FastCAT accelerates absolute quantification of proteins by using multiple short non-purified chimeric standards

Ignacy Rzagalinski[1], Aliona Bogdanova[1], Bharath Kumar Raghuraman[1], Eric R. Geertsma[1], Lena Hersemann[1], Tjalf Ziemssen[2], Andrej Shevchenko[1*]

[1]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany
[2]Center of Clinical Neuroscience, Department of Neurology, University Hospital Carl Gustav Carus, Technical University of Dresden, 01307 Dresden, Germany

[*]Corresponding author:

Dr. Andrej Shevchenko
Tel: +49 351 210 2615
Email: shevchenko@mpi-cbg.de

ABSTRACT

Absolute (molar) quantification of proteins provides the analytical rationale for system-level modelling of diverse molecular mechanisms. FastCAT method employs multiple short (<50 kDa) stable-isotope labeled chimeric proteins (CPs) composed of concatenated quantotypic (Q-) peptides representing the quantified proteins. Each CP also comprises scrambled sequences of reference (R-) peptides that relate its abundance to a single protein standard (BSA). FastCAT not only alleviates the need in purifying CP or using SDS-PAGE, but also improves the accuracy, precision and dynamic range of the absolute quantifications by grouping Q-peptides according to the expected abundance of target proteins. We benchmarked FastCAT against the reference method of MS Western and tested it in the direct molar quantifications of neurological markers in human cerebrospinal fluid at the low ng/mL level.

Keywords: absolute quantification of proteins; MS Western; QconCAT; targeted quantitative proteomics; cerebrospinal fluid; neurodegeneration; neuroinflammation

INTRODUCTION

The role of absolute (molar) quantification of proteins is multifaceted. It determines stoichiometric ratios within molecular assemblies and metabolic pathways[1] and relates them to non-proteinous compounds *e.g.* co-factors, lipids or metabolites; provides reference values and ranges for diagnostically important proteins in liquid and solid biopsies.[2] Last but not least, it estimates the expression levels in cells and tissues serving as a quantitative denominator common to all omics sciences. In contrast to popular immunodetection methods (ELISA or Western blotting),[3] mass spectrometry quantifies proteins by comparing the abundance of endogenous quantotypic peptides with corresponding synthetic peptides having with exactly known concentration. Protein concentration is then inferred from the concentrations of quantotypic peptides. However, protein and peptide properties are unique and may vary significantly.[4] To support clinical diagnostics, it is necessary to quantify a selection of disease-related proteins whose molar abundance differs by several orders of magnitude. It is therefore not surprising that, in contrast to relative quantification, absolute quantitative methods lack generality and unification.

Methods for absolute quantification of proteins (reviewed in[5]) rely upon different types of internal standards, including (but not limited to) synthetic peptides (*e.g.* AQUA)[6]; full-length or partial protein sequences (*e.g.* PSAQ[7] and QPrEST[8], respectively) and also chimeric proteins composed of concatenated proteotypic peptides from many different proteins (QconCAT)[9] (reviewed in[10–13]). Because of the advances in gene synthesis, QconCAT offers several appealing qualities *e.g.* ease of multiplexing, support of targeted and mid- to large-scale quantification applied to individual proteins and complexes, biochemical pathways[14] or clinically relevant selection of proteins.[15,16] QconCAT chimeras are expressed in *E.coli*, enriched and purified by affinity chromatography and their stock concentration is determined

by amino acid analysis or some protein assays[17]. Alternatively, an additional (secondary) peptide concatenated standard (PCS) was used to quantitatively characterize multiple primary PCSs.[18] Also, a QconCAT chimera could include the sequence of [Glu1]-Fibrinopeptide B[19] that enabled its one-peptide quantification, although synthetic peptide standards should be used with caution.[20]

By using GeLC-MS/MS, MS Western workflow alleviated the need in making and standardizing a purified stock of CP. Also, CP standard was designed such that it included not only quantotypic (Q-)peptides, but also several reference (R-)peptides.[21] The bands of CP and of reference protein (BSA) were co-digested with gel slabs containing target proteins and the recovered peptides analyzed by LC-MS/MS. The abundance of CP was referenced to exactly known amount of BSA (available as NIST certified standard) and then the abundance of the target protein was calculated from the abundance of CP assuming that its full cleavage produced proteotypic peptides in the equimolar amount. MS Western quantification was based on the concordant values (CV <10 %) obtained from multiple (usually 2 to 4) Q-peptides per each protein of interest.[22,23]

Because of the use of SDS, GeLC-MS/MS helped to detect more membrane proteins. SDS PAGE also alleviated the need in purifying CP, yet it limited the analyses throughput. Although collecting hundreds of quantotypic peptides in large (up to 290 kDa) chimeras is appealing, it is practically inflexible because new proteins and / or peptides could not be added at will. Furthermore, the yielded Q-peptides are strictly equimolar, which hampers the quantification of proteins having drastically (more than 100-fold) different abundance. This, however, is often required for monitoring protein biomarkers.[24]

Here, we report on FastCAT (**Fast**-track Qcon**CAT**) method that preserves the accuracy and consistency of MS Western quantification, yet is faster, more flexible and easier to use particularly in translational proteomics applications. In contrast to MS Western, FastCAT

workflow relies on the parallel use of many relatively short (<50 kDa) non-purified CPs that comprise Q-peptides for the target proteins, but also scrambled R-peptides to reference the CP concentrations to the same BSA standard.

EXPERIMENTAL SECTION

Chemicals and reagents

LC-MS grade solvents (water, acetonitrile and isopropanol), formic acid (FA) and trifluoroacetic acid (TFA) were purchased from Fisher Scientific (Waltham, MA) or Merck (Darmstadt, Germany). Trypsin and trypsin/Lys-C proteases (MS grade) were from Promega (Madison, WI); RapiGest detergent from Waters (Eschborn, Germany); other common chemicals from Sigma-Aldrich (Munich, Germany). Polyacrylamide gradient gels (4-20 %) were from Serva Electrophoresis GmbH (Heidelberg, Germany). Protein standards: glycogen phosphorylase (GP), alcohol dehydrogenase (ADH), enolase (ENO) and ubiquitin (UBI) were purchased as a lyophilized powder from Sigma-Aldrich. Reference protein standard (BSA, Pierce grade, in ampoules) was from Thermo Fisher Scientific (Waltham, MA). Isotopically labeled amino acids ($^{13}C_6$,$^{15}N_4$-L-arginine (R) and $^{13}C_6$-L-lysine (K)) were purchased from Silantes GmbH (Munich, Germany).

Design and expression of chimeric protein standards

In total, six CP standards of different molecular weights (MW) were designed and expressed. General scheme of CP design is presented in Figure S1, while other details including MWs, isotopic enrichment labeling efficiency and full-length sequences are in Table S4. DNA sequences encoding CPs were codon-optimized for *E.coli* by Genscript on-line tool and

synthesized by Genscript. CPs were produced using pET backbone (Novagenn) and *E.coli* BL21 (DE3) (argA lysA) strain auxotrophic for arginine and lysine supplemented with $^{13}C_6$,$^{15}N_4$-L-arginine and $^{13}C_6$-L-lysine as described.[21]

Sample preparation for proteomics analyses

Polyacrylamide gels were stained with Coomassie CBB R250 and gel slabs corresponding to the targeted range of MW were excised. In-gel digestion with trypsin was performed as previously reported[21] with enzyme-to-substrate ratio of 1:50. *E.coli* background with spiked 4 standard proteins was on-pellet digested with trypsin (1:20) after protein precipitation with isopropyl alcohol according to the previously described protocol.[25] Aliquots of cerebrospinal fluid (CSF) of 20 µL volume were in-solution digested with trypsin/Lys-C protease mix (1:20) in the presence of RapiGest (Waters) detergent.[26] CSF samples were obtained from patients diagnosed with multiple sclerosis and stored as freshly frozen aliquots. All patients gave their prior written consent. The study was approved by the institutional review board of the University Hospital Dresden (EK348092014).

