TMTpro: Design, synthesis and initial evaluation of a Proline-based isobaric 16-plex Tandem Mass Tag reagent set

Andrew Thompson[1*], Nikolai Wölmer[2], Sasa Koncarevic[2], Stefan Selzer[2], Gitte Böhm[2], Harald Legner[2], Peter Schmid[2], Stefan Kienle[2], Petra Penning[2], Claudia Höhle[2], Antje Berfelde[2], Roxana Martinez-Pinna[2], Vadim Farztdinov[2], Stephan Jung[2], Karsten Kuhn[2], Ian Pike[1]

[1]Proteome Sciences Plc, Hamilton House, Mabledon Place, London, WC1H 9BB, United Kingdom

[2]Proteome Sciences R&D GmbH & Co. KG, Altenhöferallee 3, 60438 Frankfurt am Main, Germany

Supplementary Figures and Information Contents:

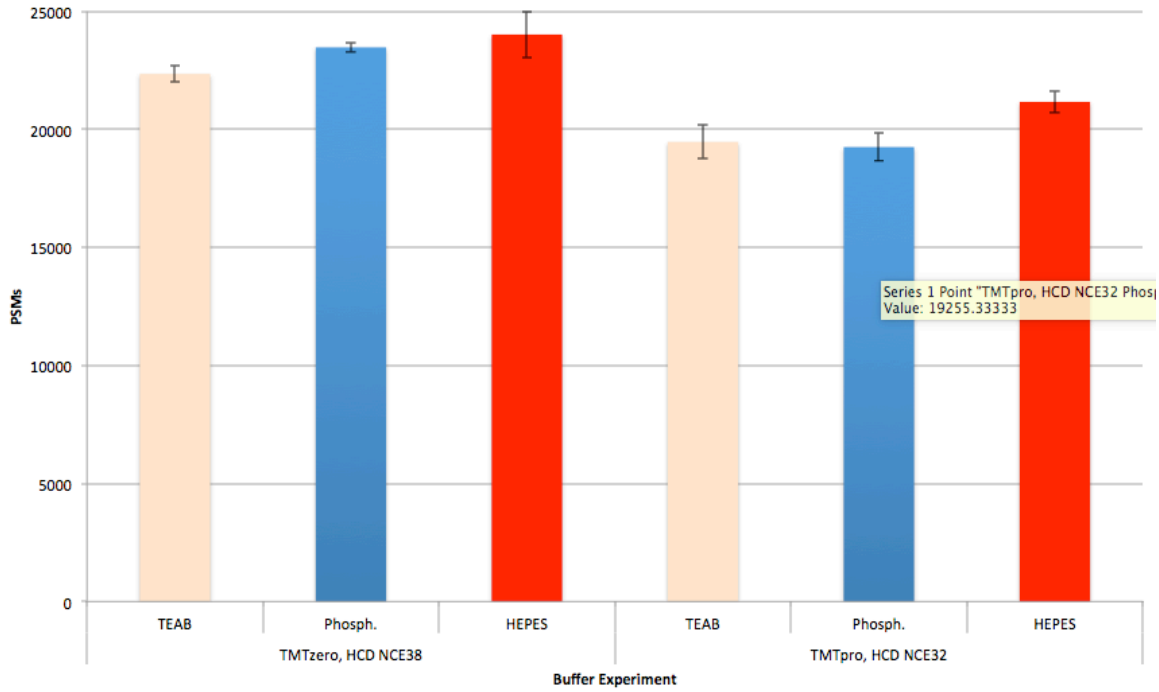
Optimization of collision energies for TMTpro analysis

We carried out some initial experiments on relative ID rates using the undoped tag structures to optimize the mass spectrometry conditions for the analysis of the TMTpro reagents. A first observation about the new tags is the larger size and isobutyl side-chain results in a consistent later elution of all labeled peptides (1 to 2 minutes) but this has no significant impact on overall ID rates. Initial experiments suggested a Normalized Collision Energy (NCE) of 32 (data not shown) for the TMTpro structure in contrast to an optimal NCE around 38 for TMT. Following this initial experiment we also compared buffers for labeling. The standard protocol in the TMT10/11 kit uses triethyl ammonium carbonate (TEAB) buffer for labeling of peptides with TMT N-hydroxysuccinimide (NHS) esters and these are conditions that we frequently use for TMT-labeling [47]. However, various groups have also used different buffers including HEPES [48] and basic Phosphate buffers [12]. We found that HEPES performed slightly better than TEAB or Phosphate buffer for TMTpro amongst the basic buffers and likewise in parallel experiments with TMT10/11, there is a modest benefit from using HEPES (Supplementary Figure 1A). In addition we have previously shown that low pH labelling in Phosphate buffer at pH 4.5 with TMT10/11 reagents can be quite beneficial [34] and so we compared the labeling of 5 replicates of a rat liver sample using the best high pH conditions with HEPES buffer and TEAB buffer at pH 8.0 for TMT10/11 with HEPES buffer at pH 8.0 and low pH Phosphate buffer labeling at pH 4.5 for TMTpro (Figure S1B). In our hands low pH labeling appears to perform the best for TMTpro. We also did an analysis of over-labelling rates for TMTpro and TMT10/11. It can be seen in Supplementary Figure 2 that TMTpro seems to suffer slightly less from over-labelling than TMT10/11. In particular, low pH labelling generated very low levels of over-labeling with TMTpro consistent with our previous report on low pH solid phase labeling with TMT10/11 [34]. We speculate that the lower over-labelling seen with TMTpro may be due to slightly lower reactivity of the larger tag resulting from steric hindrance of the larger structure but we have no compelling data to support this speculation at this stage.

