Comparison of CID, ETD, and HCD for Top-down Characterization of Histones

Zhixin(Michael) Tian

Nikola Tolić, Rui Zhao, Shawna, Hengel,

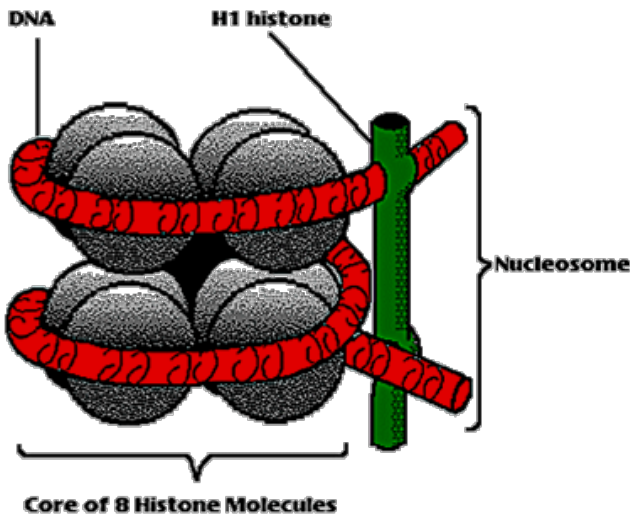
Si Wu, Ronald J. Moore, Errol W. Robinson,

Richard D. Smith, Ljiljana Paša-Tolić

Environmental Molecular Sciences Laboratory

at Pacific Northwest National Laboratory

Histones and post-translational modifications (PTMs)

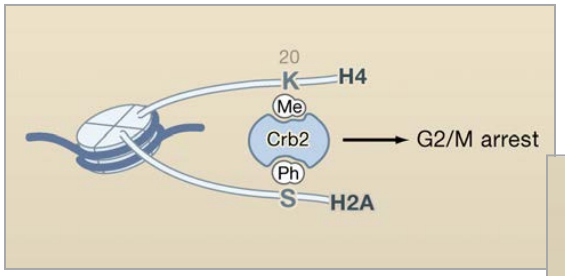


Nucleosome

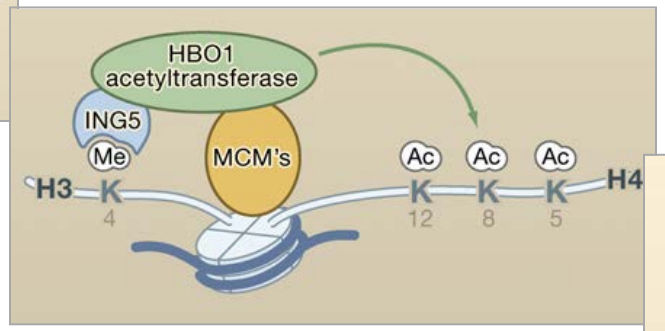
http://www.accessexcellence.com/AB/GG/nucleosome.gif

Core histones: H4, H2B, H2A, H3

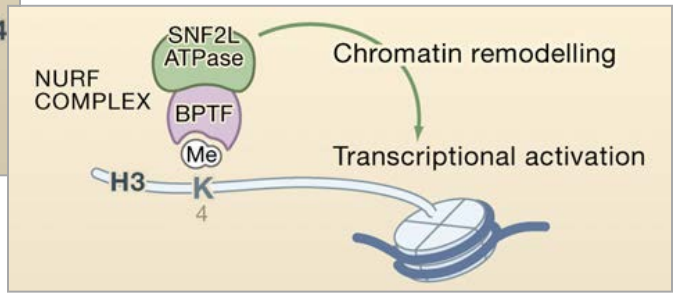
DNA Repair[1]



DNA Transcription[3]

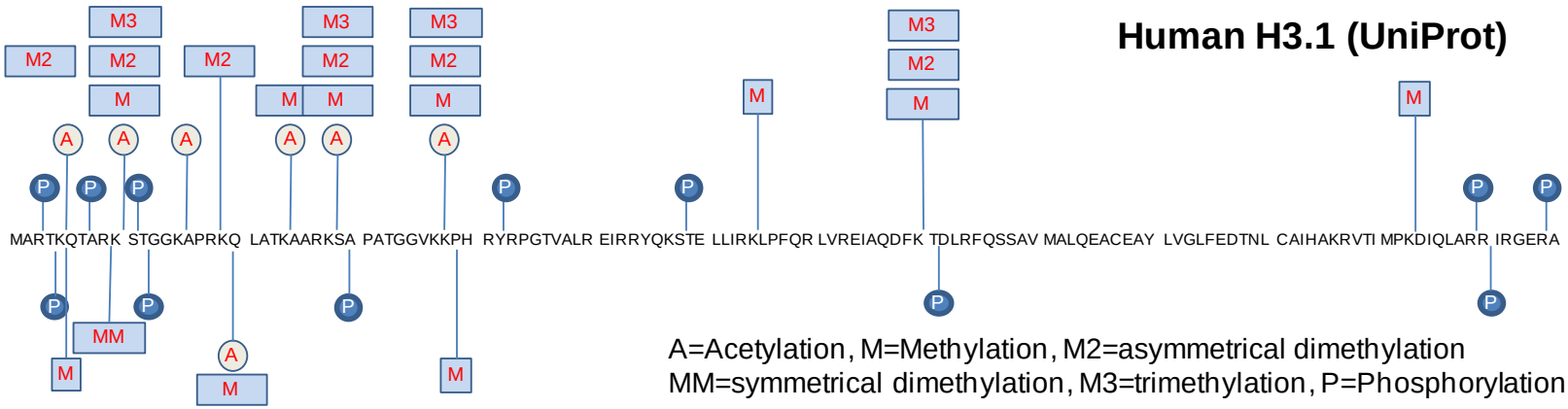


DNA Replication[2]

