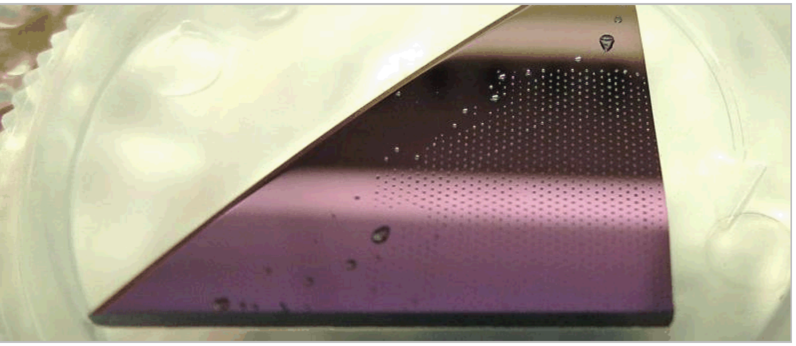
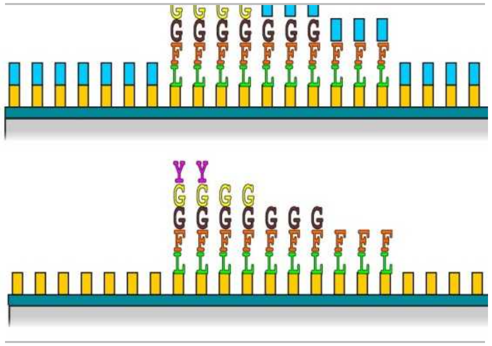
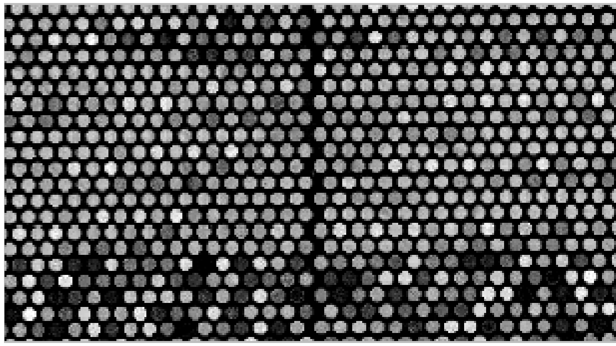
SAND2014-19416C

Exceptional service in the national interest







Analysis of Peptide Microarrays on Si Using ToF-SIMS

James (Tony) Ohlhausen and Conrad James; Sandia National Laboratories

David Smith; HealthTell, Stephen Johnston and Neal Woodbury; Arizona State University

AVS 61st International Symposium and Exhibition
Baltimore, MD, November 9-14, 2014, BI+ASTuA9
https://www.avs.org/Meetings-Exhibits/AVS-61st-International-Symposium-Exhibition/Information

Outline

- Acknowledgements
- Introduction
- Amino Acid Fragments
- Peptide Array Introduction
  - Manufacturing
  - ToF-SIMS Analysis Conditions
  - Data Processing
- Peptide Array ToF-SIMS Results
  - Comparison of Ion maps to peptide maps
  - Relationship of intensity with amino acid location
- Amino Acid Fragment Summary
- Summary

Acknowledgements

Sandia National Laboratories

Sandia National Laboratories

Peptide Design
Cathy Branda, Team Lead
Peter Anderson
Joe Schoeniger

Microfabrication
Stephen Casalnuovo, PI
David Santos
Jason Harper
Adrian Schiess
Michael Brumbach
David Wheeler
Mark Kinnan
Jamie Howell
Thayne Edwards
Komandoor Achyuthan
Tony Ohlhausen
Edward Heller
Jaclyn Murton
Katlin Schroeder

Systems Analysis
Julie Fruetel, Team Lead
Margot Kimura
David Reichmuth
Trisha Miller
Paula Krauter
Todd West

Lawrence Livermore National Laboratory
Tom Slezak, LLNL Team Lead
Beth Vitalis
Kevin Mcloughlin
Adam Zemla
Jimmy Thissen
Shea Gardner
Ana Paula De Oliveira Sales
Dan Faisso
Tom Edmunds
Carl Melius
Richard Wheeler

HealthTell™
Bill W. Colston
David Smith
Kathy Sykes
Robert Kumelis

Funding provided by:

Homeland Security
Science and Technology

ASU
Prof. Neal Woodbury
Prof. Stephen Johnston
ASU Biodesign Institute
Arizona State University

Innovations In Medicine:
Stephen Johnston
Neal Woodbury
Chris Diehnelt
Phillip Stafford
Alex Carpenter
Bart Legutki
Penny Gwynne
Douglas Daniel
Lucas Restrepo
Nidhi Gupta
Zhan-Gong Zhao
Donnie Shepard
Andrey Loskutov
Jinglin Fu
Rebecca Halperin
Valeriy Domenyuk
John Lainson
Meredith Blynn
Kari Kotlarczyk
Kristen Seifert
Loren Howell

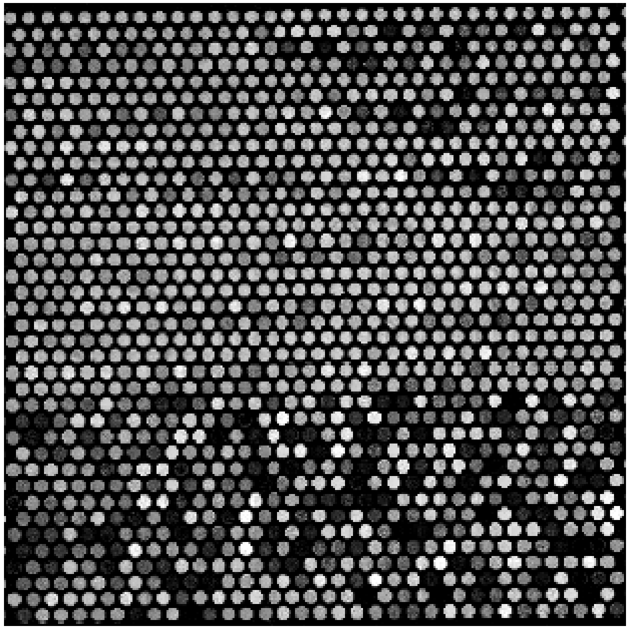
Army Research Office 1951
Defense Threat Reduction Agency
NSF
Bridging the Gap DARPA Powered by Ideas

John Galgiani: U. of Arizona
Hoda Anton-Culver: UC Irvine
Sam Hanash: Fred Hutchinson Cancer Center
Adi Gazdar: UT Southwestern
Adrienne Scheck: Mayo Clinic
Dawn E. Jaroszewki: Mayo Clinic
Matthew Grieving: Nextval, Inc.
George Poste: Complex Adaptive Systems Initiative

Introduction

Team would like a quick and robust quality control analysis tool.

Given a "known" monolayer peptide microarray on Si……



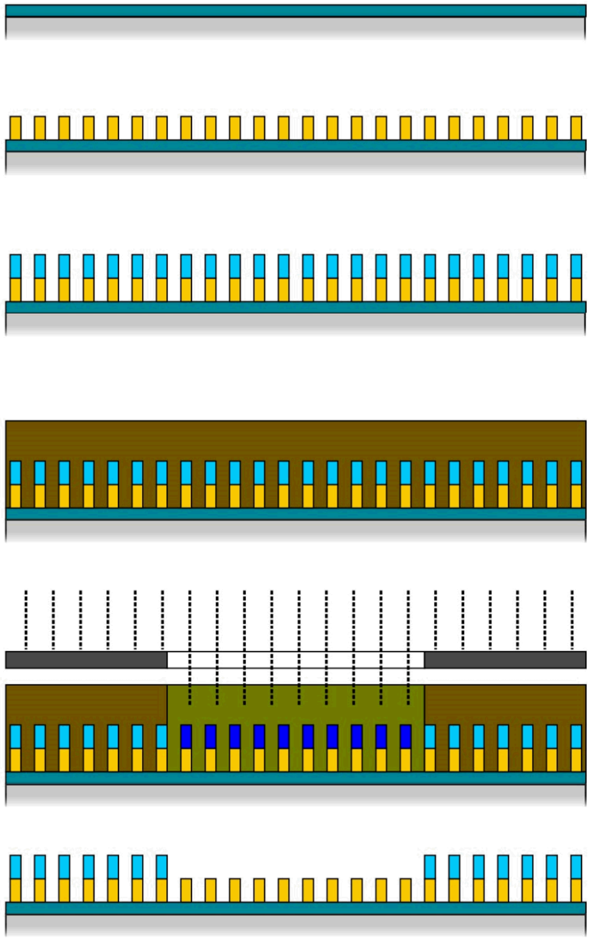
What can ToF-SIMS tell us?

