

New Protein Level Quantitation Features in Protein Prospector

Peter R. Baker¹, Robert J. Chalkley¹, Jonathan C. Trinidad and Alma L. Burlingame¹

¹ Mass Spectrometry Facility, Dept. of Pharmaceutical Chemistry, University of California, San Francisco, USA

Protein Quantitation

Protein Prospector supports most isotopic labelling strategies:

- SILAC
- ICAT
- iTRAQ (4 and 8-plex)
- O16/O18
- BEMAD

Raw Type Quantitation
Intensity Threshold Resolution L/H Int Area Threshold L/H Area
SNR Threshold Noise Mean Noise SD RT Interval (secs)
¹³C% ¹⁵N% ¹⁸O% iTRAQ Purity Corr
Median IQR Mean Std Dev Num

Protein Prospector Quantitation Options

Protein Report (SILAC L and R)

project name

Rank	Acc #	CAF_B14_C6toC11										Protein MW	Species	Protein Name	
		Num Unique	% Cov	Best Disc Score	Best Expect Val	Med L/H A	Q1	Q3	Mn L/H A	-2.0σ	+2.0σ				Num
1	P02787	67	67.3	6.41	2.8e-11	14.2	8.12	25.9	14.6	3.00	71.4	46	77050.6	HUMAN	Serotransferrin precursor (Transferrin) (Siderophilin) (Beta-1-metal-binding globulin)
2	P13796	44	62.1	5.85	3.2e-10	2.32	2.00	2.50	2.17	1.04	4.51	39	70158.8	HUMAN	Plastin-2 (L-plastin) (Lymphocyte cytosolic protein 1) (LCP-1) (LC64P)
3	P63103	36	69.0	5.80	3.8e-10	1.55	1.39	1.66	1.42	0.725	2.78	28	27745.3	BOVIN	14-3-3 protein zeta/delta (Protein kinase C inhibitor protein 1) (KCIP-1) (Factor activating exoenzyme S) (FAS)
4	P62157	36	93.9	6.86	4.4e-12	1.74	1.43	2.28	1.87	0.855	4.10	29	16706.5	BOVIN	Calmodulin (CaM)
5	P12763	22	37.6	7.00	2.3e-12	47.1	23.6	100	41.3	6.47	264	16	38419.1	BOVIN	Alpha-2-HS-glycoprotein precursor (Fetuin-A) (Asialofetuin)
6	P11142	17	34.7	4.88	1.9e-8	2.02	1.39	2.59	1.71	0.349	8.44	14	70898.7	HUMAN	Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8)
7	P18669	16	64.0	5.70	6.1e-10	1.83	1.47	2.19	1.81	1.000	3.29	13	28673.0	HUMAN	Phosphoglycerate mutase 1 (EC 5.4.2.1) (EC 5.4.2.4) (EC 3.1.3.13) (Phosphoglycerate mutase isozyme B) (PGAM-B) (BPG-dependent PGAM 1)
8	P30101	15	31.1	5.82	3.5e-10	1.73	1.41	2.09	1.61	0.777	3.33	11	56782.9	HUMAN	Protein disulfide-isomerase A3 precursor (EC 5.3.4.1) (Disulfide isomerase ER-60) (ERp60) (58 kDa microsomal protein) (p58) (ERp57) (58 kDa glucose-regulated protein)

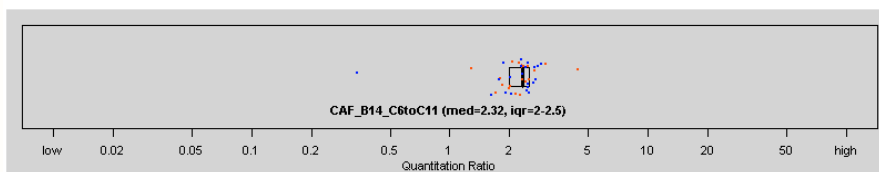
Protein level quantitation statistics

Link to single protein peptide report

Peptide Report (SILAC L and R)