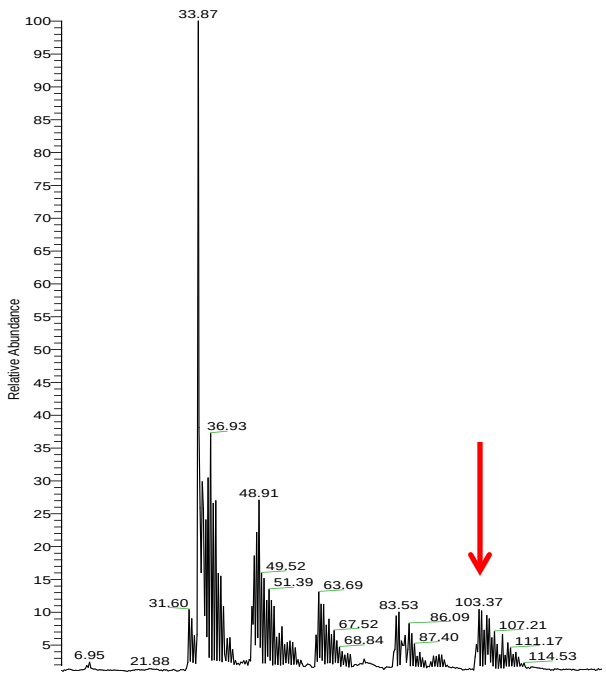


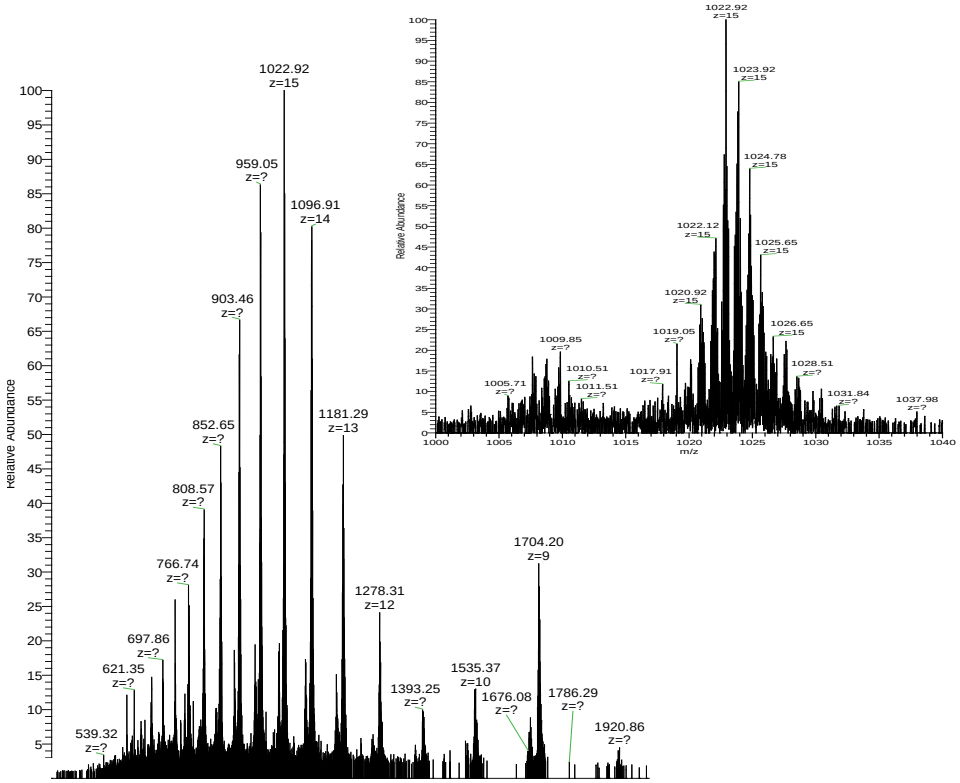
1. Nature Reviews Mol. Cell Bio. 6, 757 (2005); 2. Molecules and Cells 28, 149 (2009); 3. Science 272, 371 (1996)

