Special Topic
Analysis of ARTEMIS-derived MHC-I Peptides by Mass Spectrometry

Phil Gafken

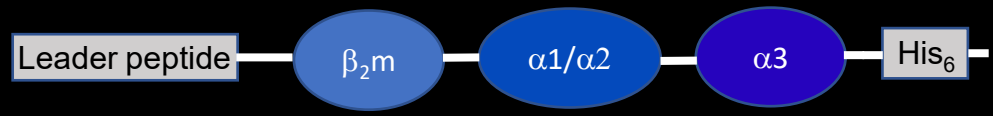
Fred Hutch Proteomics & Metabolomics Resource

July 24, 2023

MHC-I Peptides Isolated from Single Chain Dimers

(*ARTEMIS reagents*)

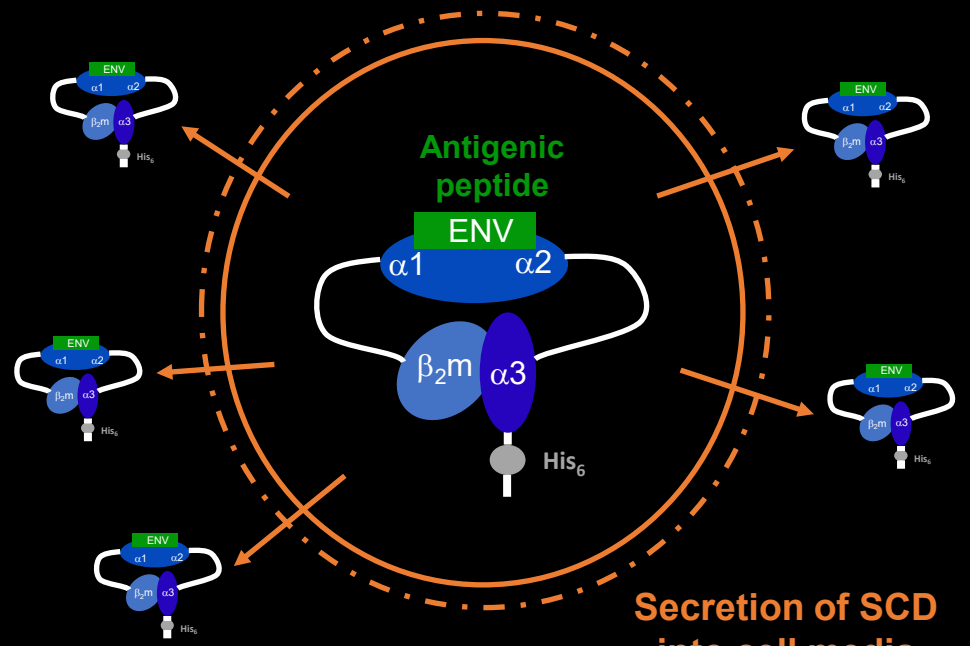
Virus #1 Single chain dimer (SCD)



Virus #2 HIV SF162 ENV [peptide source]

Transduce into HEK293 cells +TNFα & INFγ

- Transfect/transduce cells with HLA allele of choice
- ~ 2-4 x 10^8 cell required
- SCD collected from media using Ni++ column



Developed by Roland Strong and Kathryn Finton
Front Immunol 12:658372

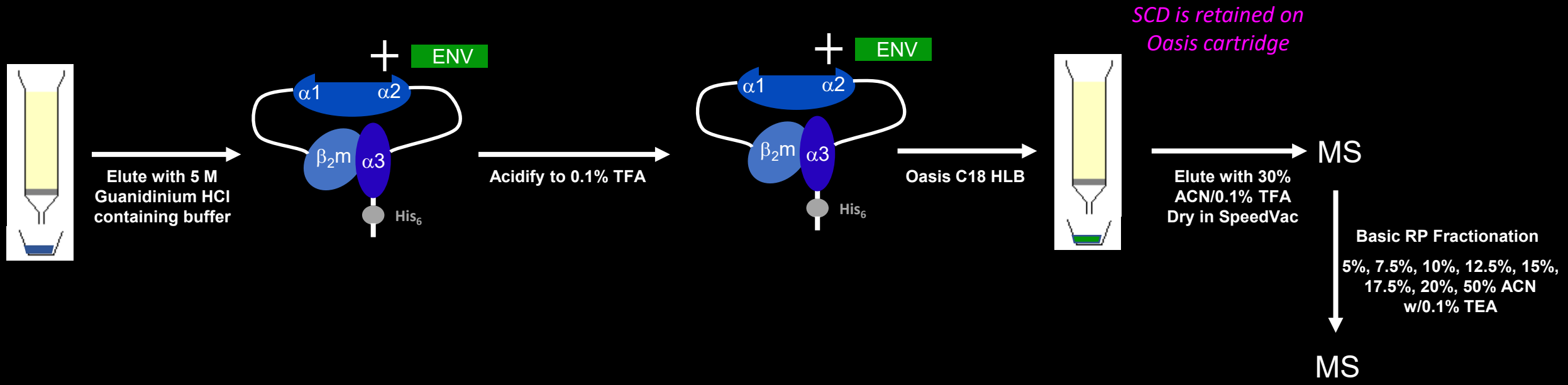
MHC-I Peptides Isolated from Single Chain Dimers

(*Biochemical isolation of MHC-peptides*)

Isolate SCDs with Ni++ affinity chromatography

Dissociate antigenic peptide from SCD

Isolate antigenic peptides by C18 reverse phase chromatography

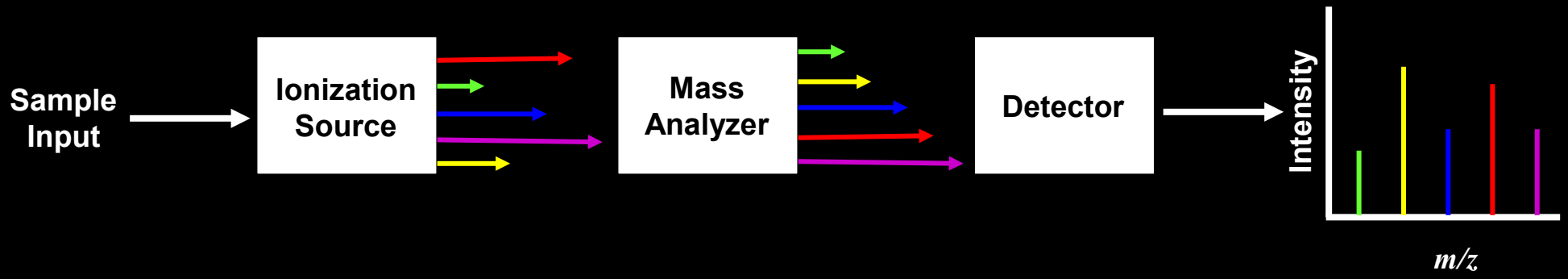


Developed by Roland Strong and Kathryn Finton
Front Immunol 12:658372

Components of a Mass Spectrometer

1. *Source*- produces gas-phase ions from the sample
2. *Mass analyzer*- resolves ions based on their m/z ratio
3. *Detector*- detects ions resolved by the mass analyzer

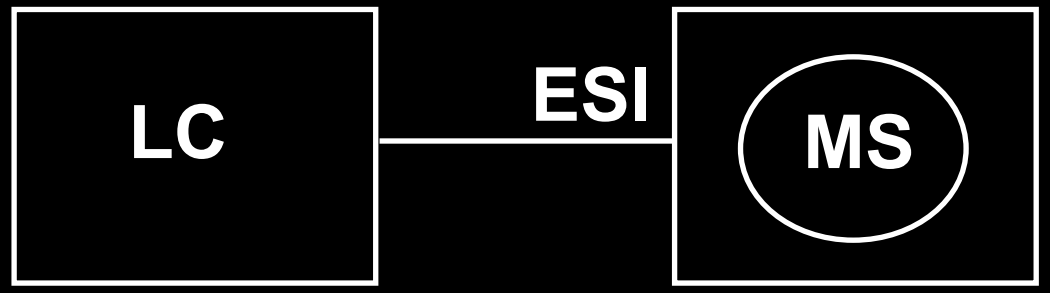
All mass spectrometers contain these three components



