA) values were then calculated for amino acids in the sample from the equation:

$$SRTnA = \frac{SRTx - RTA}{RTL - RTD}$$

where SRTx = R.T. of sample amino acid x, and the other terms are the values determined from the standards (see above). SRTnA values for each sample amino acid were then averaged for each type of sequencer.

This procedure adjusted for differences in overall length of a chromatographic run. Standard deviations from the averages were small.

Construction of scaled timelines for PTH-amino-acid elution

For each type of sequencer, scaled R.T. values for each amino acid were calculated as follows:

Scaled R.T. = SRTnA x (AvRTL-AvRTD) + AvRTA

where AvRTL, AvRTD, and AvRTA are the average retention times for Leu, Asp, and Ala.

Relative Yield

Peak area of an amino acid in the sample as a per cent of the area of a peak whose log would lie on the trend line at the cycle where the amino acid was observed

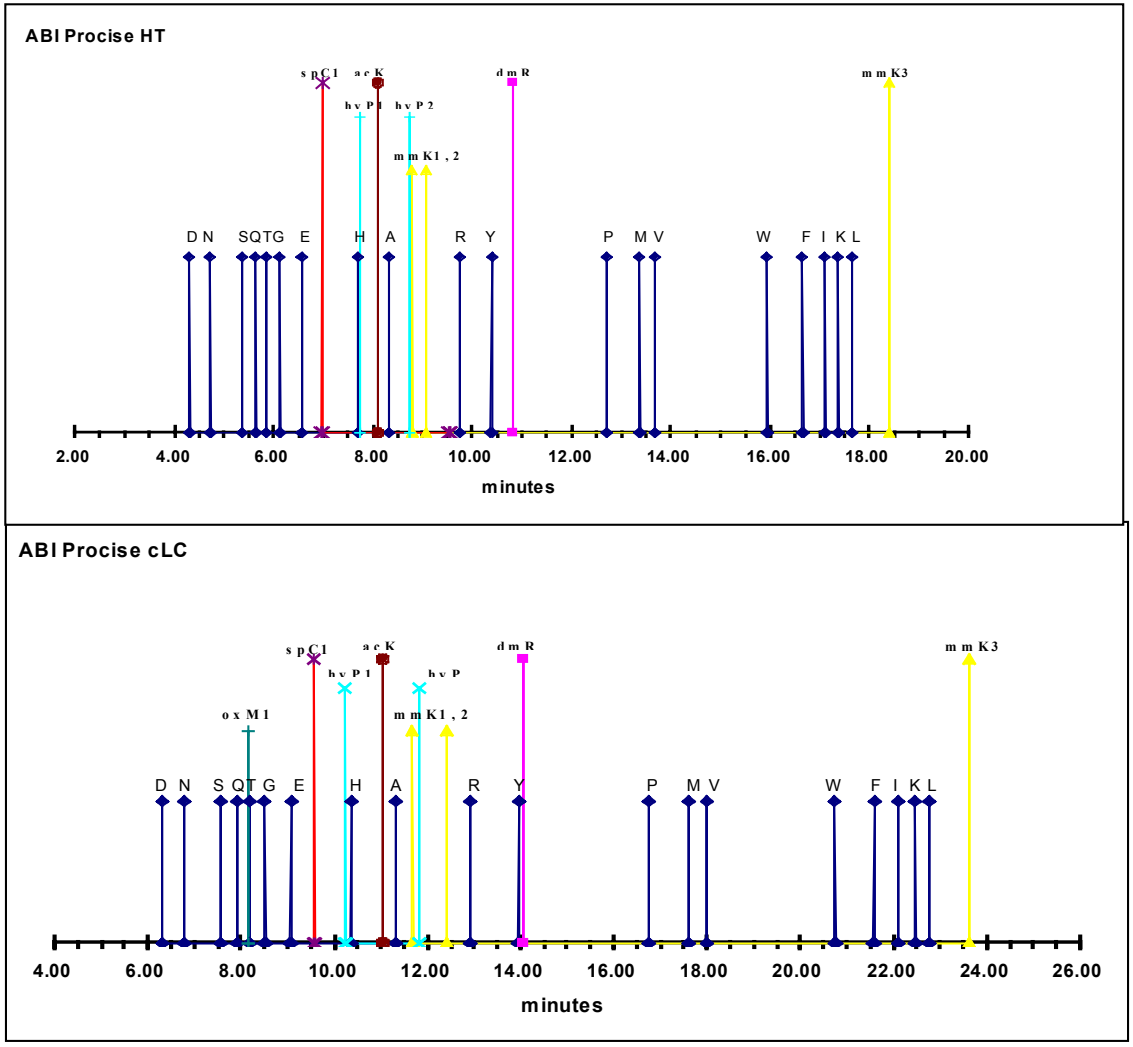
PTH Amino Acid Elution Parameters for the ABI Procise HT

