Kai A. Reidegeld

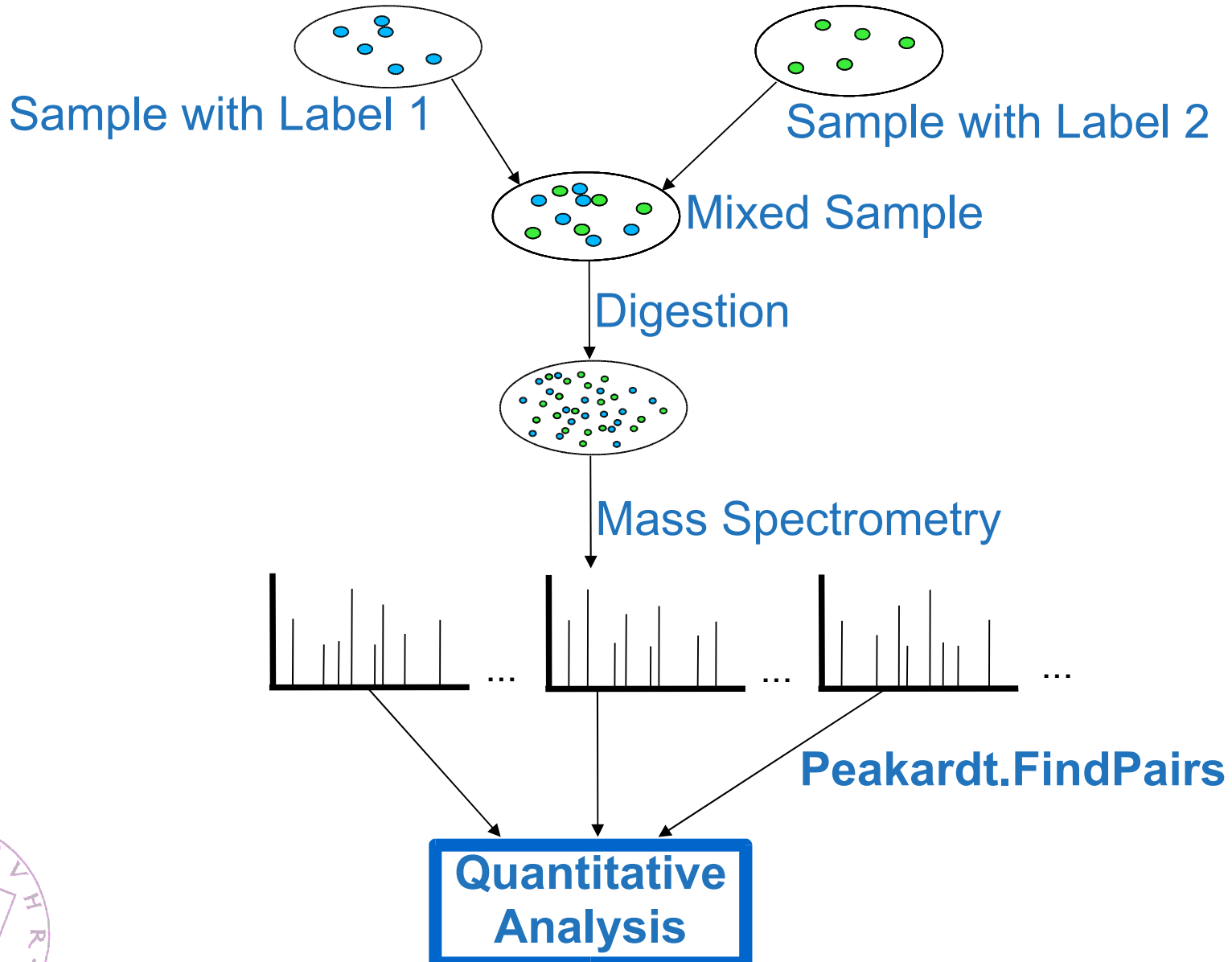
Peakardt.FindPairs

software for automatic quantitative evaluation of stable isotope-coded peptide mass spectra

Medical Proteom-Center
Prof. Dr. Helmut E. Meyer
Ruhr-University Bochum

Kai A. Reidegeld

Sample Processing Steps in Stable Isotope Labeling Experiments



Kai A. Reidegeld

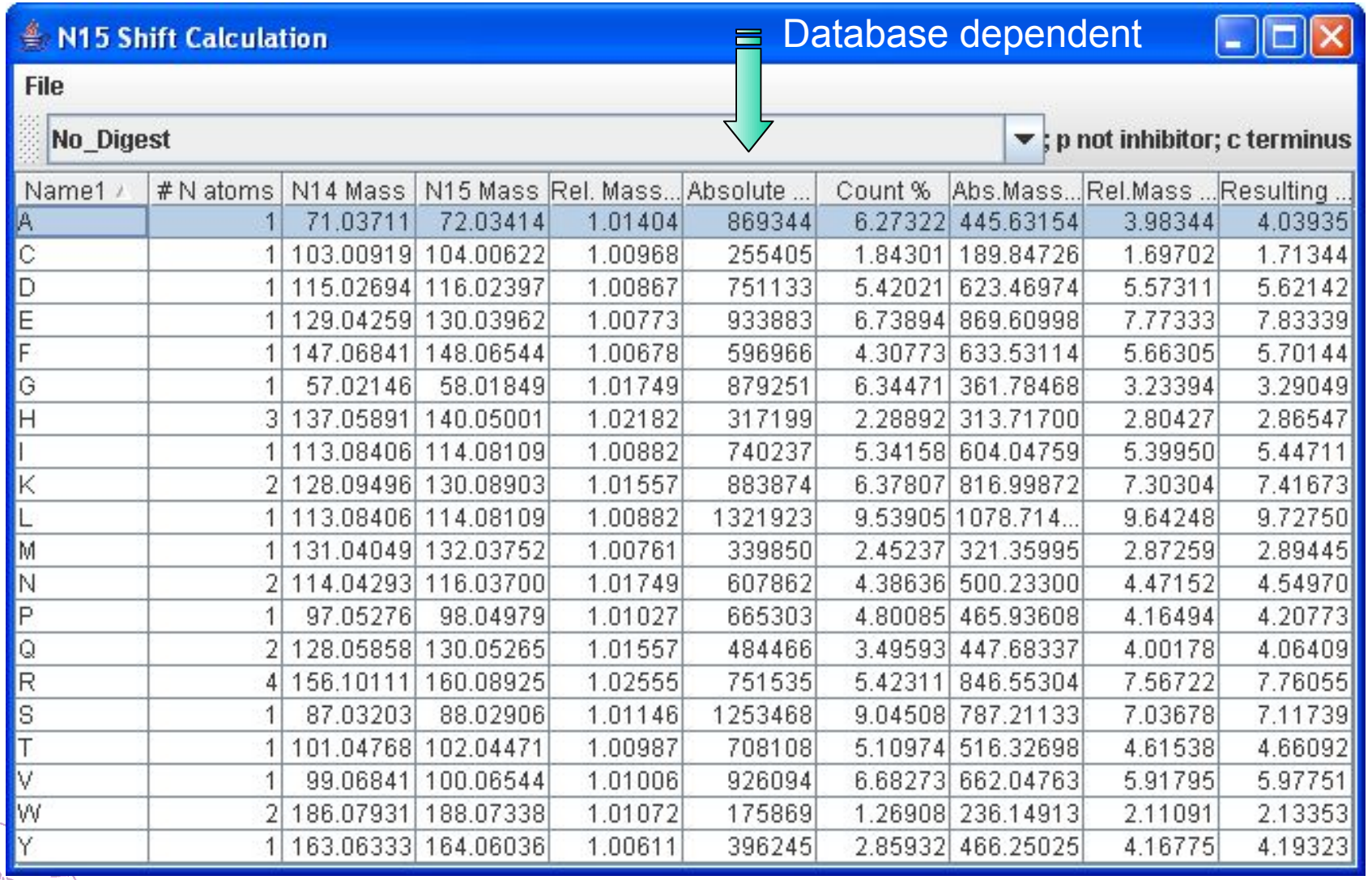
Estimated Mass Shifts for Stable Isotope Labeling Techniques