LC-MS/MS analyses

LC-MS/MS was performed on a Q Exactive HF (Thermo Scientific, Germany) hybrid tandem mass spectrometer coupled with Eksigent 400 nanoLC system (Sciex, Germany) using the Nanospray Flex source (Thermo Fisher Scientific, Germany). Protein digests were loaded onto a trap column for 5 min at 7 µL/min and separated on the Acclaim PepMap 100 column (C18, 3 µm, 75 um x 150 mm) using 120 min gradients (5-45% B) at 300 nL/min in data-dependent acquisition (DDA) or parallel reaction monitoring (PRM) modes (as specified). DDA and PRM methods consisted of MS1 scan from *m/z* 350 to 1700 with automatic gain control (AGC) target value of 3×10^6; maximum injection time (IT) of 60 ms and targeted mass resolution $R_{m/z=200}$ of

60,000. Top-12 DDA method employed the precursor isolation window of 1.6 Th; AGC of 1×10^5; maximum IT of 50 ms; $R_{m/z=200}$ of 15,000; normalized collision energy (NCE) of 25% and dynamic exclusion of 30 s. The scheduled PRM method acquired MS/MS using 10-min retention time (RT) windows with the inclusion list of 97 precursors; precursor isolation window of 1.6 Th; AGC of 1×10^6; a maximum IT of 80 ms; $R_{m/z=200}$ of 30,000 and NCE of 25%.

Data processing and analysis

Raw LC-MS/MS data from DDA experiments were processed by Progenesis LC-MS v.4.1 (Nonlinear Dynamics, UK) software for RT alignment, peak picking and extracting the peptide features. To match peptides to target proteins, MS/MS spectra were searched by Mascot v.2.2.04 software (Matrix Science, London, UK) against a customized database containing sequences of all target proteins and relevant (either *E.coli* or human) background proteome. Precursor mass tolerance of 5 ppm; fragment mass tolerance of 0.03 Da; fixed modification: carbamidomethyl (C); variable modifications: acetyl (protein N-terminus) and oxidation (M); labels: $^{13}C_6$(K) and $^{13}C_6{}^{15}N_4$ (R), cleavage specificity: trypsin with up to 2 missed cleavages allowed). All PRM data sets were analyzed with Skyline 21.1.0.278 software.[27] Peak integration was inspected manually. Mass transitions in labeled and unlabeled peptides and matching retention time and peak boundaries were confirmed the peptide identities. A minimum of five transitions were required for the correct identification of the targeted peptides. In addition, the comparison of the measured fragment spectrum to the *in-silico* Prosit-derived[28] library spectrum by the normalized spectral contrast angle was used that resulted in the library dot product (dotp) correlation values of 0.87 or higher.

RESULTS AND DISCUSSION

Using crude CP standards for protein quantification

Trypsin cleavage of purified CP produces Q- and R-peptides in a strictly equimolar concentration.[21,29,30] However, in crude extracts the concentration balance between individual Q-peptides could be affected by non-proportional contribution of products of intra-cell proteolysis of CP and/or incomplete translation of the CP gene. Also, despite CP are highly expressed,[21] their purification reduces the load of an expression host (*E.coli*) proteins and their contribution to overall compositional complexity of the analyzed sample is unclear.

Therefore, we set out to test under what conditions the full-length CP standards could be employed with no further purification. To this end, we selected three CP standards spanning a wide MW range (CP01 ~265 kDa, CP02 ~79kDa and CP03 ~42 kDa) (Supplementary Table S4). We loaded protein extracts onto 1D SDS-PAGE, excised the CP bands, sliced the entire gel slab in several MW ranges below the band of CPs and analyzed them separately by GeLC-MS/MS (Figure 1). We observed that the relative abundance of Q-peptides depends on molecular weights of the full-length CP and its truncated forms, but also on the peptide location within the CP sequence. In 265 kDa CP peptides located closer to its N-terminus were overrepresented and constituted 40 to 80 % of total peptide abundances (Figure 1A). We also observed the same trend for the middle-size CP with 20 to 40 % fraction of truncated forms and slight prevalence of N-terminally located peptides (Figure 1B). However, in the shortest (42 kDa) CP the contribution of truncated forms was minor (ca 10-15 %) and independent of peptides location (Figure 1C).

Because of lower expression and high fraction of proteoforms having incomplete sequence truncated elsewhere at the C-terminal side, larger CP required extensive purification. Shorter CPs (less than *ca.* 50 kDa) are highly expressed with much lower fraction of truncated

forms having no location prevalence. Consistently, the relative abundances of Q-peptides at different locations within the CP sequence (N-terminal *vs.* middle ("M") *vs.* C-terminal) in the purified CP and non-purified CP (CP band together with truncation products) were very close (Figure 1E). Therefore, we concluded that spiking the total *E.coli* extract without isolating the full-length CP should not bias the quantification.

Since short CPs are highly expressed in *E.coli* they are usually diluted more than 100-fold down to *ca.* 1 µM concentration. Therefore, we wondered if adding the extract of *E.coli* containing an appropriate amount of a non-purified CP elevated protein background. For reliable comparison, we mixed a volume of metabolically labeled extract containing typical working amounts of CP (*ca.* 100 fmol) with an equivalent volume of unlabeled extract containing *ca.* 500 ng of total protein. We then compared relative abundances of labeled and unlabeled forms of nine major *E.coli* proteins (Table S1) and concluded that adding an extract with unpurified CP increased protein background by as little as 2 % (Figure S2).

We therefore concluded that short ~50 kDa CPs could be spiked into quantified samples as total (crude) *E.coli* lysates with no prior purification. With that size, they would be typically encoding for 25 to 30 Q-peptides and 3 to 5 R-peptides within a typical CP construct.[21]

FastCAT workflow: the concept and its validation

We reasoned that, by employing short non-purified CP, a targeted absolute quantification workflow (termed FastCAT for **Fast**-track Qcon**CAT**) could significantly accelerate the analysis. Besides Q-peptides used to quantify target proteins, a typical FastCAT construct contains multiple (usually 3 to 5) R-peptides used to determine the *in-situ* CP concentration by referencing it to the known amount of spiked-in BSA. Hence, FastCAT workflow requires neither CP purification (externally or by gels) nor an external determination

of its concentration, while the multi-peptide quantification procedure is the same as in MS Western.

We cross-validated FastCAT by comparing it against MS Western.[21] To this end, we prepared an approximately equimolar (*ca.* 1 µM) mixture of the 4 standard proteins (GP, UBI, ADH and ENO) and spiked it (*ca.* 20 pmol of each protein) into *E.coli* background (*ca.* 50 µg of total protein). We determined their exact quantities by MS Western using 42 kDa (CP03) as an internal standard.[21] In parallel, we quantified them by FastCAT protocol using the same 42 kDa construct, but without 1D SDS-PAGE. Importantly, we did the quality check of the digestion completeness for both methods by comparing relative abundances of labeled Q-peptides in CP03 and corresponding unlabeled peptides in the endogenous proteins. The difference in protein quantities determined by FastCAT and MS Western were below 15 % for all proteins (Table 1).

Since in FastCAT workflow CP was not purified, we additionally checked if the positioning of Q- and R-peptides within the CP backbone affected the quantification. To this end, we designed another CP standard, CP04, having the same Q-peptides as in CP03, but distributing the same five R-peptides (R1-R5) over the entire CP sequence compared to a single block of R-peptides at the C-terminus of CP03 (see Figure 2).