Histones are heavily modified and very complex



RPLC-MS/MS of HeLa core histones

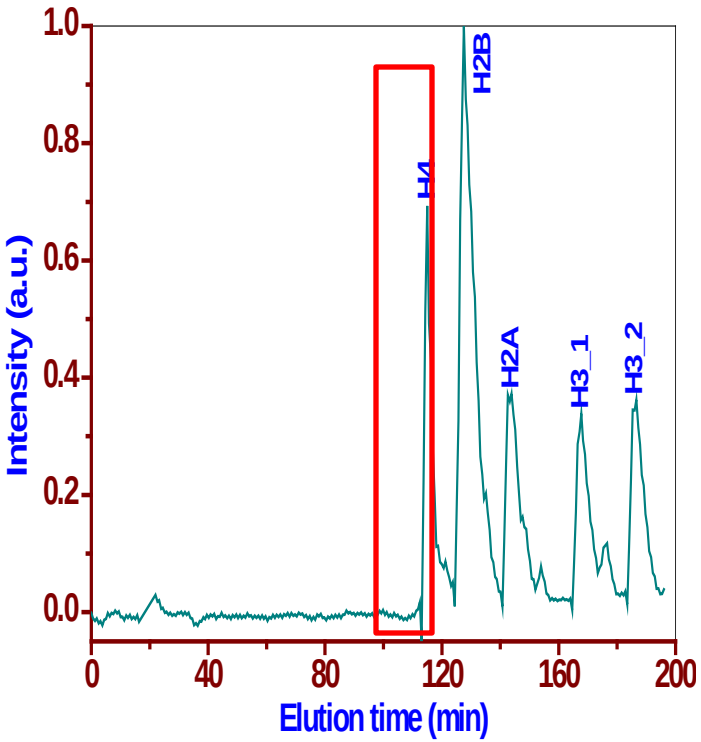




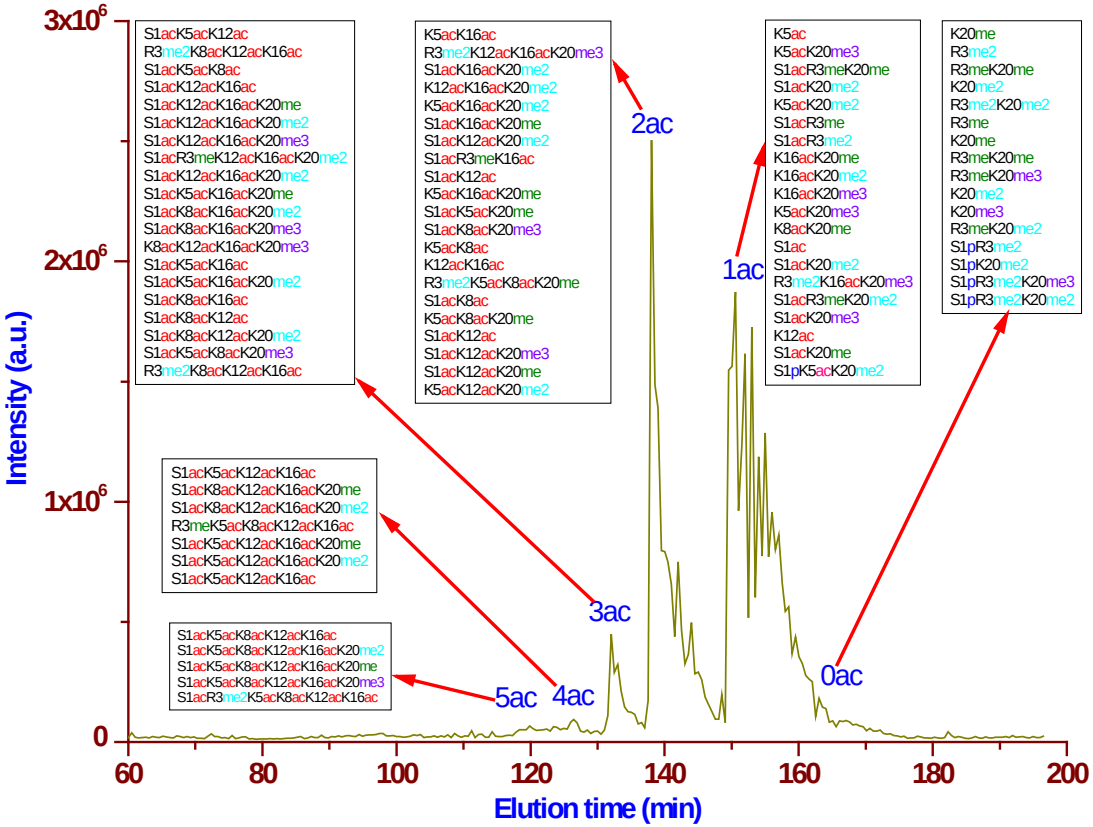
Online RPLC/WCX-MS/MS HeLa core histones