Labeling Technique	Mass Shift ΔM
ICAT / Cleavable ICAT	+8 Da (+442 / +227)
ALICE	+10 Da
MCAT	+42 Da
GIST	+4 Da
^{18}O	+4 Da (resp. 2 Da)
ICPL	+4 Da
iTRAQ	+1, +2, +3 or +4 Da
SILAC	+6 Da (resp. 3 Da)
^{15}N stable isotope	*1,0122...

- the ^{15}N mass shift is depending on sample, tissue, organism…
- calculation based on database analysis

Kai A. Reidegeld

Mass-Shift Factor Calculation from Protein DB



Name1	# N atoms	N14 Mass	N15 Mass	Rel. Mass...	Absolute ...	Count %	Abs.Mass...	Rel.Mass ...	Resulting ...
A	1	71.03711	72.03414	1.01404	869344	6.27322	445.63154	3.98344	4.03935
C	1	103.00919	104.00622	1.00968	255405	1.84301	189.84726	1.69702	1.71344
D	1	115.02694	116.02397	1.00867	751133	5.42021	623.46974	5.57311	5.62142
E	1	129.04259	130.03962	1.00773	933883	6.73894	869.60998	7.77333	7.83339
F	1	147.06841	148.06544	1.00678	596966	4.30773	633.53114	5.66305	5.70144
G	1	57.02146	58.01849	1.01749	879251	6.34471	361.78468	3.23394	3.29049
H	3	137.05891	140.05001	1.02182	317199	2.28892	313.71700	2.80427	2.86547
I	1	113.08406	114.08109	1.00882	740237	5.34158	604.04759	5.39950	5.44711
K	2	128.09496	130.08903	1.01557	883874	6.37807	816.99872	7.30304	7.41673
L	1	113.08406	114.08109	1.00882	1321923	9.53905	1078.714...	9.64248	9.72750
M	1	131.04049	132.03752	1.00761	339850	2.45237	321.35995	2.87259	2.89445
N	2	114.04293	116.03700	1.01749	607862	4.38636	500.23300	4.47152	4.54970
P	1	97.05276	98.04979	1.01027	665303	4.80085	465.93608	4.16494	4.20773
Q	2	128.05858	130.05265	1.01557	484466	3.49593	447.68337	4.00178	4.06409
R	4	156.10111	160.08925	1.02555	751535	5.42311	846.55304	7.56722	7.76055
S	1	87.03203	88.02906	1.01146	1253468	9.04508	787.21133	7.03678	7.11739
T	1	101.04768	102.04471	1.00987	708108	5.10974	516.32698	4.61538	4.66092
V	1	99.06841	100.06544	1.01006	926094	6.68273	662.04763	5.91795	5.97751
W	2	186.07931	188.07338	1.01072	175869	1.26908	236.14913	2.11091	2.13353
Y	1	163.06333	164.06036	1.00611	396245	2.85932	466.25025	4.16775	4.19323