* Size of the arrows indicate the m/z of the ions, and not the flight or order of detection

LC/ESI MS

Liquid Chromatography/Electrospray Ionization Mass Spectrometry



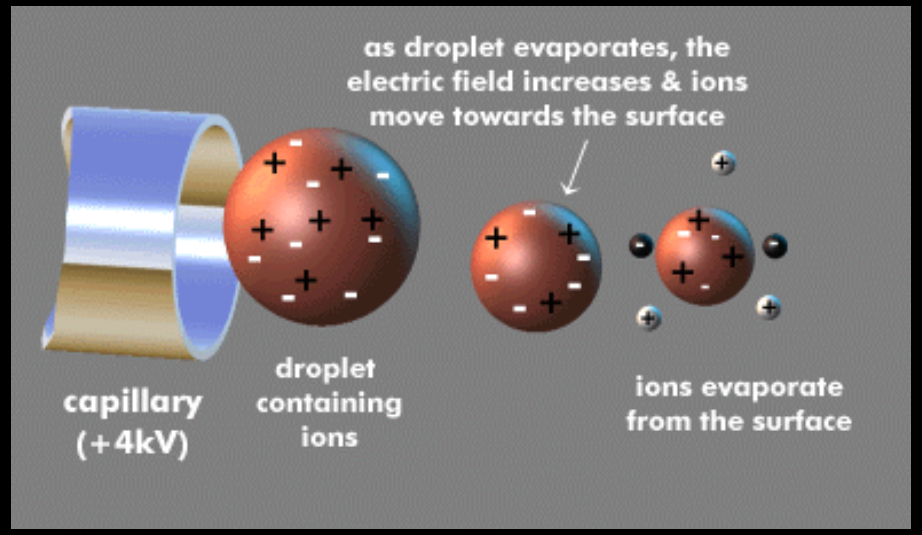
Chromatographic separations are in-line with the mass spectrometer

ESI converts ions in the liquid phase to ions in the gas phase

MS acts as a detector (identical to a UV detector in typical HPLC)

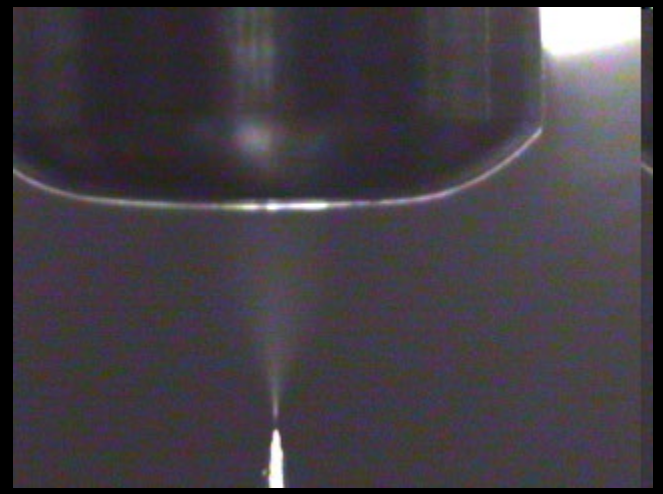
ESI

Electrospray Ionization

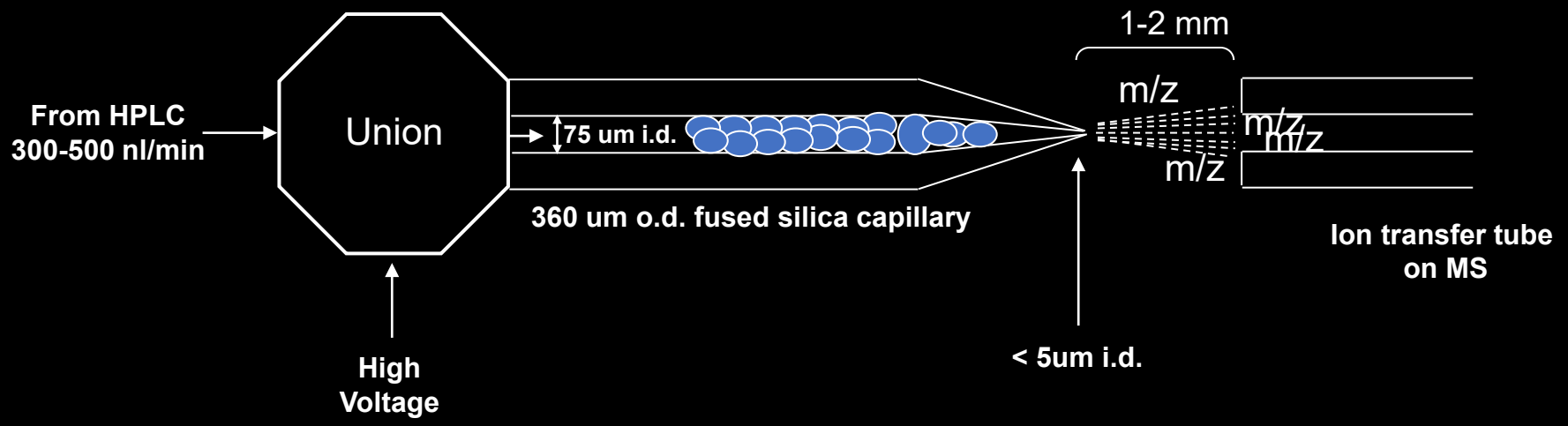


John Fenn awarded 2002 Nobel Prize in Chemistry for contribution to developing ESI

ESI


Ion transfer tube of MS
Electrospray capillary
Taylor cone

Common LC-MS Configuration

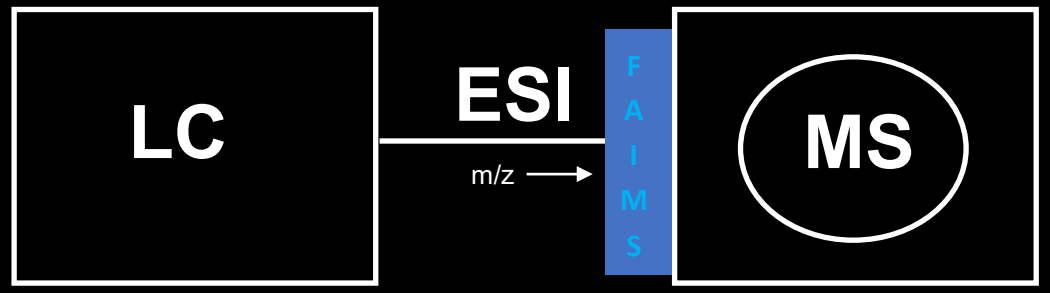


- Column lengths are usually 25-30 cm in length
- The smaller the column i.d., the higher the chromatographic resolution
- Currently we can identify approx. 6000 proteins in a single injection from 200 ng HeLa lysate
- Current configurations have FAIMS device between capillary and ion transfer tube