Methods	IDs
RPLC-MS/MS	127
WCX-MS/MS	135
RPLC-WCX-MS/MS	708

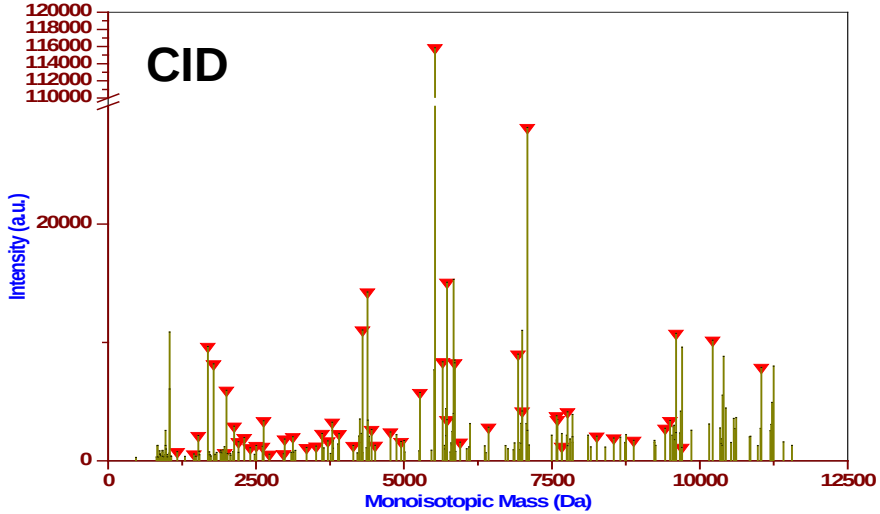
RPLC separation of histones

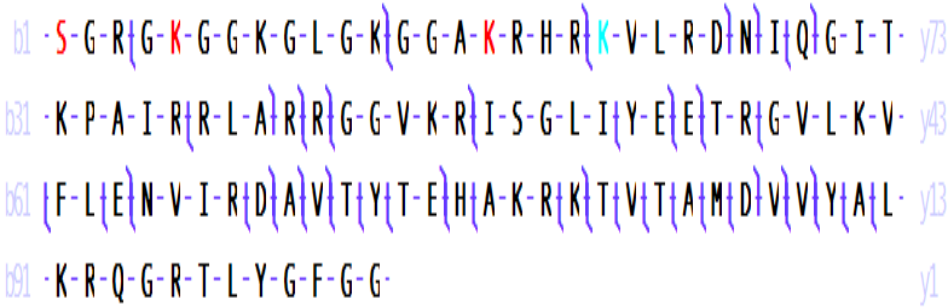


WCX-MS/MS of H4

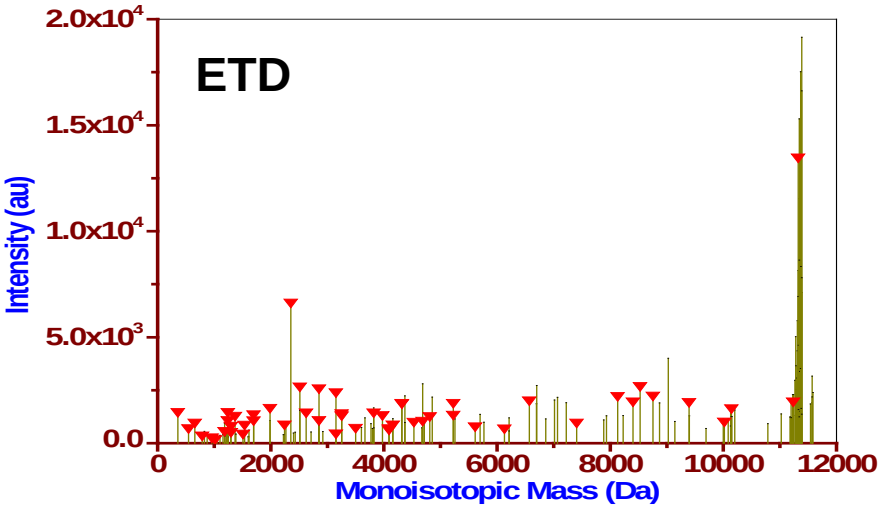


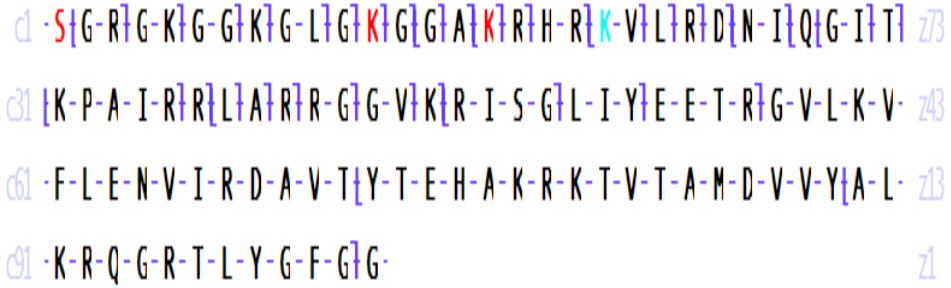
Unambiguous localization of K12ac in histone H4





Ambiguous identifications:
H4_S1ac**K{5, 8, 12}ac**K16acK20me2



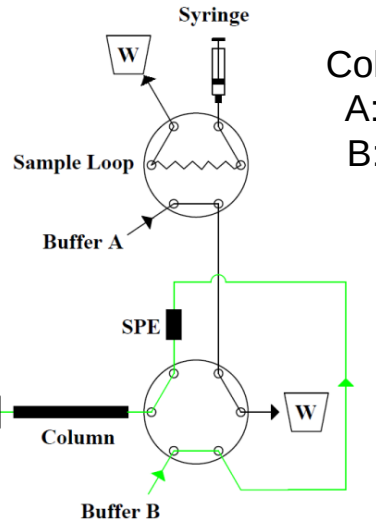


Unique identification:
H4_S1ac**K12ac**K16acK20me2

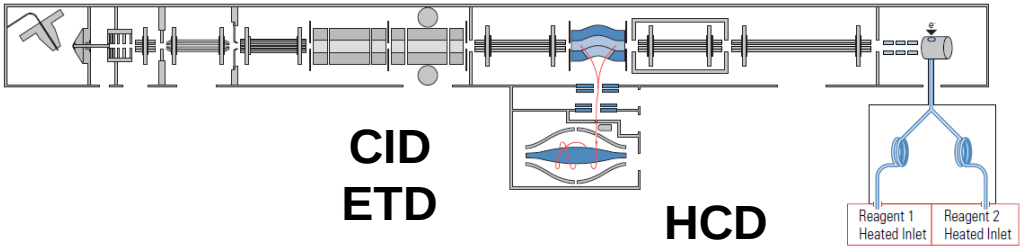
Methylation (mono) Acetylation Phosphorylation Dimethylation Trimethylation