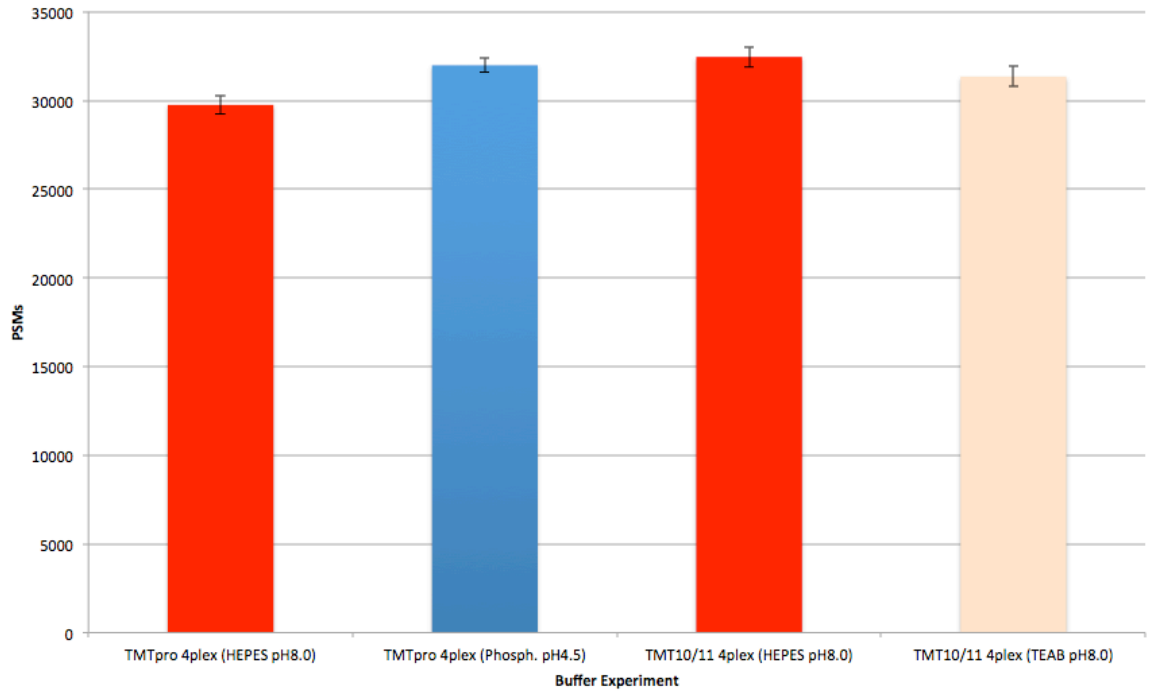
Upon completion of a set of 11 TMTpro tags for comparison with the 11 TMT10/11 reagents that are available, we re-checked the NCE conditions in a multiplexed setting. We carried out an initial NCE scouting experiment around NCE32 (Supplementary Figure 3), which suggested that NCE30 might be slightly better than our initial NCE32 condition. However, NCE30 was the lowest value tested in this experiment so we repeated the experiment with some lower NCE

values (Supplementary Figure 4A) to assess whether an even lower NCE value might be optimal. In this experiment an NCE of 29 appeared to be optimal. However, since these were n=1 experiments and the differences in ID rates were in the range of run-to-run variability we carried out a further repeat of the experiment with the NCE values 29, 30 and 32 in triplicate, shown in Supplementary Figure 4B. It can be seen that overall, the peptide ID rates do not vary much between NCE29 and NCE32. It does appear that there is some tolerance around the optimal collision energy but this may vary from instrument to instrument slightly and we urge end-users to check NCE settings for their particular instrument setup.