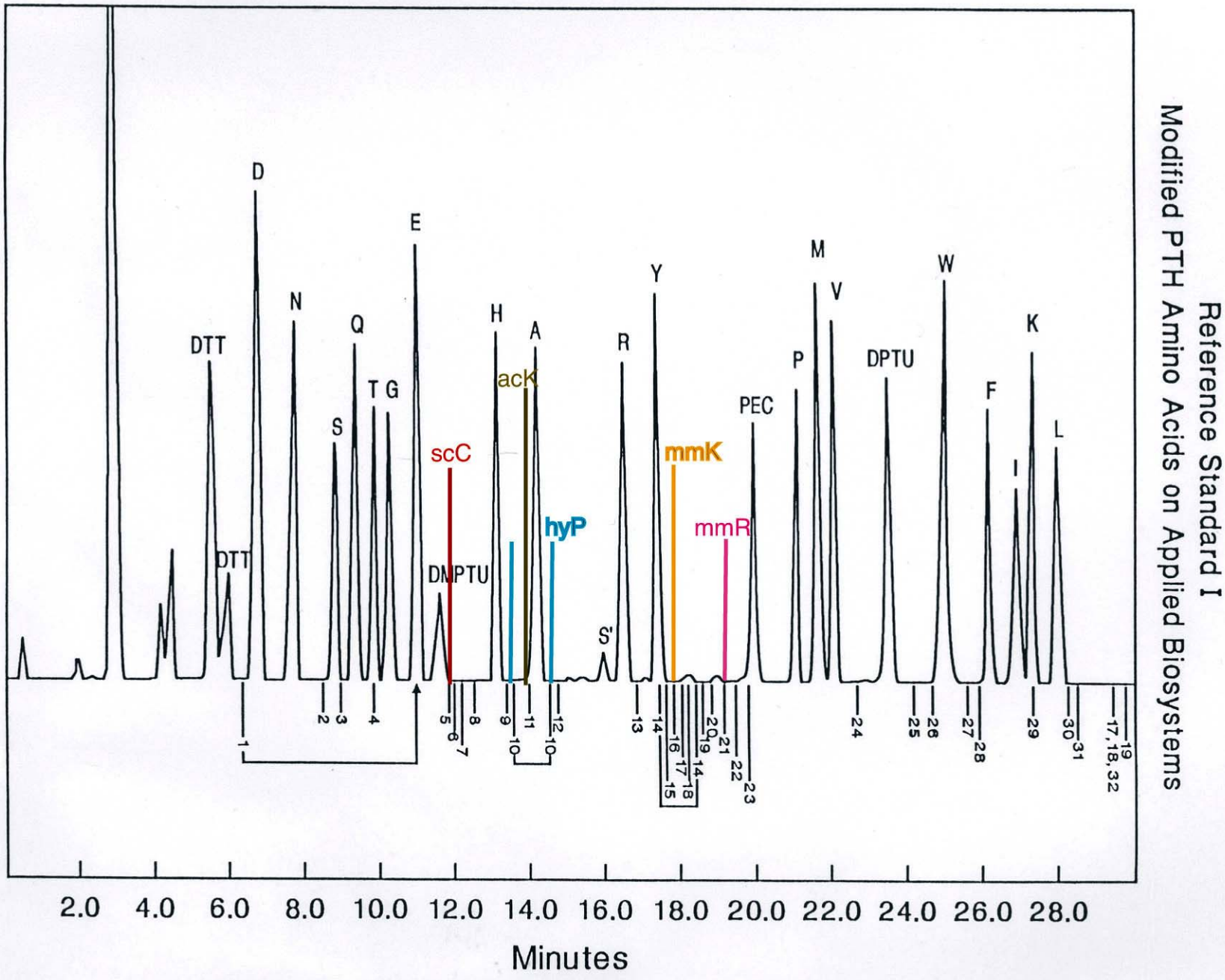
Cycle		n =	Relative Yield	Std Dev	Av SRTnA's	Std Dev SRTnA's	Scaled Values
1	Tyr	15	104%	10%	0.16	0.01	10.39
2	dmR	14	98%	20%	0.19	0.04	10.81
3	Ala	15	123%	10%	0.00	0.01	8.34
4	Trp	15	39%	16%	0.57	0.01	15.95
5	mmK1	2	8%	1%	0.04	0.02	8.77
	mmK2	11	7%	3%	0.06	0.02	9.05
	mmK3	11	82%	22%	0.75	0.04	18.37
6	Pro	15	65%	14%	0.33	0.01	12.74
7	pS1	8	5%	2%	-0.22	0.02	5.43
	pS2	7	8%	2%	0.11	0.04	9.79
	pS3	6	3%	1%	0.17	0.04	10.62
	pS4	6	3%	1%	0.31	0.30	12.41
8	Ala	15	78%	13%	0.00	0.01	8.34
9	oxM1	1	19%		-0.14	0.07	6.40
	oxM2	15	83%	16%	0.38	0.01	13.39
10	Tyr	15	94%	14%	0.16	0.01	10.41
11	spC1	14	79%	17%	-0.10	0.01	6.96
	spC2	3	11%	5%	0.09	0.18	9.52
12	Arg	15	68%	19%	0.11	0.02	9.76
13	acK	14	111%	21%	-0.02	0.01	8.08
14	Ser1	15	46%	17%	-0.22	0.01	5.37
	Ser2	1	9%		0.10	0.01	9.57
	Ser3	1	3%		0.21		11.04
	Ser4	1	3%		0.23		11.30
15	Trp	15	38%	19%	0.57	0.01	15.96
16	hyP1	13	30%	13%	-0.04	0.02	7.72
	hyP2	14	49%	21%	0.03	0.01	8.73
17	Tyr	15	109%	23%	0.16	0.01	10.42
18	Ala	15	118%	26%	0.00	0.01	8.34
19	Lys	15	154%	48%	0.68	0.02	17.40
20	Arg	14	61%	26%	0.11	0.02	9.75