Example: IPI Database for Arabidopsis Thaliana v3.04

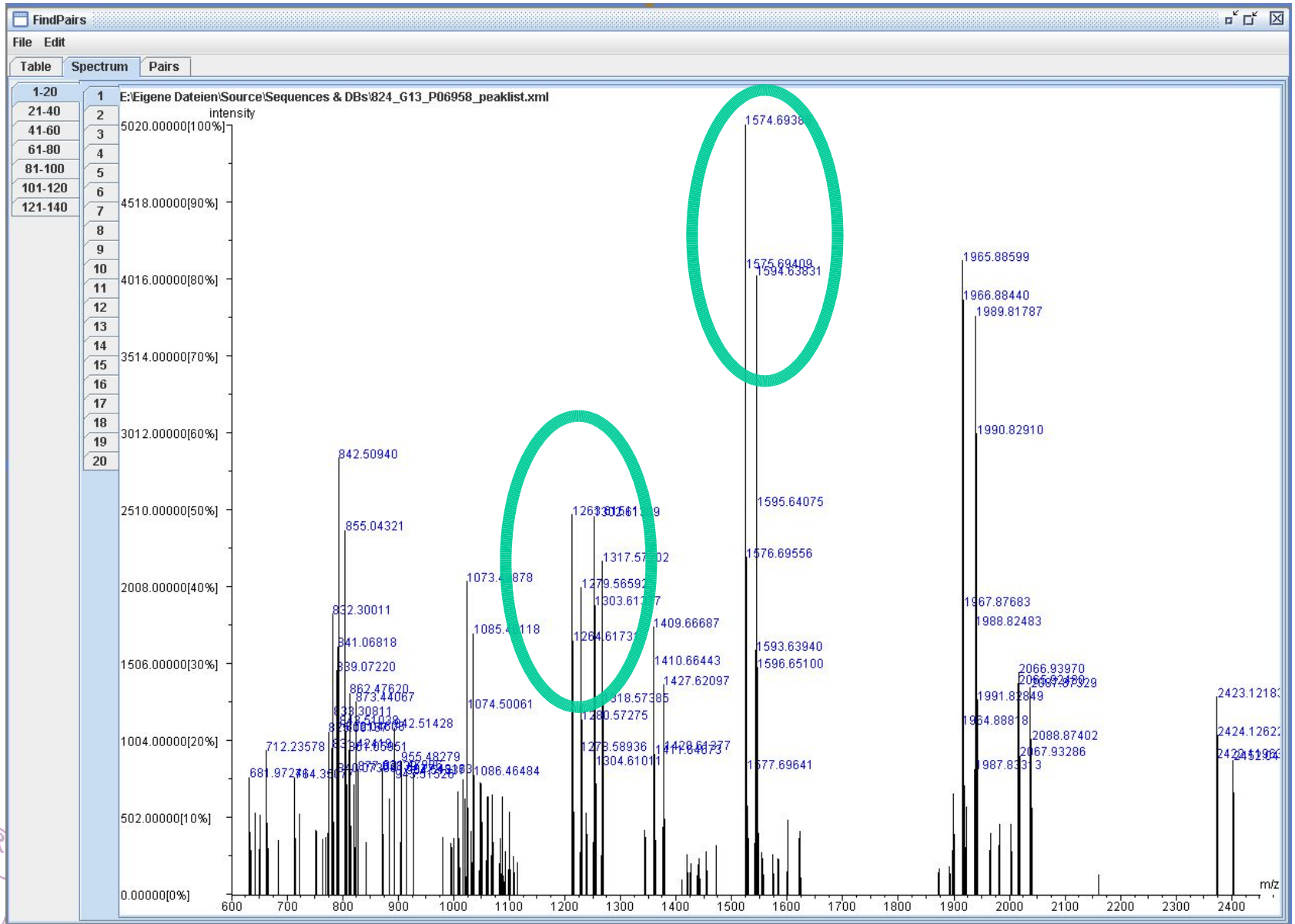
Kai A. Reidegeld

Protein Database Mass-Shift Factor

Organism	IPI Protein DB	Mass-shift factor %
Arabidopsis Thaliana	ARATH v3.01	1.21545398
Arabidopsis Thaliana	ARATH v3.04	1.21543732
Brachydanio rerio	BRARE v3.02	1.22314878
Brachydanio rerio	BRARE v3.05	1.22313178
Gallus gallus	CHICK v3.00	1.23005704
Homo sapiens	HUMAN v3.03	1.23799997
Homo sapiens	HUMAN v3.06	1.23651426
Mus musculus	MOUSE v3.03	1.23223444
Mus musculus	MOUSE v3.06	1.23165359
Rattus norvegicus	RAT v3.03	1.22956131
Rattus norvegicus	RAT v3.06	1.23147794
	Arithmetic mean:	**1.2278791**