LC/ESI-FAIMS MS

Liquid Chromatography/Electrospray Ionization Mass Spectrometry

Field Asymmetric Ion Mobility Spectrometry



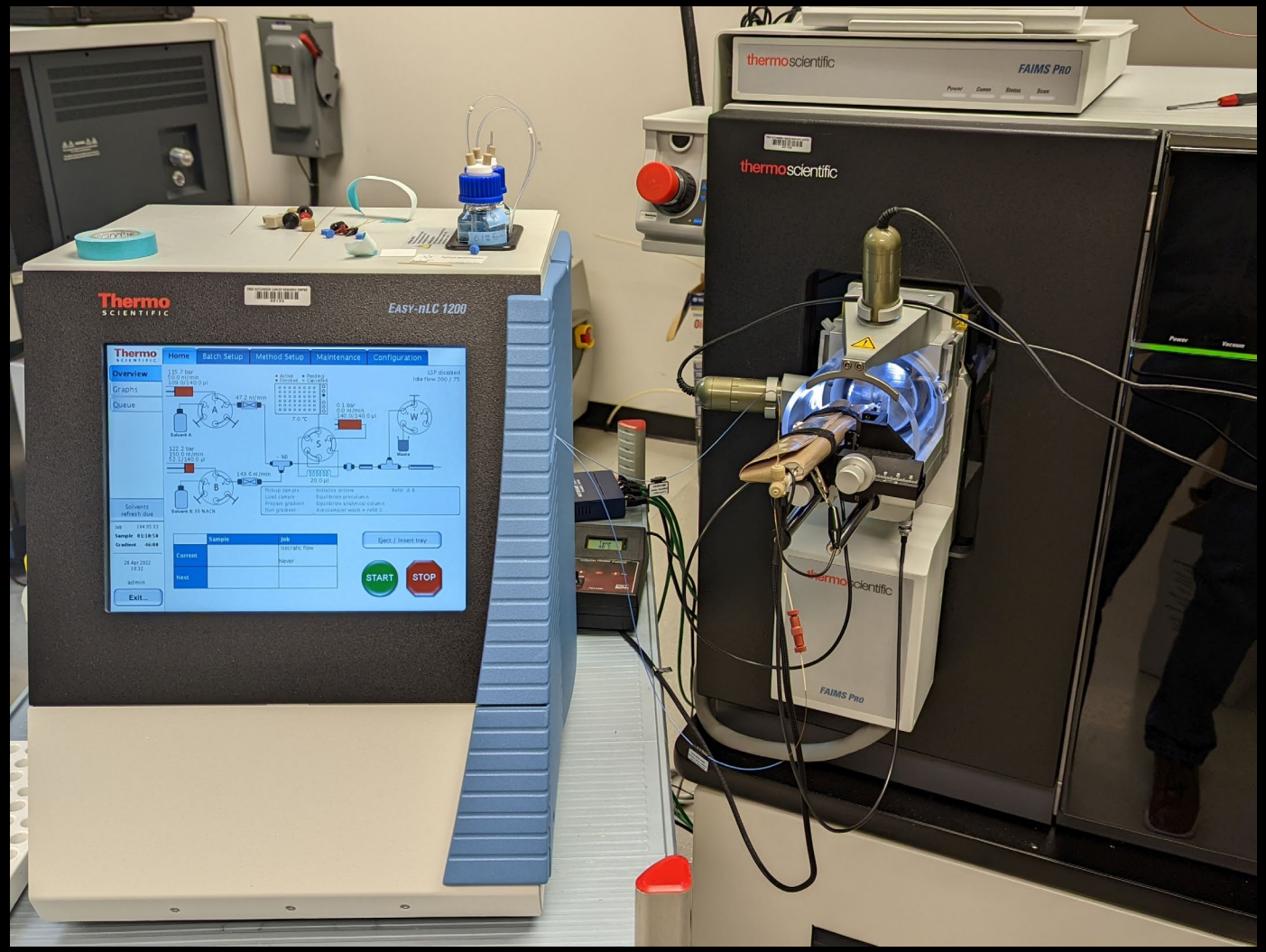
Chromatographic separations are in-line with the mass spectrometer

ESI converts ions in the liquid phase to ions in the gas phase

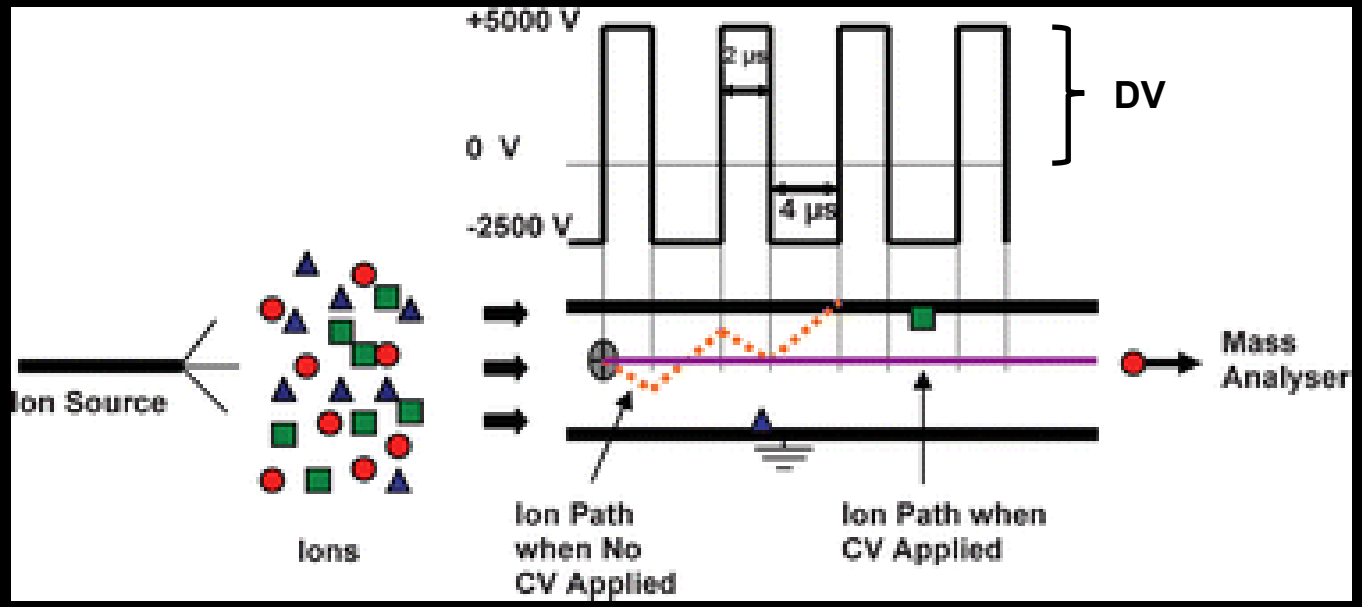
FAIMS provides gas-phase separation of ions before MS

MS acts as a detector (identical to a UV detector in typical HPLC)

FAIMS at Fred Hutch



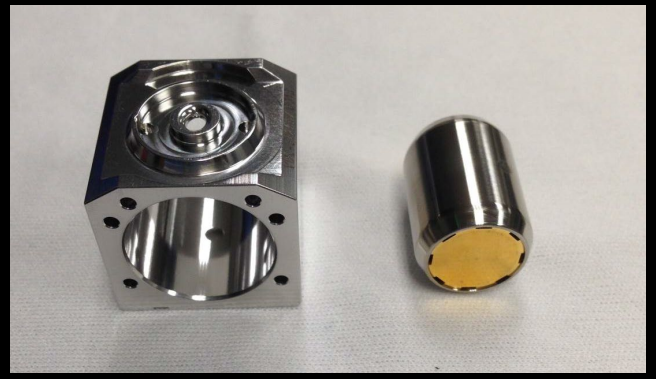
General Principle Behind FAIMS



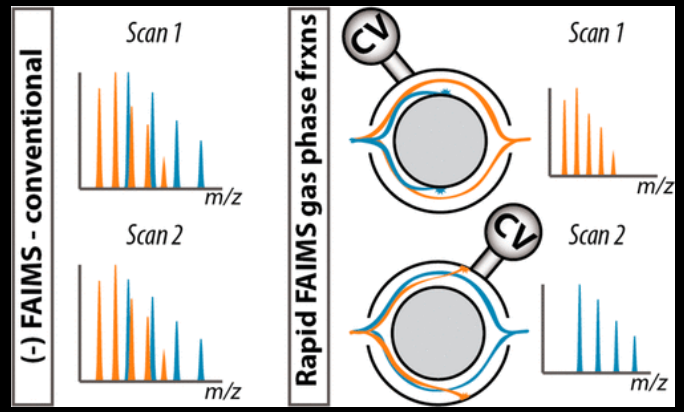
- Example depicts FAIMS with parallel plates
- DV = Dispersion Voltage
- CV= Compensation Voltage (typically between -10 V and -100 V)

Analyst 132, 842-864 (2007).