- Peptide sequencing?
- Amino acid maps?
- Peptide density?
- Other?

- Can ToF-SIMS be used to verify microarray?

In this study, we are attempting to determine what information ToF-SIMS can provide.

Microarray fabrication photolithography process creates patterned peptide sequences



SiO_2 on Si

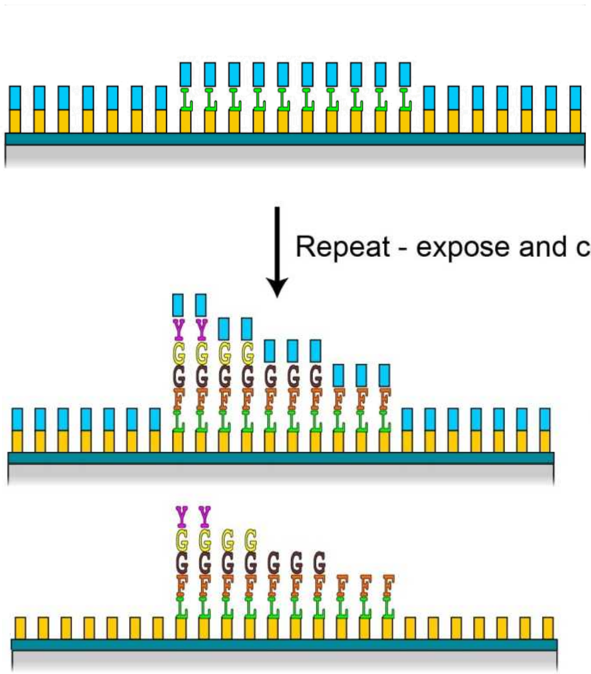
Aminosilane (ATPES)

Boc'd linker

Photoresist coating

Light exposure

Photoresist removal

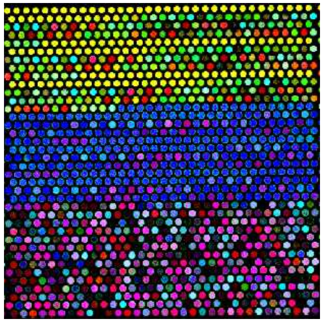


L Coupling

Repeat - expose and couple AAs

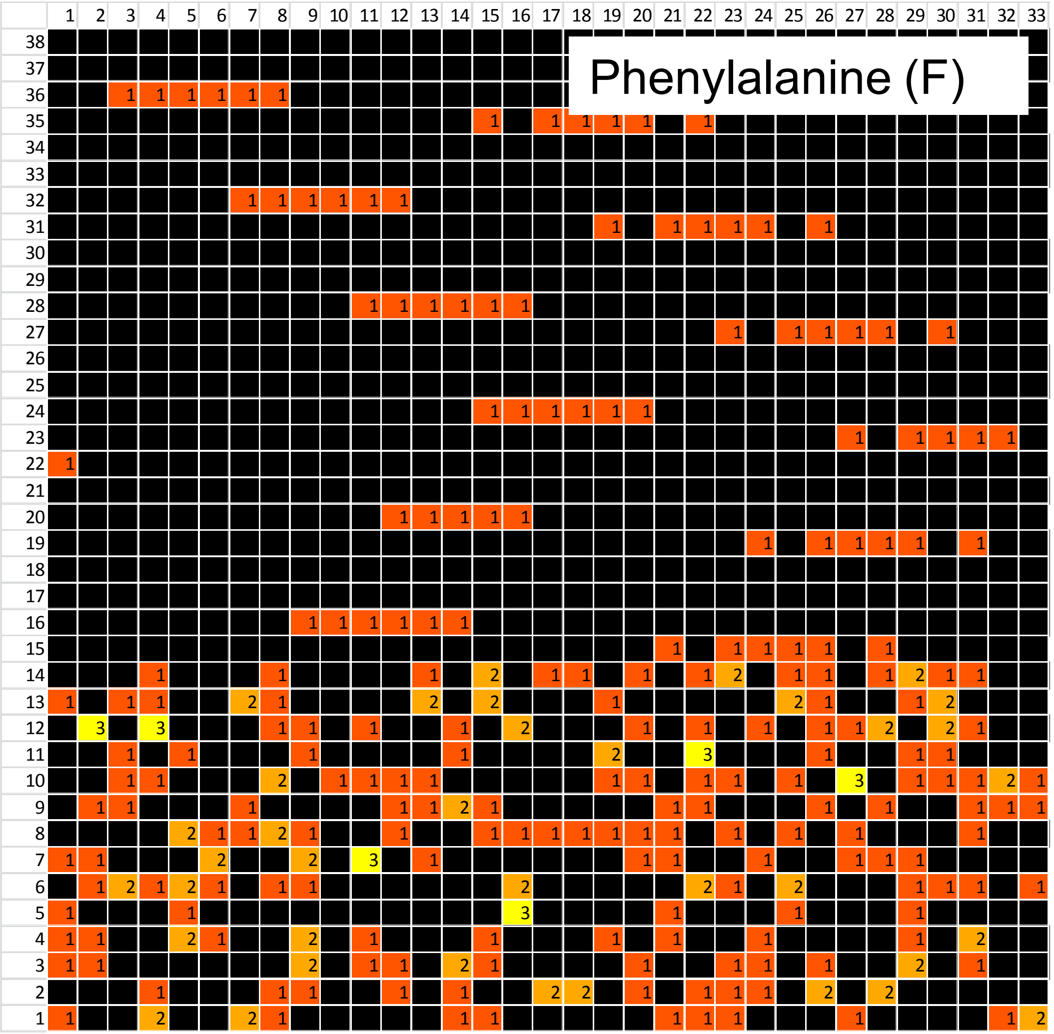
YGGFL 5mer build

Chemical deprotection



The result: A monolayer peptide array with custom sequences at each dot.

10x10mm peptide array consists of 38x33 (1254) 200µm diameter dots containing a peptide monolayer on Silicon



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
38																																	
37																																	
36			1	1	1	1	1	1																									
35															1		1	1	1	1		1											
34																																	
33																																	
32							1	1	1	1	1	1																					
31																			1		1	1	1	1		1							
30																																	
29																																	
28											1	1	1	1	1	1																	
27																							1		1	1	1	1		1			
26																																	
25																																	
24															1	1	1	1	1	1													
23																											1		1	1	1	1	
22	1																																
21																																	
20												1	1	1	1	1																	
19																								1		1	1	1	1		1		
18																																	
17																																	
16									1	1	1	1	1	1																			
15																					1		1	1	1	1		1					
14				1				1					1		2		1	1		1		1	2		1	1		1	2	1	1		
13	1		1	1			2	1					2		2				1						2	1			1	2			
12		3		3				1	1		1			1		2				1		1		1		1	1	2		2	1		
11			1		1				1					1					2			3				1			1	1			
10			1	1				2		1	1	1	1						1	1		1	1		1		3		1	1	1	2	1
9		1	1				1					1	1	2	1						1	1				1		1			1	1	1
8					2	1	1	2	1			1			1	1	1	1	1	1	1		1		1		1				1		
7	1	1				2			2		3		1							1	1			1			1	1	1				
6		1	2	1	2	1		1	1							2						2	1		2				1	1	1		1
5	1				1											3					1				1				1				
4	1	1			2	1			2		1				1				1		1			1					1		2		
3	1	1							2		1	1		2	1					1			1	1		1			2		1		
2				1				1	1			1		1			2	2		1		1	1	1		2		2					
1	1			2			2	1						1	1						1	1	1				1					1	2