Optimizing fragmentation for histone characterization

RPLC



Column (200 µid*70cm, C5(J), 5µ, 300Å)
A: 20%ACN, 5%IPA, 0.6%FA
B: 45%ACN, 45%IPA, 0.6%FA



LTQ Orbitrap Velos

HeLa core histones (3.75 µg)

	NCE (%)	AT (ms)
HCD_45_30	45	30
HCD_45_60	45	60
HCD_60_60	60	60
CID_35_30	35	30
CID_45_60	45	60
CID_60_60	60	60
ETD_10	/	10
ETD_15	/	15
ETD_20	/	20

.RAW

ProSightPC 2.0

IDs

Different fragmentation conditions

	NCE (%)	RT (ms)	IDs	Avg [–log(P_Score)]
HCD_45_30	45	30	123	16.9
HCD_45_60	45	60	136	15.3
HCD_60_60	60	60	144 **±8**	20.4 **±0.3**
CID_35_30	35	30	62	10.3
CID_45_60	45	60	56	16.5
CID_60_60	60	60	65	21.2
ETD_10	/	10	83	23.3
ETD_15	/	15	160	34.5
ETD_20	/	20	138	28.9

HCD conditions identification of H4_S1acK16ack20me2

HCD_45_30: Matching fragments=**27**, P_Score=**6.5E-31**

```
b1  - S - G - R - G - K - G - G - K - G - L - G - K - G - G - A - K - R - H - R - K - V - L - R - D - N - I - Q - G - I - T -  y73
b31 - K - P - A - I - R - R - L - A - R - R - G - G - V - K - R - I - S - G - L - I - Y - E - E - T - R - G - V - L - K - V -  y43
b61 - F - L - E - N - V - I - R - D - A - V - T - Y - T - E - H - A - K - R - K - T - V - T - A - M - D - V - V - Y - A - L -  y13
b91 - K - R - Q - G - R - T - L - Y - G - F - G - G -                                                                          y1
```

HCD_45_60: Matching fragments=**38**, P_Score=**5.9E-39**

```
b1  - S - G - R - G - K - G - G - K - G - L - G - K - G - G - A - K - R - H - R - K - V - L - R - D - N - I - Q - G - I - T -  y73
b31 - K - P - A - I - R - R - L - A - R - R - G - G - V - K - R - I - S - G - L - I - Y - E - E - T - R - G - V - L - K - V -  y43
b61 - F - L - E - N - V - I - R - D - A - V - T - Y - T - E - H - A - K - R - K - T - V - T - A - M - D - V - V - Y - A - L -  y13
b91 - K - R - Q - G - R - T - L - Y - G - F - G - G -                                                                          y1
```