Num Unique	% Cov	Best Disc Score	Best Expect Val	Med L/H A	Q1	Q3	Mn L/H A	-2.0σ	+2.0σ	Num
44	62.1	5.85	3.2e-10	2.32	2.00	2.50	2.17	1.04	4.51	39

Protein level quantitation statistics



Scatterplot of quantitation ratios

<input type="checkbox"/>	m/z	z	ppm	Peptide	Fraction	S	Area Pk 1	Area Pk 2	L/H Area	Score	Expect	# in DB
<input type="checkbox"/>	924.0940	3	20	EGICAIGGTSEQSSVGTQHSYSEEEK	L6101204	23.73				65.2	3.2e-10	2
<input type="checkbox"/>	764.8853	2	-5.3	SLVGIGGQDLNEGNR	L6101204	25.53	101.9	50.68	2.48	53.8	2.4e-9	1
<input type="checkbox"/>	477.9255	3	-0.88	AYYHLEQVAPK	L6101204	27.05	45.89	27.64	1.91	41.5	6.9e-9	2
<input type="checkbox"/>	838.4271	2	4.2	FSLVGIGGQDLNEGNR	L6101204	29.41	31.36	15.42	2.50	44.3	1.1e-8	1
<input type="checkbox"/>	789.9125	4	7.8	AYYHLEQVAPKGFEEGVPVAVIDMSGLR	L6101204	33.3	14.11	10.86	1.72	48.9	1.5e-8	1
<input type="checkbox"/>	751.8756	2	-5.3	MINLSVPTIDER	L6101204	29	169.3	80.48	2.41	42.2	4.3e-8	5
<input type="checkbox"/>	884.7813	3	-6.3	HVIPMNPNTNDFNAVGDGIVLCK HVIPMNPNTNDF(SILAC NC L)FNAVGDGIVL(SILAC NC L)CK	L6101204	32.32	76.55	32.97	2.67	42.5	4.4e-8	1
<input type="checkbox"/>	1081.0146	2	6.3	VDTDGNGYISFNLNDFK	L6101204	35.17	55.20	27.43	2.32	41.1	5.5e-8	1
<input type="checkbox"/>	686.3527	2	-9.6	INLSVPTIDER	L6101204	27.68	49.71	22.80	2.50	42.3	8.3e-8	5
<input type="checkbox"/>	758.3849	2	-18	MINLSVPTIDER MINL(SILAC NC L)SVPDTIDER(SILAC R)	L6101204	29.14	211.4	98.52	2.46	40.5	1.0e-7	5
<input type="checkbox"/>	711.6855	3	8.4	VREITENLMATGDLDODGR	L6101204	26.99	23.79	15.49	2.02	50.7	2.0e-7	1
<input type="checkbox"/>	793.4329	2	7.1	VYALPEDLVEVNPK	L6101204	29.62	33.15	14.36	2.66	37.0	2.4e-7	2

Link to raw data display

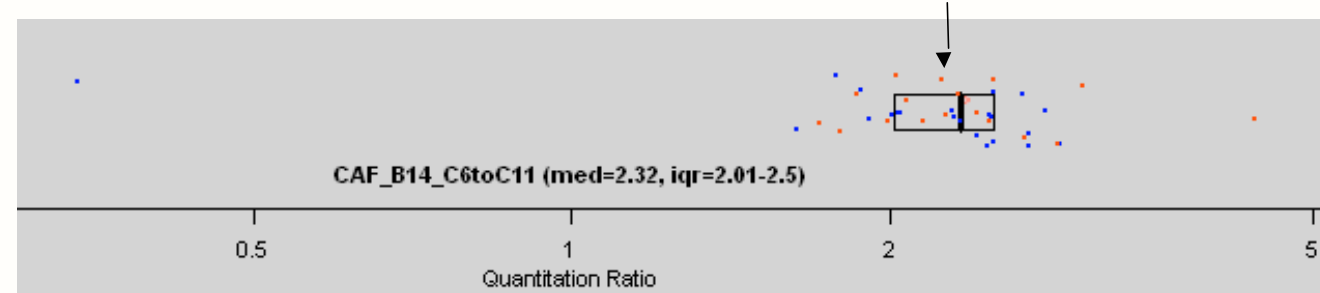
Checkbox to eliminate outliers

Number of times the peptide occurs in the database

Details of Quantitation Ratio Scatterplot

In quantitation experiments the ratios measured for a given protein often vary widely for a number of reasons. Scatterplots are one way of showing how reliable the protein level quantitation ratio is.