This experiment resulted in two major findings. Firstly, using CP03 and CP04 independently led to concordant quantification. With all R-peptides used for calculating CPs abundances, the quantities of target proteins differed by less than 15%. Secondly, the quantification was practically unaffected by positioning of R-peptides. Indeed, using differentially positioned R-peptides (N (R1/R5) *vs.* “M” (R2) *vs.* C (R3/R4)) in CP04 led to the quantification of all target protein with CV *ca.* 9 % (Figure 2A). This minor variability could not be solely attributed to the different location of R-peptides. A similar value of CV *ca.* 7 %

was observed when target proteins were quantified using CP03 with similarly positioned R-peptides (R1/R5 *vs.* R2 *vs.* R3/R4, all at C-terminus) as in CP04 (Figure 2B).

We therefore concluded that in the FastCAT workflow the design of CP and relative positioning of R-peptides have no major impact on the proteins quantification.

Multiplexing of FastCAT

Relatively short CP standards having the MW of 40 to 50 kDa will typically comprise 20 to 30 Q-peptides. Considering that the robust protein quantification would require 3 to 5 peptides per each protein, one CP would enable the quantification of 5 to 8 individual proteins of, preferably, similar abundance. Hence, there is a need in multiplexing the FastCAT quantification capacity.

We therefore proposed to group proteotypic peptides into CPs according to the expected abundance of target proteins and then to use multiple CPs in parallel to cover the desired number of proteins (Figure 3). However, the abundance of different CPs should be referenced to the same spiked-in BSA standard. We achieved it by including reference peptides with the scrambled sequences (R_s-peptides)[31] that, nevertheless, elicit very similar response in MS1 spectra compared to corresponding native R_n-peptides. However, we reasoned that, for better sensitivity, the targeted analysis would require PRM also for quantifying CPs in the same LC-MS/MS run. Therefore, for each R_n/R_s peptides pair, we selected a representative combination of fragment ions that adequately reflected the peptides abundance[32,33] and, hence, enable parallel quantification of multiple CPs.

To this end, we designed CP05 and CP06 proteins (see Supplementary Figure S3 for amino acid sequences and Figure S4 for the distributions of their truncated forms) containing 42 Q-peptides from 10 selected human proteins (these CPs were further used for the case study in the section below). They also comprised 10 R-peptides as 5 pairs of native (R_n) and

scrambled (R_s) sequences. However, R-peptides were placed into CP05 and CP06 such that they contained 3 native plus 2 scrambled, and 2 native plus 3 scrambled R-peptides, respectively (Figure 4). To emulate the impact of protein background, we spiked these CPs in *ca.* 1 µM concentration into a 20 µL aliquot of human cerebrospinal fluid (CSF) and analyzed by the method of PRM.

For each R_n/R_s peptide pair we compared the distribution of abundances of the most intense fragments within y- and b-ions (see Supplementary Figure S5 for MS2 spectra). Out of 5 pairs, three pairs produced very similar profiles, while two other pairs mismatched (Figure 5).

Next, we assessed the concordance of the *in-situ* PRM-based quantification of CPs using each R_s-peptide with the reference values obtained by R_n-peptides. To do this, we considered three quantification scenarios based on the selection of different combinations of fragment ions: Top1, Top3(y) and Top6(y,b) (as exemplified in Figure 5); and compared them to the values computed from the abundance of MS1-peaks (see Supplementary Table S2 for details). As expected, the Top1 approach not only resulted in the highest error (>35 %) for the two mismatching R_n/R_s (#2 and #3) pairs, but also revealed the highest overall (for all R_n/R_s pairs) discordance with MS1 (Figure 6). In contrast, summing up the abundances of more fragments (in Top3(y) or Top6(y,b) scenarios) compensated minor differences in the fragmentation patterns of R_n/R_s-peptides. We note that for low resolution selected reaction monitoring (SRM) experiments, that cannot follow many transitions and using b-ions should rather be avoided,[34] the Top3(y) approach could be the most practical compromise.

However, for PRM-based quantification summing up the abundance of the 6 most abundant y- and b-fragment ions will lead to the most consistent estimates with lower than 20 % difference to MS1-based determinations (Figure 6). Eventually, within the two CPs the

relative abundances of all used (both native and scrambled) R-peptides measured via Top6(y,b) fragments differed by less than 5% (Supplementary Figure S6).

We therefore concluded that multiplexing of FastCAT workflow could be achieved by using several CPs comprising multiple scrambled peptides of BSA. Quantities of individual CPs could be referenced to the same BSA standard by PRM analysis that relies on the summation of the six most intense y- and b-ions for each R_s (in the CP) and R_n (in BSA standard) peptide pair.

Case study: absolute quantification of neurological protein markers in the human CSF

Multiple sclerosis is an immune-mediated demyelinating and neurodegenerative disease of the central nervous system (CNS), which is usually accompanied by blood-brain barrier disruption, infiltration of immune cells into the CNS, and nerve fibers demyelination and axonal loss.[35] It alters the CSF proteome and monitoring the levels of protein markers by common clinical chemistry methods (*e.g.* ELISA) aids in its molecular diagnostics.[36] However, these methods suffer from low concordance, limited scope and substantial costs. Here we employed FastCAT to determine the molar concentrations of a selection of protein markers having a broad range of physicochemical properties, molar abundance and magnitude of response towards the disease.

We obtained 11 samples of CSF from five (four female and one male) 30 to 61 years of age patients that were diagnosed with relapsing-remitting multiple sclerosis. CSF was drawn from each patient at two time points: the first puncture was performed during the initial diagnostics while the second (and, for one patient, also the third) puncture was performed *ca.* 2 years later prior to a planned treatment switch to validate that no significant inflammation and neurodestruction occurred. Based on clinical indications, ten protein markers were selected out of *ca.* 700 proteins detected in a pooled CSF sample in the course of preliminary experiment.

Those included two major lipoproteins (APOE and APOD); inflammation-related glycoproteins (AACT, ZAG and LRG1); markers of axonal (CNTN1) and synaptic (NPTX1 and VGF) related disorders; a member of the granins family (SCG1) and, finally, a neuroinflammatory marker (CH3L1) typically increased in patients with multiple sclerosis.[37,38] We then selected 42 Q-peptides and assembled them in CP05 and CP06 (also mentioned above) according to the arbitrary abundance of target proteins.

The method precision was evaluated by processing and analyzing the pooled CSF sample in triplicate (Figure S7). For both CP and target proteins peptide and protein levels the median coefficient of variation was below 6 %. Importantly, peptides originating from the same protein led to their highly concordant quantification as exemplified by median CV of *ca.* 12 %.

The median values and the ranges of variation of target proteins are reported in Table 2, which also include Q-peptides and the estimates of quantification concordance for the independent quantification by multiple peptides. Concentration ranges determined by FastCAT showed particularly good concordance with previously reported SRM and PRM determinations. For instance, similar concentration ranges were obtained for APOE,[39,40] AACT,[41] SCG1[42] and CH3L1.[43] At the same time, the concordance with ELISA measurements was limited for both FastCAT and published SRM/PRM values. While ranges determined by ELISA for AACT,[44] LRG1,[45,46] CH3L1[43,47,48] and CNTN1[44] were close to FastCAT, levels for both apolipoproteins were discordant (yet, again, concordant with SRM/PRM).[49–52] This, however, is consistent with the previously described discrepancies between ELISA and mass spectrometry.[39] Molar concentration of ZAG, VGF and NPTX1 was not reported previously.

Concentrations determined in individual patients (Supplementary Table S3) were stable over the two-year period of treatment and, except one patient, they clustered together at the PCA plot (supplementary Figure S8). PCA plot singled out one female patient presumably

because of her higher age, although the proteins concentration in both of her biopsies were concordant.

