1. Tip-based fractionation of peptides according to theoretical net charge states at pH 2.7 (scalability experiments)

1.1 Sample preparation

HeLa cells were lysed with 950 µL of lysis buffer containing: 1% SDS, 50 mM Tris, 150 mM NaCl, and cOmplete™, Mini EDTA-free protease inhibitor (Roche) at pH 7.8. To digest nucleic acids, benzonase nuclease and $MgCl_2$ were added, followed by incubation at 37°C for 30 min and centrifugation at 18 000 g at 4°C for 15 min to remove debris. The supernatant was carefully collected and a bicinchoninic acid (BCA) assay (Thermo Scientific) was performed, in order to determine protein concentration. Cysteines were reduced with 10 mM dithiothreitol (DTT) for 30 min at 56°C and subsequently alkylated with 30 mM iodoacetamide (IAA) for 30 min at room temperature (RT) in the dark, then again 10 mM DTT were added for quenching. The sample (corresponding to 2.7 mg of protein) was subjected to ethanol protein precipitation by 10-fold dilution with ice-cold ethanol and storage at -40°C for 1 h, followed by centrifugation at 4°C at 15 000 g for 30 min and a washing step of 10 min with acetone. The supernatant was carefully removed and the protein pellets were resolubilized with 150 µL of 6 M GuHCl. GuHCl concentration was lowered to 0.2 M with 50 mM ammonium bicarbonate, pH 7.8. Next, ACN and $CaCl_2$ were added to final concentrations of 5% and 2 mM, respectively. Finally, trypsin was added in 1:20 (w/w, enzyme to substrate) and incubated at 37°C for 14 h. Proteolytic digestion was controlled using monolithic reversed phase chromatography, as described previously(1) . Generated peptides were desalted using C18 SPE columns (SPEC 3 mL C18AR 30 mg, Agilent Technologies) according to the manufacturer's recommendations. Eluted peptides were dried to completeness under vacuum, followed by resolubilization in 520 µL SCX buffer A (10 mM KH_2PO_4, 20% ACN, pH 2.7). Different amounts of this sample, namely 10 µg, 50 µg, and 200 µg (based on BCA) were used in triplicate and subjected to SCX tip-based fractionation.

1.2 Tip-based SCX fractionation

Originally, we optimized the conditions for the tip-based SCX fractionation of peptides according to theoretical net charge states at pH 2.7. We used *(i)* different frits (C8 Empore Disc (3M) and Whatman Glass microfiber filter (Sigma Aldrich) and compared the peptide recocvery, *(ii)* different SCX particle to peptide ratios, as well as for the individual elution steps: *(iii)* different buffer compositions (varying % of SCX buffer B: 250 mM KCl, 10 mM KH_2PO_4, 20% ACN, pH 2.7; in SCX buffer A: 10 mM KH_2PO_4, 20% ACN, pH 2.7), *(iv)* different volumes (from 50 µL to 150 µL) *(v)* with and without a subsequent wash step with 100% SCX buffer A, *(vi)* different centrifugation times from 5 to 15 min, as well as *(vii)* the combination of subsequent elution steps with different % of SCX buffer B for higher charge states. All experiments were conducted step-by-step in duplicate or triplicate using aliquots of unlabeled and labeled peptides aliquots from the same HeLa digest. To assess the performance of the different conditions, samples were analyzed as batches in randomized order by LC-MS and identified peptide-spectrum-matches as 1% FDR were used to compare the recovery (number of identified PSM and peptides) as well as the separation of peptides based on theoretical net charge states (again, taking into account both peptides and PSMs). For this, each Arg, Lys (unlabelled and dimethyl as well as iTRAQ labeled), and His residue and N-terminus (unlabelled, dimethyl as well as iTRAQ labeled) were counted as one positive charge. Thus, we finally came up with the optimized protocol as follows (see also tables S1 and S2).

A plug of Whatman Glass microfiber filter (Sigma Aldrich) was added as frit into a conventional 2-200 µL pipette tip (Eppendorf). 100 mg of SCX particles (Polysulfoethyl A, 5 µm particle size, 200 Å pore size, PolyLC Inc.) were resuspended in 500 µL SCX buffer A (10 mM KH_2PO_4, 20% ACN, pH 2.7). After thorough mixing, 25 µL (~5 mg particles) were loaded onto the tip in case of 10 µg and 50 µg sample, whereas 50 µL were loaded for the 200 µg sample. The microcolumn was conditioned with 100 µL of SCX buffer A. The buffer was removed by centrifugation at 1000 g for 8 min and discarded. Samples with 10, 50, 200 µg starting amount were resolubilized in 20, 20, 50 µL of SCX buffer A, respectively, and loaded onto the tip

followed by centrifugation at 600 g for 3 min (10 and 50 µg) or 6 min (200 µg). The flow through was collected in an Protein LoBind tube (Eppendorf). The step was repeated once for 10 µg and 50 µg and twice for 200 µg sample. The flow through was discarded. SCX buffers with different salt concentrations at pH 2.7 were successively used to elute peptides with theoretical net charge states of +1, +2, +3,+4 (see table S1). Peptides were eluted by centrifugation with 1000 g for 8 min (10 µg and 50 µg sample) and 1000 g for 10 min (200 µg sample), respectively. Eluted fractions were collected in individual Protein LoBind tubes (fractions 1, 2, 3, 4) and evaporated completely under vacuum. Whereas fractions 1-3 were resolubilized in 150 µL 0.1% trifluoroacetic acid (TFA), fraction 4 was resolubilized in 50 µL of 0.1% TFA and desalted using self-made microcolumns packed with Poros Oligo R3 reversed-phase material (Applied Biosystems) (2). Briefly, the microcolumn was conditioned with 100 µL ACN, followed by equilibration with 100 µL of 0.1% TFA. Fraction 4 was loaded twice. Bound peptides were washed two times with 0.1% of TFA and afterwards eluted with 100 µL of 60% (v/v) ACN, 0.1% TFA. The sample was dried completely under vacuum and resolubilized in 150 µL 0.1% TFA.