General Principle Behind FAIMS



- ThemoScientific's implementation of FAIMS uses a cylindrical electrode
- Nitrogen is used as a carrier gas
- Compensation voltages applied to center electrode correct ions' trajectories allowing them to enter the mass spectrometer
- Preset compensation voltages allow groups of ions to traverse the cylindrical path, typically based on charge state; typically use 2 or 3 CVs in a method



Anal. Chem. 90, 15, 9529-9537 (2018)

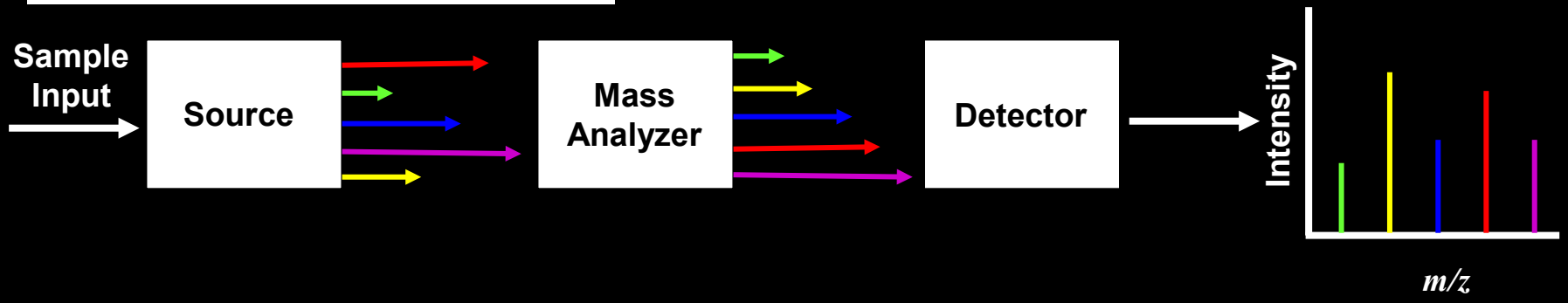
- Erichment of yellow ions with CV1
- Enrichment of blue ions with CV2

Why Use FAIMS?

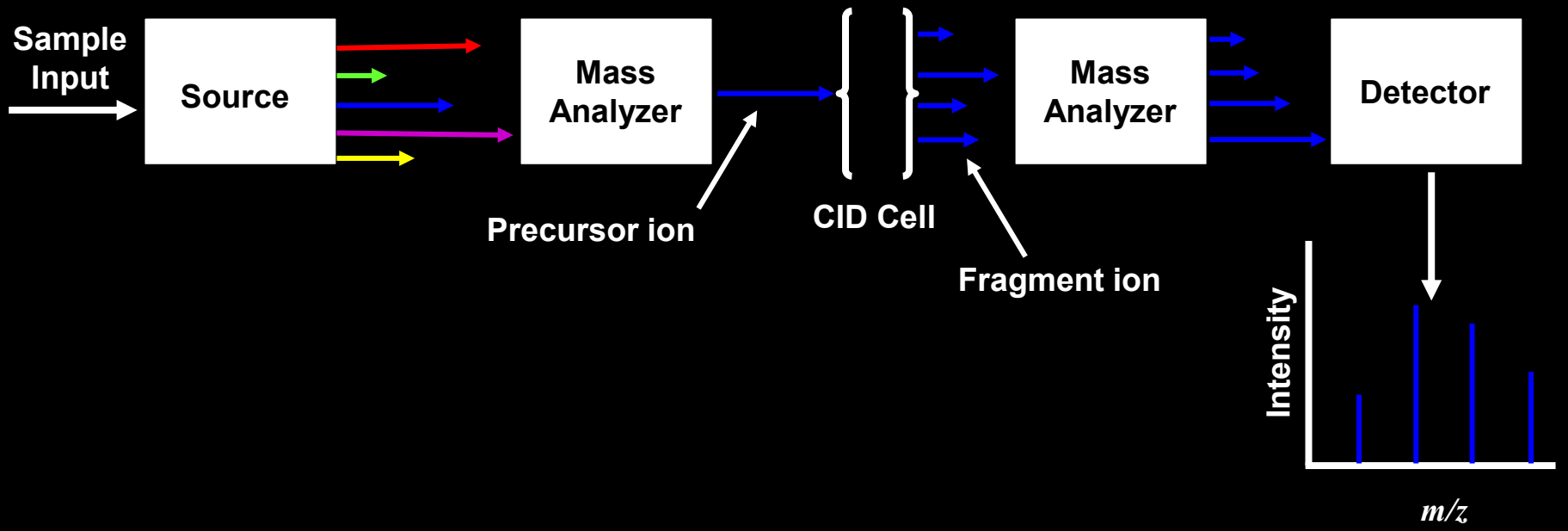
- FAIMS provides gas phase fractionation to simplify ion packets entering the mass spectrometer for analysis; the reduced complexity allows the mass spectrometer to “see” more
- Neutrals are blocked by the FAIMS electrode, preventing them from entering the instrument and increasing the S/N of detected ions
- Blocking neutrals keeps the instrument cleaner for longer periods of time
- Simplified ion packets and increased S/N can increase protein identifications by 5%-15%; that could increase a result of 8,000 proteins when not using FAIMS to 8400 to 9200 proteins when using FAIMS!

Multistage Mass Analysis

Single-stage analysis (MS)