Several trends (*e.g.* increase in APOD and ZAG; decrease in NPTX1) corroborated previous reports.[53–55] At the same time, there was no consistent change in the levels of prospective multiple sclerosis markers SCG1 and LRG1.[46,56]

Notably, the amount of spiked CP05 standard was ca 10-fold higher than of CP06 (309 fmol/column *vs.* 33 fmol/column, respectively). For better consistency, BSA was spiked at some intermediate amount (100 fmol/column). In this way, PRM covered 100-fold range of concentrations from *ca.* 20 ng/mL for CH3L1 to *ca.* 2000 ng/mL for APOD without breaking inter-peptide quantification consistency. Considering the observed signal-to-noise ratios, PRM was not close to the limit of detection and, in principle, should allow us to reach 10-fold higher sensitivity if the appropriate amount of yet another CP standard is spiked into CSF sample. Taken together, we demonstrated that FastCAT supported direct molar quantification of 10 neurological protein markers is CSF at the low ng/mL levels, better that 100-fold dynamic range and good (CV <20 %) inter-peptide quantification consistency based on two to four peptides per each protein.

CONCLUSION AND PERSPECTIVES

Absolute quantification of proteins offers heavily missing data on the concentrations (or molar abundances) of proteins in cells, tissues and biofluids. FastCAT workflow combines flexibility and ease of use of peptide-based quantitation. It relies on a large number of peptide standards having the exactly known equimolar concentration that are produced *in-situ* and quantified in the same LC-MS/MS run together with peptides from target proteins. Controlled

scrambling of R-peptides together with alternative strategies in metabolic labeling offer practically infinite capabilities for multiplexing. The quantification could be organized in a "modular" fashion by combining CPs of the desired composition and concentration.

In the future, it might be practical to set up a repository of plasmids encoding CPs and make them publicly available at some nominal fees. This will improve the analyses consistency and, eventually, brining absolute quantification closer to clinical chemistry requirements. While throughput of LC-MS/MS quantification could hardly compete with ELISA, its accuracy, independence of antibodies, quantification transparency and analytical flexibility, including compatibility with major protocols for biochemical enrichment and robotic sample preparation, might be appealing in many translational applications. We also argue that the interested laboratories should work together towards benchmarking and validating the absolute quantification methods by ring trials that are now common in neighboring omics fields, *e.g.* lipidomics.[57]

ACKNOWLEDGMENTS

Work in Shevchenko group is supported by Max Planck Society. The authors are grateful for other group members for expert technical support and useful discussions.

REFERENCES

(1) Bantscheff, M.; Lemeer, S.; Savitski, M. M.; Kuster, B. Quantitative Mass Spectrometry in Proteomics: Critical Review Update from 2007 to the Present. *Anal. Bioanal. Chem.* **2012**, *404* (4), 939–965. https://doi.org/10.1007/s00216-012-6203-4.

(2) Füzéry, A. K.; Levin, J.; Chan, M. M.; Chan, D. W. Translation of Proteomic Biomarkers into FDA Approved Cancer Diagnostics: Issues and Challenges. *Clin. Proteomics* **2013**, *10* (1). https://doi.org/10.1186/1559-0275-10-13.

(3) Ren, A. H.; Diamandis, E. P.; Kulasingam, V. Uncovering the Depths of the Human Proteome: Antibody-Based Technologies for Ultrasensitive Multiplexed Protein Detection and Quantification. *Mol. Cell. Proteomics* **2021**, *20*, 100155. https://doi.org/10.1016/J.MCPRO.2021.100155.

(4) Boja, E. S.; Fehniger, T. E.; Baker, M. S.; Marko-Varga, G.; Rodriguez, H. Analytical Validation Considerations of Multiplex Mass-Spectrometry-Based Proteomic Platforms for Measuring Protein Biomarkers. *J. Proteome Res.* **2014**, *13* (12), 5325–5332. https://doi.org/10.1021/pr500753r.

(5) Calderón-Celis, F.; Encinar, J. R.; Sanz-Medel, A. Standardization Approaches in Absolute Quantitative Proteomics with Mass Spectrometry. *Mass Spectrom. Rev.* **2018**, *37* (6), 715–737. https://doi.org/10.1002/mas.21542.

(6) Gerber, S. A.; Rush, J.; Stemman, O.; Kirschner, M. W.; Gygi, S. P. Absolute Quantification of Proteins and Phosphoproteins from Cell Lysates by Tandem MS. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (12), 6940–6945. https://doi.org/10.1073/pnas.0832254100.

(7) Picard, G.; Lebert, D.; Louwagie, M.; Adrait, A.; Huillet, C.; Vandenesch, F.; Bruley, C.; Garin, J.; Jaquinod, M.; Brun, V. PSAQ^{TM} Standards for Accurate MS-Based Quantification of Proteins: From the Concept to Biomedical Applications. *J. Mass Spectrom.* **2012**, *47* (10), 1353–1363. https://doi.org/10.1002/jms.3106.

(8) Zeiler, M.; Straube, W. L.; Lundberg, E.; Uhlen, M.; Mann, M. A Protein Epitope Signature Tag (PrEST) Library Allows SILAC-Based Absolute Quantification and Multiplexed Determination of Protein Copy Numbers in Cell Lines. *Mol. Cell. Proteomics* **2012**, *11* (3), O111.009613. https://doi.org/10.1074/mcp.O111.009613.

(9) Beynon, R. J.; Doherty, M. K.; Pratt, J. M.; Gaskell, S. J. Multiplexed Absolute Quantification in Proteomics Using Artificial QCAT Proteins of Concatenated Signature Peptides. *Nat. Methods* **2005**, *2* (8), 587–589. https://doi.org/10.1038/nmeth774.

(10) Brun, V.; Masselon, C.; Garin, J.; Dupuis, A. Isotope Dilution Strategies for Absolute Quantitative Proteomics. *J. Proteomics* **2009**, *72* (5), 740–749. https://doi.org/10.1016/j.jprot.2009.03.007.

(11) Villanueva, J.; Carrascal, M.; Abian, J. Isotope Dilution Mass Spectrometry for Absolute Quantification in Proteomics: Concepts and Strategies. *J. Proteomics* **2014**, *96*, 184–199. https://doi.org/10.1016/j.jprot.2013.11.004.

(12) Shuford, C. M.; Walters, J. J.; Holland, P. M.; Sreenivasan, U.; Askari, N.; Ray, K.; Grant, R. P. Absolute Protein Quantification by Mass Spectrometry: Not as Simple as Advertised. *Anal. Chem.* **2017**, *89* (14), 7406–7415. https://doi.org/10.1021/acs.analchem.7b00858.

(13) Scott, K. B.; Turko, I. V.; Phinney, K. W. Quantitative Performance of Internal Standard Platforms for Absolute Protein Quantification Using Multiple Reaction Monitoring-Mass Spectrometry. *Anal. Chem.* **2015**, *87* (8), 4429–4435. https://doi.org/10.1021/acs.analchem.5b00331.

(14) Carroll, K. M.; Simpson, D. M.; Eyers, C. E.; Knight, C. G.; Brownridge, P.; Dunn, W. B.; Winder, C. L.; Lanthaler, K.; Pir, P.; Malys, N.; et al. Absolute Quantification of the

Glycolytic Pathway in Yeast: Deployment of a Complete QconCAT Approach. *Mol. Cell. Proteomics* **2011**, *10* (12), 1–15. https://doi.org/10.1074/mcp.M111.007633.