1A

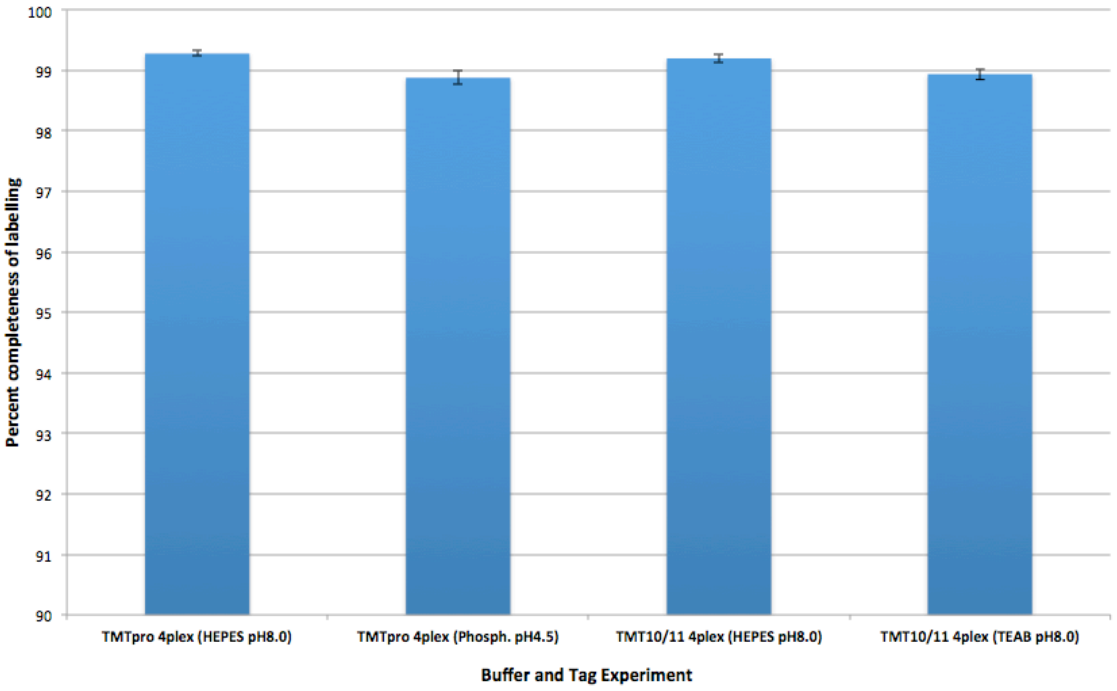


1B

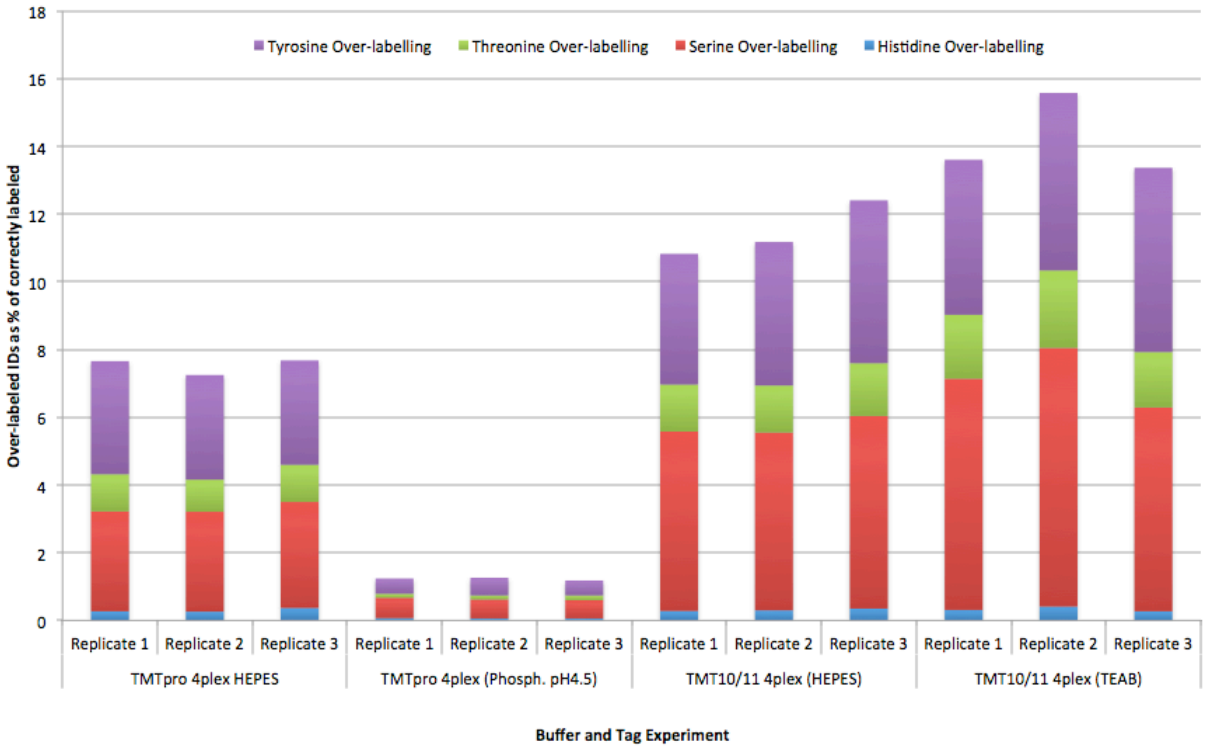


Supplementary Figure 1 - 1A: pH 8.0 buffer optimization experiment for TMTpro and TMT10/11. 1B: comparison of low pH labeling at pH 4.5 with pH 8.0 for TMTpro and TMT10/11.

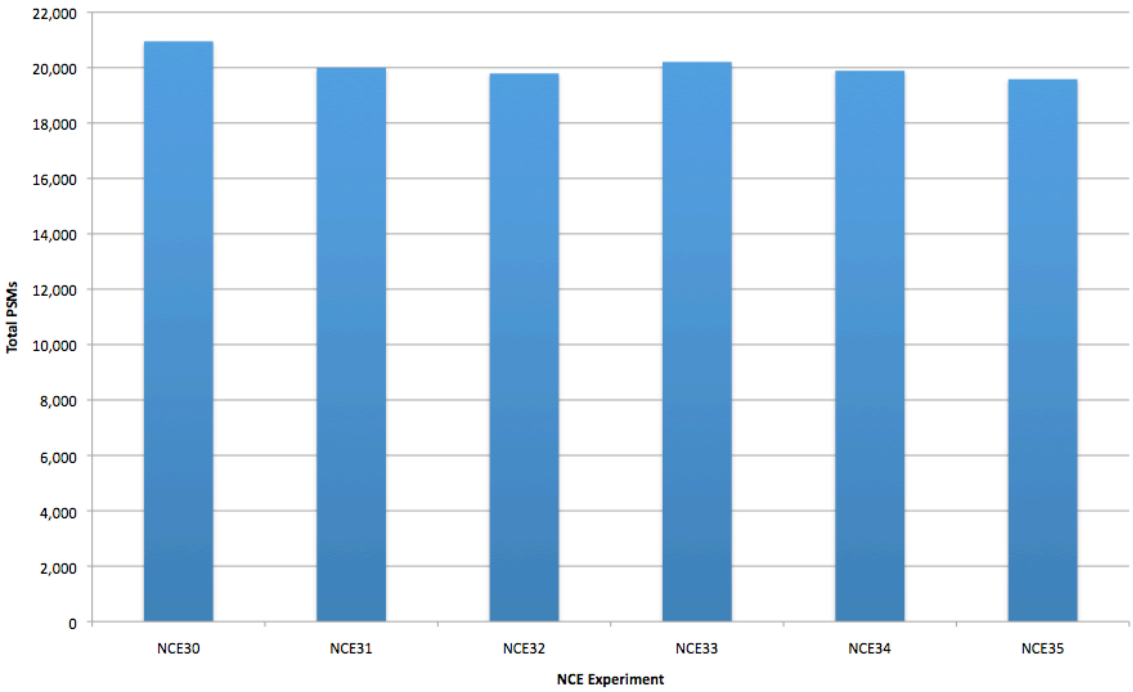
A



B

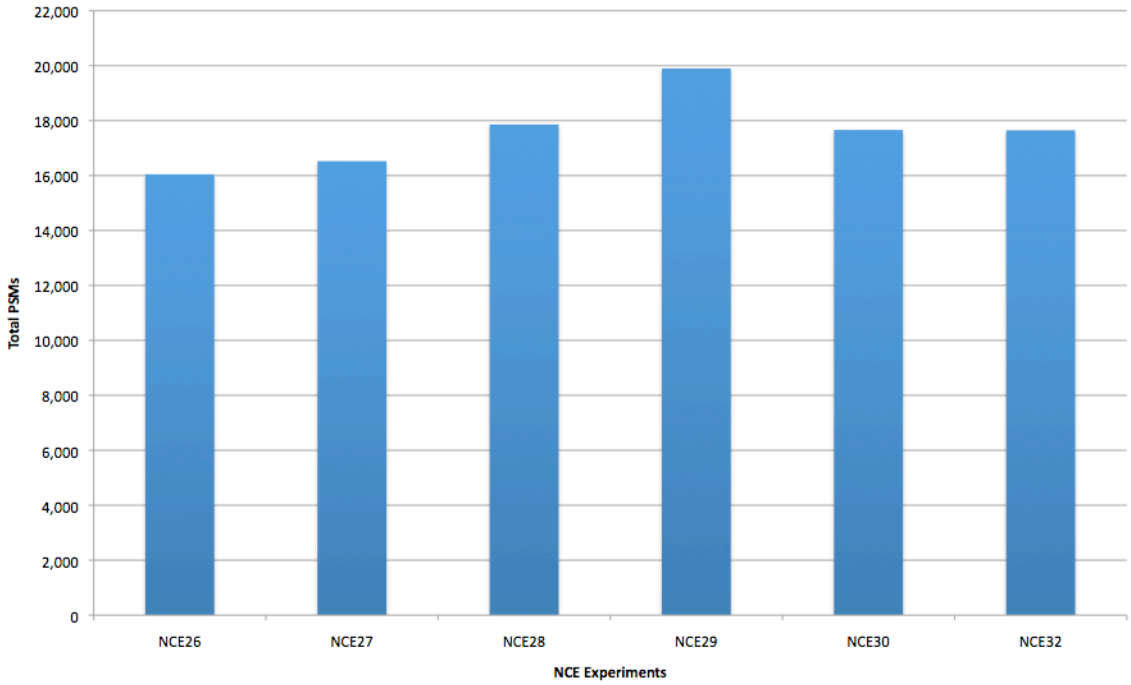


Supplementary Figure 2: 2A: Chart of labeling completeness determined by comparison of numbers of completely labeled peptides with unlabeled. 2B: Assessment of over-labelling with TMTpro.