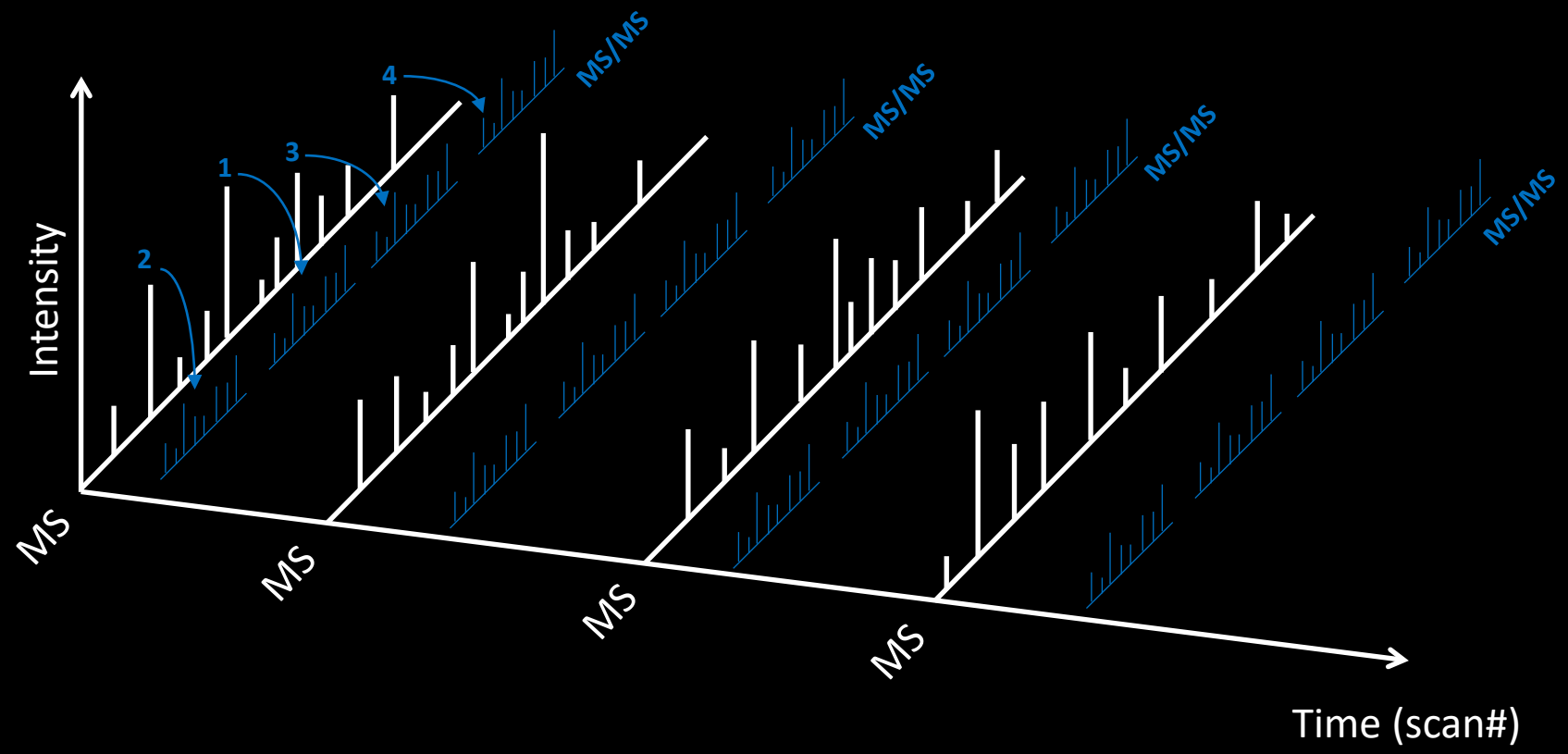


Dual-stage or tandem analysis (MS/MS)



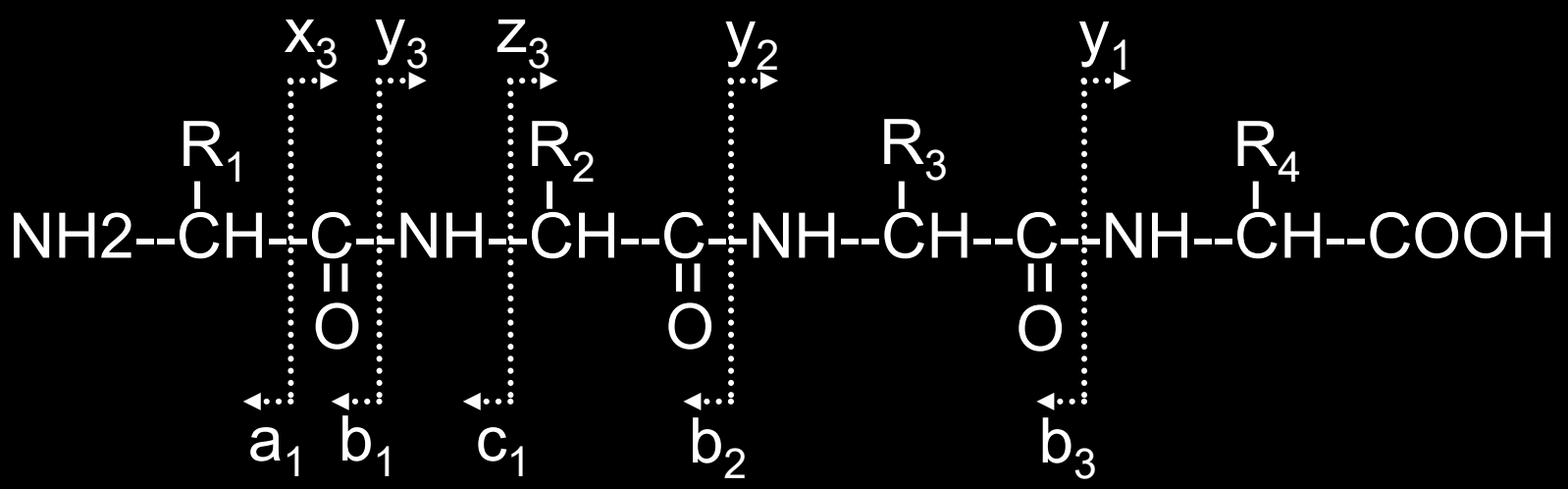
Representation of LC-MS/MS Data

(MS with MS/MS data)



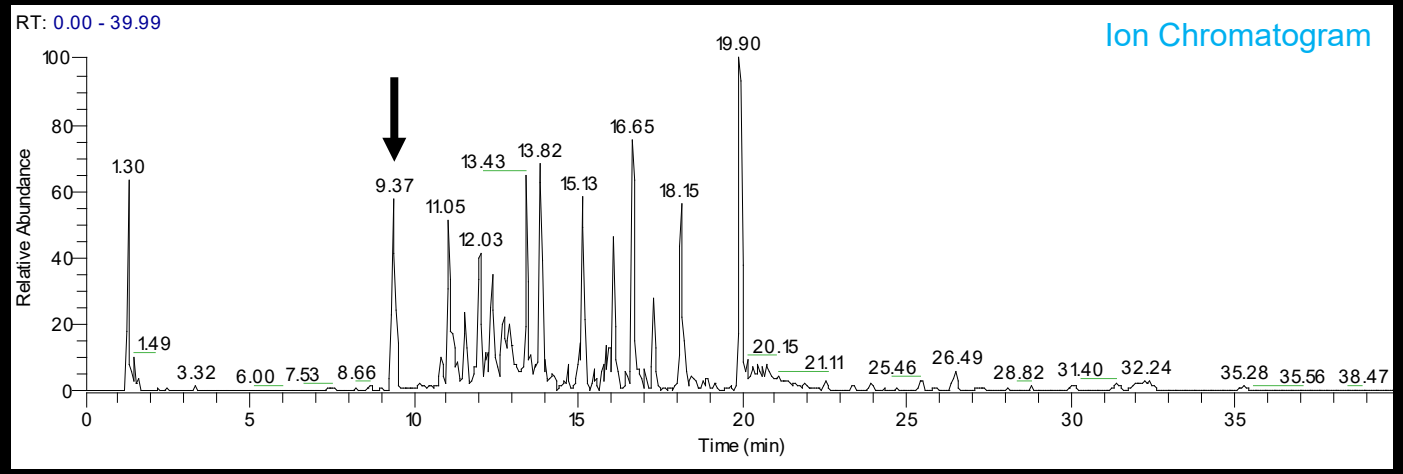
- Current Orbitrap instruments can collect ~ 45 MS/MS spectra per second
- Current data collection is typically based on top 1 second data collection