(15) Al-Majdoub, Z. M.; Al Feteisi, H.; Achour, B.; Warwood, S.; Neuhoff, S.; Rostami-Hodjegan, A.; Barber, J. Proteomic Quantification of Human Blood-Brain Barrier SLC and ABC Transporters in Healthy Individuals and Dementia Patients. *Mol. Pharm.* **2019**, *16* (3), 1220–1233. https://doi.org/10.1021/acs.molpharmaceut.8b01189.

(16) Chen, J.; Wang, M.; Turko, I. V. Quantification of Amyloid Precursor Protein Isoforms Using Quantification Concatamer Internal Standard. *Anal. Chem.* **2013**, *85* (1), 303–307. https://doi.org/10.1021/ac3033239.

(17) Pratt, J. M.; Simpson, D. M.; Doherty, M. K.; Rivers, J.; Gaskell, S. J.; Beynon, R. J. Multiplexed Absolute Quantification for Proteomics Using Concatenated Signature Peptides Encoded by QconCAT Genes. *Nat. Protoc.* **2006**, *1* (2), 1029–1043. https://doi.org/10.1038/nprot.2006.129.

(18) Kito, K.; Okada, M.; Ishibashi, Y.; Okada, S.; Ito, T. A Strategy for Absolute Proteome Quantification with Mass Spectrometry by Hierarchical Use of Peptide-Concatenated Standards. *Proteomics* **2016**, *16* (10), 1457–1473. https://doi.org/10.1002/pmic.201500414.

(19) Bennett, R. J.; Simpson, D. M.; Holman, S. W.; Ryan, S.; Brownridge, P.; Eyers, C. E.; Colyer, J.; Beynon, R. J. DOSCATs: Double Standards for Protein Quantification. *Sci. Rep.* **2017**, *7* (February), 1–12. https://doi.org/10.1038/srep45570.

(20) Shuford, C. M.; Sederoff, R. R.; Chiang, V. L.; Muddiman, D. C. Peptide Production and Decay Rates Affect the Quantitative Accuracy of Protein Cleavage Isotope Dilution Mass Spectrometry (PC-IDMS). *Mol. Cell. Proteomics* **2012**, *11* (9), 814–823. https://doi.org/10.1074/mcp.O112.017145.

(21) Kumar, M.; Joseph, S. R.; Augsburg, M.; Bogdanova, A.; Drechsel, D.; Vastenhouw, N. L.; Buchholz, F.; Gentzel, M.; Shevchenko, A. MS Western, a Method of Multiplexed Absolute Protein Quantification Is a Practical Alternative to Western Blotting. *Mol. Cell. Proteomics* **2018**, *17* (2), 384–396. https://doi.org/10.1074/mcp.O117.067082.

(22) Penkov, S.; Raghuraman, B. K.; Erkut, C.; Oertel, J.; Galli, R.; Ackerman, E. J. M.; Vorkel, D.; Verbavatz, J.-M.; Koch, E.; Fahmy, K. A Metabolic Switch Regulates the Transition between Growth and Diapause in C. Elegans. *BMC Biol.* **2020**, *18* (1), 1–20.

(23) Raghuraman, B. K.; Hebbar, S.; Kumar, M.; Moon, H. K.; Henry, I.; Knust, E.; Shevchenko, A. Absolute Quantification of Proteins in the Eye of Drosophila Melanogaster. *Proteomics* **2020**, *20* (23), 1–10. https://doi.org/10.1002/pmic.201900049.

(24) Borg, J.; Campos, A.; Diema, C.; Omeñaca, N.; De Oliveira, E.; Guinovart, J.; Vilaseca, M. Spectral Counting Assessment of Protein Dynamic Range in Cerebrospinal Fluid Following Depletion with Plasma-Designed Immunoaffinity Columns. *Clin. Proteomics* **2011**, *8* (1). https://doi.org/10.1186/1559-0275-8-6.

(25) Knittelfelder, O.; Traikov, S.; Vvedenskaya, O.; Schuhmann, A.; Segeletz, S.; Shevchenko, A.; Shevchenko, A. Shotgun Lipidomics Combined with Laser Capture Microdissection: A Tool to Analyze Histological Zones in Cryosections of Tissues. *Anal. Chem.* **2018**, *90* (16), 9868–9878. https://doi.org/10.1021/acs.analchem.8b02004.

(26) Barkovits, K.; Tönges, L.; Marcus, K. CSF Sample Preparation for Data-Independent Acquisition. In *Methods in Molecular Biology*; 2019; Vol. 2044, pp 61–67. https://doi.org/10.1007/978-1-4939-9706-0_4.

(27) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: An Open Source Document Editor for Creating and Analyzing Targeted Proteomics Experiments. *Bioinformatics* **2010**, *26* (7), 966–968. https://doi.org/10.1093/bioinformatics/btq054.

(28) Gessulat, S.; Schmidt, T.; Zolg, D. P.; Samaras, P.; Schnatbaum, K.; Zerweck, J.; Knaute, T.; Rechenberger, J.; Delanghe, B.; Huhmer, A.; et al. Prosit: Proteome-Wide Prediction of Peptide Tandem Mass Spectra by Deep Learning. *Nat. Methods* **2019**, *16* (6), 509–518. https://doi.org/10.1038/s41592-019-0426-7.

(29) Simpson, D. M.; Beynon, R. J. QconCATs: Design and Expression of Concatenated Protein Standards for Multiplexed Protein Quantification. *Anal. Bioanal. Chem.* **2012**, *404* (4), 977–989. https://doi.org/10.1007/s00216-012-6230-1.

(30) Scott, K. B.; Turko, I. V.; Phinney, K. W. *QconCAT: Internal Standard for Protein Quantification*, 1st ed.; Elsevier Inc., 2016; Vol. 566. https://doi.org/10.1016/bs.mie.2015.09.022.

(31) Raghuraman, B. K.; Bogdanova, A.; Moon, H.; Rzagalinski, I.; Geertsma, E. R.; Hersemann, L.; Shevchenko, A. Median-Based Absolute Quantification of Proteins Using Fully Unlabeled Generic Internal Standard (FUGIS). *J. Proteome Res.* **2021**. https://doi.org/10.1021/acs.jproteome.1c00596.

(32) Winter, D.; Seidler, J.; Kugelstadt, D.; Derrer, B.; Kappes, B.; Lehmann, W. D. Minimally Permutated Peptide Analogs as Internal Standards for Relative and Absolute Quantification of Peptides and Proteins. *Proteomics* **2010**, *10* (7), 1510–1514. https://doi.org/10.1002/pmic.200900695.

(33) Remily-Wood, E. R.; Koomen, J. M. Evaluation of Protein Quantification Using Standard Peptides Containing Single Conservative Amino Acid Replacements. *J. Mass Spectrom.* **2012**, *47* (2), 188–194. https://doi.org/10.1002/jms.2053.

(34) Lange, V.; Picotti, P.; Domon, B.; Aebersold, R. Selected Reaction Monitoring for Quantitative Proteomics: A Tutorial. *Mol. Syst. Biol.* **2008**, *4* (222). https://doi.org/10.1038/msb.2008.61.

(35) Dendrou, C. A.; Fugger, L.; Friese, M. A. Immunopathology of Multiple Sclerosis. *Nat. Rev. Immunol.* **2015**, *15* (9), 545–558. https://doi.org/10.1038/nri3871.

(36) Novakova, L.; Zetterberg, H.; Sundström, P.; Axelsson, M.; Khademi, M.; Gunnarsson, M.; Malmeström, C.; Svenningsson, A.; Olsson, T.; Piehl, F.; et al. Monitoring Disease Activity in Multiple Sclerosis Using Serum Neurofilament Light Protein. *Neurology* **2017**, *89* (22), 2230–2237. https://doi.org/10.1212/WNL.0000000000004683.