Peptide Sequence

(5-22 amino acid units long)

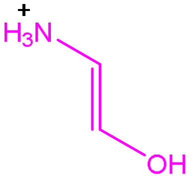
Si--(ATPES)--L…GSGG

Row	Column	Sequence
38	1	LGSGG

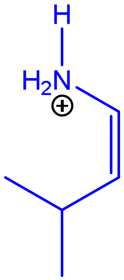
29	3	LGSGG
28	3	LAVRGSGG
27	3	LVPGGSGG
26	3	LWHKGSGG
25	3	LRDSHNYLRHSVVGSGG
24	3	LRDYADSHNYLRHSVVGSGG
23	3	LRKHVDSHNYLRHSVVGSGG
22	3	LRRVWDSHNYLRHSVVGSGG
21	3	LRDSHNYLRHSVVGSGG
20	3	LRSAWDSHNYLRHSVVGSGG

3	33	LWDYGQGHYLSWGEGSGG
2	33	LEPYPHNEVYWWGSGG
1	33	LWHWFFQLGSGG

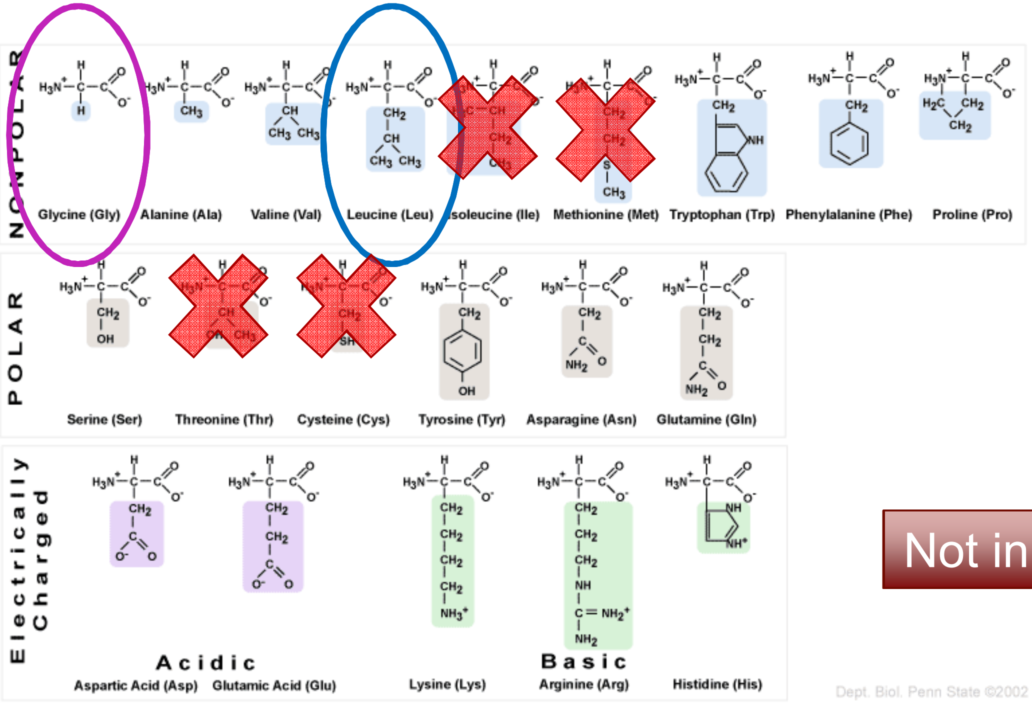
Appropriate peaks must be used: Immonium Ions are common



Chemical Formula: $C_2H_6NO^+$
m/z: 60.04 (100.0%), 61.05 (2.3%)
G-Immonium Ion



Chemical Formula: $C_5H_{12}N^+$
L-Immonium Ion:



Dept. Biol. Penn State ©2002

Positive Ion Amino Acid Peaks

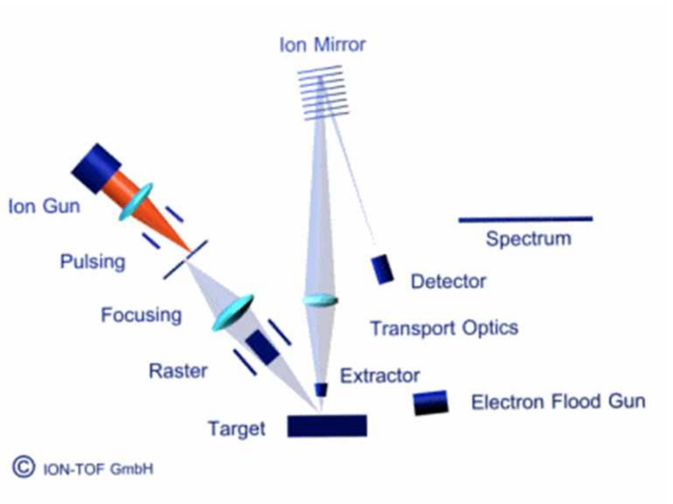
			Immonium Ion	
Amino Acid	**Abbreviation**	**Symbol**	**Formula**	**Mass**
Alanine	Ala	A	C2H6N	44.051
Arginine	Arg	R	none	
Asparagine	Asn	N	C3H7N2O	87.0598
Aspartic acid	Asp	D	C3H6NO2	88.0436
Cysteine	Cys	C	none	
Glutamine	Gln	Q	none	
Glutamic acid	Glu	E	none	
Glycine	Gly	G	CH4N	30.034
Histidine	His	H	C5H8N3	110.0767
Isoleucine	Ile	I	C5H12N	86.101
Leucine	Leu	L	C5H12N	86.101
Lysine	Lys	K	C5H10N	84.082
Methionine	Met	M	none	
Phenylalanine	Phe	F	C8H10N	120.08
Proline	Pro	P	C4H8N	70.0653
Serine	Ser	S	C2H6NO	60.0447
Threonine	Thr	T	C3H8NO	74.0606
Tryptophan	Trp	W	C10H11N2	159.09
Tyrosine	Tyr	Y	C8H10NO	136.0757
Valine	Val	V	C4H10N	72.0835

🟥 Not in peptide array

TOF-SIMS Analysis Conditions

- Data were acquired using an Ion-TOF ToF.SIMS 5 Instrument
  - Stage scanning: 11x11mm (121mm^2) total image size, 5µm/pixel, 30 primary ion shots/pixel
  - 25kV Bi_4^+ primary ion, HC-BUNCHED mode, acquiring positive secondary ions
  - Acquisition time: 7 Hours, 34 minutes
- Result: 2204x2204 pixel image at full spectral resolution
- Peak images reconstructed to show proper array:

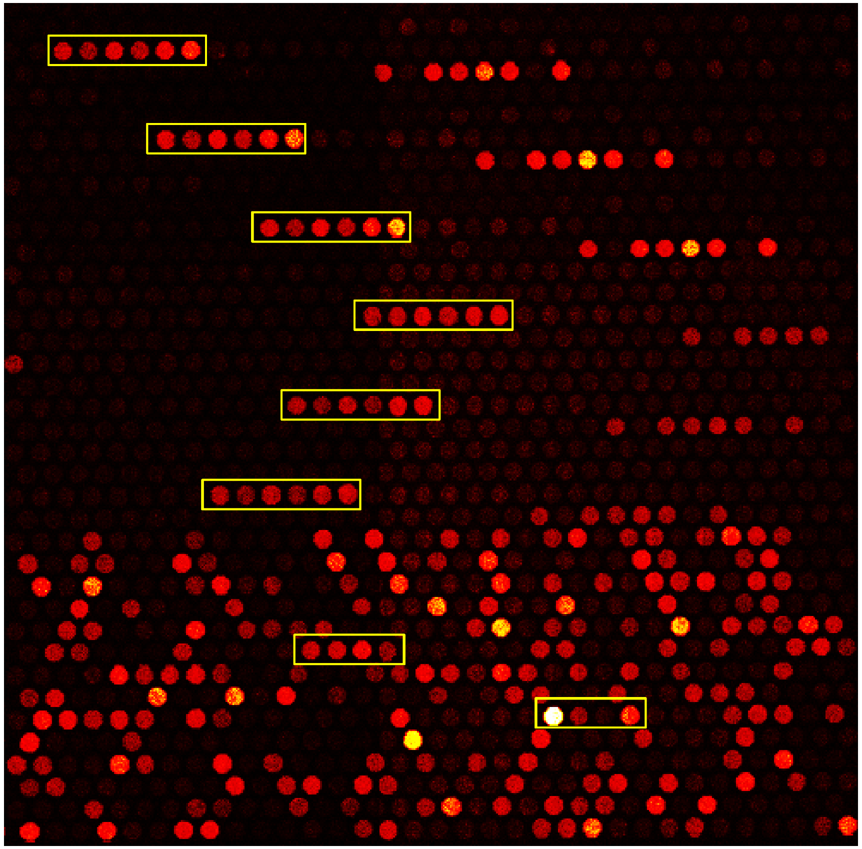
ION-TOF TOF-SIMS.5



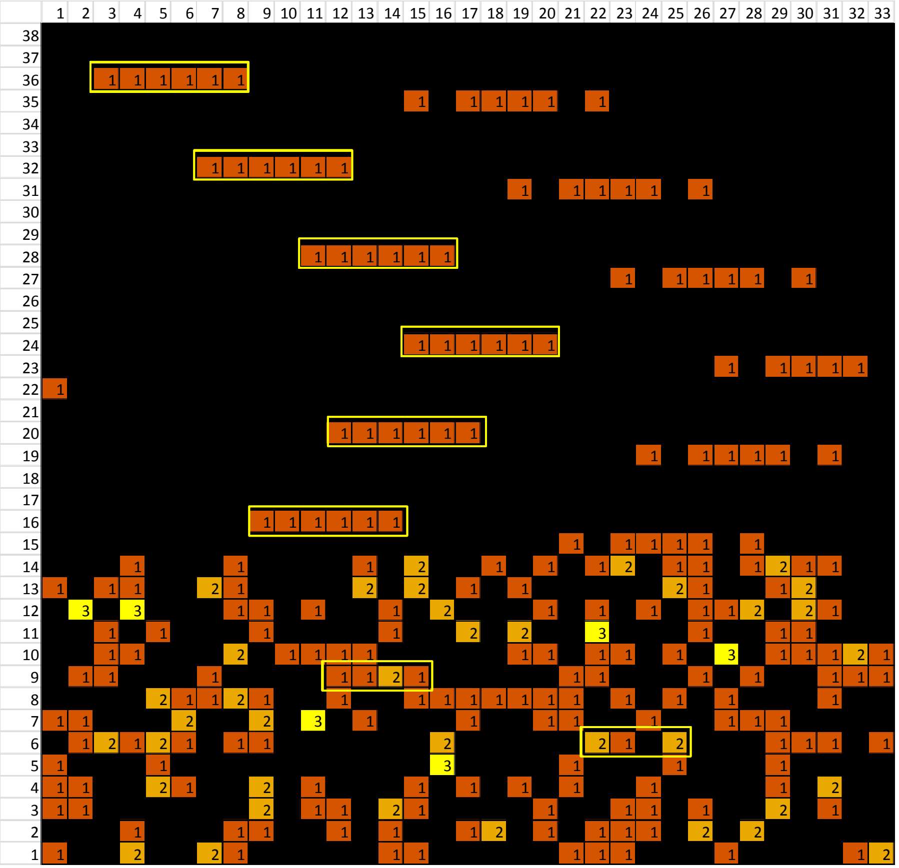
- Key ToF-SIMS Features:
  - Extreme surface sensitivity
  - Molecular and elemental identification
  - Trace element analysis on Si surfaces
  - Imaging from sub micron areas to full stage

Ion signatures correlate to peptide array map

TOF-SIMS scan for $C_8H_{10}N^+$ 120D
(F Immonium Ion)



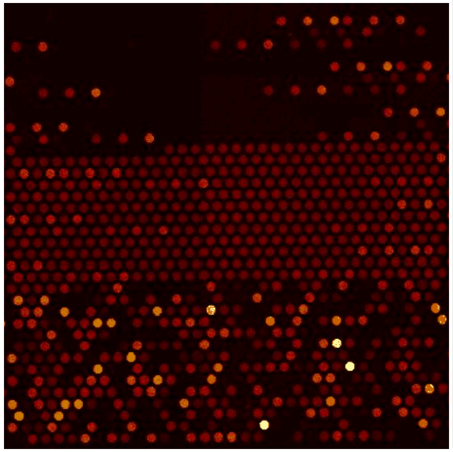
Phenylalanine map of peptide array:



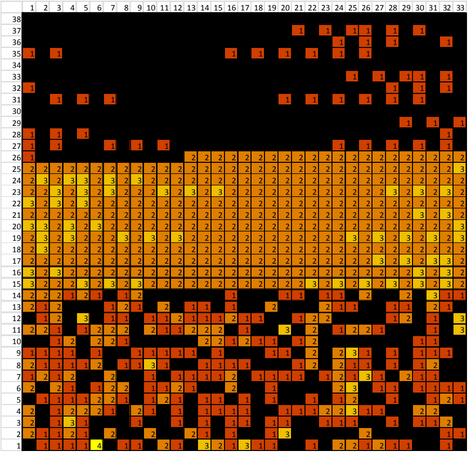
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
38																																	
37																																	
36			1	1	1	1	1	1																									
35															1		1	1	1	1		1											
34																																	
33																																	
32							1	1	1	1	1	1																					
31																			1		1	1	1	1		1							
30																																	
29																																	
28											1	1	1	1	1	1																	
27																							1		1	1	1	1		1			
26																																	
25																																	
24															1	1	1	1	1	1													
23																											1		1	1	1	1	
22	1																																
21																																	
20												1	1	1	1	1	1																
19																								1		1	1	1	1		1		
18																																	
17																																	
16									1	1	1	1	1	1																			
15																					1		1	1	1	1		1					
14				1				1					1		2			1		1		1	2		1	1		1	2	1	1		
13	1		1	1			2	1					2		2		1		1						2	1			1	2			
12		3		3				1	1		1			1		2				1		1		1		1	1	2		2	1		
11			1		1				1					1			2		2			3				1			1	1			
10			1	1				2		1	1	1	1						1	1		1	1		1		3		1	1	1	2	1
9		1	1				1					1	1	2	1						1	1				1		1			1	1	1
8					2	1	1	2	1			1			1	1	1	1	1	1	1		1		1		1				1		
7	1	1				2			2		3		1				1			1	1			1			1	1	1				
6		1	2	1	2	1		1	1							2						2	1		2				1	1	1		1
5	1				1											3					1				1				1				
4	1	1			2	1			2		1				1		1		1		1			1					1		2		
3	1	1							2		1	1		2	1					1			1	1		1			2		1		
2				1				1	1			1		1			1	2		1		1	1	1		2		2					
1	1			2			2	1						1	1						1	1	1				1					1	2

Other signatures match well with array maps

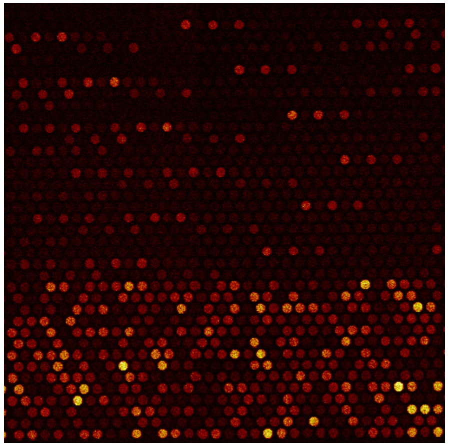
$C_4H_{10}N^+$ 72D (V Immonium Ion)