Boxplot indicating median and interquartile range

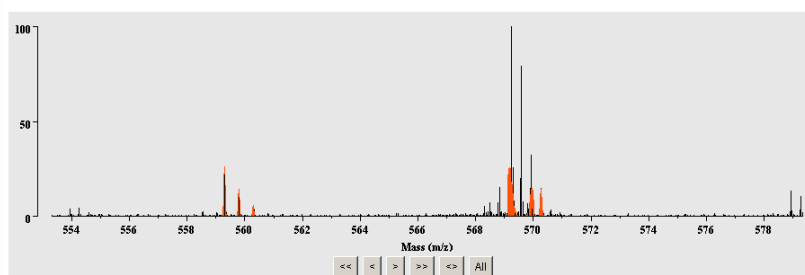


Red: Peptide is only found in this database accession
Blue: Peptide is found in multiple database accessions
<>: 1 measurement peak is below a defined SNR

Elimination of Outliers

Access to the raw data allows outliers to be investigated. They can then be eliminated via the checkboxes on the peptide report.

MS-Display



← Outlier probably caused by coeluting peptides

m/z Pk 1	m/z Pk 2	Noise Mean	Noise St Dev
559.2913	569.1757	8.16	3.64
559.7918	569.9488	8.16	3.64
560.2924	570.2653	8.16	3.64

Int Pk 1	Int Pk 2	SNR Pk 1	SNR Pk 2	Resolution Pk 1	Resolution Pk 2	FWHM Pk 1	FWHM Pk 2	L/H Intrinsic	L/H Intensity	Area Pk 1	Area Pk 2	L/H Intrinsic	L/H Area
554.0	531.8	152.1	146.0	9558	2587	0.05851	0.2200	0.8124	1.28	34.57	124.8	0.8124	0.341
301.8	302.8	82.86	83.12	10392	3943	0.05387	0.1446	1.210	0.824	17.34	46.69	1.210	0.307
117.8	308.6	32.35	84.73	9837	7755	0.05696	0.07354	1.370	0.279	7.160	24.21	1.370	0.216

Elemental Composition Pk 1: [C50 H81 N14 O13 S1](#)

Elemental Composition Pk 2: [C32 H81 N12 O13 S1 13C18 15N2](#)