PTH Amino Acid Elution Parameters for the ABI Procise cLC

Cycle		n =	Relative Yield	Std Dev	Av SRTnA's	Std Dev SRTnA's	Scaled Values
1	Tyr	7	103%	19%	0.16	0.00	13.95
2	dmR	7	90%	19%	0.17	0.04	14.04
3	Ala	7	121%	6%	0.00	0.00	11.30
4	Trp	7	37%	16%	0.58	0.01	20.76
5	mmK1	1	10%		0.02		11.68
	mmK2	1	3%		0.07		12.40
	mmK3	7	69%	29%	0.75	0.02	23.62
6	Pro	7	57%	12%	0.33	0.01	16.77
7	pS1	3	4%	1%	-0.23	0.00	7.49
	pS2	3	4%	4%	0.14		13.63
	pS3	1	2%		0.21		14.76
	pS4	1	4%		0.23		15.05
8	Ala	7	76%	13%	0.00	0.00	11.33
9	oxM1	3	24%	13%	-0.19	0.01	8.13
	oxM2	7	73%	13%	0.38	0.01	17.61
10	Tyr	7	97%	8%	0.16	0.00	13.95
11	spC1	7	85%	12%	-0.11	0.01	9.56
	spC2	0	0%				11.30
12	Arg	7	73%	18%	0.09	0.02	12.86
13	acK	7	111%	19%	-0.02	0.01	11.03
14	Ser1	6	42%	12%	-0.23	0.01	7.54
	Ser2	1	6%		0.03		11.84
	Ser3	1	8%		0.22		14.96
	Ser4	0	0%				
15	Trp	7	33%	19%	0.57	0.01	20.76
16	hyP1	6	21%	5%	-0.07	0.02	10.23
	hyP2	7	49%	4%	0.03	0.00	11.82
17	Tyr	7	102%	9%	0.16	0.01	13.91
18	Ala	7	111%	6%	0.00	0.01	11.26
19	Lys	7	84%	22%	0.68	0.01	22.48
20	Arg	7	62%	19%	0.09	0.02	12.85