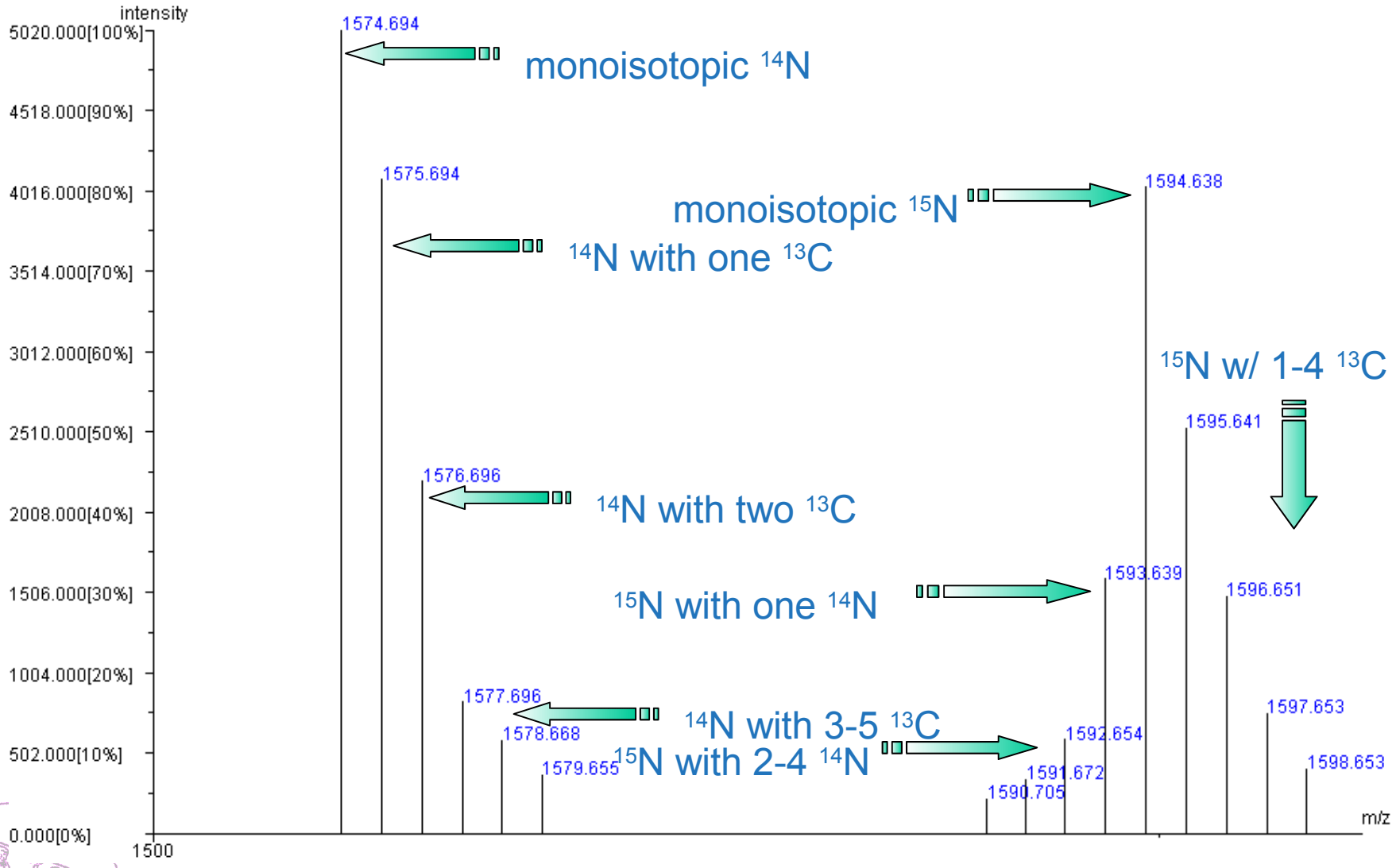
Kai A. Reidegeld

MALDI MS spectrum of ^{14}N-^{15}N labeled *E.Coli* sample



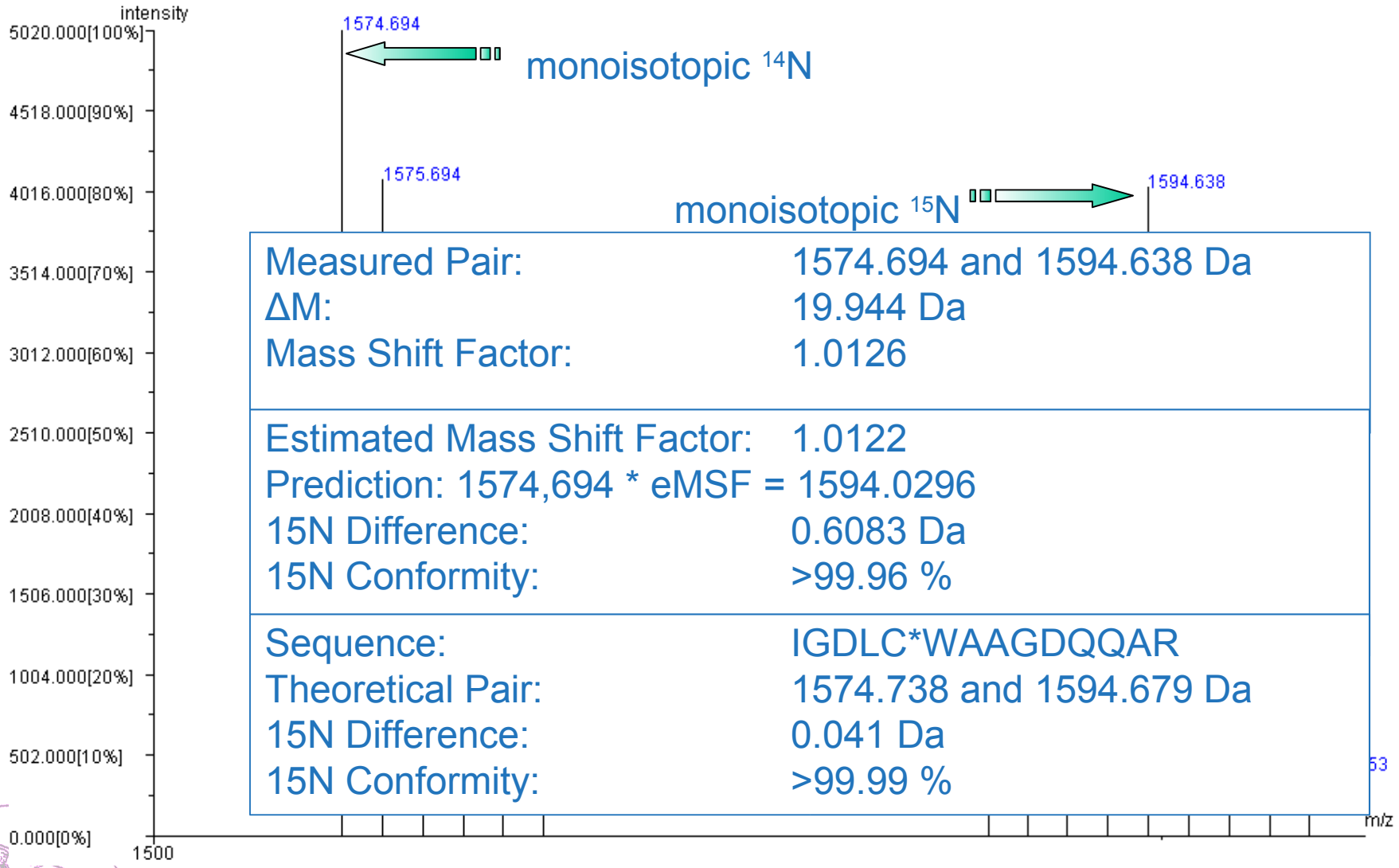
Kai A. Reidegeld

MALDI PMF of a $^{15}N/^{14}N$ labeled Peptide



Kai A. Reidegeld

MALDI PMF of a $^{15}N/^{14}N$ labeled Peptide



Kai A. Reidegeld

What is the quantitative difference?

Peakset	Intensity ^{14}N	Intensity ^{15}N	Intensity Ratio
Monoisotopic	5022	4038	1 : 0.8
Mono+1*^{13}C	9107	6572	1 : 0.72
Mono+2*^{13}C	11305	8052	1 : 0.71
Mono+2*^{13}C+^{15}N Sat.	11305	9643	1 : 0.85
Mono+all ^{13}C	13066	9201	1 : 0.7
All	13066	11923	1 : 0.91

Mathematical:

use all peaks, because all contribute to the isotopic distribution

Ease of detection:

just use the monoisotopic peaks

Reproducibility and a good compromise:

require a pair to have at least two ^{13}C Peaks
and the ^{15}N peak to have one satellite

Kai A. Reidegeld

Two possible ways for quantitation:

- **use proposed heuristic:**
  - Fixed mass shift (summand or factor)
  - any value possible
    - (even for not yet invented labeling techniques)

- **use protein sequence information:**
  - choose labeling reagent as secondary modification
  - let Peakardt do an *in-silico* digestion of the sequence
  - only those peak pairs are quantitated which come from peptides digested from the protein sequence

Kai A. Reidegeld

How to do quantitative analysis of mass spectrometric data with Peakardt.FindPairs