(37) Canto, E.; Tintore, M.; Villar, L. M.; Costa, C.; Nurtdinov, R.; Alvarez-Cermeno, J. C.; Arrambide, G.; Reverter, F.; Deisenhammer, F.; Hegen, H.; et al. Chitinase 3-like 1: Prognostic Biomarker in Clinically Isolated Syndromes. *Brain* **2015**, *138* (4), 918–931. https://doi.org/10.1093/brain/awv017.

(38) Comabella, M.; Fernández, M.; Martin, R.; Rivera-Vallvé, S.; Borrás, E.; Chiva, C.; Juli, E.; Rovira, A.; Cantó, E.; Alvarez-Cermeño, J. C.; et al. Cerebrospinal Fluid Chitinase 3-like 1 Levels Are Associated with Conversion to Multiple Sclerosis. *Brain* **2010**, *133* (4), 1082–1093. https://doi.org/10.1093/brain/awq035.

(39) Martínez-Morillo, E.; Nielsen, H. M.; Batruch, I.; Drabovich, A. P.; Begcevic, I.; Lopez, M. F.; Minthon, L.; Bu, G.; Mattsson, N.; Portelius, E.; et al. Assessment of Peptide Chemical Modifications on the Development of an Accurate and Precise Multiplex Selected Reaction Monitoring Assay for Apolipoprotein e Isoforms. *J. Proteome Res.* **2014**, *13* (2), 1077–1087. https://doi.org/10.1021/pr401060x.

(40) Minta, K.; Brinkmalm, G.; Janelidze, S.; Sjödin, S.; Portelius, E.; Stomrud, E.; Zetterberg, H.; Blennow, K.; Hansson, O.; Andreasson, U. Quantification of Total Apolipoprotein e and Its Isoforms in Cerebrospinal Fluid from Patients with Neurodegenerative Diseases. *Alzheimer's Res. Ther.* **2020**, *12* (1), 1–11. https://doi.org/10.1186/s13195-020-00585-7.

(41) Andersson, A.; Remnestål, J.; Nellgård, B.; Vunk, H.; Kotol, D.; Edfors, F.; Uhlén, M.; Schwenk, J. M.; Ilag, L. L.; Zetterberg, H.; et al. Development of Parallel Reaction

Monitoring Assays for Cerebrospinal Fluid Proteins Associated with Alzheimer's Disease. *Clin. Chim. Acta* **2019**, *494* (March), 79–93. https://doi.org/10.1016/j.cca.2019.03.243.

(42) Zhu, S.; Wuolikainen, A.; Wu, J.; Öhman, A.; Wingsle, G.; Moritz, T.; Andersen, P. M.; Forsgren, L.; Trupp, M. Targeted Multiple Reaction Monitoring Analysis of CSF Identifies UCHL1 and GPNMB as Candidate Biomarkers for ALS. *J. Mol. Neurosci.* **2019**, *69* (4), 643–657. https://doi.org/10.1007/s12031-019-01411-y.

(43) Stoop, M. P.; Singh, V.; Stingl, C.; Martin, R.; Khademi, M.; Olsson, T.; Hintzen, R. Q.; Luider, T. M. Effects of Natalizumab Treatment on the Cerebrospinal Fluid Proteome of Multiple Sclerosis Patients. *J. Proteome Res.* **2013**, *12* (3), 1101–1107. https://doi.org/10.1021/pr3012107.

(44) Ottervald, J.; Franzén, B.; Nilsson, K.; Andersson, L. I.; Khademi, M.; Eriksson, B.; Kjellström, S.; Marko-Varga, G.; Végvári, Á.; Harris, R. A.; et al. Multiple Sclerosis: Identification and Clinical Evaluation of Novel CSF Biomarkers. *J. Proteomics* **2010**, *73* (6), 1117–1132. https://doi.org/10.1016/j.jprot.2010.01.004.

(45) Miyajima, M.; Nakajima, M.; Motoi, Y.; Moriya, M.; Sugano, H.; Ogino, I.; Nakamura, E.; Tada, N.; Kunichika, M.; Arai, H. Leucine-Rich A2-Glycoprotein Is a Novel Biomarker of Neurodegenerative Disease in Human Cerebrospinal Fluid and Causes Neurodegeneration in Mouse Cerebral Cortex. *PLoS One* **2013**, *8* (9), 1–12. https://doi.org/10.1371/journal.pone.0074453.

(46) Chong, P. F.; Sakai, Y.; Torisu, H.; Tanaka, T.; Furuno, K.; Mizuno, Y.; Ohga, S.; Hara, T.; Kira, R. Leucine-Rich Alpha-2 Glycoprotein in the Cerebrospinal Fluid Is a Potential Inflammatory Biomarker for Meningitis. *J. Neurol. Sci.* **2018**, *392* (May), 51–55. https://doi.org/10.1016/j.jns.2018.07.006.

(47) Kušnierová, P.; Zeman, D.; Hradílek, P.; Zapletalová, O.; Stejskal, D. Determination of Chitinase 3-like 1 in Cerebrospinal Fluid in Multiple Sclerosis and Other Neurological Diseases. *PLoS One* **2020**, *15* (5), 1–14. https://doi.org/10.1371/journal.pone.0233519.

(48) Gil-Perotin, S.; Castillo-Villalba, J.; Cubas-Nuñez, L.; Gasque, R.; Hervas, D.; Gomez-Mateu, J.; Alcala, C.; Perez-Miralles, F.; Gascon, F.; Dominguez, J. A.; et al. Combined Cerebrospinal Fluid Neurofilament Light Chain Protein and Chitinase-3 Like-1 Levels in Defining Disease Course and Prognosis in Multiple Sclerosis. *Front. Neurol.* **2019**, *10* (September), 1–11. https://doi.org/10.3389/fneur.2019.01008.

(49) Batruch, I.; Lim, B.; Soosaipillai, A.; Brinc, D.; Fiala, C.; Diamandis, E. P. Mass Spectrometry-Based Assay for Targeting Fifty-Two Proteins of Brain Origin in Cerebrospinal Fluid. *J. Proteome Res.* **2020**, *19* (8), 3060–3071. https://doi.org/10.1021/acs.jproteome.0c00087.

(50) Wahrle, S. E.; Shah, A. R.; Fagan, A. M.; Smemo, S.; Kauwe, J. S. K.; Grupe, A.; Hinrichs, A.; Mayo, K.; Jiang, H.; Thal, L. J.; et al. Apolipoprotein E Levels in Cerebrospinal Fluid and the Effects of ABCAI Polymorphisms. *Mol. Neurodegener.* **2007**, *2* (1), 1–9. https://doi.org/10.1186/1750-1326-2-7.

(51) Roher, A. E.; Maarouf, C. L.; Sue, L. I.; Hu, Y.; Wilson, J.; Beach, T. G. Proteomics-Derived Cerebrospinal Fluid Markers of Autopsy-Confirmed Alzheimer's Disease. *Biomarkers* **2009**, *14* (7), 493–501. https://doi.org/10.3109/13547500903108423.

(52) Kuiperij, H. B.; Hondius, D. C.; Kersten, I.; Versleijen, A. A. M.; Rozemuller, A. J. M.; Greenberg, S. M.; Schreuder, F. H. B. M.; Klijn, C. J. M.; Verbeek, M. M. Apolipoprotein D: A Potential Biomarker for Cerebral Amyloid Angiopathy. *Neuropathol. Appl. Neurobiol.* **2020**, *46* (5), 431–440. https://doi.org/10.1111/nan.12595.

(53) Reindl, M.; Knipping, G.; Wicher, I.; Dilitz, E.; Egg, R.; Deisenhammer, F.; Berger, T. Increased Intrathecal Production of Apolipoprotein D in Multiple Sclerosis. *J.*

Neuroimmunol. **2001**, *119* (2), 327–332. https://doi.org/10.1016/S0165-5728(01)00378-2.