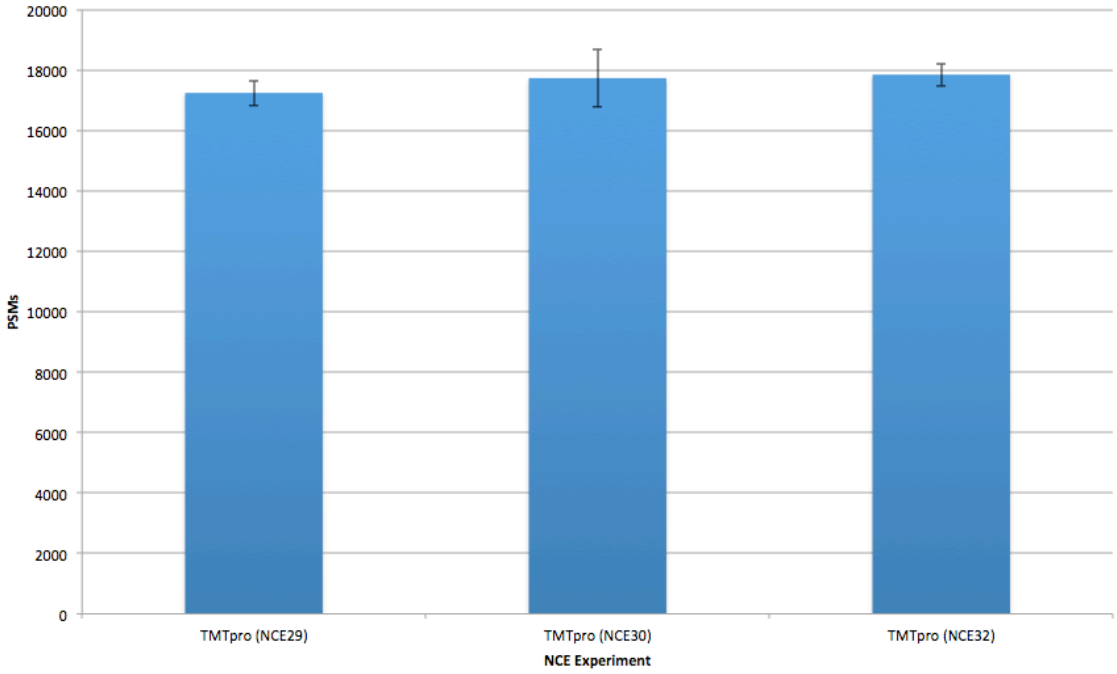


Supplementary Figure 3: first NCE optimization experiment with TMTpro 11plex.

A

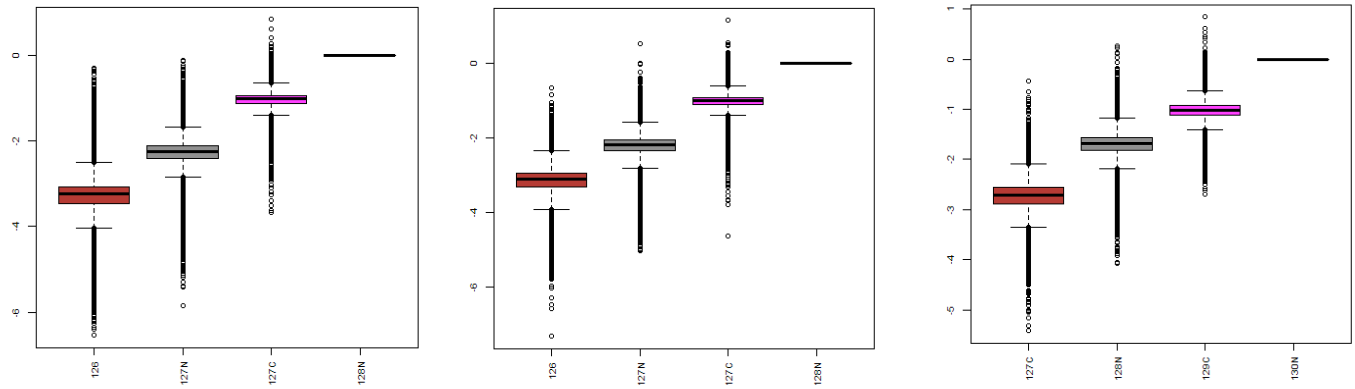


B



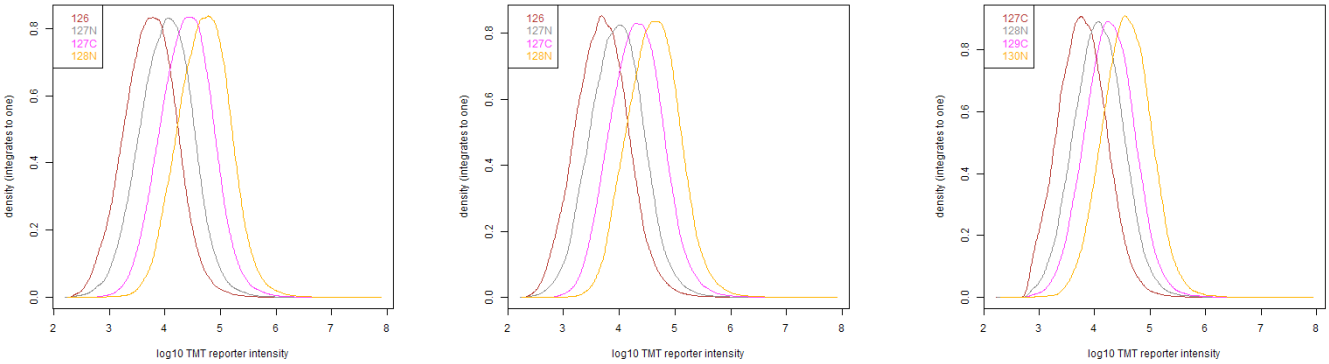
Supplementary Figure 4A: second NCE optimization experiment with TMTpro 11plex. Supplementary Figure 4B: Final NCE optimization experiment. Replicate measurements at each NCE show some variation but overall it can be seen that there is little difference in behavior at different collision energies.

TMT10/11 TEAB **TMT10/11 HEPES** **TMTpro HEPES (8 frac)**



Supplementary Figure 5: Box plots of the phosphopeptide dilution series. TMTpro shows only the 8 fraction data (6 fraction data is essentially the same). Overall the TMTpro data is quantitatively comparable to TM10/11.

TMT10/11 TEAB **TMT10/11 HEPES** **TMTpro HEPES (8 frac)**



Supplementary Figure 6: Reporter intensity distributions for 4-plex phosphopeptide experiments.