1. Open Mass Spectra
(2. Adjust Parameters)
3. Press  -Button

Kai A. Reidegeld

1. Open Mass Spectra

PMF – Peptide Mass Fingerprint or
PFF – Peptide Fragmentation Fingerprint

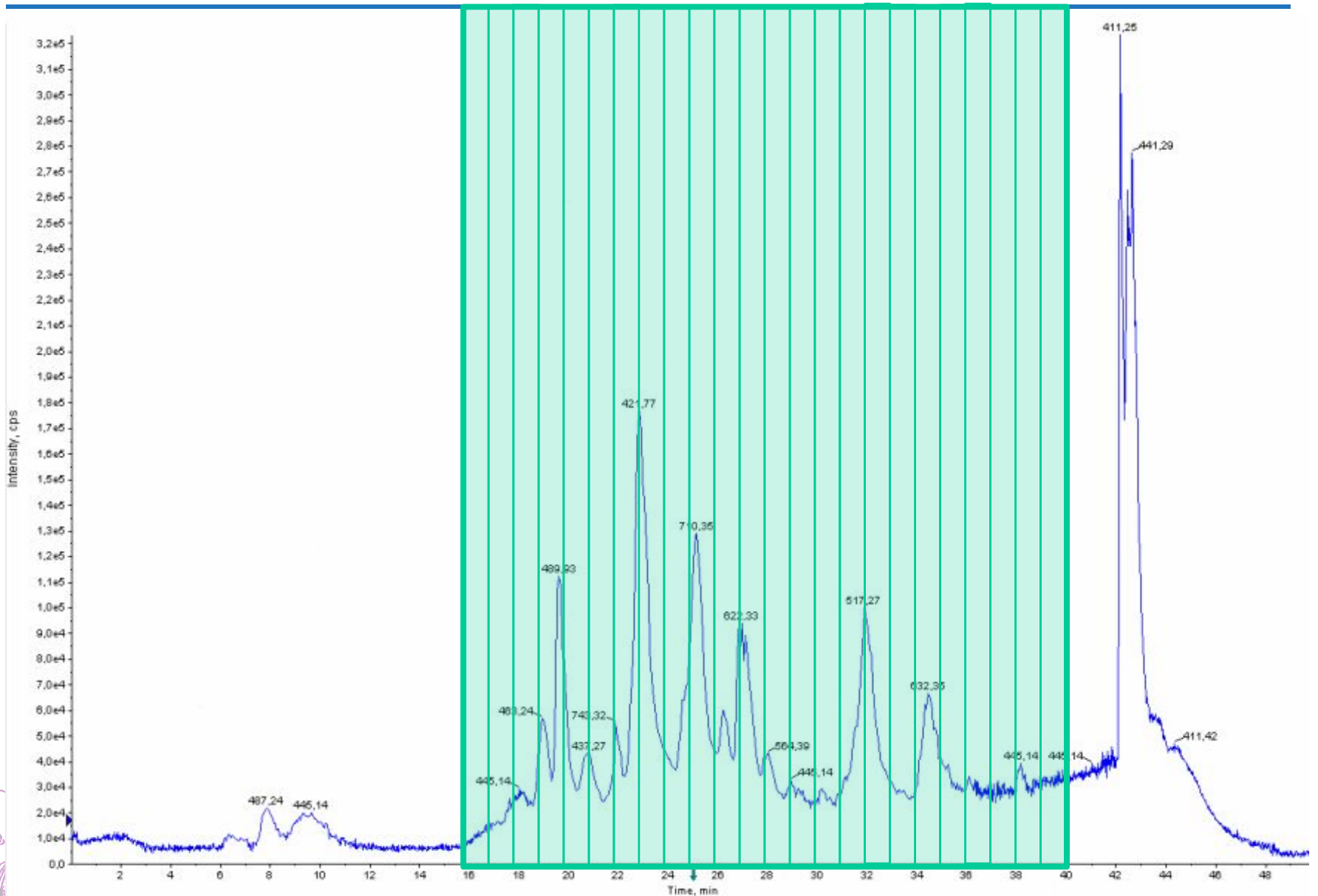
choose among:

- Applied Biosystems - QStarXL peaklists
- Bruker - Ultraflex - Xmass XML and peaklist files
- Mascot Generic Format (MGF files)
- Thermo Electron - LCQ Xcalibur DTA files

open recursively multiple (thousands) spectra at once

Kai A. Reidegeld

Quantitation of nHPLC MS runs

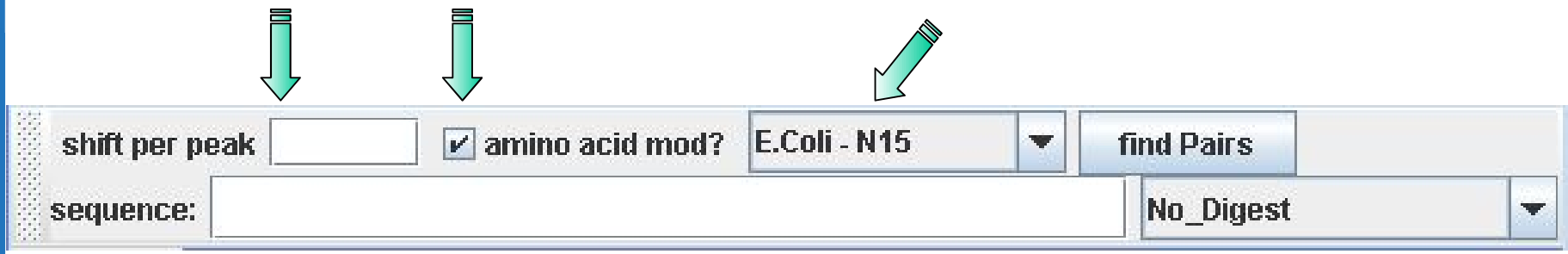


Kai A. Reidegeld

2. Parameters

- enter **mass shift value** if protein is unknown

mass shift value summand or factor? predefined mass shift values?



Kai A. Reidegeld

2. Parameters

- enter **sequence** if protein is known

enter protein sequence

choose protease

Kai A. Reidegeld

2. Parameters (cntd.)

General Parameters

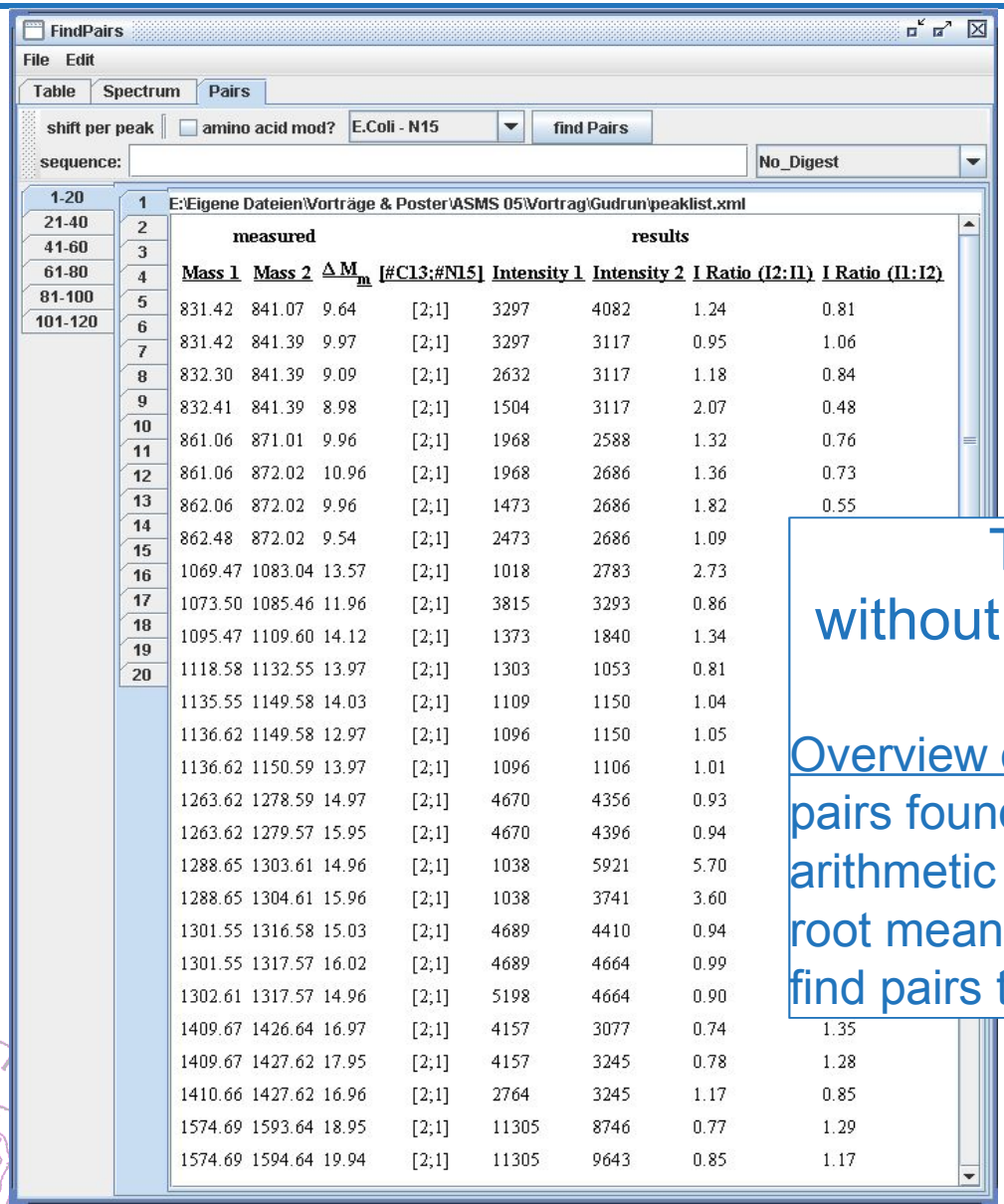
- expected mass shift (unknown sequence)
- # of ^{13}C peaks, # of ^{15}N satellites
- mass range
- mass tolerance

