Supporting information S-1

PAcIFIC goes faster, quantitative and accurate

Alexandre Panchaud[1,2], Sunhee Jung[1,3], Scott A. Shaffer[1,4], John D. Aitchison[2], David R. Goodlett[1*]

[1] Department of Medicinal Chemistry, University of Washington, Seattle, WA

[2] Nestlé Research Center, Lausanne, Switzerland

[3] Institute for Systems Biology, Seattle, WA

[4] University of Massachusetts, Worcester, MA, USA

Table of content:

Experimental Procedures

For the optimization of PAcIFIC, all data were acquired in duplicate as previously described [13] on either an LTQ XL or LTQ Velos ion trap. Only a limited *m/z* range was used consisting of 10 injections and the number of channels recorded per injection was changed with values ranging from 10 to 30 data-independent MS2 events per cycle as well as the *m/z* increment from one channel to the next one (1.5, 2.0, 2.5 or 3.0). See **Figure S-1** for more details.

For quantitative PAcIFIC (qPAcIFIC), all data were acquired in duplicate on an LTQ XL ion trap. A collision induced dissociation (CID) spectrum was acquired every 1.5 *m/z* value followed by a pulse-Q dissociation (PQD) spectrum in the low *m/z* range where TMT reporter ions are expected (121 to 132 *m/z*). Thus, in a single LC-MS/MS experiment 30 such scans are acquired (i.e. 15 pairs of CID/PQD) across a range of 22.5 *m/z*. Repeated injections were preformed in an identical fashion until the desired *m/z* mass range was achieved (400-1400 *m/z*). For CID spectrum, ion population was set to 1 x 10^4, precursor isolation width to 2.5 Th, activation Q to 0.250, activation time to 30 ms and collision energy to 35%. For PQD spectrum, ion population was set to 1 x 10^4, precursor isolation width to 2.5 Th, activation Q to 0.550, activation time to 0.4 ms and collision energy to 35%. Identical parameters were used for the comparative data-dependent runs with a cycle set to pairs of CID/PQD scans on the three most intense precursor ions with a dynamic exclusion set to 30 second after a repeat count of 2. See **Figure S-2** for more details.

For PAcIFIC at high mass accuracy (aPAcIFIC), all data were acquired on a LTQ Orbitrap XL as previously described [13] using a cycle of 15 data-independent CID spectra with a precursor ion survey scan inserted every 5 tandem mass spectrum for optimal sampling rate. Each survey scan was acquired from 400-2000 *m/z* in the Orbitrap analyzer at 60'000 resolution (at 400 Th) using an optimal ion population of 5 x 10^5 controlled by automatic gain control. See **Figure S-3** for more details.

Database search pre-processing. For each analysis, RAW files were converted to mzXML files using ReAdW 4.3.1 (SPC, Seattle, WA, USA) for further database search. In the case of the qPAcIFIC experiment, CID and PQD spectrum pairs were first combined into a single tandem mass spectrum using an in-house Perl software (qPAcIFIC.pl). The output consists of a new mzXML file containing CID-PQD combined spectra.

In the case of aPAcIFIC, a feature detection step was first performed on the high resolution survey scans using Hardklor [17]. This information was further used to correct the precursor mass of the data-independent spectrum using an in-house Perl program (aPAcIFIC.pl). Briefly, this program looks for a precursor ion in the hardklor output file within: (i) a specified retention window (± 12 sec.), (ii) a specified *m/z* range (± 1.25 *m/z*). If no precursor ion is found, the tandem mass spectrum is left unmodified. If a precursor ion is found, the precursor mass of the tandem mass spectrum is changed. In the case of multiple precursor ions found (up to 3), replicate tandem mass spectrum are written with their respective mass to account for co-fragmenting peptides. The output consists of two separate mzXML files containing either modified tandem MS spectra (to be searched at high mass accuracy) or raw tandem MS spectra (to be searched at low mass accuracy). See **Figure S-3** for more details.