TMT10/11 (TEAB Labelling)		TMT10/11 (HEPES Labelling)		TMTpro (HEPES Labelling)	
TMT10-126	300µg	TMT10-126	300µg	TMTpro–127C	300µg
TMT10-127N	600µg	TMT10-127N	600µg	TMTpro-128N	600µg
TMT10-127C	1200µg	TMT10-127C	1200µg	TMTpro-129C	1200µg
TMT10-128N	2400µg	TMT10-128N	2400µg	TMTpro-130N	2400µg

Supplementary Table 1: Labeling plan for the 4-plex phosphopeptide enrichment experiments comparing TMTpro labeling in HEPES with TMT labeling in HEPES and TEAB

Sample	HeLa Amount	Yeast Amount	Ratio Y:H
S1	100µg	2.0µg	1 : 50
S2	100µg	4.5µg	1 : 22
S3	100µg	6.0µg	1 : 16.7
S4	100µg	10.4µg	1 : 9.6
S5	100µg	18.0µg	1 : 5.6

Supplementary Table 2: HeLa to Yeast ratios for double proteome experiments

Reporter	TMT11plex 1	TMT11plex 2	TMT11plex 3
TMT10-126	S1A1	Ref	S3A5
TMT10-127N	S2A1	S5A3	S5A5
TMT10-127C	S3A1	S4A3	S1A5
TMT10-128N	S4A1	S3A3	S4A5
TMT10-128C	S5A1	S2A3	S2A5
TMT10-129N	Ref	S1A3	S2A6
TMT10-129C	S1A2	S5A4	S1A6
TMT10-130N	S2A2	S4A4	S5A6
TMT10-130C	S3A2	S3A4	S3A6
TMT10-131	S4A2	S2A4	S4A6
TMT11-131C	S5A2	S1A4	Ref

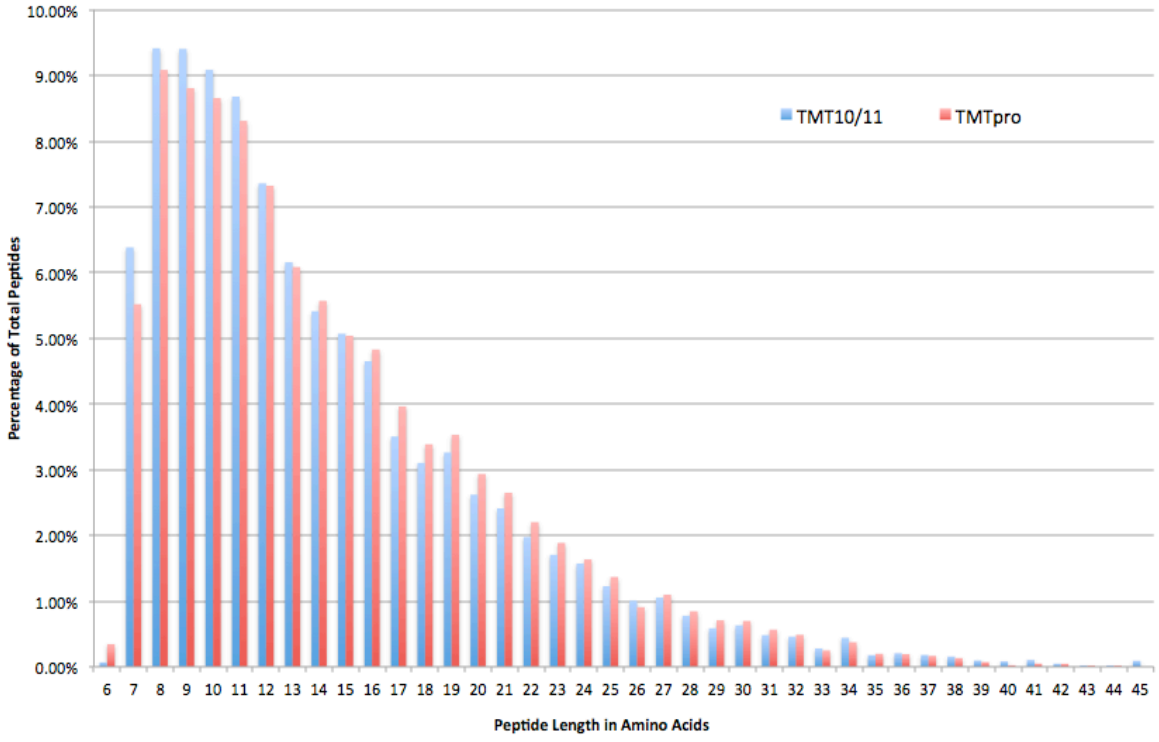
Supplementary Table 3: TMT10/11 30-plex Labelling plan

Reporter	TMTpro 16plex 1	TMTpro 16plex 2
TMTpro-126	S1A7	S5A10
TMTpro-127N	S2A7	S4A10
TMTpro-127C	S3A7	S3A10
TMTpro-128N	S4A7	S2A10
TMTpro-128C	S5A7	S1A10
TMTpro-129N	Ref	S5A11
TMTpro-129C	S1A8	S4A11
TMTpro-130N	S2A8	S3A11
TMTpro-130C	S3A8	S2A11
TMTpro-131N	S4A8	S1A11
TMTpro-131C	S5A8	Ref
TMTpro-132N	S1A9	S5A12
TMTpro-132C	S2A9	S4A12
TMTpro-133N	S3A9	S3A12
TMTpro-133C	S4A9	S2A12
TMTpro-134N	S5A9	S1A12

Supplementary Table 4: TMTpro 30-plex Labelling plan

	Workflow	Phosphopeptide		MS2		MS3	
	Label	TMT10/11	TMTpro	TMT10/11	TMTpro	TMT10/11	TMTpro
Full Scan (MS1)	Detector type	Orbitrap	Orbitrap	Orbitrap	Orbitrap	Orbitrap	Orbitrap
	Orbi resolution	120k	120k	120k	120k	120k	120k
	Max IT (ms)	100	100	100	100	50	100
	AGC Target	2.0E+05	2.0E+05	5.0E+04	2.0E+05	4.0E+05	2.0E+05
	Scan range (m/z)	400-1400	400-1400	400-1400	400-2000	350-1500	400-2000
MS2	Isolation window	1.2	1.2	1.2	1.2	1.2	1.2
	Activation type	HCD	HCD	HCD	HCD	CID	CID
	NCE (%)	38	32	38	32	35	35
	Detector Type	Orbitrap	Orbitrap	Orbitrap	Orbitrap	Ion trap	Ion trap
	Orbi resolution	60k	60k	30k	30k	Rapid	Rapid
	Max IT (ms)	100	100	100	100	50	50
	AGC Target	1.00E+05	1.00E+05	1.00E+05	1.00E+04	1.40E+04	1.40E+04
MS3	Isolation window	NA	NA	NA	NA	2	2
	Activation type	NA	NA	NA	NA	HCD	HCD
	NCE (%)	NA	NA	NA	NA	65	65
	Detector Type	NA	NA	NA	NA	Orbitrap	Orbitrap
	Orbi resolution	NA	NA	NA	NA	50k	50k
	Max IT (ms)	NA	NA	NA	NA	105	105
	AGC Target	NA	NA	NA	NA	1.00E+05	1.00E+05

Supplementary Table 5: MS Settings in each MS workflow



Supplementary Figure 7: Peptide Length Distribution for peptides identified in the 3x11plex and 2x16plex experiments with TMT10/11 and TMTpro respectively

Data Assembly and Pre-processing in detail:

Data integration, pre-processing and statistical testing were conducted with scripts written in house but using standard analysis tools. A standard workflow was intentionally used in order to guarantee that the results obtained could be reproduced in any other laboratory if desired.