Table S1. Conditions for charge-based separation of peptides using SCX tips at pH 2.7. Given are buffer volumes and centrifugation steps. SCX A buffer: 10 mM KH_2PO_4, 20% ACN, pH 2.7. SCX buffer B: 250 mM KCl, 10 mM KH_2PO_4, 20% ACN, pH 2.7. Note: For 50 µL of SCX material, centrifugation steps need to be extended to 10 min, each.

+1 fraction	**+2 fraction**	**+3 fraction**	**+4 fraction**
1. load 100 µL A 100%	1. load 100 µL B15%	1. load 100 µL B30%	1. load 100 µL B65%
2. Centrifuge/1000 g/8 min	2. Centrifuge/1000 g/8 min	2. Centrifuge/1000 g/8 min	2. Centrifuge/1000 g/8 min
Collect step 2	3. load 100 µL A100%	3. load 100 µL A100%	**Collect step 2**
	4. Centrifuge/1000 g/8 min	4. Centrifuge/1000 g/8 min	
	Collect steps 2 and 4, pool	5. load 100 µL B35%	
		6. Centrifuge/1000 g/8 min	
		Collect steps 2, 4, 6, pool	

1.3 LC-MS analysis

For 10 µg and 50 µg samples, 10% of fraction 1, 1% of fractions 2 and 3, and 10 % of fraction 4 were used for further nano-LC-MS/MS analysis. In case of the 200 µg sample, eluted fractions were not dried under vacuum, instead 4 µL from each fraction were used for further nano-LC-MS/MS analysis. All fractions were analyzed by nano-LC-MS/MS using an Orbitrap XL mass spectrometer online- coupled to a nano RSLC HPLC system (both Thermo Scientific). Samples were loaded onto a trap column (C18, 5 µm, 100 Å, Acclaim PepMap™ RSLC, 100 µm x 2 cm, Thermo Scientific) at a flow rate of 20 µL/min 0.1% TFA and subsequently separated on a main column (C18, 3 µm, 100 Å, Acclaim PepMap® RSLC, 75 µm x 25 cm, Thermo Scientific) using a binary gradient consisting of solvent A (0.1% formic acid (FA)) and solvent B (0.1% FA, 84% ACN) at a flow rate of 300 nL/min at 45°C. Gradient increased linearly from 3-42% B in 60 min. The mass spectrometer was operated in data-dependent acquisition (DDA) mode acquiring full MS Scans at a resolution of 60,000, at a scan range 300-2000 m/z, followed by MS/MS of the 10 most abundant ions (Top 10) using collision induced dissociation (CID) with a normalized collision energy of 35 and acquisition in the ion trap. The polysiloxane ion at m/z 371.1012 was used as lock mass (3). AGC target values were $5x10^5$ for MS and 10^4 for MS/MS scans, allowing for a maximum injection time of 100 ms for both MS and MS/MS scans.

Fraction 4 was analyzed using an Orbitrap Fusion mass spectrometer (Thermo Scientific) online-coupled to a nano RSLC HPLC system. Samples were loaded onto a trap column (C18, 5 µm, 100 Å, Acclaim PepMap™ RSLC, 100 µm x 2cm, Thermo SCIENTIFIC) at a flow rate of 20 µL/min 0.1% TFA and subsequently separated on a main column (C18, 3 µm, 100 Å, Acclaim PepMap® RSLC, 75 µm x 50 cm) using a binary gradient ranging from 3-50% B in 45 min at a flow rate of 250 nL/min at 60°C. The mass spectrometer was operated in DDA mode acquiring full MS from 300-1500 m/z at a resolution of 60,000 in the Orbitrap. Precursor ions with charge states 3-8 were isolated in the quadrupole with an isolation window of 2.0 m/z and fragmented using electron transfer dissociation (ETD) and a dynamic exclusion of

30 s. MS/MS scans were acquired in the ion trap in rapid scan mode and the first mass was set to 120 m/z. AGC target values were 2 x 10^5 for MS and 1 x 10^4 for MS/MS, maximum injection times were 50 ms and 60 ms for MS and MS/MS, respectively. A 10% NH_4OH solution was placed in front of the ion source for charge state reduction, as described by Thingholm et al (4).

1.4 Data Analysis

Raw data were searched against the human Uniprot database (July 2015, 20,207 target sequences) using Mascot 2.4 (Matrix Science) and the Proteome Discoverer software version 1.3 (Thermo Scientific). Trypsin was set as enzyme, allowing a maximum of two missed cleavage sites. Mass tolerances were set to 10 ppm for MS and 0.5 Da for MS/MS. N-terminal acetylation (+ 42.0105 Da) of protein N-termini and oxidation (+ 15.9949) of methionine were set as variable and carbamidomethylation of Cys (+ 57.0214 Da) as fixed modification. High confidence corresponding to a false discovery rate (FDR) < 1% on the peptide spectrum match (PSM) level, and search engine rank 1 were used as filters.

Theoretical charge states of high confidence PSMs were calculated based on the identified peptide sequences, counting the occurrence of the following amino acid residues representing positive charges at pH 2.7: Free N-termini, Arg, Lys, His. For each amount of starting material (10 µg, 50 µg, 200 µg) and fraction (1, 2, 3, 4) the charge state distribution of PSMs was plotted, considering the average of three replicates (figure 1b). All relative standard deviations were below 10%.