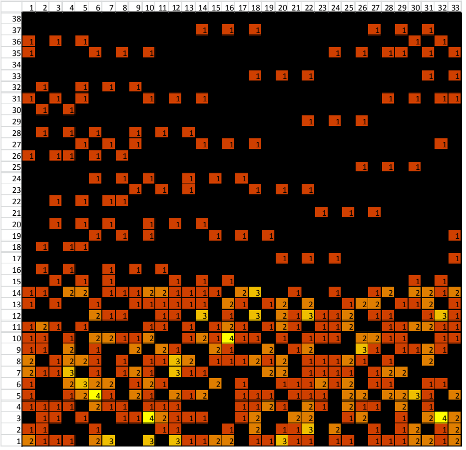
Valine map from Sequence File



$C_9H_8N^+$ 130D (W related Ion)



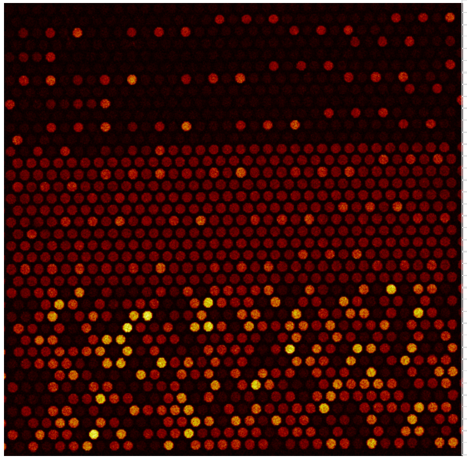
Tryptophan map from Sequence File



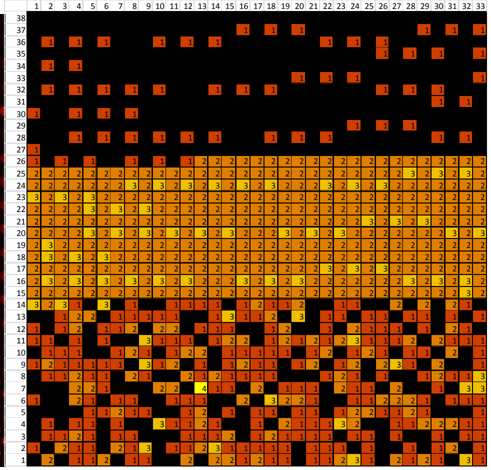
Valine

Tryptophan

$C_5H_8N_3^+$ 110D (H Immonium Ion)



Histidine map from Sequence File

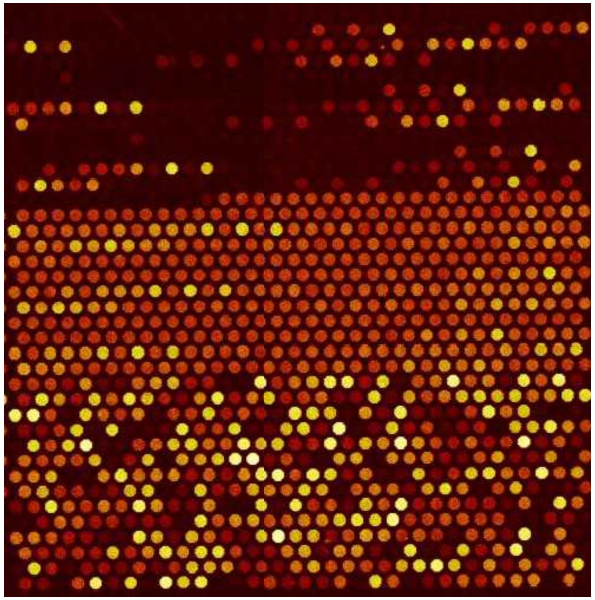


Hystidine

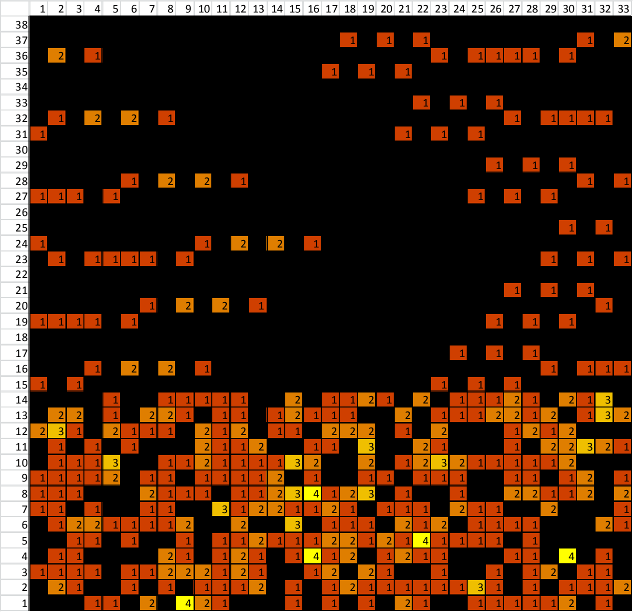
Some ions represent multiple amino acids.