(54) Tumani, H.; Rau, D.; Lehmensiek, V.; Guttmann, I.; Tauscher, G.; Palm, C.; Hirt, V.; Süssmuth, S. D.; Brettschneider, J. Liquorproteomanalyse Bei Klinisch Isoliertem Syndrom Und MS. *Nervenarzt* **2009**, *80* (S1), 30–31. https://doi.org/10.1007/s00115-009-2777-2.

(55) Duits, F. H.; Brinkmalm, G.; Teunissen, C. E.; Brinkmalm, A.; Scheltens, P.; Van Der Flier, W. M.; Zetterberg, H.; Blennow, K. Synaptic Proteins in CSF as Potential Novel Biomarkers for Prognosis in Prodromal Alzheimer's Disease. *Alzheimer's Res. Ther.* **2018**, *10* (1), 1–9. https://doi.org/10.1186/s13195-017-0335-x.

(56) Kroksveen, A. C.; Jaffe, J. D.; Aasebø, E.; Barsnes, H.; Bjørlykke, Y.; Franciotta, D.; Keshishian, H.; Myhr, K. M.; Opsahl, J. A.; van Pesch, V.; et al. Quantitative Proteomics Suggests Decrease in the Secretogranin-1 Cerebrospinal Fluid Levels during the Disease Course of Multiple Sclerosis. *Proteomics* **2015**, *15* (19), 3361–3369. https://doi.org/10.1002/pmic.201400142.

(57) Bowden, J. A.; Heckert, A.; Ulmer, C. Z.; Jones, C. M.; Koelmel, J. P.; Abdullah, L.; Ahonen, L.; Alnouti, Y.; Armando, A. M.; Asara, J. M.; et al. Harmonizing Lipidomics: NIST Interlaboratory Comparison Exercise for Lipidomics Using SRM 1950-Metabolites in Frozen Human Plasma. *J. Lipid Res.* **2017**, *58* (12), 2275–2288. https://doi.org/10.1194/jlr.M079012.

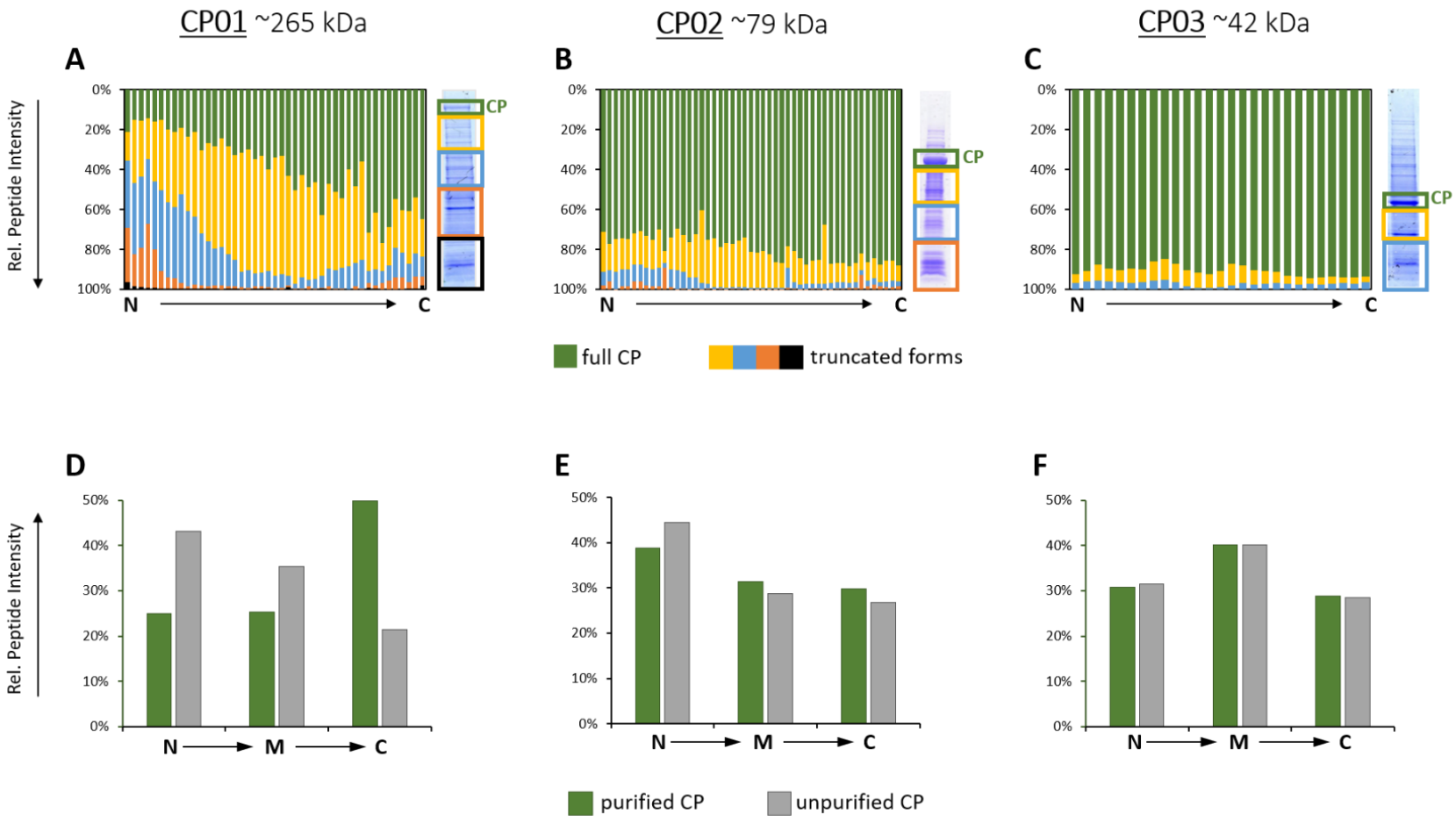


Figure 1. Truncation pattern for the chimeric proteins: CP01 ~265 kDa (panels: **A**, **D**), CP02 ~79 kDa (panels: **B**, **E**), CP03 ~42 kDa (panels: **C**, **F**). The upper panels (**A, B, C**) present relative abundance of peptides in SDS-PAGE slabs (y-axis) versus peptides position in CP sequence (x-axis); color-coding is exemplified at the right-hand side panel. Lower panels (D, E, F) present relative abundance of peptides from the purified CP (only CP band) versus the non-purified CP (sum of all bands) for selected Q-peptides located at the N- and C-termini as well as for the middle of the CP sequence ("M").

Table 1. Comparison of protein quantitation by FastCAT and MS Western.

Protein	Calculated amount[1] [fmol/column]		Quantification error %
	MS Western	FastCAT	
GP	197.5	189.9	3.8 %
UBI	nq	231.2	nq
ADH	205.9	183.0	11.1 %
ENO	222.9	208.9	6.3 %

[1]Proteins were quantified by averaging the amounts calculated using 3 to 5 Q-peptides. UBI was not detected/quantified (nq) by MS Western.