2. Tip-based fractionation of iTRAQ- labeled peptides

2.1 Sample preparation and SCX tip-based fractionation

HeLa lysate was prepared as above. 6 x 150 µg (based on BCA) of protein were processed using filter-aided sample preparation (FASP) (5, 6). Briefly, cell lysates corresponding to 150 µg of protein were diluted up to 450 µL with freshly prepared 8 M urea/100 mM Tris-HCl, pH 8.5(7) and placed on a Nanosep centrifugal device (30 KDa, PALL). The devices were centrifuged at 13,900 g at RT for 20 min. All the following centrifugation steps were performed under similar conditions i.e. 13,900 g, RT, 15 min. To remove residual SDS, three washing steps were carried out using 100 µL of 8 M urea/100 mM Tris-HCl, pH 8.5 and for the buffer exchange; the devices were washed thrice with 100 µL of 50 mM Triethylammonium bicarbonate (TEAB), pH 7.8. To the concentrated proteins, 100 µL of digestion buffer comprising 1:20 (w/w) sequencing grade trypsin, 0.2 M GuHCl, 2 mM $CaCl_2$ in 50 TEAB, pH 7.8 were added and incubated at 37°C for 14 h. The generated tryptic peptides were recovered by centrifugation followed by two consequent washing steps with 50 µL of 50 mM TEAB, pH 7.8 and 50 µL of ultra-pure water. Finally, peptides were acidified with 10% TFA (pH < 3.0) and digests were quality controlled as described previously (1). Afterwards, peptides were desalted (SPEC Pt C18 Ar, 4 mg sorbent, Agilent Technologies) as described above. Before labeling with iTRAQ 8plex reagents(8) each sample (~1 µg) was analyzed by nano-LC-MS/MS to adjust minor differences in sample amounts due to pipetting errors or incorrect protein/peptide concentration estimation based on the alignment of total ion chromatograms (TICs). Dried pellets were resolubilized in 30 µL of iTRAQ(8) dissolution buffer (0.5 M triethylammonium bicarbonate pH, 8.5). 80 µL of isopropanol was added to each 8-plex iTRAQ reagent (*i.e.* 113, 114, 115, 116, 119, 121) and added to the samples, followed by an incubation at RT for 2 h. Afterwards, samples were pooled and the reaction was quenched by adding 100 µL of pure water. An aliquot of 1 µg was measured by nano-LC-MS/MS in order to control label efficiency, as described previously (9). Labeling efficiency was determined based on the share of labeled lysine residues on the PSM level at 1% FDR and was ~99%. From the pooled sample, 3 aliquots à

50 µg were taken, dried to completeness under vacuum, resolubilized in 50 µL of SCX A and subjected to SCX tip-based fractionation, as described above (table S1). Fractions 1-4 were prepared for nano-LC-MS/MS as described for the unlabeled peptides.

2.2 LC-MS analysis:

10% of fractions 1 and 4 and 1% of fractions 2 and 3 were analyzed by nano-LC-MS/MS. Fraction 1-3 were analyzed using an Orbitrap Velos mass spectrometer online-coupled to a nano RSLC HPLC system (both Thermo Scientific) using columns and LC gradients as described above. The mass spectrometer was operated in DDA mode acquiring full MS scans from 300-2000 m/z at a resolution of 30,000, followed by MS/MS of the 5 most abundant ions (Top 5) using higher energy collisional dissociation (HCD) fragmentation with a normalized collision energy of 42 and the first mass set to 105 m/z. The polysiloxane ion at m/z 371.1012 was used as lock mass (3). MS/MS were acquired at a resolution of 7500. Fraction 4 was measured as described above on an Orbitrap Fusion. For fractions 2, 3 and 4, 10% NH_4OH solution was placed in front of the ion source to compensate for elevated charge states, as described by Thingholm et al (4).

2.3 Data analysis

Raw data were searched as above. Trypsin was set as the enzyme type, with an allowed maximum of two missed cleavage sites. For fractions 1-3 mass tolerances were set to 20 ppm for MS and 0.02 Da for MS/MS and for fraction 4 to 10 ppm for MS and 0.5 Da for MS/MS, respectively. Filters were applied as above. For fractions 1-3, N-terminal acetylation (+ 42.0105 Da) of protein N-termini, oxidation (+ 15.9949) of methionine and iTRAQ 8-plex (+304.2053 Da) of peptide N-termini were set as variable modification; carbamidomethylation of Cys (+ 57.0214 Da) and iTRAQ 8-plex (+304.2053 Da) at Lys were set as fixed

modifications. For fraction 4, carbamidomethylation of Cys (+ 57.0214 Da), iTRAQ 8-plex (+304.2053 Da) of peptide N-termini and Lys were set as fixed modifications.

Theoretical charge states of high confidence PSMs were calculated based on the identified peptide sequences, counting the occurrence of the following amino acid residues representing positive charges at pH 2.7: iTRAQ labeled N-termini, Arg, iTRAQ-labeled Lys, His. The charge state distribution of PSMs was plotted for all three replicates (figure 1c).

3. Comparison of HPLC and tip-based ChaFRADIC

3.1 Cell culture and sample preparation

Six biological replicates of SH-SY5Y cells were grown to more than 90% confluency in Dulbecco's Modified Eagle's Medium (PAN-Biotech). The medium was replaced by fresh medium containing either 1 µM staurosporine in DMSO (Sigma Aldrich) or DMSO (both 10 µL) to three replicates, each, followed by incubation for 6 hours at 37°C. The medium supernatant of each flask which contained floating cells was discarded and cells were washed with ice-cold PBS, scraped and collected by centrifugation with 300 g for 5 min at RT. Afterwards, cell pellets were washed with 1 mL PBS and collected by another centrifugation step. The final cell pellets were lysed, carbamidomethylated and subjected to BCA assay for protein concentration, as described above. Afterwards, 100 µg (according to BCA) of protein were taken from each sample and subjected to ethanol precipitation, as described above. Protein pellets were then resolubilized with 10 µL of 10% SDS and 40 µL of iTRAQ labeling buffer (0.5 M TEAB in 20 % (v/v) of isopropanol) were added in order to have a final concentration of 2% SDS. Next, 80 µL of isopropanol was added to six iTRAQ 8-plex (113, 114, 115, 116, 119, 121) reagents, each label was added to a different sample, and incubated at RT (25°C) for 2 hours. After incubation, reactions were quenched with 60 mM

glycine, for 10 min at RT, followed by addition of 130 mM hydroxylamine for 10 min at RT. Afterwards, samples were pooled and directly subjected to another ethanol precipitation step. Protein pellets were resolubilized with 150 µL of 6 M GuHCl. GuHCl concentration was reduced to 0.2 M with 50 mM TEAB, pH 7.8. Next, ACN (100%) and $CaCl_2$ were added to a final concentration of 5% and 2 mM, respectively. Finally, trypsin was added in 1:20 (w/w, enzyme to substrate) ratio and incubated at 37°C for 14 h. Proteolytic digests were controlled using monolithic reversed phase chromatography, as described previously (1). Peptides were desalted by SPE (Oasis HLB Extraction Catridge, 10 mg). After desalting, 1 µg aliquot was analyzed by nano-LC-MS/MS for a global analysis in order to *(i)* determine systematic errors from sample preparation based on iTRAQ protein ratios and to *(ii)* determine labeling efficiency, as described above. The rest of the sample was dried, resolubilized in 310 µL SCX buffer A and split in 6 aliquots à 51 µL corresponding to ~100 µg of peptide, each.