Database search. mzXML files were searched with SEQUEST v.27 against the Pseudomonas PAO1 protein database (release 2001-09-17 / 5727 proteins) (www.pseudomonas.com/dowload.jsp), the yeast ORF database (release 2006-05-07 / 6717 proteins) (www.yeastgenome.org) or the 14 standard protein database with sequences retrieved from UniProtKB (www.uniprot.org). For regular PAcIFIC searches, precursor tolerance was set to 3.75 Da with alkylated cysteine set as a fixed modification and oxidized methionine as variable. For quantitative experiments, the TMT labeling was added as a fixed modification to lysine side chains and peptide N-termini. In the case of accurate PAcIFIC searches, the precursor ion tolerance was set to 10 ppm and the isotopic window feature used

to look -1/+1/+2/+3 from the monoisotopic mass to account for misassigned precursor masses [18]. For all searches, up to two missed cleavages were allowed.

Database search post-processing. Sequest results were converted to pepXML files and probability assessments of identified peptides were computed with PeptideProphet and ProteinProphet (ISB, Seattle, WA, USA). For all individual searches, peptides with probability scores equal or higher to 0.99 were used (estimated error rate less than 0.5%). For quantification purposes, Libra (ISB, Seattle, WA, USA) was used to calculate each individual peptide ratios as well as compute final protein ratios. The following parameters were used for the Libra condition file: reagent *m/z* set to 126.1 and 127.1, mass tolerance to 0.3, minimum threshold to 1000 (14 protein mixture) or 500 (*P. aeruginosa* digest). The isotopic contribution values were added according to the notice received in the TMT 2plex kit. Only proteins with multiple hits were considered in the quantification process.

Results

LTQ Velos results

LTQ XL and LTQ Velos instruments have different sensitivity and scanning speed, it is difficult to compare samples separately on each instrument especially with different LC systems. Therefore, in order to have strictly the identical hardware and parameters but the number of CID scan events, we conducted all experiments on the LTQ Velos but used as a reference an experiment on the LTQ Velos peformed with "Enhanced" scan mode which mimics scan speed of an LTQ XL at "normal" scan speed (See **Figure S-2.A**). As well, because of the higher sensitivity of the LTQ Velos, injected sample amount was decreased approx. 5 times to obtain identical total ion current (TIC) than was achieved on the LTQ XL. Five m/z ranges were used for these experiments with identical m/z starts (610, 655, 700, 745, 790). Following experiments were tested: 15 (experiment B), 20 (experiment C), 25 (experiment D) and 30 (experiment E) with an isolation width of 2.5 and a window increment of 1.5 *m/z*.As shown in Figure S-2.B, comparable number of unique peptides is achieved with a cycle of 25 CID scans. This change in rate decreases the time needed for a full PAcIFIC experiment over 1000 u to ~ 1.5 days (22 x 100min run).

Supplementary Table:

Table S-1: Composition of the 14 proteins standard mixture.

Solution	Description	From	Mr [Da]	Solutions [ug/ml]	Tot [nmol]	Concentration [pmol/ul]	STD14_TMT126 [ug]	STD14_TMT126 [pmol]	STD14_TMT127 [ug]	STD14_TMT127 [pmol]	TMT 126:127
1	Carbonic anhydrase	Bovine erythrocytes	29000	1000	34.48	34.48	9.66	333.33	29.00	1000.00	1:3
2	Lysozyme	Chicken egg	14700	1000	68.03	68.03	14.70	1000.00	14.70	1000.00	1:1
3	Ovalbumin	Chicken egg	44287	1000	22.58	22.58	44.29	1000.00	14.76	333.33	3:1
4	Albumin	Bovine serum	66000	1000	15.15	15.15	66.00	1000.00	13.20	200.00	5:1
5	Beta lactoglobulin	Bovine milk	18400	1000	54.35	54.35	9.20	500.00	18.40	1000.00	1:2
6	Catalase	Bovine liver	60000	1000	16.67	16.67	12.00	200.00	60.00	1000.00	1:5
7	Alpha casein	Bovine milk	25000	1000	40.00	40.00	25.00	1000.00	12.50	500.00	2:1
8	Beta casein	Bovine milk	25000	1000	40.00	40.00	25.00	1000.00	12.50	500.00	2:1
9	Insulin	Bovine pancreas	5800	1000	172.41	172.41	5.80	1000.00	5.80	1000.00	1:1
10	Myoglobine	Horse heart	17000	1000	58.82	58.82	17.00	1000.00	17.00	1000.00	1:1
11	Serotransferrin	Bovine	77000	1000	12.99	12.99	7.70	100.00	77.00	1000.00	1:10
12	Alpha amylase	Aspergillus oryzae	51000	1000	19.61	19.61	17.00	333.33	51.00	1000.00	1:3
13	Serotransferrin	Human	77000	1000	12.99	12.99	77.00	1000.00	7.70	100.00	10:1
14	Cytochrome C	Horse	12000	1000	83.33	83.33	12.00	1000.00	12.00	1000.00	1:1
15	STD14_TMT126	all					342.35	10466.66			
16	STD14_TMT127	all							345.56	10633.33	