- Degenerate immonium ions produced by multiple amino acids

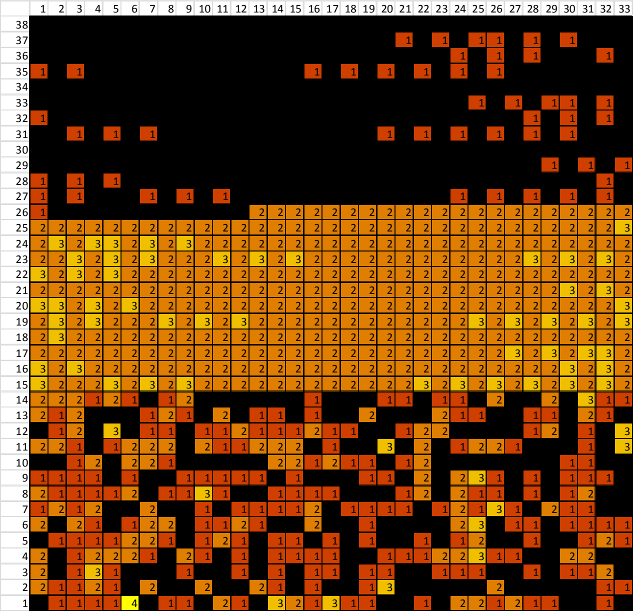
$C_4H_8N^+$ 70D



Proline map

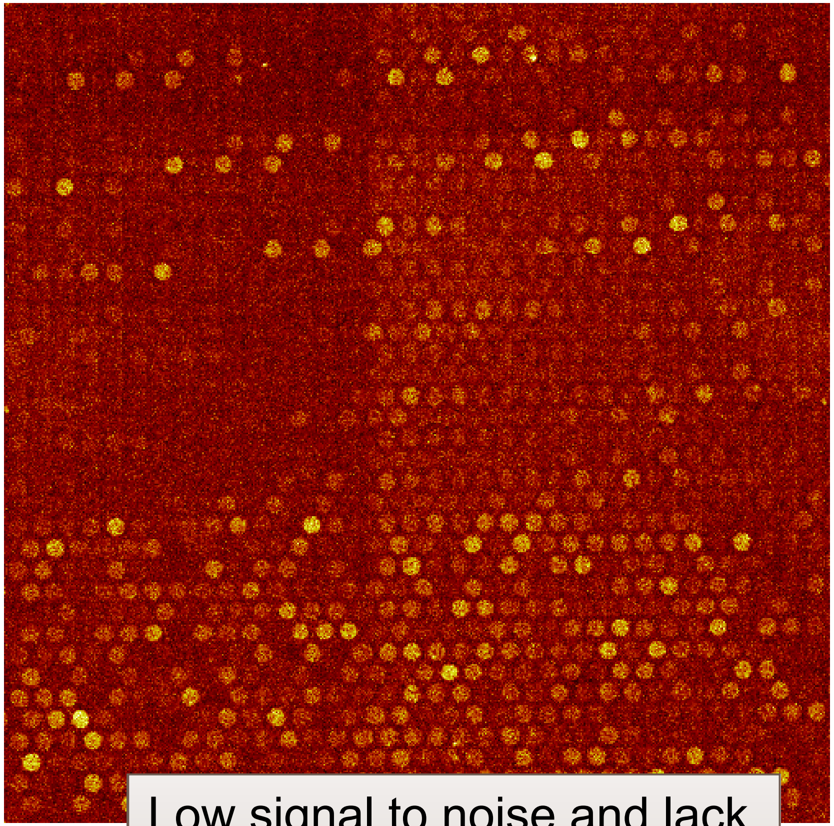


Valine map



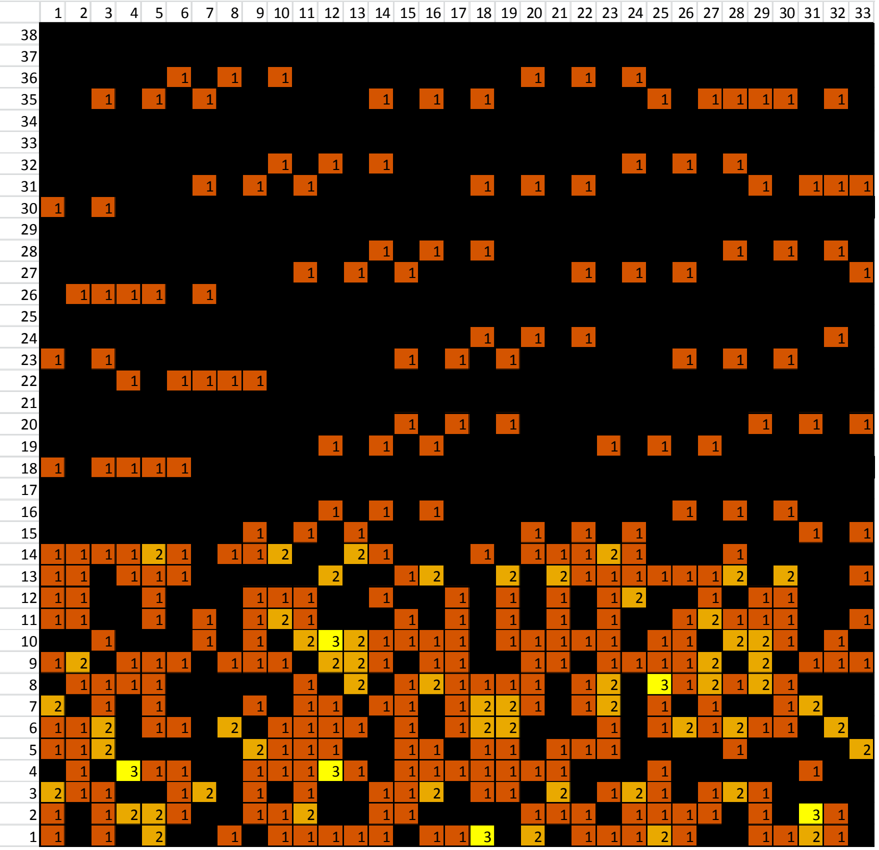
Lysine (K) Maps

$C_5H_{10}N^+$ 84D (K Immonium Ion)



Low signal to noise and lack of discrimination from the background makes this a poor ion choice.

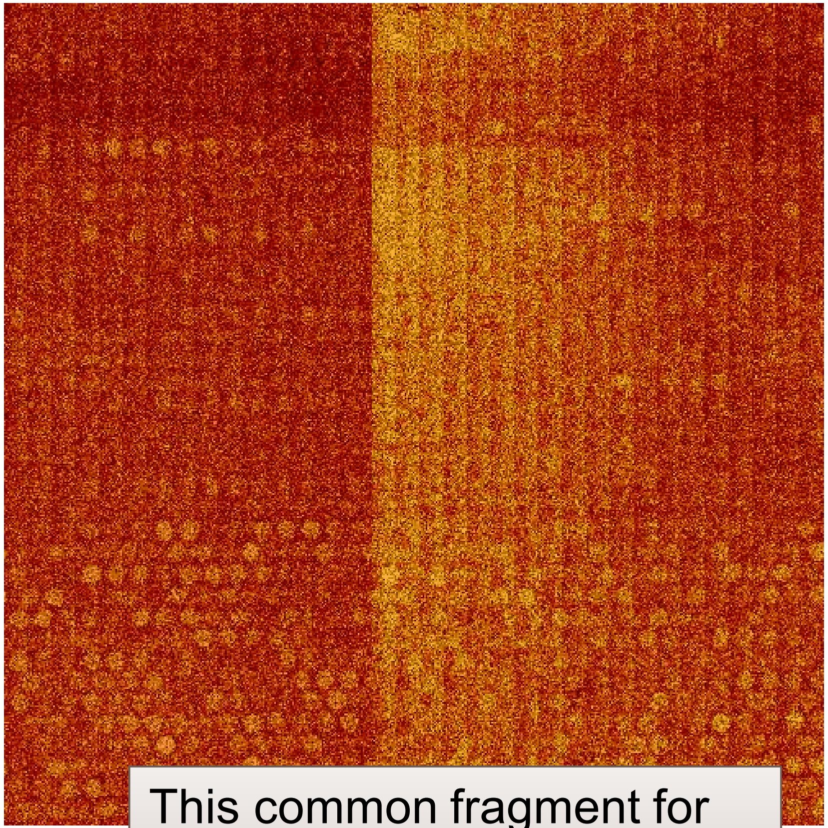
Lysine map from Sequence File
(Number of “K” in each spot sequence)



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
38																																	
37																																	
36						1		1		1										1		1		1									
35			1		1		1							1		1		1							1		1	1	1	1		1	
34																																	
33																																	
32										1		1		1										1		1		1					
31							1		1		1							1		1		1							1		1	1	1
30	1		1																														
29																																	
28														1		1		1										1		1		1	
27											1		1		1							1		1		1							1
26		1	1	1	1		1																										
25																																	
24																		1		1		1										1	
23	1		1												1		1		1							1		1		1			
22				1		1	1	1	1																								
21																																	
20															1		1		1										1		1		1
19												1		1		1							1		1		1						
18	1		1	1	1	1																											
17																																	
16												1		1		1										1		1		1			
15									1		1		1							1		1		1							1		1
14	1	1	1	1	2	1		1	1	2			2	1				1		1	1	1	2	1				1					
13	1	1		1	1	1						2			1	2			2		2	1	1	1	1	1	1	2		2			1
12	1	1			1				1	1	1			1			1		1		1		1	2			1		1	1			
11	1	1			1		1		1	2	1				1		1		1		1		1			1	2	1	1	1			1
10			1				1		1		2	3	2	1	1	1	1		1	1	1	1	1		1	1		2	2	1		1	
9	1	2		1	1	1		1	1	1		2	2	1		1	1			1	1		1	1	1	1	2		2		1	1	1
8		1	1	1	1						1		2		1	2	1	1	1	1		1	2		3	1	2	1	2	1			
7	2		1		1				1		1	1		1	1		1	2	2	1		1	2		1		1			1	2		
6	1	1	2		1	1		2		1	1	1	1		1		1	2	2				1		1	2	1	2	1	1		2	
5	1	1	2						2	1	1	1			1	1		1	1		1	1	1					1					2
4		1		3	1	1			1	1	1	3	1		1	1	1	1	1	1	1				1						1		
3	2	1	1			1	2		1		1			1	1	2		1	1		2		1	2	1		1	2	1				
2	1		1	2	2	1			1	1	2			1	1					1	1	1		1	1	1	1		1		3	1	
1	1		1		2			1		1	1	1	1	1		1	1	3		2		1	1	1	2	1			1	1	2	1	