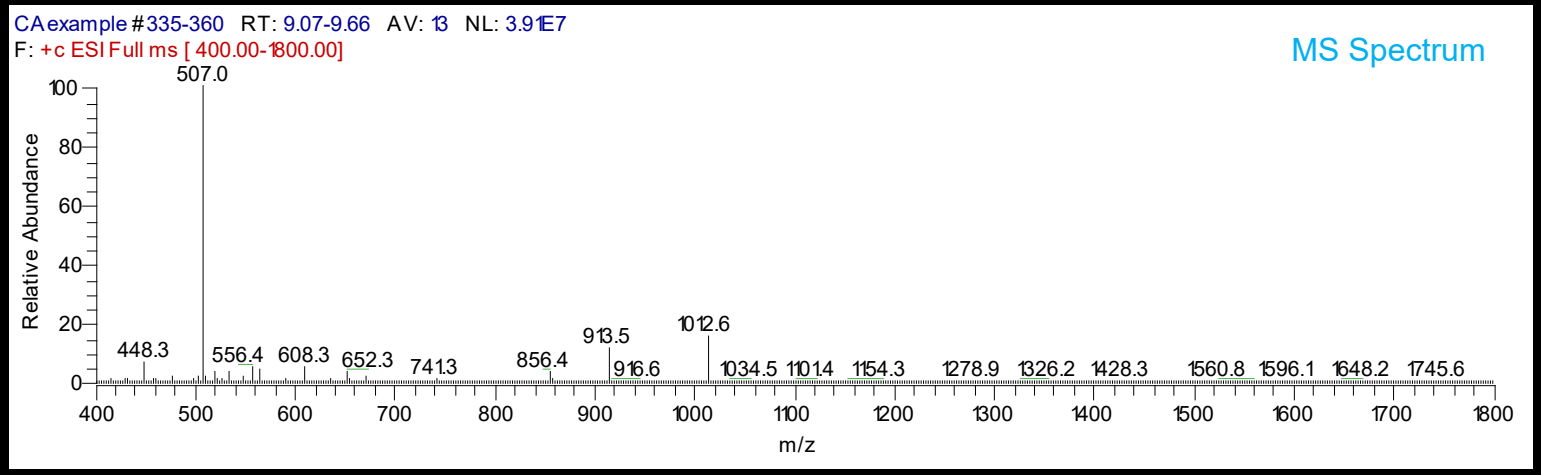
Peptide Fragmentation Nomenclature



- **a, b, c, indicates the type of N-terminal fragment**
- **x, y, z, indicates the type of C-terminal fragment**
- **1, 2, 3, indicates the location of the fragment**
- **The type of fragment produced depends on the type of fragmentation being used**

Precursor Ion (MS1) Mass Spectrum

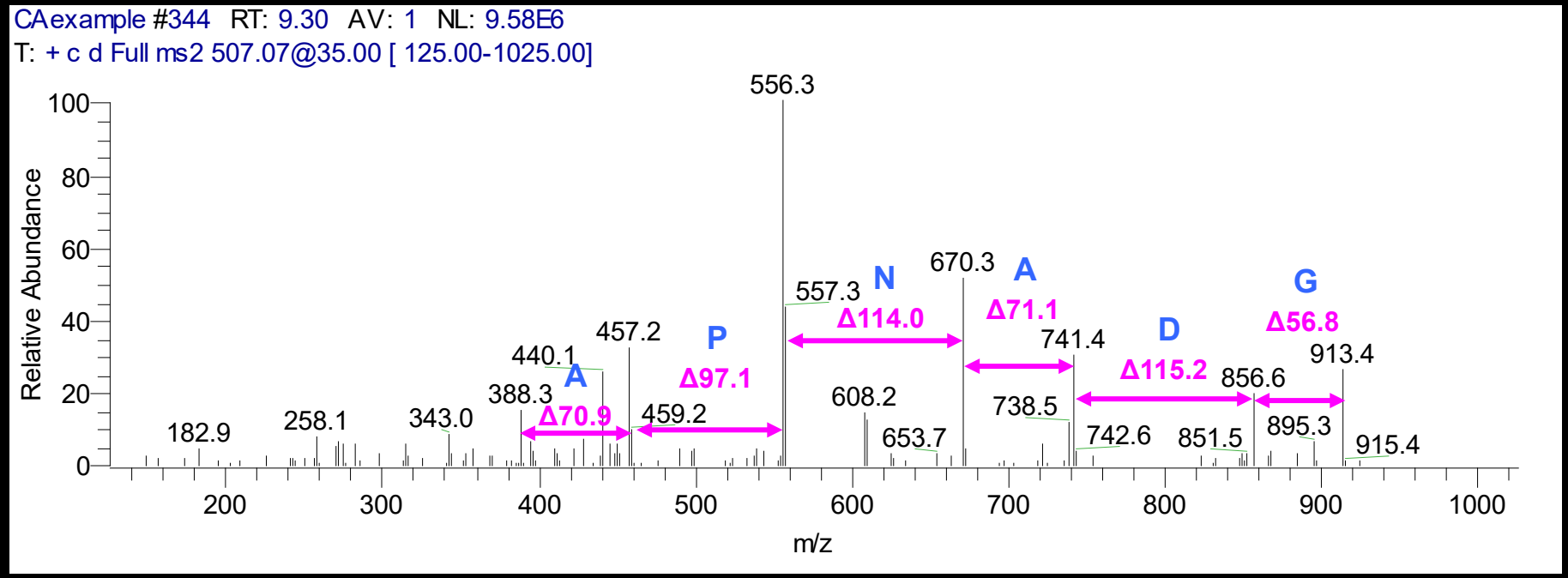




At each data point, the ion's intensity is also recorded.

Fragmentation (MS/MS or MS2) Spectrum of *m/z* 507

(de novo peptide sequencing with tandem MS data)



…G-D-A-N-P-A…

or

…A-P-N-A-D-G…

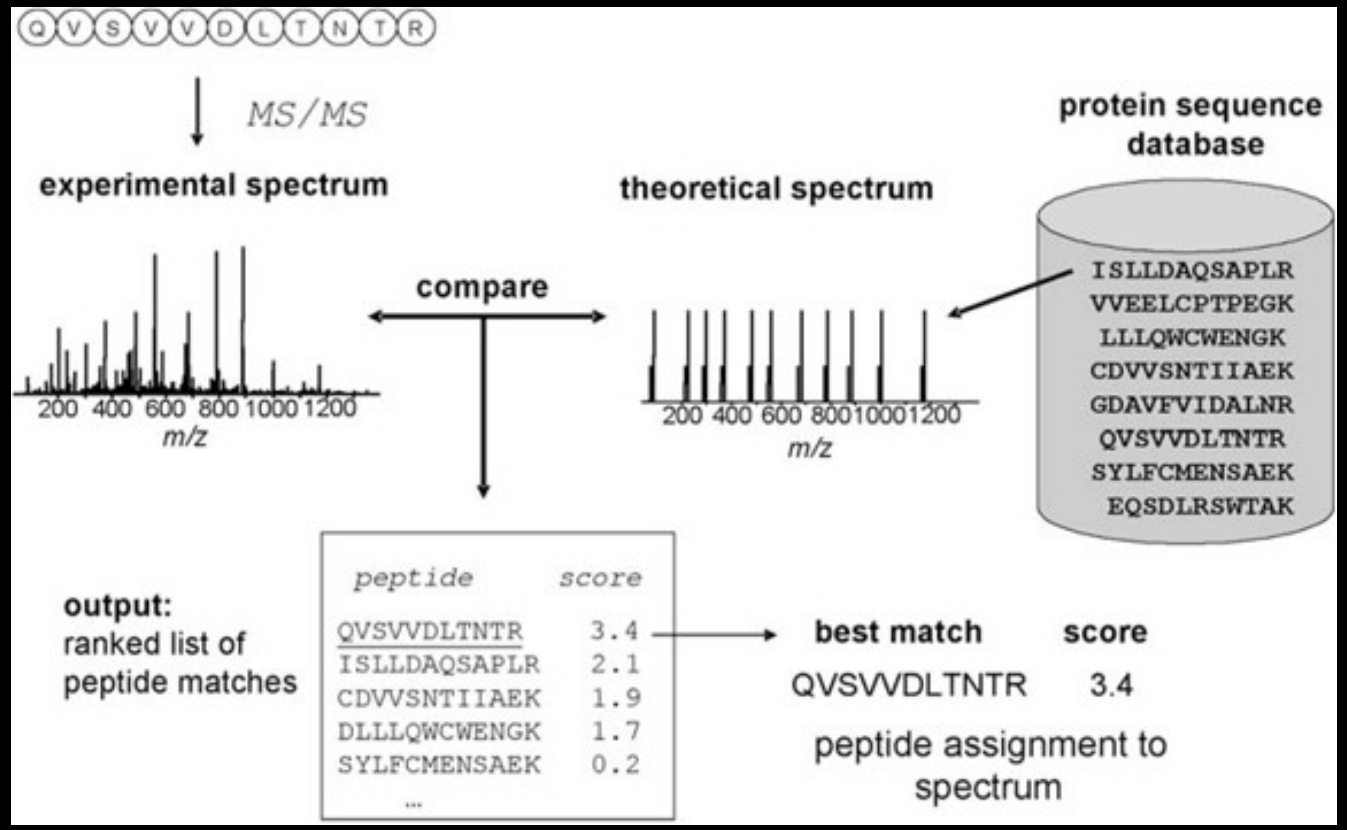
Mass differences must be associated with an ion series; difficult to tell which fragment ions are associated with which ion series.