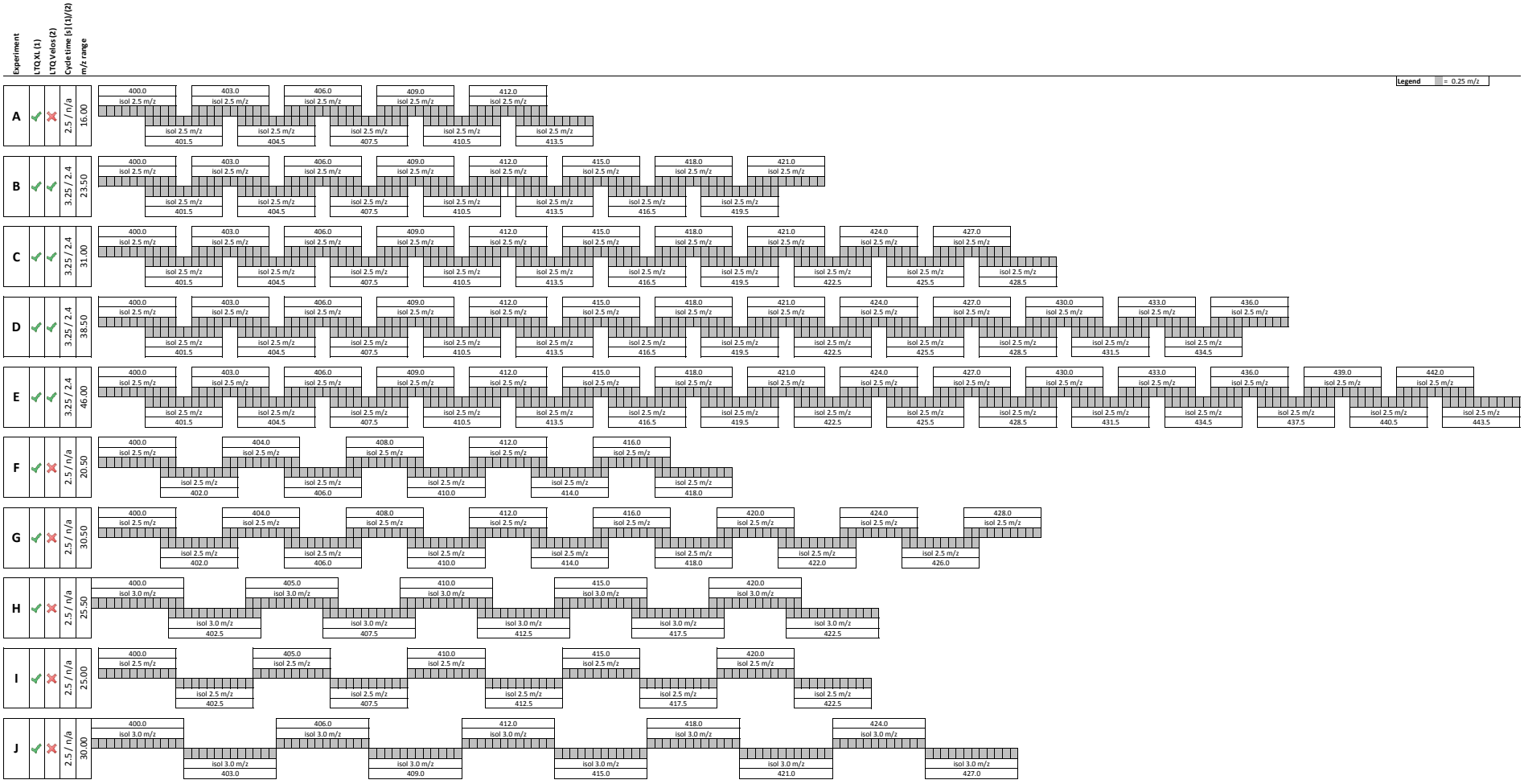