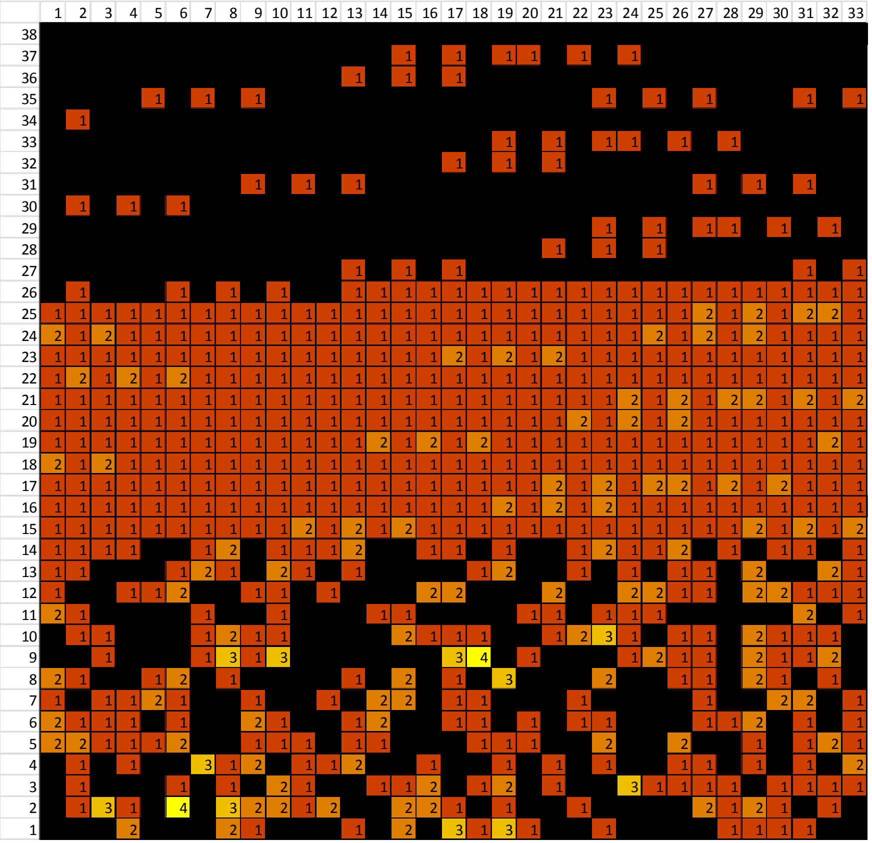
Aspartic Acid (D) Maps

$C_3H_6NO_2^+$ 88D (D Immonium Ion)



This common fragment for aspartic acid is not useful at all in this analysis!

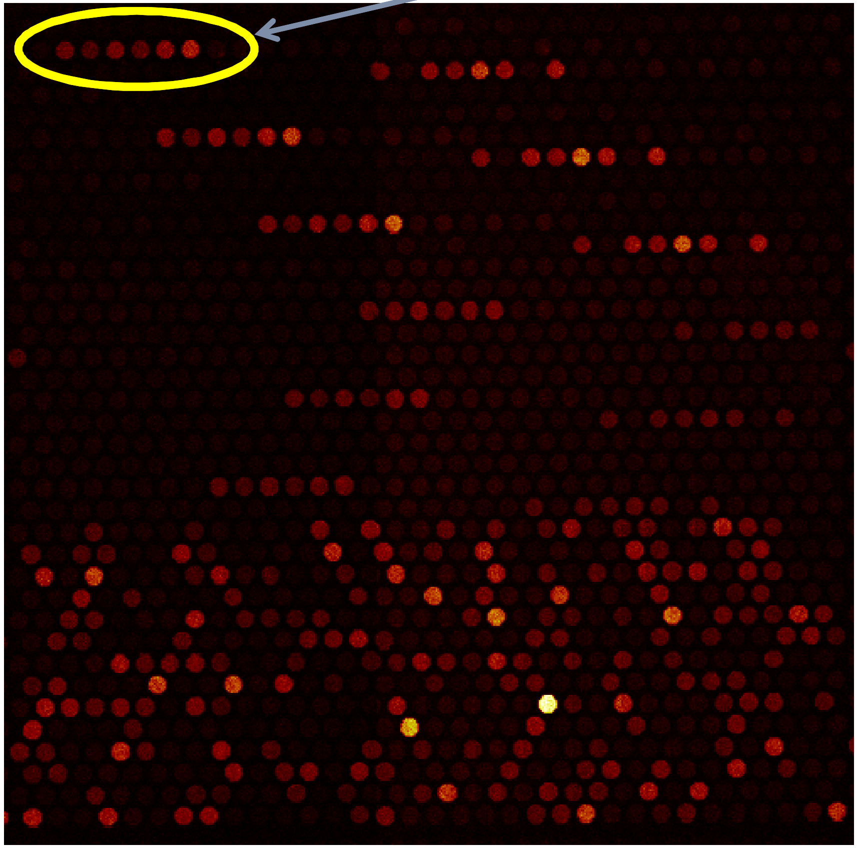
Aspartic Acid map from Sequence File
(Number of "D" in each spot sequence)

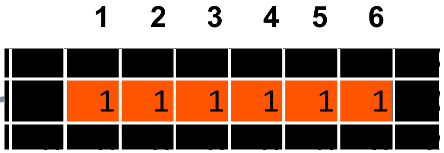


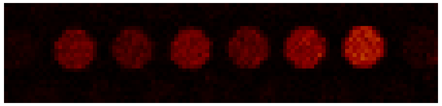
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
38																																	
37															1		1		1	1		1		1									
36													1		1		1																
35					1		1		1														1		1		1				1		1
34		1																															
33																			1		1		1	1		1		1					
32																	1		1		1												
31									1		1		1														1		1		1		
30		1		1		1																											
29																							1		1		1	1		1		1	
28																					1		1		1								
27													1		1		1														1		1
26		1				1		1		1			1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1	2	2	1
24	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1	2	1	1	1	1
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1
22	1	2	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1	2	2	1	2	1	2
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1	2	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1
18	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1	2	2	1	2	1	2	1	1	1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1	2	1	1	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1	1	1	2	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1	2
14	1	1	1	1			1	2		1	1	1	2			1	1		1			1	2	1	1	2		1		1	1		1
13	1	1				1	2	1		2	1		1					1	2			1		1		1	1		2			2	1
12	1			1	1	2			1	1		1				2	2				2			2	2	1	1		2	2	1	1	1
11	2	1					1			1				1	1					1	1		1	1	1						2		1
10		1	1				1	2	1	1					2	1	1	1			1	2	3	1		1	1		2	1	1	1	
9			1				1	3	1	3							3	4		1				1	2	1	1		2	1	1	2	
8	2	1			1	2		1					1		2		1		3				2			1	1		2	1		1	
7	1		1	1	2	1			1			1		2	2		1	1				1					1			2	2		1
6	2	1	1	1		1			2	1			1	2			1	1		1		1	1				1	1	2		1		1
5	2	2	1	1	1	2			1	1	1		1	1				1	1	1			2			2			1		1	2	1
4		1		1			3	1	2		1	1	2			1			1		1		1			1	1		1		1		2
3		1				1		1		2	1			1	1	2		1	2		1			3	1	1	1	1		1	1	1	1
2		1	3	1		4		3	2	2	1	2			2	2	1		1			1					2	1	2	1		1	
1				2				2	1				1		2		3	1	3	1			1					1	1	1	1		