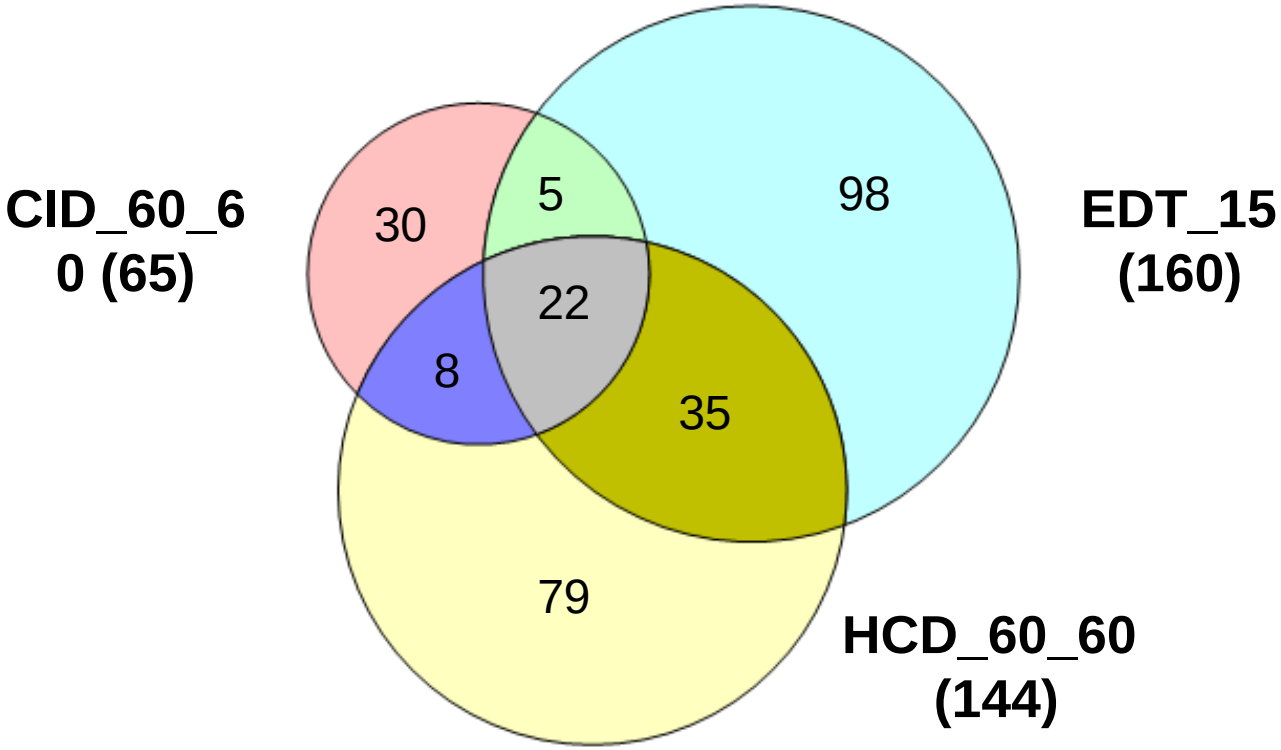
HCD_60_60: **Matching fragments=42, P_Score=1.3E-50**

```
b1  - S - G - R - G - K - G - G - K - G - L - G - K - G - G - A - K - R - H - R - K - V - L - R - D - N - I - Q - G - I - T -  y73
b31 - K - P - A - I - R - R - L - A - R - R - G - G - V - K - R - I - S - G - L - I - Y - E - E - T - R - G - V - L - K - V -  y43
b61 - F - L - E - N - V - I - R - D - A - V - T - Y - T - E - H - A - K - R - K - T - V - T - A - M - D - V - V - Y - A - L -  y13
b91 - K - R - Q - G - R - T - L - Y - G - F - G - G -                                                                          y1
```

Methylation (mono) Acetylation Phosphorylation Dimethylation Trimethylation

Optimal CID, ETD, and HCD

	NCE (%)	RT (ms)	IDs	Avg [–log(P_Score)]
HCD_60_60	60	60	144	20.4
CID_60_60	60	60	65	21.2
ETD_15	/	15	160	34.5



ETD, HCD, and CID: H4_S1acK16ack20me2

ETD: Matching fragments=**56**, P_Score=**1.1E-76**

```
c1  -S-G-R-G-K-G-G-K-G-L-G-K-G-G-A-K-R-H-R-K-V-L-R-D-N-I-Q-G-I-T-  z73
c31 -K-P-A-I-R-R-L-A-R-R-G-G-V-K-R-I-S-G-L-I-Y-E-E-T-R-G-V-L-K-V-  z43
c61 -F-L-E-N-V-I-R-D-A-V-T-Y-T-E-H-A-K-R-K-T-V-T-A-M-D-V-V-Y-A-L-  z13
c91 -K-R-Q-G-R-T-L-Y-G-F-G-G-                                      z1
```

HCD: Matching fragments=**46**, P_Score=**3.3E-56**

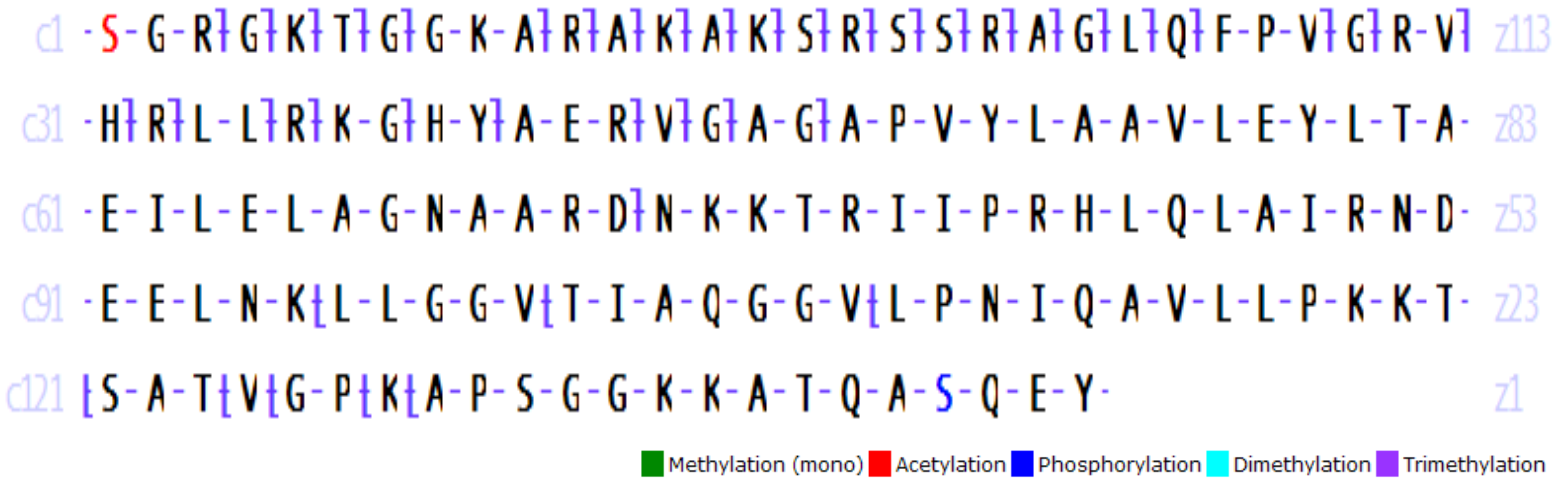
```
b1  -S-G-R-G-K-G-G-K-G-L-G-K-G-G-A-K-R-H-R-K-V-L-R-D-N-I-Q-G-I-T-  y73
b31 -K-P-A-I-R-R-L-A-R-R-G-G-V-K-R-I-S-G-L-I-Y-E-E-T-R-G-V-L-K-V-  y43
b61 -F-L-E-N-V-I-R-D-A-V-T-Y-T-E-H-A-K-R-K-T-V-T-A-M-D-V-V-Y-A-L-  y13
b91 -K-R-Q-G-R-T-L-Y-G-F-G-G-                                      y1
```