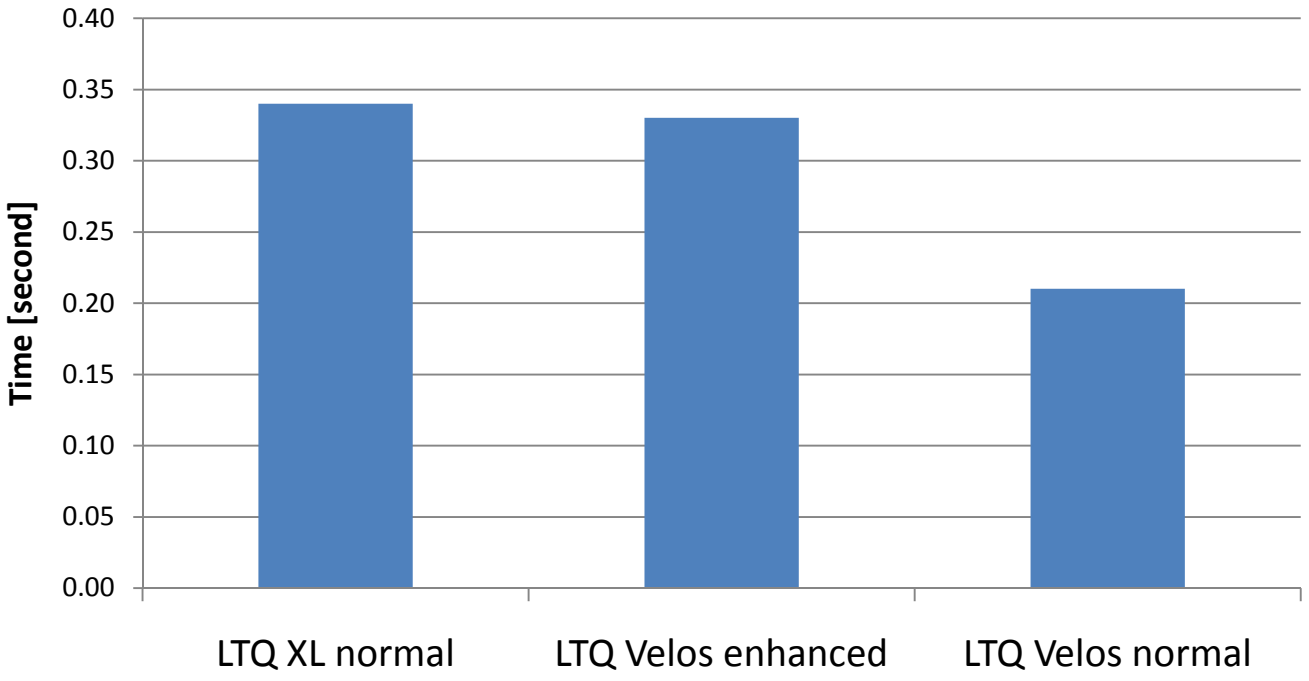