3.2 Enrichment of N-Terminal Peptides Using HPLC ChaFRADIC

Three technical aliquots (100 µg each, BCA results) were fractionated using HPLC-based ChaFRADIC and the other three (100 µg each, BCA) using ChaFRA*tip*. HPLC separations were performed using a U3000 HPLC system (Thermo Scientific) and a 150 x 1 mm POLYSULFOETHYL A column (PolyLC, Columbia, US, 5 µm particle size, 200 Å pore size), in combination with a tertiary buffer system consisting of SCX buffer A (10 mM KH_2PO_4, 20% ACN, pH 2.7), SCX buffer B (10 mM KH_2PO_4, 250 mM KCl, 20% ACN, pH 2.7) and SCX buffer C (10 mM KH_2PO_4, 600 mM NaCl, 20% ACN, pH 2.7). 51 µL of the resolubilized peptides were separated at a flow rate of 80 µL/min. Peptides were separated with an optimized gradient to efficiently separate different charge states, and fractions were automatically collected using the U3000 fractionation option. The gradient was as follows: 100% A for 10 min followed by a linear increase from 0% to 15% B in 9.3 min. Afterwards, B was kept at 15% for 8.7 min and then the gradient linearly increased from 15% to

30% B for 8 min. Then B was kept constant at 30% for 11 min, and linearly increased to 100% in 5 min. After 5 min at 100% B, C was increased to 100% in 1 min and kept constant for 5 min. Finally, A was increased to 100% in 1 min and the column was re-equilibrated at 100% A for 20 min.

Per replicate, four fractions corresponding to charge states +1, +2, +3 and +4 were collected, as indicated in figure S1 below.

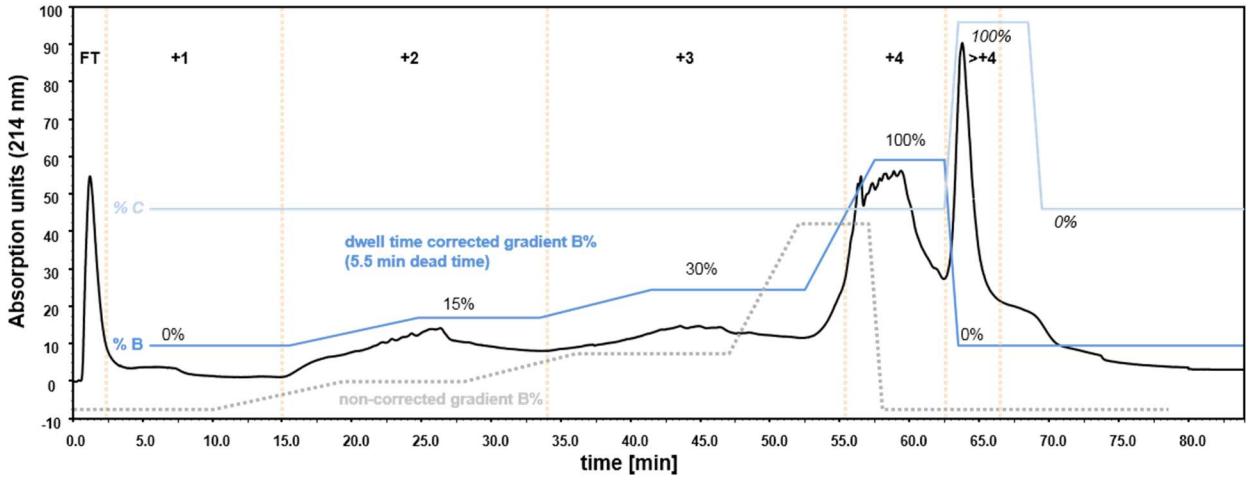

Figure S1. Scheme of the HPLC gradient used for the first and second ChaFRADIC dimensions. Fractions are indicated by vertical dashed lines. The gradients for B% and C% are indicated as blue straight lines, considering the dead time of 5.5 min of the system. Additionally, the HPLC gradient as programmed for the pump is indicated as dashed grey line. The UV trace is depicted as black line.

Next, collected fractions were completely dried under vacuum and brought to 300 µL with 200 mM Na_2HPO_4, pH 8.0 to a final pH of ~7.0. Next, free N-termini of internal peptides were derivatized with NHS-trideutero acetate in two steps, modified from Staes et al (10). Initially NHS-trideutero acetate was added to a final concentration of 20 mM and samples were incubated at 37°C for 1 h, followed by addition of another 10 mM NHS-trideutero acetate under the same reaction conditions. After 2 h of total incubation

time the reaction was quenched using 60 mM glycine for 10 min at RT, followed by incubation for 5 min at 95°C. After quenching, fractions were acidified with 10% TFA and desalted using Poros Oligo R3 reversed-phase material as described above. Bound peptides were washed two times with 0.1% of FA and afterwards eluted with 100 µL of 70% (v/v) ACN, 0.1% FA. After vacuum drying peptides were resolubilized in 51 µL of SCX buffer A. Each 1st dimension fraction was separated and fractionated in a 2nd SCX dimension under exactly the same conditions as mentioned above. Whereas N-terminal peptides retained their retention time window (*i.e.* charge state window), internal peptides shifted to earlier retention times based on the reduction in theoretical net charge induced by the N-terminal acetylation.

After collection, fractions were subjected to desalting as described above. Fractions were dried and resolubilized in 22 µL TFA 0.1%. From each fraction 2/3 were taken for nano-LC-MS/MS.