The following standard data integration methods were applied:

> Filtering of PSMs using an Isolation interference filter with a threshold of 50%. Isotope impurity correction. Median-scaling normalization of TMTpro/TMT reporter ion intensities. After normalisation, PSM level ratios of reporter ion intensities were calculated for experimental samples relative to the reference sample and log2-transformed. The PSM level log2 ratios were summarised to peptide level log2 ratios by taking median of PSM ratios specific to a given peptide sequence ID. Peptide log2Ratio data matrix was used further for pre-processing.

Next, the following data pre-processing methods were applied:

> The peptide logRatios were additionally normalized using weighted mean-scaling technique. Only peptides from human proteins were used for scaling and the weight was proportional to square of feature's presence in all samples. Missing values were analysed. Peptides with more than 35% missing quantitative values in a particular experimental group were removed from the data set. The remaining missing quantitative ratios were replaced by values imputed using an iterative PCA method [49,50].

Exploratory analysis of the obtained data was carried out in order to reveal the strongest factor driving the variance within the data. It was discovered that "plex" had rather strong effect on the data and therefore the batch correction procedure was applied to the data as described previously [51,52] .

General filtering by the standard deviation (SD) of the log ratios combined with hierarchical clustering revealed that SD can be used as a very effective filter to select yeast peptides (see Supplementary Figures 5 and 6). Selection of the threshold was based on the analysis of SD distribution (see below).

Data Analysis

Linear modeling, using the functionality of the LIMMA package [52], was applied to find fold-changes of regulated peptides. In order to run comparative analysis of labeling approaches ability to evaluate features regulation we decided to apply the same feature selection criteria for both labeling approaches, MS methods (MS2 and MS3) and to all features independent of sample group comparison.

From analysis of the distribution of logratio standard deviations (SD) we discovered that the threshold SD = 0.5 reliably separates regulated peptides which we expect to be yeast peptides (Blue points in Supplementary Figure 7) from non-regulated human peptides. Heatmaps of reduced data matrices after application of the filter SD ≥ 0.5 shows very clean separation of all five sample groups (Supplementary Figure 8) and confirms the ability of both labeling methods to differentiate between sample groups even at very low concentrations (groups S1 and S2).

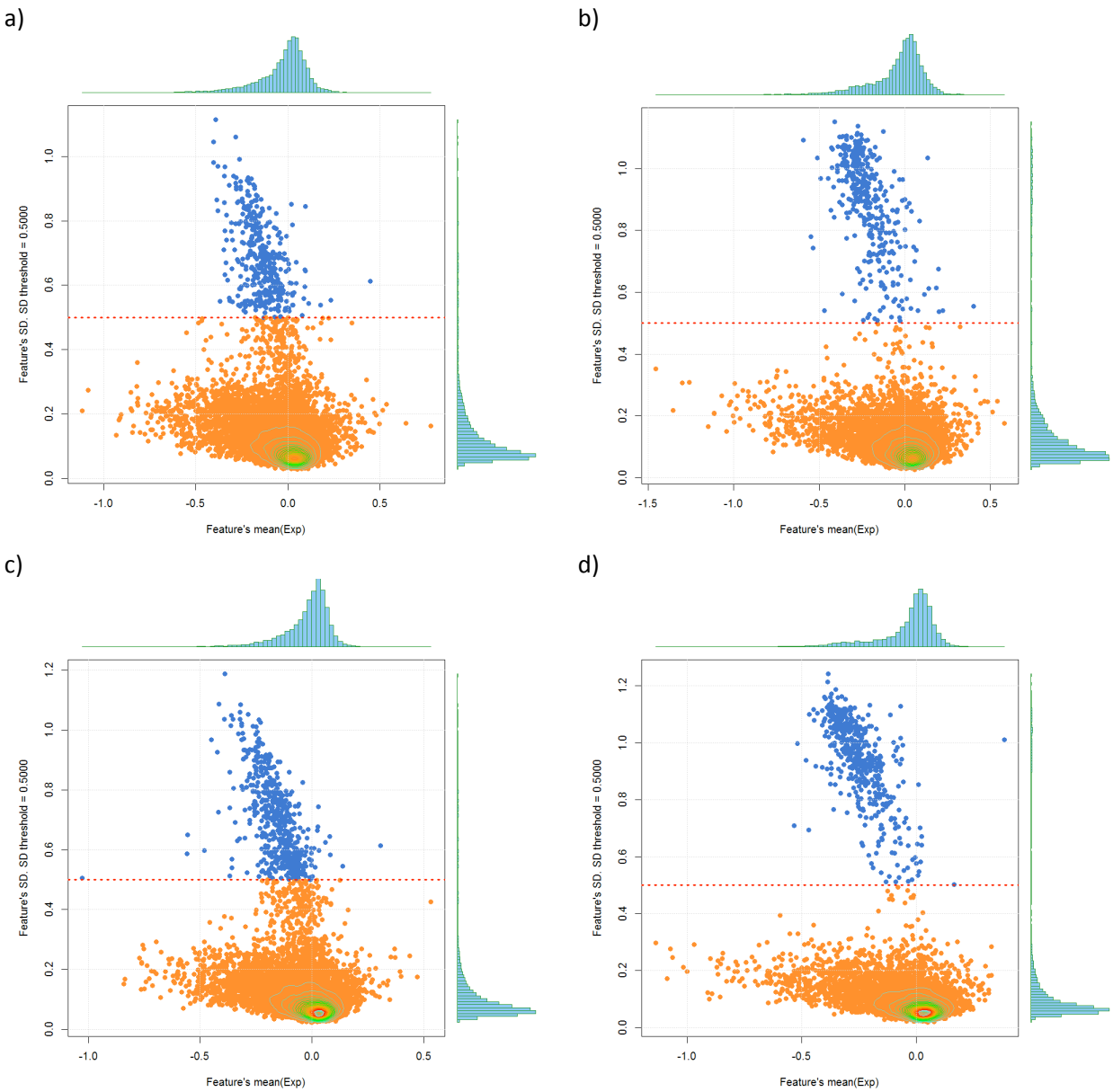
In order to evaluate precision and accuracy of log fold change (logFC) measurements between different sample groups we decided to exclude human peptides and apply more stringent isolation interference ≤ 33% filter.