Trypsin Digestion Map of Carbonic Anhydrase

Num.	From-To	MH+(expt.)	Sequence
1	1-8	971.449	SHHWGYGK
2	9-15	898.444	HGPHWHK
3	16-24	1018.496	DFPIANGER
4	25-33	1001.527	QSPVNIKTK
5	34-54	2198.219	AVVQDPALKPLALVYGEATSR
6	55-55	175.120	R
7	56-73	2098.878	MVNNGHSFNVEYDDSQDK
8	74-77	430.303	AVLK
9	78-86	979.485	DGPLTGTYR
10	87-106	2354.107	LVQFHFHWGSSQGSEHTVDR
11	107-107	147.113	K
12-13	108-121	1709.910	KYAAELHLVHWNTK
13	109-121	1581.818	YAAELHLVHWNTK
14	122-143	2253.156	YGDFGTAAQQPDGLAVVGVFLK
15	144-153	1012.543	VGDANPALQK
16	154-162	973.557	VLDALDSIK
17	163-164	248.161	TK
18	165-166	204.135	GK
19	167-207	4593.349	STDFPNFDPGSLLPNVLDYWTYPGS LTTPPLLESVTWIVLK
20	208-219	1346.699	EPISVSSQQMLK
21	220-221	322.188	FR
22	222-246	2852.477	TLNFNAEGEPELLMLANWRPAQPLK
23	247-248	289.164	NR
24	249-251	402.247	QVR
25	252-255	448.256	GFPK

Automated Database Search Algorithm

For tryptic peptides



A confident match is called a "peptide to spectrum match" or "PSM"

For MHC-I peptides

- We use Proteome Discoverer v2.5
- Workflow is identical for MHC peptides
- Major difference is protein database is not processed to tryptic peptides
- Protein database is processed to all peptides between 6 and 25 a.a. in length
- MHC peptide database is gigantic

A.I. Nesvizhskii (2007) Mass Spectrometry Data Analysis in Proteomics, pp 87-119

Maximum Destroyer Ultra Workstation

By Omics Computing

- Specifically designed for automated protein database searching
- 36 total processors
- 72 Threads
- 192 GB RAM
- 17 Tbyte hard drive space

Analysis of MHC Peptides

(*Results*)

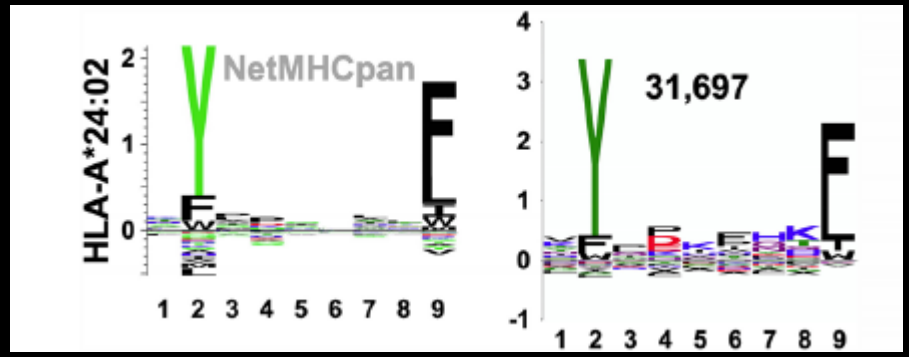
KF111716_112316_KF3_18uL_Uniprot_Hum_031914.msf

Proteins | Peptides | Search Input | Result Filters | Peptide Confidence | Search Summary

		Accession	Description	Score	Coverage	# Proteins	# Unique Peptides	# Peptides	# PSMs	MW [kDa]
1		P05534	HLA class I histocompatibility antigen, A-24 alpha chain O...	657.63	51.51 %	49	52	97	202	40.7
2		P10321	HLA class I histocompatibility antigen, Cw-7 alpha chain O...	337.86	20.22 %	15	1	46	107	40.6
3		Q31612	HLA class I histocompatibility antigen, B-73 alpha chain O...	312.85	17.63 %	13	1	45	98	40.4
4		P07437	Tubulin beta chain OS=Homo sapiens GN=TUBB PE=1 SV...	270.09	17.34 %	29	1	10	110	49.6
5		Q6P2Q9	Pre-mRNA-processing-splicing factor 8 OS=Homo sapiens...	78.87	6.30 %	2	16	17	30	273.4
6		P46783	40S ribosomal protein S10 OS=Homo sapiens GN=RPS10...	77.29	10.91 %	4	2	2	30	18.9
7		H0Y842	HLA class I histocompatibility antigen, alpha chain F (Frag...	59.39	25.33 %	5	1	13	22	25.2
8		O75643	U5 small nuclear ribonucleoprotein 200 kDa helicase OS=...	58.43	3.56 %	2	9	9	21	244.4
9		HIV_SF162_ENV	HIV_SF162_ENV	58.14	9.55 %	1	10	10	22	75.2

	A2	Sequence	# PSMs	# Proteins	# Protein Groups	Protein Group Accessions	Modifications	ΔCn	XCorr	Charge	MH+ [Da]	ΔM [ppm]
1		AYDTEVHNVW	2	1	1	HIV_SF162_ENV		0.0000	3.13	2	1233.55364	0.10
2		VYYGVPVWKEATTTLF	3	1	1	HIV_SF162_ENV		0.0000	3.05	2	1873.96916	-2.33
3		VWKEATTTLF	2	1	1	HIV_SF162_ENV		0.0000	3.01	2	1195.63518	-0.52
4		VWKEATTTL	2	1	1	HIV_SF162_ENV		0.0000	2.96	2	1048.56633	-1.00
5		RYLKDQQLL	5	1	1	HIV_SF162_ENV		0.0000	2.57	2	1176.67058	-2.54
6		KMQKEYALF	1	1	1	HIV_SF162_ENV		0.0000	2.51	2	1157.60137	-0.89
7		NYTNLIYTLI	2	1	1	HIV_SF162_ENV		0.0000	2.36	2	1227.66155	-0.35
8		MYAPPIRGQI	1	1	1	HIV_SF162_ENV		0.0000	2.13	2	1145.61052	-2.71
9		YLKDQQLL	2	1	1	HIV_SF162_ENV		0.0000	2.11	2	1020.57182	-0.61
10		KmQKEYALF	1	1	1	HIV_SF162_ENV	M2(Oxidation)	0.1181	2.09	2	1173.59197	-4.55
11		NYTNLIYTL	1	1	1	HIV_SF162_ENV		0.0000	2.09	2	1114.57720	-0.64