← Link to MS-Isotope for the 2 quantitation peaks

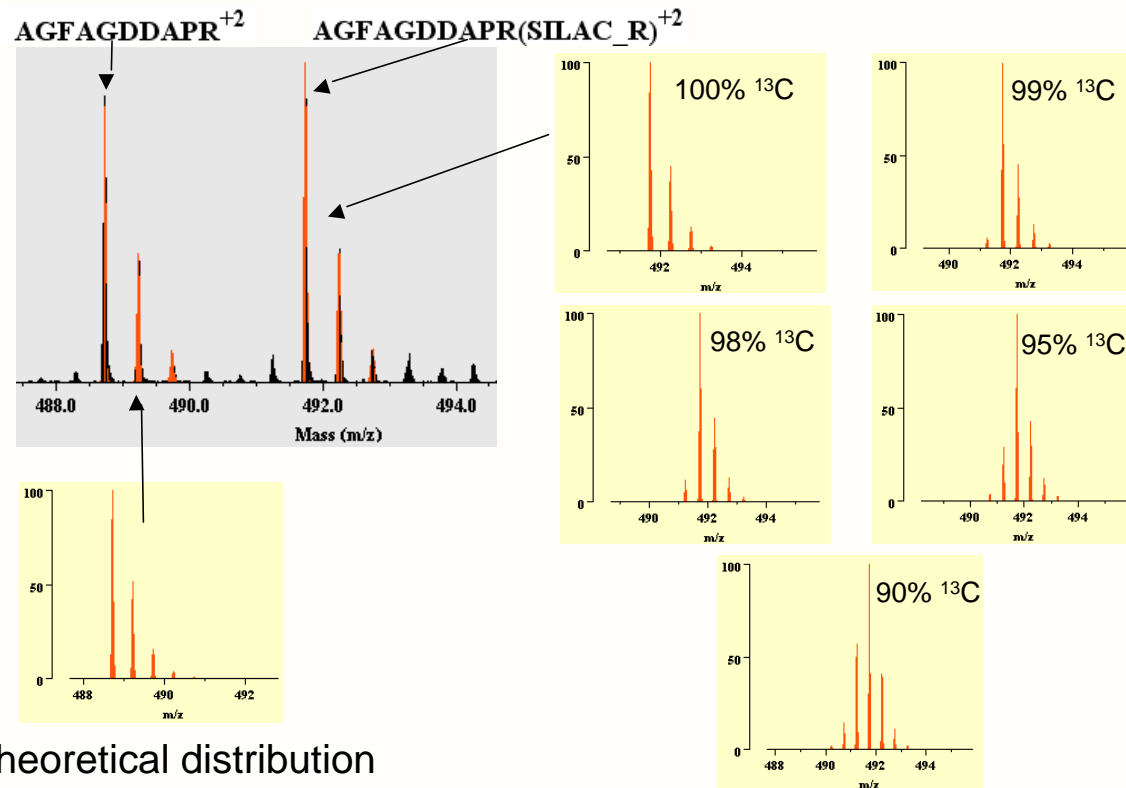
Raw Type: Quantitation RT: 26.29 ¹³C%: 100 ¹⁵N%: 100 ¹⁸O%: 100

↑ Option to account for incomplete isotopic labelling

Menu allows display of MS Full Scan or MSMS raw data

Incomplete Isotopic Labelling

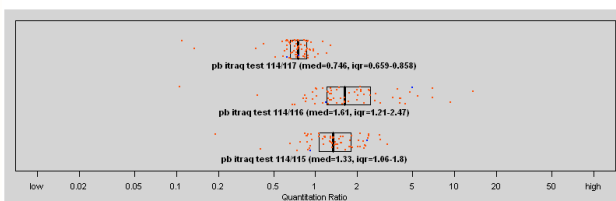
There are options on both the Search Compare form and the MS-Display report to correct for incomplete isotopic labelling.



Theoretical distribution

iTRAQ Quantitation

Num Unique	% Cov	Best Disc Score	Best Expect Val	Med 114/115 A	Q1	Q3	Mn 114/115 A	-2.0σ	+2.0σ	Num	Med 114/116 A	Q1	Q3	Mn 114/116 A	-2.0σ	+2.0σ	Num	Med 114/117 A	Q1	Q3	Mn 114/117 A	-2.0σ	+2.0σ	Num
67	47.2	6.78	5.9e-12	1.33	1.06	1.34	0.530	0.39	0.67	67	1.61	1.21	2.47	1.76	0.390	0.797	67	0.746	0.659	0.858	0.716	0.337	1.52	67

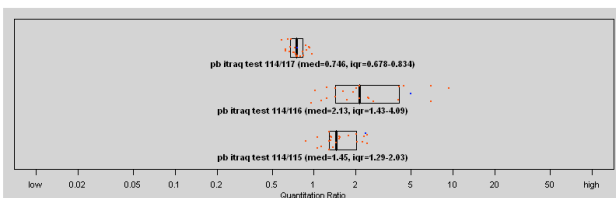


Typical iTRAQ quantitation report

m/z	#	ppm	Peptide	Fraction	S	Area 114	Area 115	Area 116	Area 117	114/115 Area	114/116 Area	114/117 Area	Score	Expect	# in DB
769.4000	5	31	ITRAQ-HFTEQWYPTLPTGSGSGMGSGGGGGGGGK	206052106	59.9	530.5	259.9	76.24	573.5	2.04	6.96	0.925	66.6	5.9e-12	1
719.8469	4	-6.2	LSQSGSSSITAAGMR	206051904	35.81	0.9677	5.051	9.123	7.134	0.192	0.106	0.136	49.2	1.8e-11	1
602.3320	4	4.8	ITRAQ-EGSISGSGSGGGGGGGLKPSITK	206052011	20.85	1275	995.3	514.4	1734	1.28	2.48	0.735	53.5	1.7e-8	1