Accuracy is calculated as relative absolute deviation of measured mean[$(\log_2 FC)_{measured}$] from the given logRatio of concentrations: $\log_2\left(\frac{c_n}{c_k}\right)_{given}$

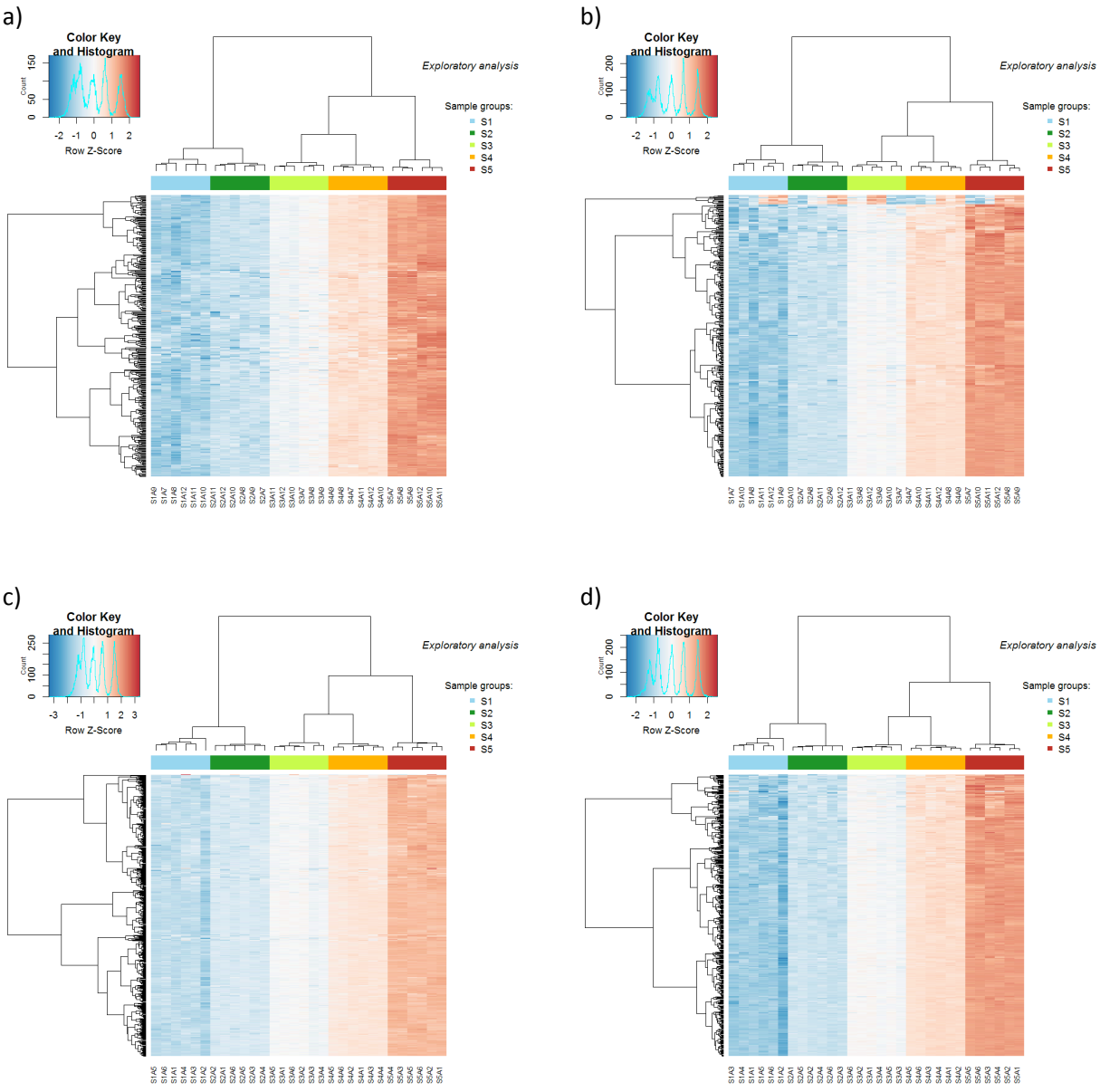
$$Accuracy = \frac{\left| mean\left[\left(\log_2 FC\right)_{measured}\right] - \log_2\left(\frac{c_n}{c_k}\right)_{given} \right|}{\log_2\left(\frac{c_n}{c_k}\right)_{given}} \times 100$$

Mean here is taken over all yeast peptides selected as described above. Precision is defined through CV as

$$CV = \frac{sd\left[\left(\log_2 FC\right)_{measured}\right]}{mean\left[\left(\log_2 FC\right)_{measured}\right]} \times 100$$



Supplementary Figure 8: Scatter plots of peptide Standard Deviations (SD) vs. logRatio in unfractionated MS data: a) TMTpro_MS2, b) TMTpro_MS3, c) TMT_MS2, d) TMT_MS3. We decided to apply one threshold for all measurements to separate features in an unbiased manner to help visualize easily the ability of both tags to detect 'regulated features'. Two main distributions are clearly seen on all plots: one centered around (SD~ 0.1, logRatio ~ 0) and the other much smaller and cigar-like (blue) around (SD~ 07-08, logRatio ~ 0.25). The important observation here is that the data obtained using TMTpro are very similar to the data obtained using TMT10/11 – compare panels a) vs c) and panels b) vs d). Red dotted line shows the threshold SD = 0.5 applied. Further analysis confirms that features below the threshold are mainly human peptides, while features above are mainly yeast peptides (blue). Therefore selection of features by the SD value turns out to be very effective tool for identifying regulated features and its quality can be seen from Supplementary Figure 8.



Supplementary Figure 9: Heatmaps of reduced dataset obtained after application of SD≥0.5 filter to unfractionated MS data: a) TMTpro_MS2, b) TMTpro_MS3, c) TMT_MS2, d) TMT_MS3. All four panels show very clean striped clustering structure which reflects step-like increase of yeast peptide abundance from group S1 to group S5 and high level of homogeneity within each sample group. This can be also seen from the Z-score histogram in the upper left corner of each panel, showing clean five peak structure. Note that after application of SD filter level of false discoveries is very low – there are only small number of features that do not follow five stripes pattern.

		TMTpro MS2				TMT10/11 MS2			
Contrast	log(Cn/Ck)	mean logFC	SD	CV [%]	Accuracy [%]	mean logFC	SD	CV [%]	Accuracy [%]
S5 vs. S1	3.17	1.86	0.36	19	41	1.92	0.40	21	39
S4 vs. S1	2.38	1.24	0.28	23	48	1.28	0.31	24	46
S3 vs. S1	1.58	0.75	0.21	28	52	0.79	0.24	30	50
S2 vs. S1	0.81	0.27	0.13	46	66	0.31	0.12	39	62
S5 vs. S2	2.36	1.59	0.26	16	33	1.62	0.30	18	32
S4 vs. S2	1.57	0.96	0.19	19	39	0.97	0.21	22	38
S3 vs. S2	0.78	0.48	0.12	25	39	0.48	0.14	29	38
S5 vs. S3	1.58	1.11	0.17	16	30	1.13	0.18	16	28
S4 vs. S3	0.79	0.48	0.10	21	39	0.49	0.10	20	38
S5 vs. S4	0.79	0.62	0.10	15	21	0.64	0.10	15	19
mean				23	41			23	39

Supplementary Table 6: Expected log fold change log(Cn/Ck), measured mean log fold change, log fold change standard deviation, precision (CV [%]) and accuracy [%] evaluated by MS2 using TMTpro (left) and TMT10/11 (right) labelled unfractionated data for indicated contrasts (first column). Provided values are the subject of applied feature selection criteria described in detail in Methods section (SD≥0.5 filter, only yeast peptides, isolation interference ≤ 33% filter).