Figure S-1: Description of the different PAcIFIC strategies applied for the optimization process. Each experiment as described in Table 1 is depicted with a single gray box representing a unit of 0.25 m/z. All figures show one cycle of MSMS events that are conducted during one LC-MSMS injection starting at 400 m/z. The upper box is the precursor mass or the center of the window. Each MSMS spectrum is depicted using gray boxes according to the isolation width used. The position of each MSMS spectrum is based on the precursor m/z increment from one MSMS spectrum to the next one (channel increment). This highlights in the different experiments how much overlap is kept from one MSMS spectrum to the next one as well as the m/z range covered by one injection or cycle.

A.



B.

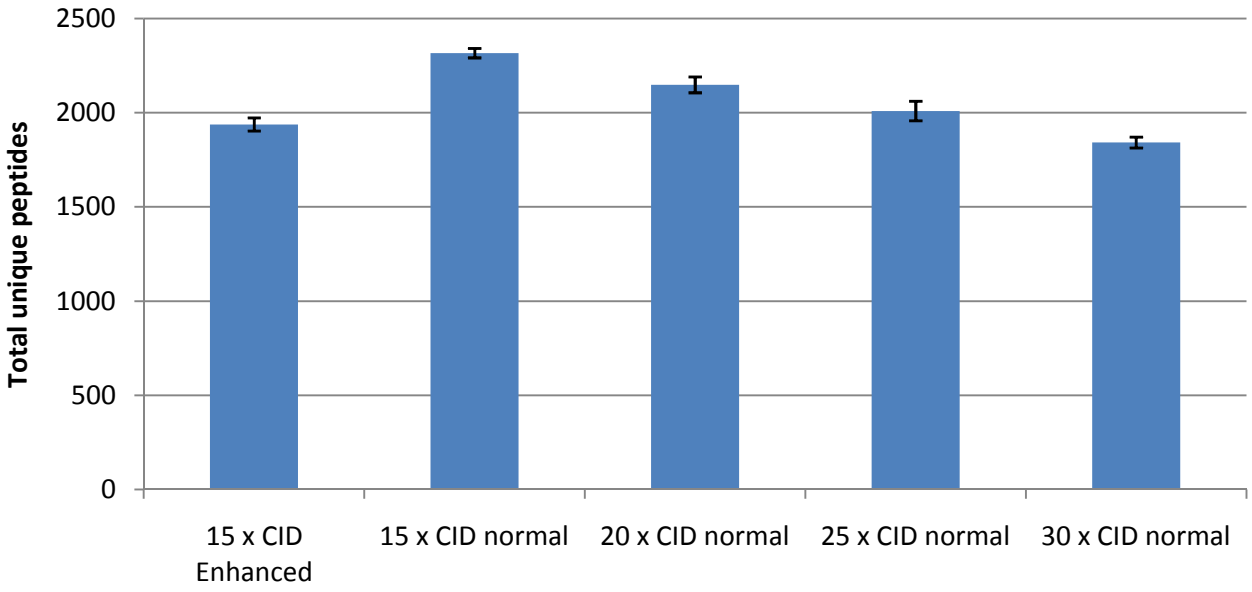


Figure S-2: PAcIFIC resuls with the LTQ Velos. In order to have identical parameters except the number of scans, all experiments were performed on the LTQ Velos. (A)To mimic scanning speed of the LTQ XL, "enhanced" scan mode was selected for each individual CID scans to obtain similar scanning time. (B) This reference experiment was then compared to experiments conducted at "normal" scan speed using 15, 20, 25 or 30 CID scans using an isolation width of 2.5 and window increment of 1.5 m/z (See **Table 1** and **Figure S-1** for more details) .