3.3 Enrichment of N-terminal peptides using ChaFRA*tip*

Column preparation (25 µL of SCX material), column conditioning and sample loading (resolubilized in 50µL SCX buffer A) were done as described in chapter 1.2. Fractionation in the first dimension was conducted as described in table S1. Afterwards, collected fractions were completely dried under vacuum and brought to 300 µL with 200 mM Na_2HPO_4, pH 8.0 to a final pH of ~7.0. Next, free N-termini of internal peptides were derivatized with NHS-trideutero acetate in two steps, modified from Staes et al (10). Initially NHS-trideutero acetate was added to a final concentration of 20 mM and samples were incubated at 37°C for 1 h, followed by addition of another 10 mM NHS-trideutero acetate under the same reaction conditions. After 2 h of total incubation time the reaction was quenched using 60 mM glycine for 10 min at RT, followed by incubation for 5 min at 95°C. After quenching, fractions were acidified with 10% TFA and desalted using Poros Oligo R3 reversed-phase material as described above. Bound peptides were washed two times with 0.1% of FA and afterwards eluted with 100 µL of 70% (v/v) ACN, 0.1% FA. After

vacuum drying peptides were resolubilized in 50 µL of SCX buffer A. Second dimension tip-based fractionation was conducted as described in table S2.

Table S2. Conditions for the 2nd ChaFRA*tip* dimension. Given are buffer volumes and centrifugation steps. SCX A buffer: 10 mM KH_2PO_4, 20% ACN, pH 2.7. SCX buffer B: 250 mM KCl, 10 mM KH_2PO_4, 20% ACN, pH 2.7. The cycles can be done in parallel for the different fractions.

cycle	+1 fraction	+2 fraction	+3 fraction	+4 fraction
I	1. load 200 µL A 100% 2. Centrifuge/1000 g/10 min **Collect step 2**	1. load 200 µL A 100% 2. Centrifuge/1000 g/10 min **Discard step 2 (washing step)**	1. load 200 µL A 100% 2. Centrifuge/1000 g/10 min **Discard step 2 (washing step)**	1. load 200 µL A 100% 2. Centrifuge/1000 g/10 min **Discard step 2 (washing step)**
II		3. load 200 µL B20% 4. Centrifuge/1000 g/10 min 5. load 200 µL A100% 6. Centrifuge/1000 g/10 min **Collect steps 4 and 6, pool**	3. load 200 µL B15% 4. Centrifuge/1000 g/10 min 5. load 200 µL A100% 6. Centrifuge/1000 g/10 min **Discard steps 4, 6 (washing steps)**	3. load 200 µL B15% 4. Centrifuge/1000 g/10 min 5. load 200 µL A100% 6. Centrifuge/1000 g/10 min **Discard steps 4, 6 (washing steps)**
III			7. load 200 µL B35% 8. Centrifuge/1000 g/10 min 9. load 200 µL A100% 10. Centrifuge/1000 g/10 min 11. load 200 µL B40% 12. Centrifuge/1000 g/10 min **Collect steps 8, 10, 12, pool**	7. load 200 µL B30% 8. Centrifuge/1000 g/10 min 9. load 200 µL A100% 10. Centrifuge/1000 g/10 min 11. load 200 µL B35% 12. Centrifuge/1000 g/10 min 13. load 200 µL A100% 14. Centrifuge/1000 g/10 min **Discard steps 8, 10, 12, 14 (washing steps)**
IV				15. load 200 µl B70% 16. Centrifuge/1000 g/10 min **Collect step 16**

After collection, fractions were subjected to desalting as described above. Fractions were dried and resolubilized in 22 µL TFA 0.1%. From each fraction 2/3 were taken for nano-LC-MS/MS.

3.4 LC-MS analysis

Fractions were analyzed by nano-LC-MS/MS using an Orbitrap Fusion Lumos (Thermo Scientific) mass spectrometer, online-coupled to a U3000 HPLC system. Samples were loaded onto a self-packed column (C18, 100 µm x 2 cm, ACE, 5 µm particle size, Advanced Chromatography Technologies Ltd) at a flow rate

of 20 µL/min 0.1 % TFA, and subsequently separated on a self-packed 25 cm main column (C18, 75 µm x 25 cm, ACE, 3 µm particle size) using a binary gradient as described above. Gradient increased linearly from 3-35 % B in 90 min. The mass spectrometer was operated in DDA mode.

Fractions 1 and 2 were measured as follows: Full MS scans from 350-1500 m/z were acquired at a resolution of 120,000 with a maximum injection time of 50 ms and an AGC target value of 2×10^5ions. MS/MS were acquired using a charge-state dependent decision tree (11) in top speed mode with a cycle time of 3 s. Only ions with a charge states of +1 to +3 were chosen for MS/MS. Per precursor, two MS/MS were acquired, one for identification (ion trap) and one for quantification (Orbitrap). The first MS/MS comprised HCD with a NCE of 35% and ion trap acquisition to generate MS/MS for peptide identification. The maximum injection time was 300 ms and the AGC target value was 2×10^3. The second MS/MS was acquired in the Orbitrap at a resolution of 15,000, using a HCD with a NCE of 60% and the first fixed mass set to 90 m/z. The maximum injection time was 100 ms, the AGC target value was 1×10^5. The 'Inject Ions for All Available Parallelizable Time' option was activated and an isolation window of 1.0 m/z was used in both scan events.

Fractions 3 and 4 were measured as follows. Full MS scans were acquired as above. MS/MS were acquired using a charge-based decision tree with three MS/MS events, all using an isolation width of 1.0 m/z. In general, two MS/MS were acquired per peptide, one primarily for identification in the ion trap and one primarily for quantification in the Orbitrap using high collision energies – information from both spectra could be aligned using the Proteome Discoverer data analysis. (i) For all charges, an HCD MS/MS was acquired in the Orbitrap for quantification of the reporter ions with the following settings: HCD CE 60%, resolution of 15,000, first mass 90 m/z, AGC target 1e5 and maximum injection time of 100 ms. (ii) For charge states 2-3 the identification MS/MS was acquired with the following settings: HCD CE 35%, rapid ion trap scan rate, first mass 120 m/z, AGC target 2e3, maximum injection time 300 ms. (iii) For charge states 4-8 the identification MS/MS was acquired as follows: EThCD fragmentation with calibrated charge

dependent ETD parameters and 30% supplemental HCD activation, rapid ion trap scan rate, first mass 120 m/z, AGC target 1e4, 100 ms maximum injection time.

For ChaFRADIC fractions +2, +3 and +4, 10% NH_4OH solution was placed in front of the ion source for charge state reduction, as described by Thingholm et al (4). For analyzing ChaFRADIC fractions +1, the fragmentation of singly charged peptides was allowed as well, while excluding m/z values of singly charged ions occurring in previous blank runs.