		TMTpro MS3				TMT10/11 MS3			
Contrast	log(Cn/Ck)	mean logFC	SD	CV [%]	Accuracy [%]	mean logFC	SD	CV [%]	Accuracy [%]
S5 vs. S1	3.17	2.46	0.41	17	22	2.60	0.42	16	18
S4 vs. S1	2.38	1.71	0.33	19	28	1.81	0.34	19	24
S3 vs. S1	1.58	1.10	0.25	22	31	1.18	0.26	22	25
S2 vs. S1	0.81	0.44	0.15	34	46	0.49	0.14	29	39
S5 vs. S2	2.36	2.03	0.29	14	14	2.10	0.30	14	11
S4 vs. S2	1.57	1.27	0.21	17	19	1.32	0.22	17	16
S3 vs. S2	0.78	0.66	0.13	20	15	0.69	0.14	21	12
S5 vs. S3	1.58	1.37	0.19	14	13	1.42	0.17	12	10
S4 vs. S3	0.79	0.61	0.11	18	22	0.63	0.09	15	20
S5 vs. S4	0.79	0.75	0.10	13	5	0.79	0.09	12	0
mean				19	22			18	18

Supplementary Table 7: Expected log fold change log(Cn/Ck), measured mean log fold change, log fold change standard deviation, precision (CV [%]) and accuracy [%] evaluated by MS3 using TMTpro (left) tand TMT10/11 (right) labelled unfractionated data for indicated contrasts (first column). Provided values are the subject of applied feature selection criteria described in detail in Methods section (SD≥0.5 filter, only yeast peptides, isolation interference ≤ 33% filter).

Tag	126	127N	127C	128N	128C	129N	129C	130N	130C	131	131C
126	0.90744102	0.00362976	0.08439201	0.00000000	0.00453721	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
127N	0.00540054	0.90009001	0.00000000	0.08910891	0.00000000	0.00540054	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
127C	0.00637233	0.00000000	0.91033227	0.00364133	0.07464725	0.00000000	0.00500683	0.00000000	0.00000000	0.00000000	0.00000000
128N	0.00000000	0.00544465	0.00816697	0.90744102	0.00000000	0.07531760	0.00000000	0.00362976	0.00000000	0.00000000	0.00000000
128C	0.00000000	0.00000000	0.01814882	0.00000000	0.90744102	0.00362976	0.06805808	0.00000000	0.00272232	0.00000000	0.00000000
129N	0.00000000	0.00044964	0.00044964	0.01708633	0.00899281	0.89928058	0.00000000	0.07104317	0.00000000	0.00269784	0.00000000
129C	0.00000000	0.00000000	0.00000000	0.00000000	0.01575533	0.00000000	0.92678406	0.00370714	0.05189991	0.00000000	0.00185357
130N	0.00000000	0.00000000	0.00000000	0.00045290	0.00045290	0.02717391	0.00905797	0.90579710	0.00000000	0.05525362	0.00000000
130C	0.00000000	0.00000000	0.00000000	0.00000000	0.00183234	0.00000000	0.03206596	0.00000000	0.91617041	0.00366468	0.04489235
131	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00046083	0.00046083	0.01843318	0.01566820	0.92165899	0.00000000
131C	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00092851	0.00000000	0.04178273	0.00000000	0.92850511

Supplementary Table 8: Isotope Correction Factors for the TMT10/11 tags used in the experiments presented in the main manuscript. Note that these numbers are batch specific and the Correction Factors supplied with each batch should be used for experiments carried out with a specific batch.

Tag	126	127N	127C	128N	128C	129N	129C	130N	130C	131N	131C	132N	132C	133N	133C	134N
126	0.92376240	0.00264094	0.07140740	0.00014905	0.00204021	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
127N	0.00628014	0.92290721	0.00000000	0.06880847	0.00000000	0.00200418	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
127C	0.00660529	0.00000000	0.92755122	0.00292319	0.06136829	0.00010915	0.00144286	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
128N	0.00001770	0.00693664	0.01035470	0.91985902	0.00000000	0.06132394	0.00000000	0.00150797	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
128C	0.00051766	0.00000000	0.01252113	0.00000000	0.93355396	0.00263696	0.04955358	0.00018140	0.00103532	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
129N	0.00000000	0.00008990	0.00093825	0.01193741	0.01039568	0.92503923	0.00000000	0.05065691	0.00000000	0.00094266	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
129C	0.00000000	0.00000000	0.00237388	0.00000000	0.02172789	0.00000000	0.92699134	0.00297716	0.04517244	0.00001180	0.00074552	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
130N	0.00000000	0.00000000	0.00000000	0.00354833	0.00020518	0.02173688	0.01076686	0.92093939	0.00000000	0.04211943	0.00000000	0.00068393	0.00000000	0.00000000	0.00000000	0.00000000
130C	0.00000000	0.00000000	0.00000000	0.00000000	0.00049821	0.00000000	0.02496534	0.00000000	0.93415180	0.00293903	0.03600035	0.00000852	0.00143674	0.00000000	0.00000000	0.00000000
131N	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00044577	0.00029248	0.02516294	0.01124516	0.92785195	0.00000000	0.03464317	0.00000000	0.00035853	0.00000000	0.00000000
131C	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00085383	0.00000000	0.03451900	0.00000000	0.93514361	0.00348036	0.02594007	0.00001390	0.00004920	0.00000000
132N	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00081435	0.00006810	0.02369015	0.00666285	0.94267034	0.00000000	0.02600981	0.00000000	0.00008440
132C	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00095363	0.00000000	0.03867478	0.00000000	0.94139878	0.00349662	0.01532320	0.00014753
133N	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00081388	0.00011539	0.02926632	0.00719115	0.94767128	0.00000000	0.01493976
133C	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00336127	0.00000000	0.04357206	0.00000000	0.94198650	0.00282181
134N	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00358518	0.00034195	0.04529179	0.00377097	0.93897609

Supplementary Table 9: Isotope Correction Factors for the TMTpro tags used in the experiments presented in the main manuscript. Note that these numbers are batch specific and the Correction Factors supplied with each batch should be used.