Unlabelled peptide

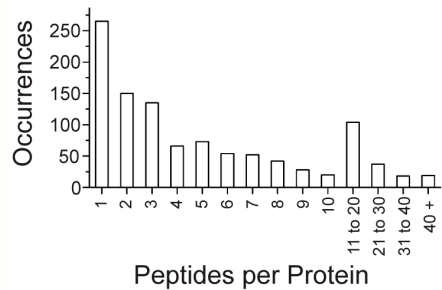
Num Unique	% Cov	Best Disc Score	Best Expect Val	Med 114/115 A	Q1	Q3	Mn 114/115 A	-2.0σ	+2.0σ	Num	Med 114/116 A	Q1	Q3	Mn 114/116 A	-2.0σ	+2.0σ	Num	Med 114/117 A	Q1	Q3	Mn 114/117 A	-2.0σ	+2.0σ	Num
23	22.1	6.78	5.9e-12	1.45	1.29	2.03	1.54	0.849	2.78	23	2.13	1.43	4.09	2.38	0.637	0.89	23	0.746	0.678	0.834	0.753	0.572	0.991	23



Same report eliminating peptides where the 114 peak intensity is less than 300

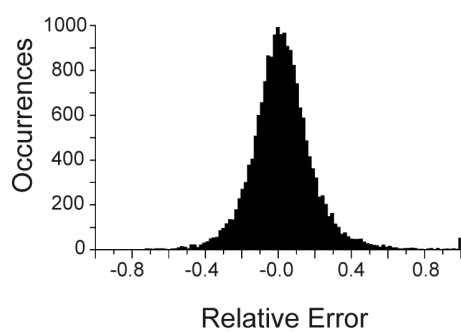
m/z	#	ppm	Peptide	Fraction	S	Area 114	Area 115	Area 116	Area 117	114/115 Area	114/116 Area	114/117 Area	Score	Expect	# in DB
769.4000	5	31	ITRAQ-HFTEQWYPTLPTGSGSGMGSGGGGGGGGK	206052106	59.9	530.5	259.9	76.24	573.5	2.04	6.96	0.925	66.6	5.9e-12	1
602.3320	4	4.8	ITRAQ-EGSISGSGSGGGGGGGLKPSITK	206052011	20.85	1275	995.3	514.4	1734	1.28	2.48	0.735	53.5	1.7e-8	1
534.6310	5	-8.5	ITRAQ-DLFYYSRPLAR	206051917	59.89	329.4	248.3	162.7	527.0	1.33	2.02	0.625	43.1	4.3e-8	1

Example Results 1



The peaklists from an iTRAQ project containing over 150 individual LC-MS/MS runs were combined into a single project file and searched. This resulted in the identification and quantification of over 10,000 unique peptides mapping to 1200 unique proteins

Example Results 2



To investigate the quantification precision for this dataset, protein expression ratios were calculated using the individual peptide ratios. For those proteins quantified using 10 or more peptides, we examined the distribution of peptide ratios relative to the calculated protein ratios. The standard deviation of relative error for peptide measurements compared to their respective protein values was 0.19

Conclusions

- Support for all major isotopic labelling strategies
- Approach has been tested on very large multi-fraction datasets
- Quantitation information available at the protein and peptide label
- Scatterplot allows easy assessment of protein level quantitation
- Easy access to raw data to detect and eliminate outliers
- Facilities for investigating and correcting for incomplete isotopic labelling
- Archive upload feature allows offsite data to be processed
- Tab delimited output allows data to be exported to external program

Acknowledgements

Funding: NIH NCRR grant 001614, 0012961, 0019934 and the Vincent Coates Foundation.

References

- 1) Lynn, A. J., Baker, P. R., Chang Y. and Chalkley, R. J. Protein Prospector as a Large Scale MS Proteomics Data Management Tool. 53rd ASMS Conference, San Antonio, Texas, USA (2005)