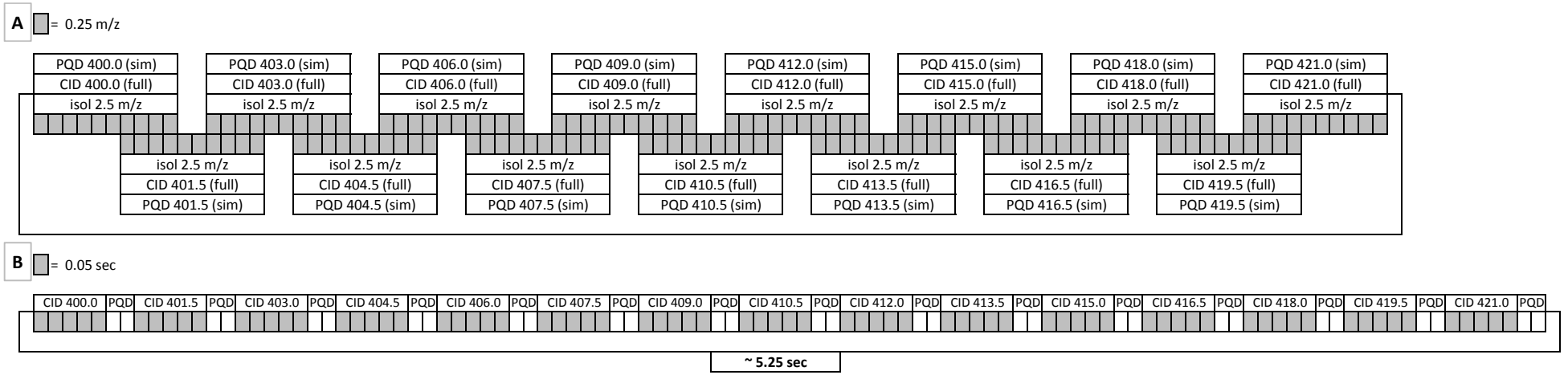
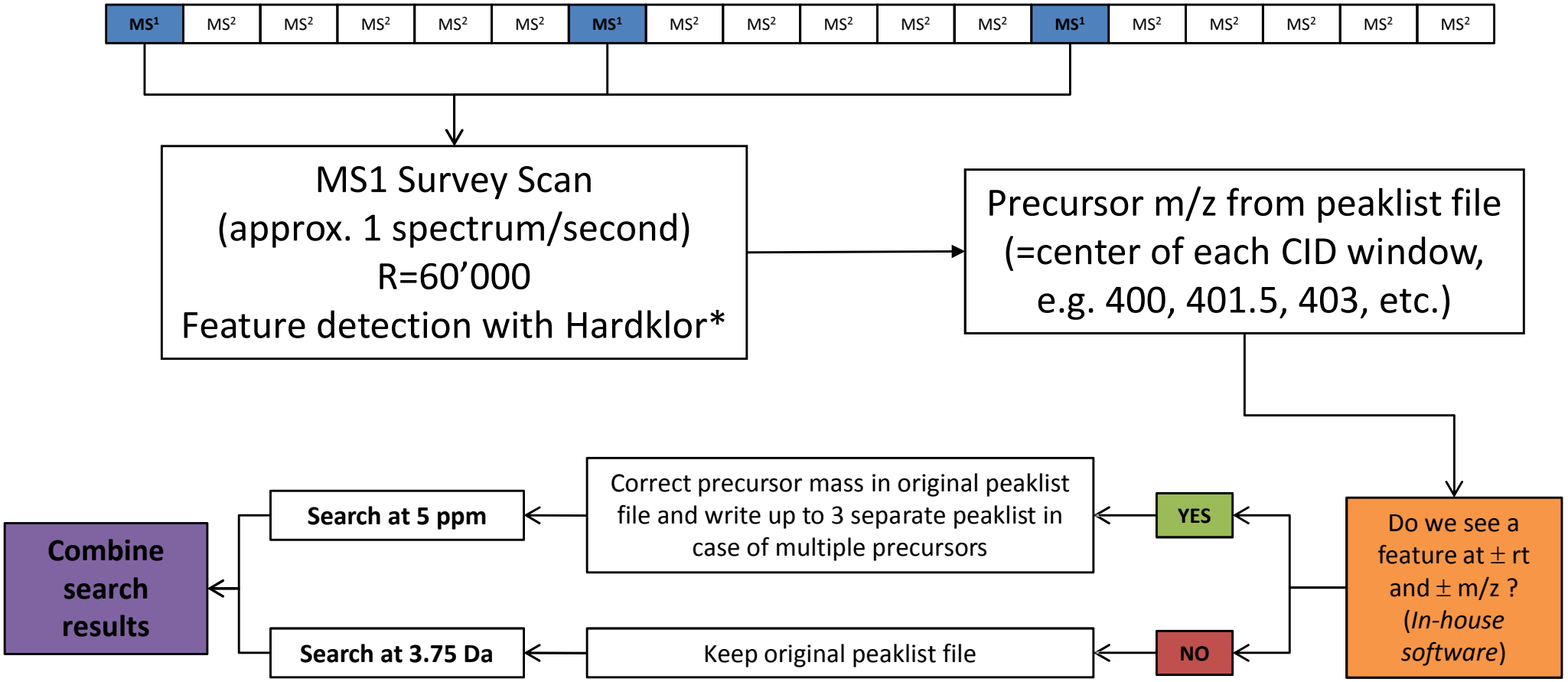


Figure S-2: Description of one acquisition cycle (one LC-MSMS injection) for the quantitative PAcIFIC analysis covering an m/z range of 400-422.5. Each channel is comprised of a CID spectrum using the full m/z range followed by a SIM PQD spectrum focused only on the TMT reporter ion region. The cycle is described first in terms of the mass-to-charge ratio covered with one gray unit corresponding to 0.25 m/z (A) and second in terms of duration with



*Hoopmann et al., Anal. Chem. **2007**;79:5620

Figure S-3: Description of the accurate PAcIFIC strategy to re-assign precursor mass of data-independent tandem mass spectrum. Briefly, survey scans are deconvoluted and features detected by the program Hardklor. For each product ion spectrum, the precursor mass corresponding to the center of the window is used to find whether a feature has been detected within a certain m/z and time range. If negative, the original peaklist is kept as such and searched with standard parameters for PAcIFIC. If positive, up to three peaklist are written with their respective new accurate precursor mass and searched with high mass accuracy parameters. Finally, both results are combined, duplicate filtered and peptide/protein lists assembled.