3.5 Data analysis

Raw data were searched against an Uniprot human database (July 2015, 20,207 target sequences) using Mascot 2.4 (Matrix Science) and the Proteome Discoverer software version 1.4 including the reporter ion quantifier and percolator nodes, employing a two-step strategy and semi-ArgC enzyme specificity with a maximum of two missed cleavage sites. To enable the quantification of both classes of N-terminal peptides, those with N-terminal iTRAQ label and those with endogenous N-acetylation, we performed a two-step search strategy: First, data were searched with iTRAQ 8-plex (+304.2053 Da) at peptide N-termini and iTRAQ 8-plex at Lys as fixed modifications. Second, N-terminal acetylation (+ 42.0105 Da) and iTRAQ 8-plex at Lys were set as fixed modifications. In both cases, carbamidomethylation of Cys (+ 57.0214 Da) was selected as fixed and oxidation (+ 15.9949 Da) of methionine as variable modifications. As filters, high confidence corresponding to an FDR < 1%, and search engine rank 1 were used. Mass tolerances were set to 10 ppm for MS and 0.5 Da for MS/MS.

PSMs representing the same *(i)* amino acid sequence and *(ii)* protein accession were grouped and their corresponding reporter ion areas summed (supplemental table 1). Summed areas were normalized using normalization factors derived from the global analysis, in order compensate for systematic differences in sample amount across channels. Normalized areas were used to calculate the ratios of staurosporine-

treated to control cells (∅ reporter ion area of three staurosporine replicates / average reporter ion area of three control replicates) and to determine p-values based on a 2-sides t-test. Data analysis was first conducted for each set individually and datasets were merged based on unique N-terminal peptides (*i.e.* concatenated sequence, Uniprot accession, modification set). Based on the corresponding unique entries, the overlap between the three HPLC and the three tip replicates was assessed (figure 2b). To compare HPLC and tip-based enrichment, only N-terminal peptides quantified at least twice with either of the two methods were considered (figure 2b). Next, *(i)* only unique N-terminal peptides (*i.e.* representing only a single Uniprot entry) that were *(ii)* at least three-fold regulated in *(iii)* at least 3 out of the 6 datasets were used to determine a consensus motif considering the amino acids p5-p5' according to the nomenclature of Schechter and Berger (12)(figure 2 c). The corresponding protein entries were used to generate a high confidence STRING protein network using Cytoscape (13)(http://www.cytoscape.org/).

4. Studying the dynamics of staurosporine-induced apoptosis using ChaFRA*tip*

4.1 Cell culture, sample preparation, ChaFRA*tip*, nano-LC-MS/MS, data analysis

Neuroblastoma SH-SY5Y cells were grown to more than 70% confluency (10^7 cells per T75-flask) in Dulbecco's Modified Eagle's Medium (PAN-Biotech). The medium was replaced by fresh medium containing either staurosporine 1 µM in DMSO (Sigma Aldrich) or DMSO (both 10 µL) and incubated for: 15 minutes, 1.5 hours, 3 hours and 6 hours in order to check for progression of apoptosis and activation of caspase proteases. At each time point, cells were examined under a microscope for changes in cell shape and detachment as indication for apoptosis (Supplemental Figure 1). The medium supernatant of each flask which contained floating cells was collected and cells was washed with ice-cold PBS. Cells were then treated with trypsin (0.025% in EDTA PBS) for 2-3 minutes at 37°C. The medium supernatant and the

trypsinized cells were pooled and cells isolated by centrifugation at 270 g for 5 minutes. Cell pellets were lysed in PBS containing 0.1% Triton X-100 and incubated with compound SV149 (an activity based probe for labelling active caspase proteases) at RT for 45 min. Labelling reactions were quenched by adding Laemmli sample buffer and half of the sample was resolved by SDS PAGE. Active caspases were visualized by fluorescent scanning of the TAMRA fluorescence using a Typhoon Trio+ (Supplemental Figure 1). For the MS samples, a similar procedure was applied. The cell pellets collected after centrifugation at 270 g were flash-frozen in liquid nitrogen and stored at -80°C for subsequent MS preparations. Only the last 3 time points (1.5, 3 and 6 hours) were harvested, including two biological replicates per time point. Two additional T75 flasks treated with DMSO for 6 hours were used as controls. Sample preparation including lysis, carbamidomethylation, and iTRAQ labeling was conducted as above, however, including 8 samples (4 conditions x 2 biological replicates) and iTRAQ labels. Protein concentrations were determined by amino acid analysis. After pooling, the sample was divided into 6 aliquots à 35 µg (amino acid analysis quantified). Then, in order to assess the robustness of the method three aliquots each were enriched for N-terminal peptides with ChaFRA*tip* by two different individuals. nano-LC-MS/MS measurements were done as described for the HPLC/tip comparison.

For data analysis endogenously acetylated peptides were excluded, even if quantified based on iTRAQ labeled Lys-residues. Each replicate was individually analyzed (supplemental table 2). PSMs quantifying the same unique entry (*i.e.* concatenated Uniprot accession and sequence) were grouped and the corresponding reporter ion intensities summed. Normalization was done as described above. For each biological replicate (A and B), the four corresponding reporter ion sums were used to calculate a median intensity (A and B). These median intensities were used to scale the corresponding reporter ion intensities.

Based on the unique entries, the six datasets were merged. For each replicate and time point, ratios compared to the corresponding control were determined using the reporter ion intensities and log2-transformed. Considering only N-termini that were quantified in at least 3 replicates per time point, the

median log2 ratio and a p-value based on a 2-sided t-test were determined per biological replicate (*i.e.* 1.5 h A and B, 3 h A and B, 6 h A and B). Only *(i)* unique N-terminal peptides that were *(ii)* at least 2-fold regulated compared to control and *(iii)* had a p-value <0.05 were considered as regulated for the respective time point. These regulated N-terminal peptides were considered to generate a heatmap using Perseus (14) and a STRING protein network using Cytoscape (figure 3).