Sample Specific Parameters

- labeling technique
- protein sequence
- digestion
- modifications

Kai A. Reidegeld

FindPairs considering Mass Shift



	measured				results			
	Mass 1	Mass 2	ΔM_m	[#C13;#N15]	Intensity 1	Intensity 2	I Ratio (I2:I1)	I Ratio (I1:I2)
1	831.42	841.07	9.64	[2;1]	3297	4082	1.24	0.81
2	831.42	841.39	9.97	[2;1]	3297	3117	0.95	1.06
3	832.30	841.39	9.09	[2;1]	2632	3117	1.18	0.84
4	832.41	841.39	8.98	[2;1]	1504	3117	2.07	0.48
5	861.06	871.01	9.96	[2;1]	1968	2588	1.32	0.76
6	861.06	872.02	10.96	[2;1]	1968	2686	1.36	0.73
7	862.06	872.02	9.96	[2;1]	1473	2686	1.82	0.55
8	862.48	872.02	9.54	[2;1]	2473	2686	1.09	
9	1069.47	1083.04	13.57	[2;1]	1018	2783	2.73	
10	1073.50	1085.46	11.96	[2;1]	3815	3293	0.86	
11	1095.47	1109.60	14.12	[2;1]	1373	1840	1.34	
12	1118.58	1132.55	13.97	[2;1]	1303	1053	0.81	
13	1135.55	1149.58	14.03	[2;1]	1109	1150	1.04	
14	1136.62	1149.58	12.97	[2;1]	1096	1150	1.05	
15	1136.62	1150.59	13.97	[2;1]	1096	1106	1.01	
16	1263.62	1278.59	14.97	[2;1]	4670	4356	0.93	
17	1263.62	1279.57	15.95	[2;1]	4670	4396	0.94	
18	1288.65	1303.61	14.96	[2;1]	1038	5921	5.70	
19	1288.65	1304.61	15.96	[2;1]	1038	3741	3.60	
20	1301.55	1316.58	15.03	[2;1]	4689	4410	0.94	
	1301.55	1317.57	16.02	[2;1]	4689	4664	0.99	
	1302.61	1317.57	14.96	[2;1]	5198	4664	0.90	
	1409.67	1426.64	16.97	[2;1]	4157	3077	0.74	1.35
	1409.67	1427.62	17.95	[2;1]	4157	3245	0.78	1.28
	1410.66	1427.62	16.96	[2;1]	2764	3245	1.17	0.85
	1574.69	1593.64	18.95	[2;1]	11305	8746	0.77	1.29
	1574.69	1594.64	19.94	[2;1]	11305	9643	0.85	1.17

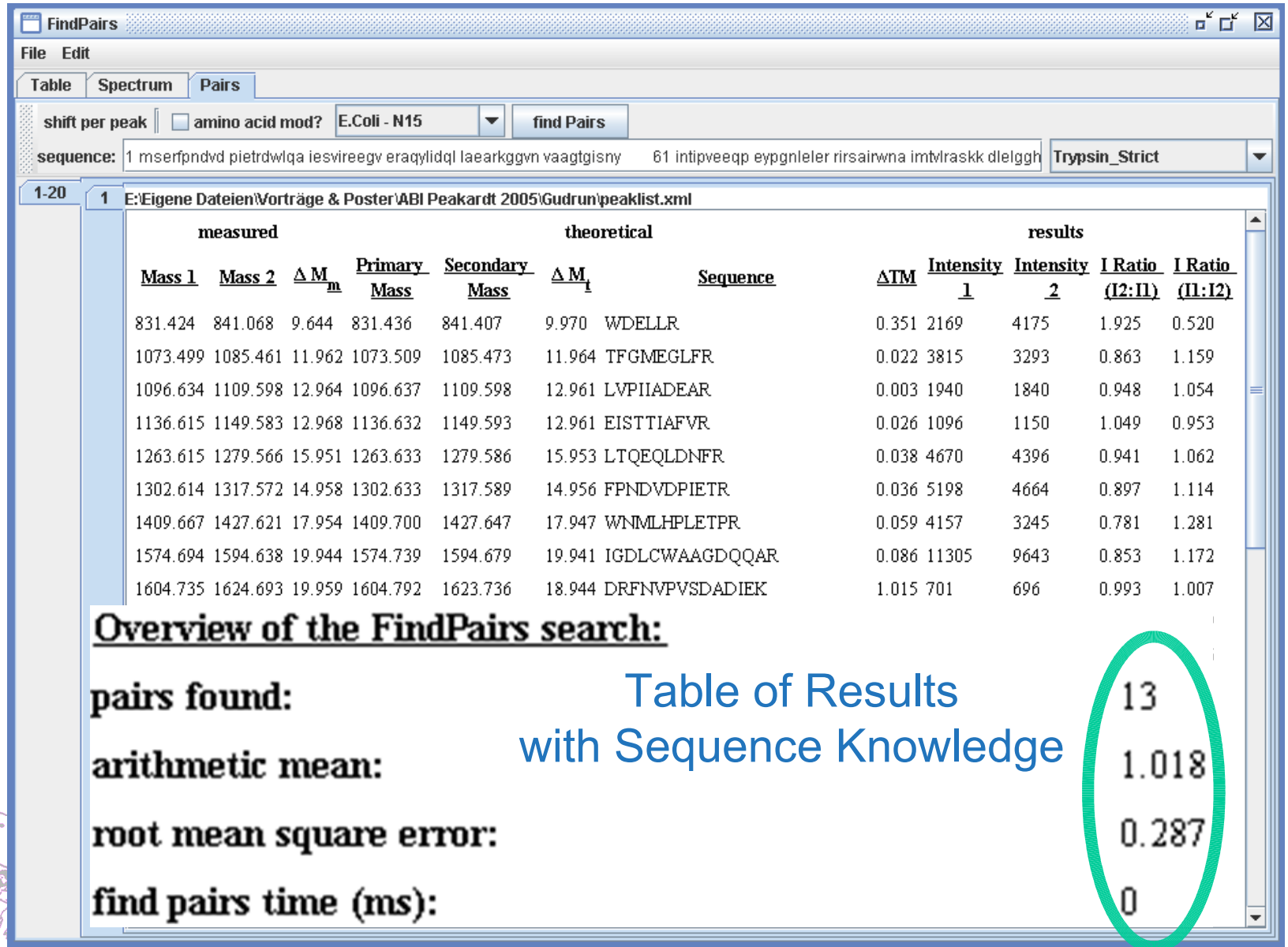
Table of Results without Sequence Knowledge