Ion signature intensity correlates to amino acid position

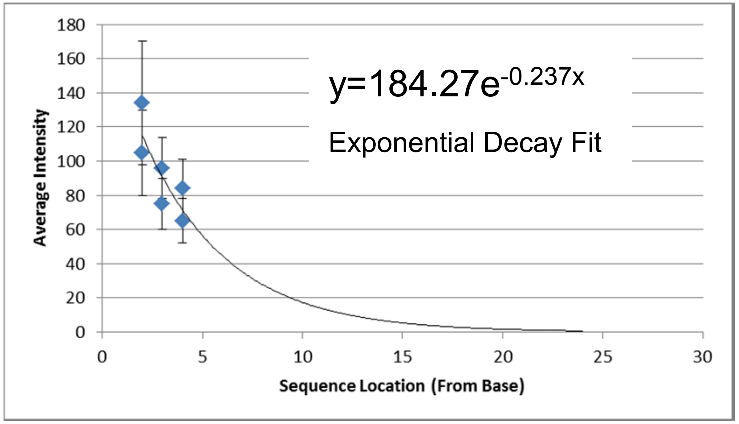
TOF-SIMS scan for $C_8H_{10}N^+$ 120D
(F Immonium Ion)







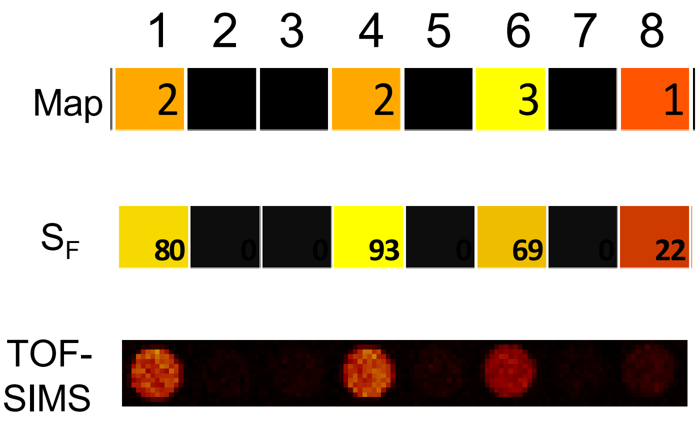
Spot	peptide	TOF-SIMS
1	LNWFGSGG	84.3 ± 17
2	LPHFGSGG	65.0 ± 13
3	LWFNGSGG	96.7 ± 18
4	LHFKGSGG	74.6 ± 15
5	LFNQGSGG	105 ± 25
6	LFKEGSGG	134 ± 36



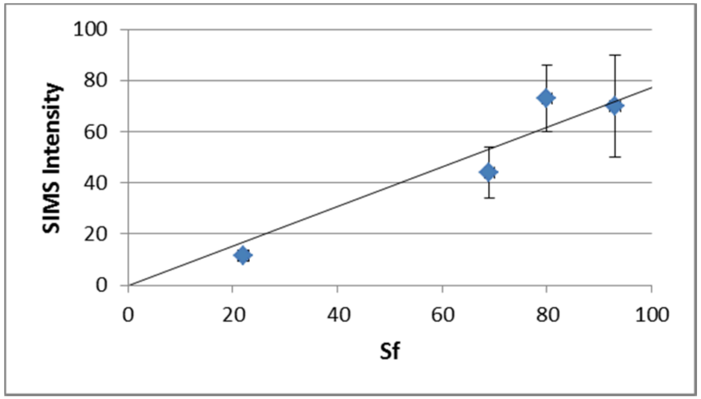
Ion intensity varies with position, copy number, and environment

- Develop an *expectation score* for the ion intensity of an amino acid as a function of position within a peptide, peptide length, and copy number

$$S_{AA} = \sum_{AA_i}^{i} fx(p_{AA_i})$$



Spot	peptide	S_F	TOF-SIMS
1	LDYAWFFYLRHRGGSGG	80	73.4 ± 13
2	LLVYVGNPRREQEGSGG	0	4.9 ± 1.2
3	LSARWRPRRYEYNGSGG	0	7.5 ± 1.5
4	LWWFKSNGFYDGSGG	93	69.8 ± 20
5	LVGGHSLHLWRGSGG	0	7.3 ± 1.5
6	LRHPNPFFKEHPNFYGSGG	69	43.8 ± 10
7	LWDYSVWPGAKGWGGSGG	0	4.0 ± 1.6
8	LHAHGNPHFNLHPQVWGSGG	22	11.4 ± 2.1



Amino Acid Fragment Summary

Positive Ion Amino Acid Peaks

					Immonium Ion				
Amino Acid	**Abbreviation**	**Symbol**	**Residue Mass**	**Free Mass**	**Formula**	**Mass**	**In Matrix?**	**In ToF-SIMS?**	**Comments**
Alanine	Ala	A	71.0786	89.0938	C_2H_6N	44.051	Yes	Yes	Overlaps with Background
Arginine	Arg	R	156.187	174.2022	none		Yes	No	$C_4H_{10}N_3$ (100), low SN
Asparagine	Asn	N	114.1036	132.1188	$C_3H_7N_2O$	87.0598	Yes	Yes	low SN, some background
Aspartic acid	Asp	D	115.0884	133.1036	$C_3H_6NO_2$	88.0436	Yes	Yes	Very poor SN
Cysteine	Cys	C	103.1386	121.1538	none		No	N/A	--Not in matrix
Glutamine	Gln	Q	128.1304	146.1456	none		Yes	No	No peak found
Glutamic acid	Glu	E	129.1152	147.1304	none		Yes	No	No peak found
Glycine	Gly	G	57.0518	75.067	CH_4N	30.034	Yes	Yes	Good SN, but high background
Histidine	His	H	137.1408	155.156	$C_5H_8N_3$	110.0767	Yes	Yes	Good SN, low background
Isoleucine	Ile	I	113.159	131.1742	$C_5H_{12}N$	86.101	No	N/A	--Not in matrix
Leucine	Leu	L	113.159	131.1742	$C_5H_{12}N$	86.101	Yes	Yes	Good SN, overlaps with I (Isoleucine)
Lysine	Lys	K	128.1736	146.1888	$C_5H_{10}N$	84.082	Yes	Yes	OK SN, background
Methionine	Met	M	131.1922	149.2074	none		No	N/A	--Not in matrix
Phenylalanine	Phe	F	147.1762	165.1914	$C_8H_{10}N$	120.08	Yes	Yes	Good SN, low background
Proline	Pro	P	97.1164	115.1316	C_4H_8N	70.0653	Yes	Yes	Good SN, but overlaps with N (Asparagine)
Serine	Ser	S	87.078	105.0932	C_2H_6NO	60.0447	Yes	Yes	OK SN, background
Threonine	Thr	T	101.1048	119.12	C_3H_8NO	74.0606	No	N/A	--Not in matrix
Tryptophan	Trp	W	186.2128	204.228	$C_{10}H_{11}N_2$	159.09	Yes	Yes	C_9H_8N (130), good SN, low background
Tyrosine	Tyr	Y	163.1756	181.1908	$C_8H_{10}NO$	136.0757	Yes	Yes	Good SN, but overlaps with H (Hystidine)
Valine	Val	V	99.1322	117.1474	$C_4H_{10}N$	72.0835	Yes	Yes	Good SN, low background
					Overlaps with others				

This list does not include many other suggested amino acid fragments from the literature.

Summary

- Peptide array was successfully imaged using ToF-SIMS
  - Structure of the array was verified (to some extent)
  - Many conventionally used amino acid peaks are compromised by overlapping peaks or low intensity
- Position in the chain has a very large effect on secondary ion yield
  - Effect on position was found to have exponential decay relationship
  - Variations in intensity (for Phenylalanine) were predicted based on this relationship.
  - Because of this decay, it was not possible to attempt sequencing of the peptide in ToF-SIMS (end of chain)
- A modified ToF-SIMS peak list for amino acids was presented

What can ToF-SIMS tell us?

- Peptide sequencing? ***No (Some position information)***
- Amino acid maps? ***Yes, for most amino acids***
- Peptide density? ***No***
- Other? ***Image-based information (spot size and distributiuon)***

- Can ToF-SIMS be used to verify microarray? ***Somewhat***