4.2 Quantitative proteome and phosphoproteome: sample preparation and nano-LC-MS/MS

Sample preparation and proteolysis were performed using the FASP(5, 6) protocol with minor changes. Briefly, cell lysates corresponding to 50 µg (amino acid analysis quantified) of protein were diluted up to 450 µL with freshly prepared 8 M urea/100 mM Tris-HCl, pH 8.5(7) and placed on a Nanosep centrifugal device (30 KDa, PALL). The devices were centrifuged at 13,900 g at RT for 20 min. All the following centrifugation steps were performed under similar conditions i.e. 13,900 g, RT, 15 min. To remove residual SDS, three washing steps were carried out using 100 µL of 8 M urea/100 mM Tris-HCl, pH 8.5 and for the buffer exchange; the devices were washed thrice with 100 µL of 50 mM TEAB, pH 7.8. To the concentrated proteins, 100 µL of digestion buffer comprising 1:20 (w/w) sequencing grade trypsin, 0.2 M GuHCl, 2 mM $CaCl_2$ in 50 mM TEAB, pH 7.8 were added and incubated at 37°C for 14 h. The generated tryptic peptides were recovered by centrifugation followed by two consequent washing steps with 50 µL of 50 mM TEAB, pH 7.8 and 50 µL of ultra-pure water. Finally, peptides were acidified with 10% TFA (pH < 3.0) and digests were quality controlled as described previously (1). Next, all samples were completely dried in a SpeedVac and were resolubilized in 0.5 M TEAB, pH 8.5. The peptide concentration was determined using a NanoDrop 2000 UV-Vis spectrophotometer (Thermo Scientific, Germany).

Before labeling with iTRAQ 8plex reagents(8) each sample (~1 µg) was analyzed by nano-LC-MS/MS to adjust minor differences in sample amounts due to pipetting errors or incorrect protein/peptide

concentration estimation based on the alignment of total ion chromatograms (TICs). Each sample was labeled with iTRAQ reagents according to the manufacturer's instructions (AB SCIEX). The multiplexed sample (~400 µg) was divided into two parts, *i.e.* 50 µg for global proteome analysis and 350 µg for phosphopeptide enrichment and both parts were dried completely in a SpeedVac.

Phosphopeptide enrichment was done using TiO_2(15, 16) beads followed by hydrophilic interaction chromatography (HILIC)(17). In total three HILIC fractions were collected, dried completely and stored at -80°C until further use.

For the global proteome analysis, dried peptides were resolubilized in 200 µL of 0.5% TFA (pH < 3.0) and desalted using C18 SPEC tips (4 mg, Varian). Eluted peptides were dried in a SpeedVac and resolubilized in 10 mM ammonium formate, pH 8.0. 25 µg of multiplexed sample was fractionated by reversed-phase chromatography at pH 8.0 on a Biobasic C18, 0.5 x 150 mm, 5 µm particle size column using an UltiMate 3000 LC system (both Thermo Scientific, Germany) with buffers A: 10 mM ammonium formate, pH 8.0, and B: 84% ACN in 10 mM ammonium formate, pH 8.0. Peptides were loaded onto the column with buffer A at a flow rate of 12.5 µL/min and separation was carried out using the following gradient: 3% B for 10 min, 3-45% B in 40 min, 45-60% B in 5 min, 60-95% B in 5 min, 95% B hold for 5 min, 95%-3% B in 5 min and finally re-equilibrate the column with 3% B for 20 min. In total 16 fractions were collected at 1 min intervals from min 5 to 80 in a concatenation mode. Finally, all the fractions were completely dried in a SpeedVac and stored at -80°C until further use.

Dried HILIC fractions were resolubilized in 15 µL of 0.1% TFA and analyzed by nano-LC-MS/MS using an Ultimate 3000 nano RSLC system coupled to a Q Exactive HF mass spectrometer (both Thermo Scientific). Peptides were preconcentrated on a 75 µm x 2 cm C18 trapping column for 5 min using 0.1% TFA (v/v) with a flow rate of 20 µL/min followed by separation on a 75 µm x 50 cm C18 main column (both Acclaim Pepmap nanoviper, Thermo Scientific) with a 125 min LC gradient ranging from 3-35% of B (84% ACN in

0.1% FA) at a flow rate of 250 nL/min. The Q Exactive HF was operated in DDA mode and MS survey scans were acquired from m/z 300 to 1,500 at a resolution of 60,000 using the polysiloxane ion at m/z 371.1012 as lock mass (3). The fifteen most intense ions were isolated with a 0.4 m/z window and fragmented by higher energy collisional dissociation (HCD) with a normalized collision energy of 33%, taking into account a dynamic exclusion of 20 s. MS/MS were acquired at a resolution of 15,000. Automated gain control (AGC) target values and fill times were set to 3×10^6 and 110 ms for MS and 1×10^5 and 200 ms for MS/MS, respectively, with a minimum intensity threshold of 5×10^4.

All 16 pH 8.0 fractions were resolubilized in 45 µL of 0.1% TFA and 1/3 of each sample was analyzed by nano-LC-MS/MS as above, with minor changes. The twenty most intense ions were isolated for MS/MS, the dynamic exclusion was 30 s, AGC target values and fill times were 3×10^6 and 50 ms for MS and 2×10^5 and 200 ms for MS/MS, respectively with a minimum intensity threshold of 1×10^5.

During both global and phosphoproteome analyses, a vial with 10% (v/v) NH_4OH solution was placed in front of the nano-ESI source for charge state reduction as described previously by Thingholm et al (4).

4.3 Quantitative proteome and phosphoproteome: iTRAQ data normalization and statistical evaluation

For the global proteome, instead of working with 7 iTRAQ ratios against one reference sample, as provided by the Proteome Discoverer software, here ratios were transformed into normalized abundance values (NAVs). Thus, for statistical comparison a hypothetical ratio (*i.e.* 113/113) was generated in order to have eight data points in an experiment. The individual channel ratios of each protein were log2 transformed and for each channel a median over all proteins was (MD1) calculated. Next, a second median (MD2) was generated by taking the medians across all channels. MD2 was then subtracted from individual MD1s to deduce the normalization values (NVs) for each channel. These NVs were subsequently used to normalize the respective log2 transformed ratios for each protein/channel (normalized ratios, NR). Next, a third

median (MD3, scaling factor) was calculated by taking the NR for each protein across all channels. This MD3 was then subtracted from the NR of each protein to get NAVs for each protein. Finally, ratios were calculated by dividing the average NAVs or each time point, *i.e.* 1.5 h staurosporine, by the average NAVs of the controls. Only proteins quantified by at least 2 unique peptides were considered (supplemental table 3).