Overview of the FindPairs search:

pairs found:	44
arithmetic mean:	1.41
root mean square error:	0.93
find pairs time (ms):	0

Kai A. Reidegeld

FindPairs with Sequence Knowledge



measured					theoretical			results				
Mass 1	Mass 2	ΔM_m	Primary Mass	Secondary Mass	ΔM_t	Sequence	ΔTM	Intensity 1	Intensity 2	I Ratio (I2:I1)	I Ratio (I1:I2)
831.424	841.068	9.644	831.436	841.407	9.970	WDELLR	0.351	2169	4175	1.925	0.520
1073.499	1085.461	11.962	1073.509	1085.473	11.964	TFGMEGLFR	0.022	3815	3293	0.863	1.159
1096.634	1109.598	12.964	1096.637	1109.598	12.961	LVPIIADEAR	0.003	1940	1840	0.948	1.054
1136.615	1149.583	12.968	1136.632	1149.593	12.961	EISTTIAFVR	0.026	1096	1150	1.049	0.953
1263.615	1279.566	15.951	1263.633	1279.586	15.953	LTQEQLDNFR	0.038	4670	4396	0.941	1.062
1302.614	1317.572	14.958	1302.633	1317.589	14.956	FPNDVDPIETR	0.036	5198	4664	0.897	1.114
1409.667	1427.621	17.954	1409.700	1427.647	17.947	WNMLHPLETPR	0.059	4157	3245	0.781	1.281
1574.694	1594.638	19.944	1574.739	1594.679	19.941	IGDLCWAAGDQQAR	0.086	11305	9643	0.853	1.172
1604.735	1624.693	19.959	1604.792	1623.736	18.944	DRFNVPVSDADIEK	1.015	701	696	0.993	1.007

Overview of the FindPairs search:

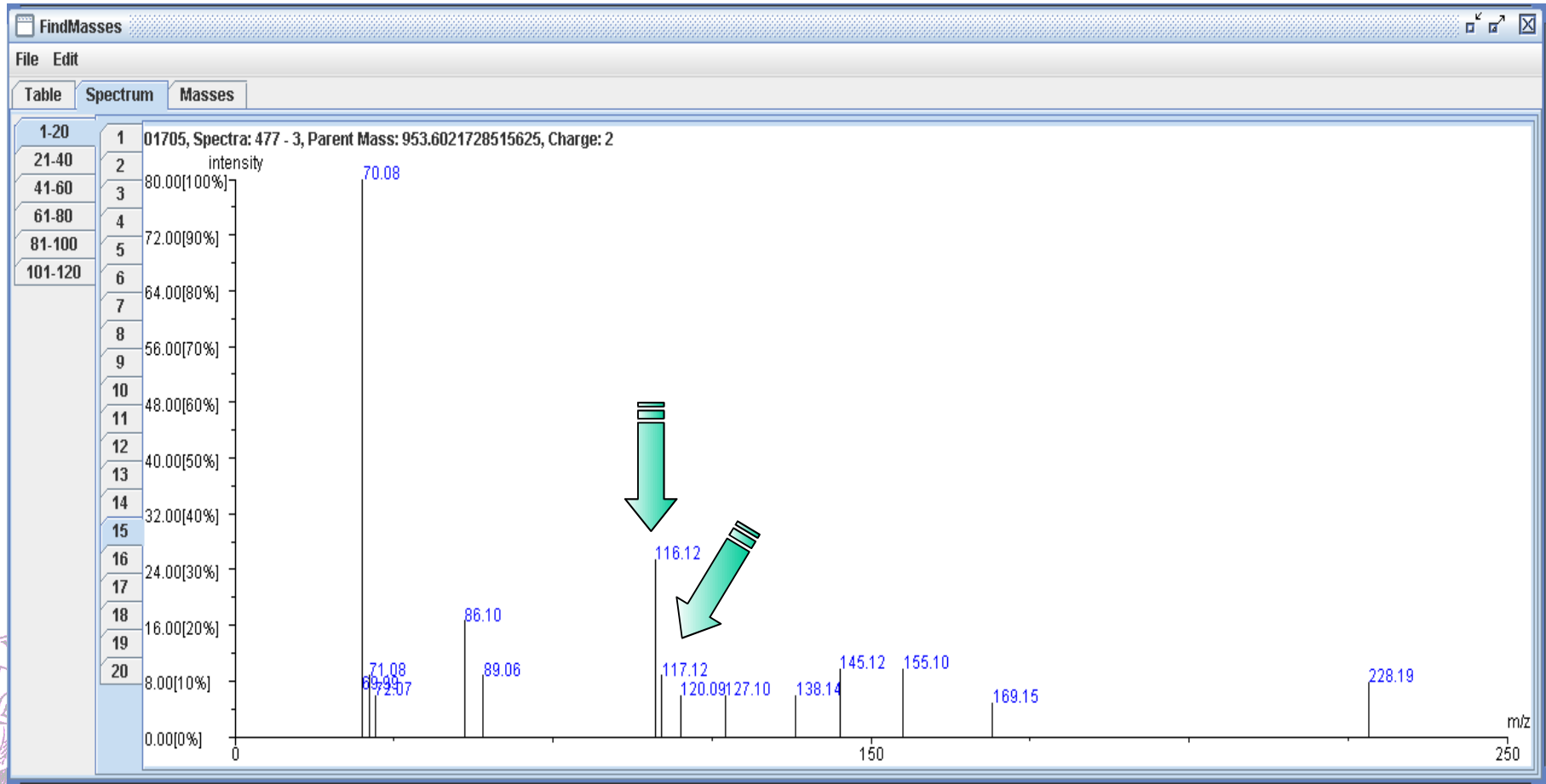
pairs found:	13
arithmetic mean:	1.018
root mean square error:	0.287
find pairs time (ms):	0