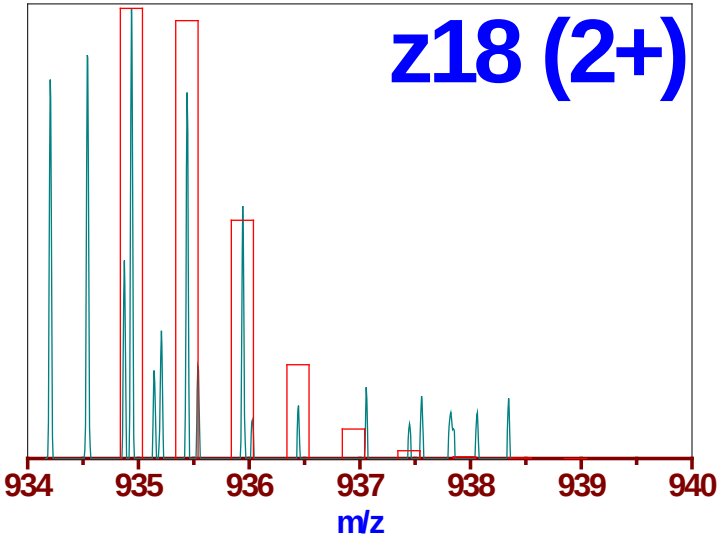
CID: Matching fragments=**31**, P_Score=**1.1E-40**

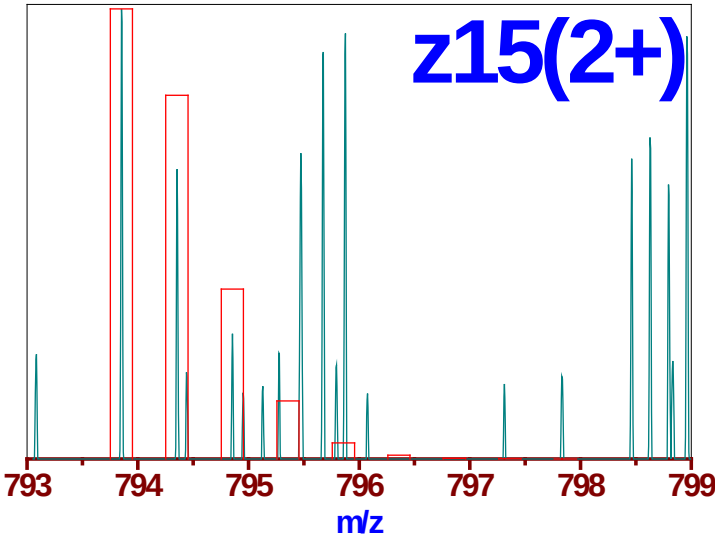
```
b1  -S-G-R-G-K-G-G-K-G-L-G-K-G-G-A-K-R-H-R-K-V-L-R-D-N-I-Q-G-I-T-  y73
b31 -K-P-A-I-R-R-L-A-R-R-G-G-V-K-R-I-S-G-L-I-Y-E-E-T-R-G-V-L-K-V-  y43
b61 -F-L-E-N-V-I-R-D-A-V-T-Y-T-E-H-A-K-R-K-T-V-T-A-M-D-V-V-Y-A-L-  y13
b91 -K-R-Q-G-R-T-L-Y-G-F-G-G-                                      y1
```

Methylation (mono) Acetylation Phosphorylation Dimethylation Trimethylation

H2AX_S1acS139p uniquely identified only by ETD







Top-down only! Not with bottom-up or middle-down!