		Accession	Description	Score	Coverage	# Proteins	# Unique Peptides	# Peptides	# PSMs	MW [kDa]
10		P62140	Serine/threonine-protein phosphatase PP1-beta catalytic s...	53.38	9.79 %	4	2	6	19	37.2
11		P04350	Tubulin beta-4A chain OS=Homo sapiens GN=TUBB4A PE...	52.79	17.34 %	27	1	10	22	49.6
12		Q68D85	Natural cytotoxicity triggering receptor 3 ligand 1 OS=Ho...	51.98	13.22 %	1	9	9	17	50.8
13		E9PMD7	Serine/threonine-protein phosphatase PP1-alpha catalytic...	49.86	8.70 %	9	1	5	17	28.9
14		E9PC52	Histone-binding protein RBBP7 OS=Homo sapiens GN=RB...	43.49	13.46 %	12	5	7	17	46.9



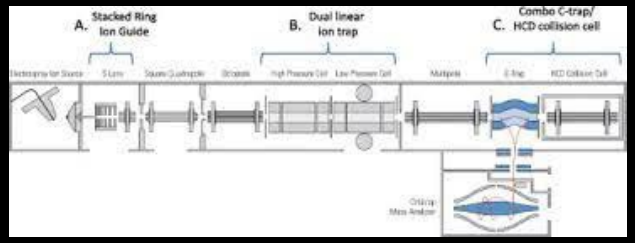
- 6500 peptides identified (+/- 5 ppm mass accuracy & 1% FDR)
- 5775 8-11 mer peptides
- 11 Env peptides detected

- High compliance with reported A24 peptide binding motif
  - Y/F at P2
  - F/I/W/L at C-term

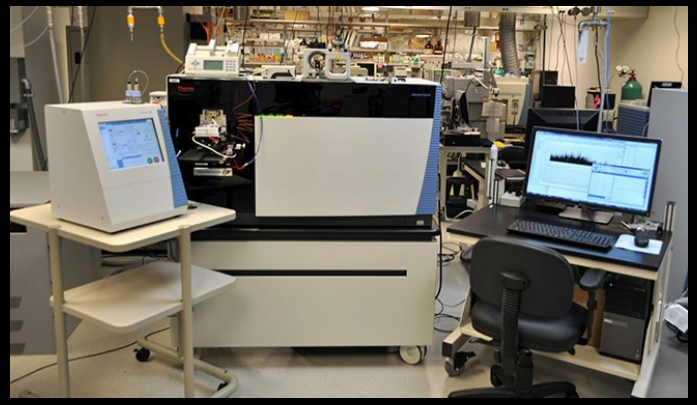
Developed by Roland Strong and Kathryn Finton: Front Immunol 12:658372

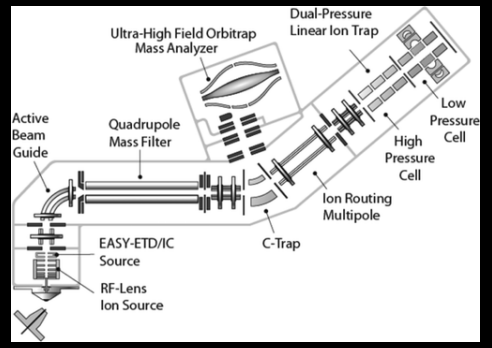
Evolution of Orbitrap Mass Spectrometers

Orbitrap Elite

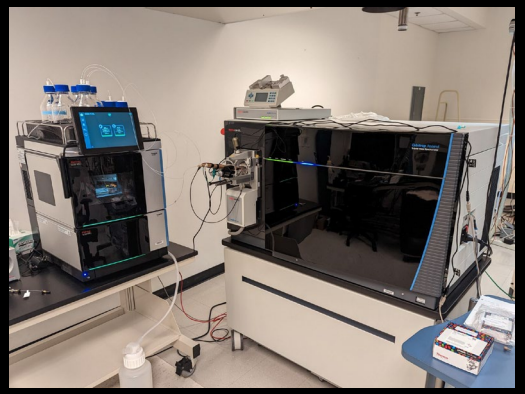


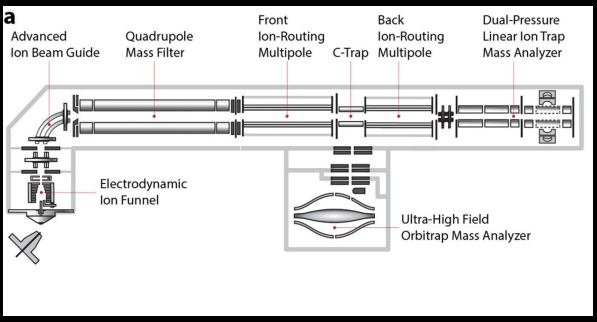
Orbitrap Fusion





Orbitrap Ascend





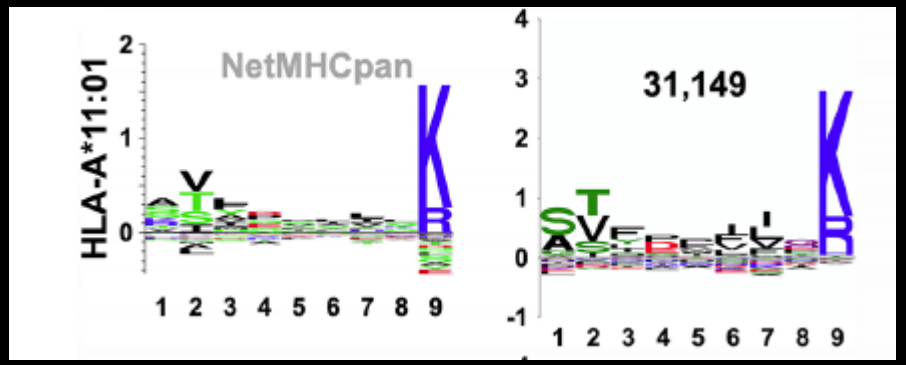
Circa:	2010	2013	2022
Configuration:	hybrid	tribrid	tribrid
MaxScan Speed:	8 Hz	20 Hz	45 Hz
Sensitivity (HeLa):	2400 proteins	4800 proteins	6200 proteins

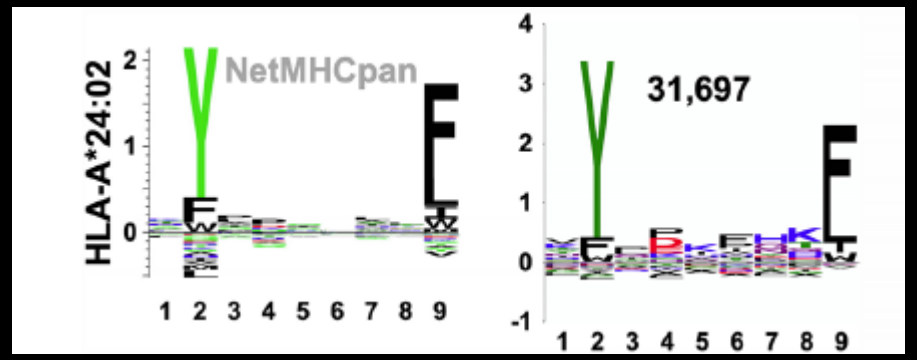
Latest Results with Orbitrap Ascend w/FAIMS

Sample	Fusion "Normal" Method	Ascend w/FAIMS -20/-40/-60	Ascend w/FAIMS -20/-40	Ascend W/FAIMS -20/-40/-60 (+1 only)	Ascend w/FAIMS - 20/-40/-60 Slower Gradient
A11_SCD_FAIMS	3502	7339	5208	39	8333
A24_SCD_FAIMS	2679	5708	4658	858	9062

results for 8 to 14 mers

?





Front Immunol 12:658372