Table of Results
with Sequence Knowledge

Kai A. Reidegeld

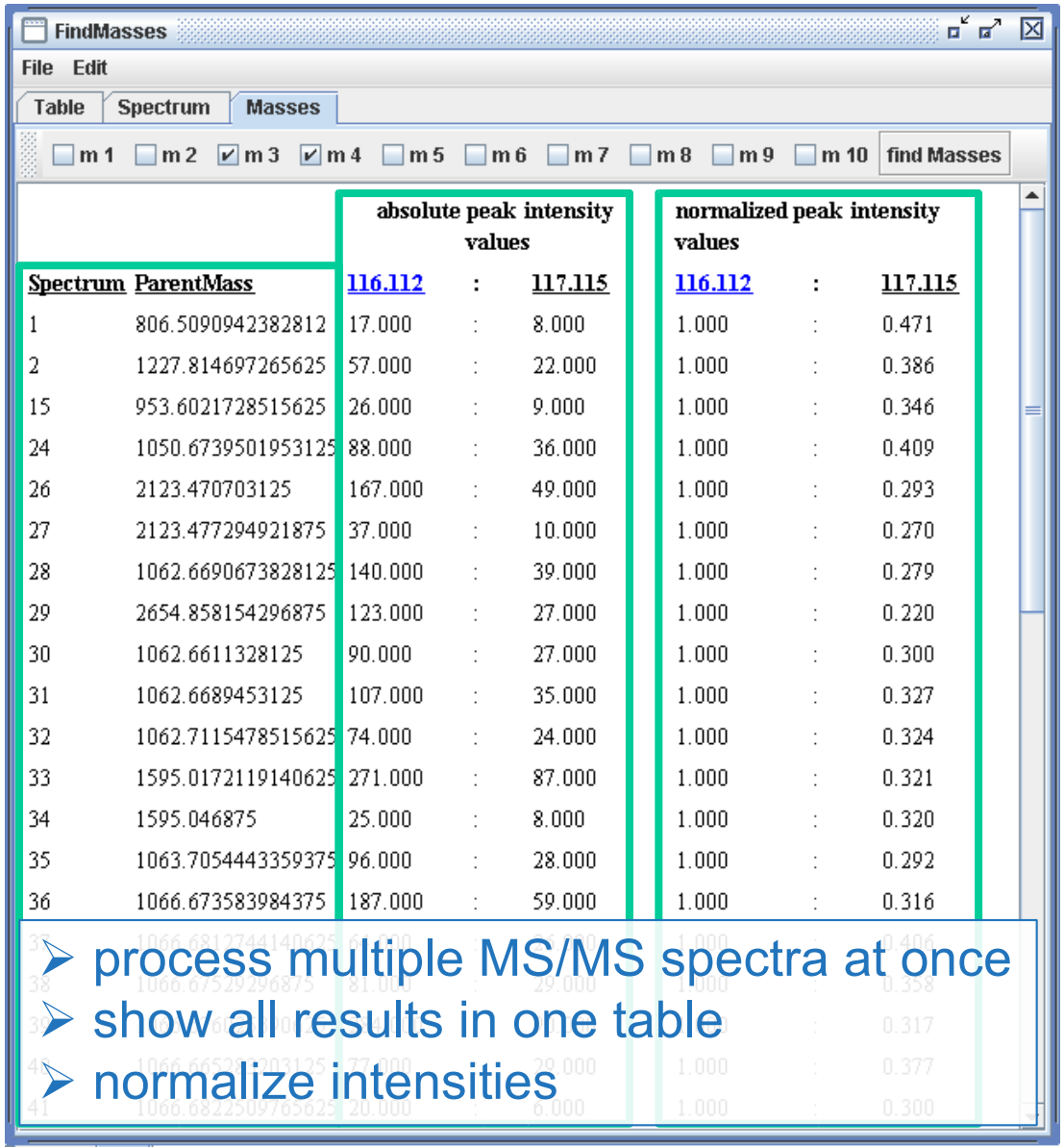
Peakardt.FindMasses: quantitation of fixed mass values

- quantitation for reagents which produce well defined masses like iTRAQ™
- predefined m/z values: e.g. iTRAQ™ reporter ions in PFF spectra
  - 114, 115, 116, 117



Kai A. Reidegeld

FindMasses: iTRAQ labeled Protein 116:117=1:0.33



		absolute peak intensity values			normalized peak intensity values		
Spectrum	ParentMass	116.112	:	117.115	116.112	:	117.115
1	806.5090942382812	17.000	:	8.000	1.000	:	0.471
2	1227.814697265625	57.000	:	22.000	1.000	:	0.386
15	953.6021728515625	26.000	:	9.000	1.000	:	0.346
24	1050.6739501953125	88.000	:	36.000	1.000	:	0.409
26	2123.470703125	167.000	:	49.000	1.000	:	0.293
27	2123.477294921875	37.000	:	10.000	1.000	:	0.270
28	1062.6690673828125	140.000	:	39.000	1.000	:	0.279
29	2654.858154296875	123.000	:	27.000	1.000	:	0.220
30	1062.6611328125	90.000	:	27.000	1.000	:	0.300
31	1062.6689453125	107.000	:	35.000	1.000	:	0.327
32	1062.7115478515625	74.000	:	24.000	1.000	:	0.324
33	1595.0172119140625	271.000	:	87.000	1.000	:	0.321
34	1595.046875	25.000	:	8.000	1.000	:	0.320
35	1063.7054443359375	96.000	:	28.000	1.000	:	0.292
36	1066.673583984375	187.000	:	59.000	1.000	:	0.316

- process multiple MS/MS spectra at once
- show all results in one table
- normalize intensities

Kai A. Reidegeld

Conclusions

Peakardt.FindPairs and FindMasses
software for quantition of stable isotope-coded mass spectra

- Reliable
- Accurate
- Customizable
- Future-proof
- Easy to use
- High Performance

Kai A. Reidegeld

Acknowledgements

Prof. Dr.
Helmut E. Meyer

Junior-Prof. Dr.
Bettina Warscheid

Dipl.-Biol.
Romano Hebeler
Poster Number: 585

Sebastian Wiese
Poster Number: 005

Dipl. Chem.
Gudrun Franke

Bouchta Lakhal

Thank you for your attention!

Interesting places to visit:
www.medizinisches-proteom-center.de
www.hbpp.org
www.arbeitstagung.de