Scaled Elution Times for Standard and Modified PTH Amino Acids for ABI Procise HT & Procise cLC




Reference Standard I
Modified PTH Amino Acids on Applied Biosystems
D
E
N
DTT
Q
H
A
T G
ackK
S
scC
R
Y
M
V
W
P
PEC
DPTU
K
F
I
L
mmK
hyP
mmR
DTT
DNPTU
S
1
2
3
4
5
6
7
8
9
10
11
10
12
13
14
16
17
18
15
19
14
20
19
21
22
23
24
25
26
27
28
29
30
31
19
17,18,32
2.0
4.0
6.0
8.0
10.0
12.0
14.0
16.0
18.0
20.0
22.0
24.0
26.0
28.0
Minutes

Results / Discussion

- **(asym) N,N - Dimethyl Arginine**
  - most incorrectly assigned (6/20)
    - not in Crankshaw/Grant handout
    - partial coelution with Tyr
      - more coelution with Tyr on cLC
      - coelutes with Arg on Agilent
  - peak area comparable to other a.a.'s

Results / Discussion

- **N-ε- Methyl Lysine**
  - Difficult to assign (8/20)
    - 1 peak (minor?) in Crankshaw/Grant handout shown after Tyr
    - With ABI instruments in this study, minor peaks eluted earlier (between Ala and Arg)
    - Major peak eluted ~ 1 min past Leu
      - missed by not looking beyond Leu?
    - Major peak eluted before Lys on Agilent
  - peak area comparable to other a.a.'s

Results / Discussion

- **Phosphoserine**
  - unstable to sequencing chemistry
    - no peaks were expected
  - Assignments were deduced at least 3 ways
    - noting low level serine peak with higher Ser' peak
    - yield of serine at cycle 14 much higher than cycle 7
    - 1 facility derivatized to PTH-S-2-aminoethanethiol

Results / Discussion

- **Methionine sulfoxide**
  - gets converted back to Met on most sequencers
    - few reporting peak around Gln/Thr
  - Most correctly identified (19/20)
    - most respondents reported as Met
    - deduced presence by MS
    - stated oxM will convert to Met
  - Edman sequencing not a reliable means to identify this modification

Results / Discussion

- **Cysteine -S-β-Propionamide**
  - result of alkylation with acrylamide
  - Many correctly identified (15/20)
  - major peak eluting between Glu / His
  - peak area comparable to other a.a.’s

Results / Discussion

- **N-ε-Acetyl-Lysine**
  - many correctly assigned (16/20)
    - elution similar to reported in Crankshaw/Grant handout
  - peak area higher than other a.a.'s
  - sharp peak eluting before Ala on ABI's and after Ala on Agilent

Results / Discussion

- **4-Hydroxyproline**
  - many correctly identified (15/20)
    - elution similar to reported in Crankshaw/Grant handout
    - most report 2 peaks; first eluting before or with His and the second after Ala
    - ratio of first peak area to second roughly 3:5
    - 4 report one peak after Ala

Conclusions

- **Relative retention times of the modified between similar instruments were very consistent.**
- **Sequencing and elution properties of the modified amino acids on the ABI Procise HT and cLC have been well characterized, along with single examples from ABI 477, Agilent, and Porton sequencers.**

Acknowledgements

Thanks to all the participating laboratories for taking the time to analyze the sample and sending in their results. Without their participation, this effort would not have been successful.