Advantage of top-down fragmentation

HCD_60_60: H31_K9me2K27me2K36meK79me

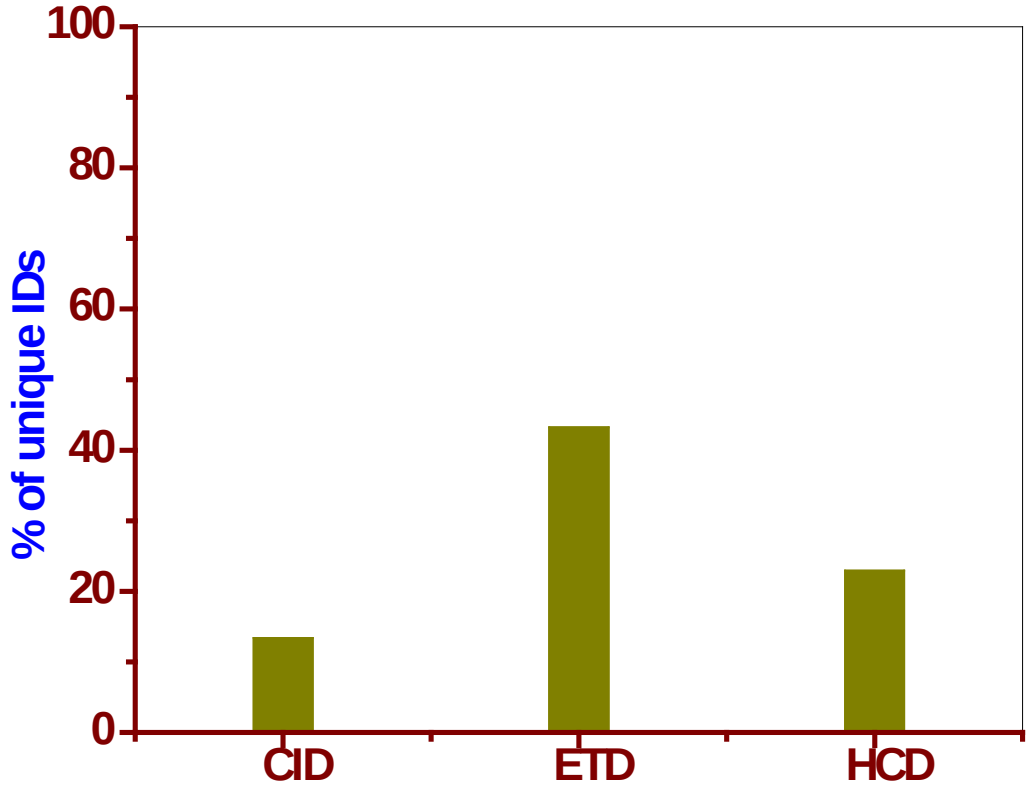
```
b1   - A - R - T - K - Q - T - A - R - K - S - T - G - G - K - A - P - R - K - Q - L - A - T - K - A - A - R - K - S - A - P -   y106
b31  - A - T - G - G - V - K - K - P - H - R - Y - R - P - G - T - V - A - L - R - E - I - R - R - Y - Q - K - S - T - E - L -   y76
b61  - L - I - R - K - L - P - F - Q - R - L - V - R - E - I - A - Q - D - F - K - T - D - L - R - F - Q - S - S - A - V - M -   y46
b91  - A - L - Q - E - A - C - E - A - Y - L - V - G - L - F - E - D - T - N - L - C - A - I - H - A - K - R - V - T - I - M -   y16
b121 - P - K - D - I - Q - L - A - R - R - I - R - G - E - R - A -   y1
```

CID_60_60: H2B1O_K46meK57me2

```
b1   - P - D - P - A - K - S - A - P - A - P - K - K - G - S - K - K - A - V - T - K - A - Q - K - K - D - G - K - E - R - K -   y96
b31  - R - S - R - K - E - S - Y - S - I - Y - V - Y - K - V - L - K - Q - V - H - P - D - T - G - I - S - S - K - A - M - G -   y66
b61  - I - M - N - S - F - V - N - D - I - F - E - R - I - A - G - E - A - S - R - L - A - H - Y - N - K - R - S - T - I - T -   y36
b91  - S - R - E - I - Q - T - A - V - R - L - L - L - P - G - E - L - A - K - H - A - V - S - E - G - T - K - A - V - T - K -   y6
b121 - Y - T - S - S - K -   y1
```

Methylation (mono) Acetylation Phosphorylation Dimethylation Trimethylation

Fragmentation efficiency of CID, ETD, HCD



$$\%_of_unique_IDs = \frac{unique_IDs}{unique_IDs + ambiguous_IDs} \times 100$$

Conclusions

- An online 2D LC-FTMS platform dramatically improved throughput and sensitivity compared to more traditional platforms, and resulted in identification of hundreds of histone isoforms from microgram levels of protein, including phosphorylation.
- For a single fragmentation method, ETD provides the most confident identiifcations on average.
- Total number of identifications more than doubled when CID was switched to ETD or HCD.
- Combinatorial modifications across the entire protein sequence could only be identified using top-down proteomics.

Acknowledgements

Ljiljana Paša-Tolić

Joshua Aldrich

Robby Robinson

Shawna Hengel

Si Wu

Bioinformatics

Gordon A. Anderson

Matthew E. Monroe

Nikola Tolić

Samuel O. Purvine

Instrumentation

Ronald J. Moore

Anil K. Shukla

Boyd Champion

Carrie D. Nicora

Daniel J. Orton

Heather M. Mottaz-Brewer

Karl Weitz

Robbie Heegel

Rui Zhao

Ryan T. Kelly

Therese R. W. Clauss

Thomas Fillmore

Michael Russcher

Randy Norheim

Richard D. Smith and other group members

Support and Funding: *EMSL at PNNL, DOE, NIH NCRR (RR018522)*