A similar strategy was employed for the phosphoproteome data analysis. However, reporter ion areas of unique phosphorylated PSMs were considered instead or ratios. Next, for unique phosphopeptides PSMs representing the same protein, peptide sequence, and phosphorylation site, areas were summed per channel, and ratios were calculated as aforementioned. Only unique phosphopeptides with phosphoRS probabilities >90% were considered (supplemental table 3).

The ratios of the 5877 and 5547 quantified high confidence phosphopeptides and proteins, respectively, were used to generate heatmaps using Perseus (figure 3).

4.4 Comparison of N-terminal peptide set from this study to Peptide Atlas

To analyze whether the N-terminal peptides identified in this study are subject to any bias as often observed for PTM enrichment methods, we compared our set of 3,086 unique N-terminal peptides with the Peptide Atlas (PA; October 2017, 1,222,862 peptide entries in total). Therefore, we applied the following strategy: *(i)* We extracted for each protein in PA protein the most N-terminal peptide entry if it had either semi-Arg-C or semi-trypsin specificity (N-terminus flexible, C-terminus specific), corresponding to our set of (neo) N-terminal peptides. This yielded 11,282 PA trypsin and 7,868 PA Arg-C peptides. *(ii)* We compared the datasets regarding physicochemical properties such as *(i)* theoretical peptide charge states as pH 2.0, *(ii)* pI values, *(iii)* GRAVY values, *(iv)* peptide length in number of amino acids, *(v)* number of acidic and *(vi)* basic amino acids. The results are summarized in Supplemental Figures 3-6.

References

1. Burkhart, J. M., Schumbrutzki, C., Wortelkamp, S., Sickmann, A., and Zahedi, R. P. (2012) Systematic and quantitative comparison of digest efficiency and specificity reveals the impact of trypsin quality on MS-based proteomics. *Journal of Proteomics* 75, 1454-1462.
2. Larsen, M. R., Graham, M. E., Robinson, P. J., and Roepstorff, P. (2004) Improved Detection of Hydrophilic Phosphopeptides Using Graphite Powder Microcolumns and Mass Spectrometry: Evidence for in Vivo Doubly Phosphorylated Dynamin I and Dynamin III. *Molecular & Cellular Proteomics* 3, 456-465.
3. Olsen, J. V., de Godoy, L. M. F., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005) Parts per Million Mass Accuracy on an Orbitrap Mass Spectrometer via Lock Mass Injection into a C-trap. *Molecular & Cellular Proteomics* 4, 2010-2021.
4. Thingholm, T. E., Palmisano, G., Kjeldsen, F., and Larsen, M. R. (2010) Undesirable Charge-Enhancement of Isobaric Tagged Phosphopeptides Leads to Reduced Identification Efficiency. *Journal of Proteome Research* 9, 4045-4052.
5. Manza, L. L., Stamer, S. L., Ham, A.-J. L., Codreanu, S. G., and Liebler, D. C. (2005) Sample preparation and digestion for proteomic analyses using spin filters. *PROTEOMICS* 5, 1742-1745.
6. Wisniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nat Meth* 6, 359-362.
7. Kollipara, L., and Zahedi, R. P. (2013) Protein carbamylation: In vivo modification or in vitro artefact? *PROTEOMICS* 13, 941-944.
8. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., and Pappin, D. J. (2004) Multiplexed Protein Quantitation in Saccharomyces cerevisiae Using Amine-reactive Isobaric Tagging Reagents. *Molecular & Cellular Proteomics* 3, 1154-1169.
9. Burkhart, J. M., Vaudel, M., Gambaryan, S., Radau, S., Walter, U., Martens, L., Geiger, J., Sickmann, A., and Zahedi, R. P. (2012) The first comprehensive and quantitative analysis of human platelet protein composition allows the comparative analysis of structural and functional pathways. *Blood* 120, e73.
10. Staes, A., Impens, F., Van Damme, P., Ruttens, B., Goethals, M., Demol, H., Timmerman, E., Vandekerckhove, J., and Gevaert, K. (2011) Selecting protein N-terminal peptides by combined fractional diagonal chromatography. *Nat. Protocols* 6, 1130-1141.
11. Mischerikow, N., van Nierop, P., Li, K. W., Bernstein, H. G., Smit, A. B., Heck, A. J., and Altelaar, A. F. (2010) Gaining efficiency by parallel quantification and identification of iTRAQ-labeled peptides using HCD and decision tree guided CID/ETD on an LTQ Orbitrap. *The Analyst* 135, 2643-2652.
12. Schechter, I., and Berger, A. (1968) On the active site of proteases. 3. Mapping the active site of papain; specific peptide inhibitors of papain. *Biochemical and biophysical research communications* 32, 898-902.
13. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 13, 2498-2504.
14. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., and Cox, J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature methods* 13, 731-740.
15. Palmisano, G., Parker, B. L., Engholm-Keller, K., Lendal, S. E., Kulej, K., Schulz, M., Schwämmle, V., Graham, M. E., Saxtorph, H., Cordwell, S. J., and Larsen, M. R. (2012) A Novel Method for the Simultaneous Enrichment, Identification, and Quantification of Phosphopeptides and Sialylated Glycopeptides Applied to a Temporal Profile of Mouse Brain Development. *Molecular & Cellular Proteomics* 11, 1191-1202.

16. Dickhut, C., Radau, S., and Zahedi, R. (2014) Fast, Efficient, and Quality-Controlled Phosphopeptide Enrichment from Minute Sample Amounts Using Titanium Dioxide. In: Martins-de-Souza, D., ed. *Shotgun Proteomics*, pp. 417-430, Springer New York.
17. Solari, F. A., Kollipara, L., Sickmann, A., and Zahedi, R. P. (2016) Two Birds with One Stone: Parallel Quantification of Proteome and Phosphoproteome Using iTRAQ. *Methods in molecular biology (Clifton, N.J.)* 1394, 25-41.