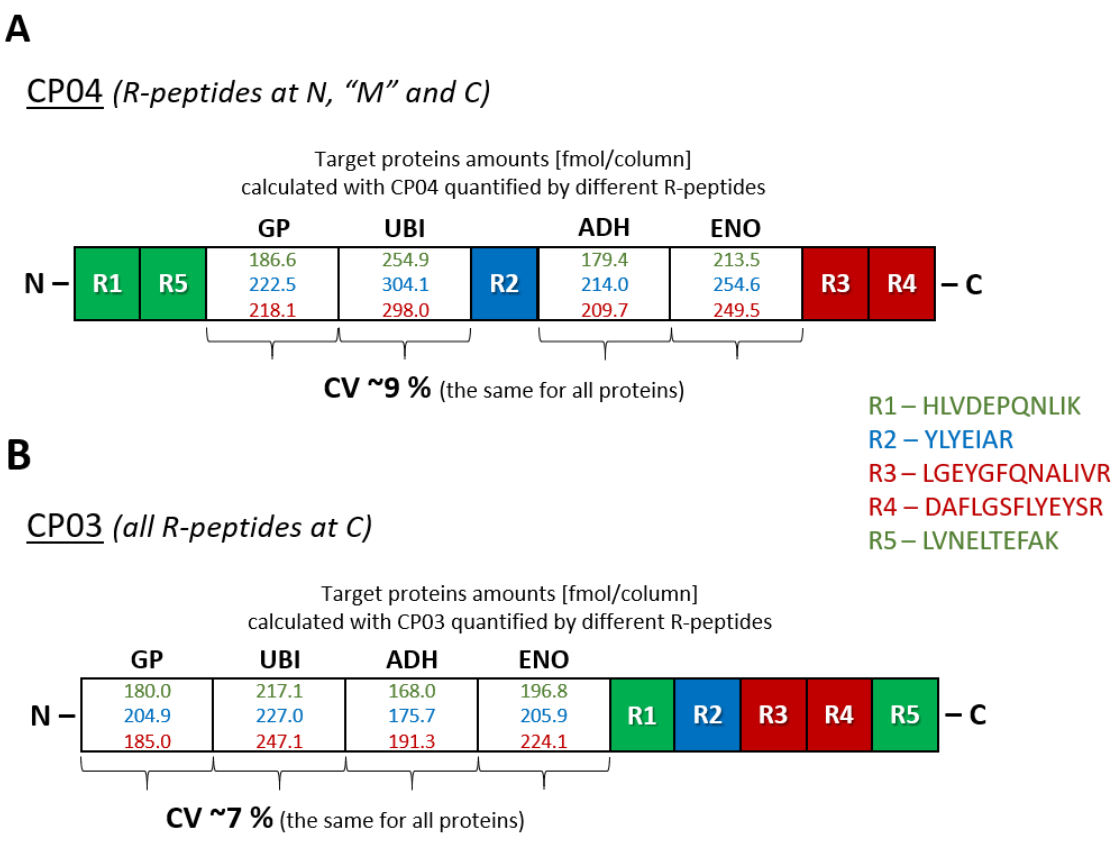


Figure 2. Impact of the position of R-peptides within CP sequence on proteins quantification. Comparison of the target proteins quantification by using CP04 and N- *vs.* "M" *vs.* C-terminus R-peptides as well as by using CP03 and the same groups of peptides but all positioned at C-terminus.

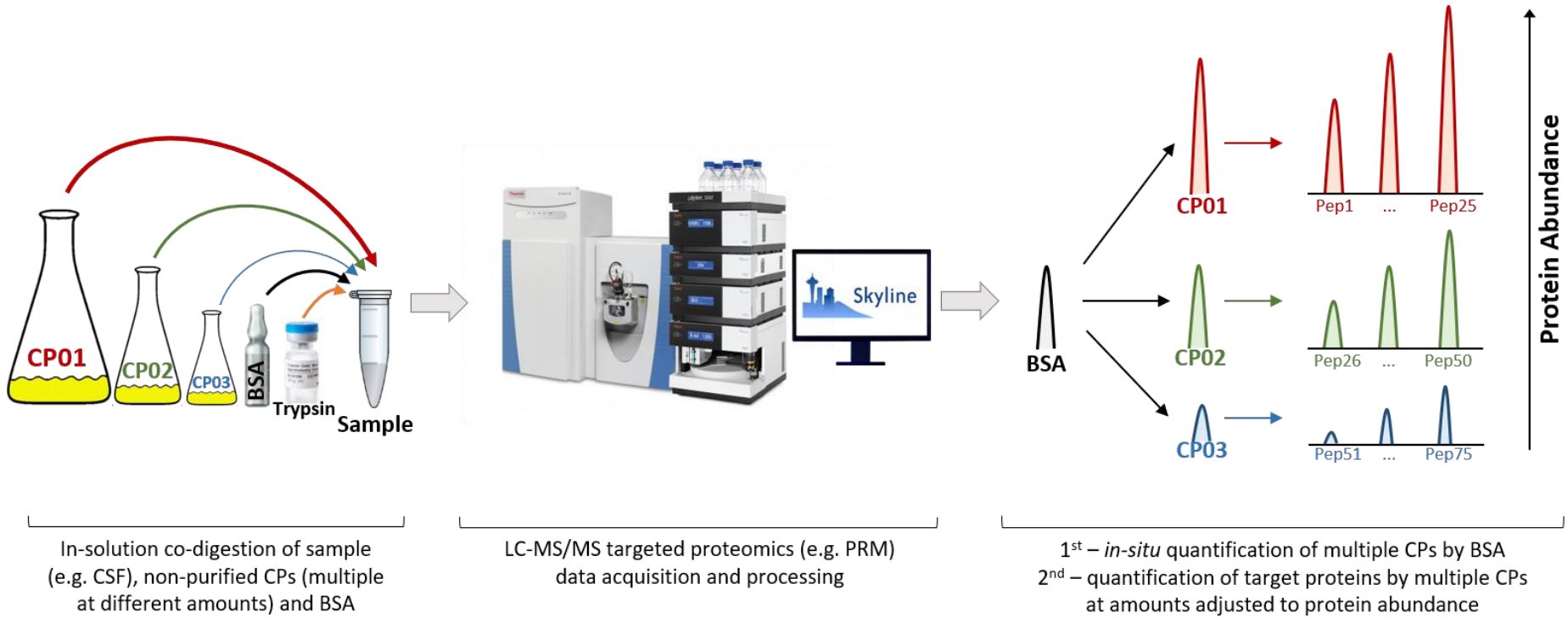


Figure 3. FastCAT workflow using multiple non-purified CP standards for absolute parallel quantification. 1st step: quantification of CPs by their reference peptides and spiked-in BSA; 2nd step: quantification of the target proteins by quantotypic peptides from CPs. Both 1st and 2nd steps are performed together during the same LC-MS/MS run.

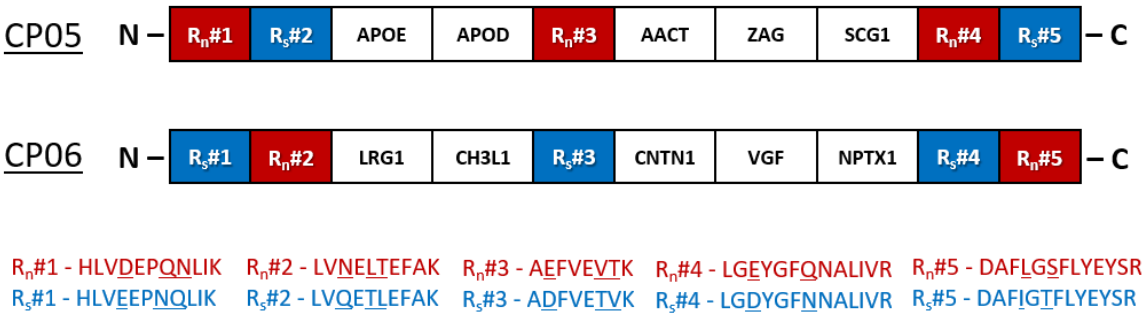


Figure 4. Scheme of CP05 and CP06 constructs comprised of the native (R_n#1 – R_n#5) and scrambled (R_s#1 – R_s#5) forms of 5 BSA peptides (shown as one block/peptide) as well as the multiple Q-peptides from 10 target proteins (shown as one block/protein).

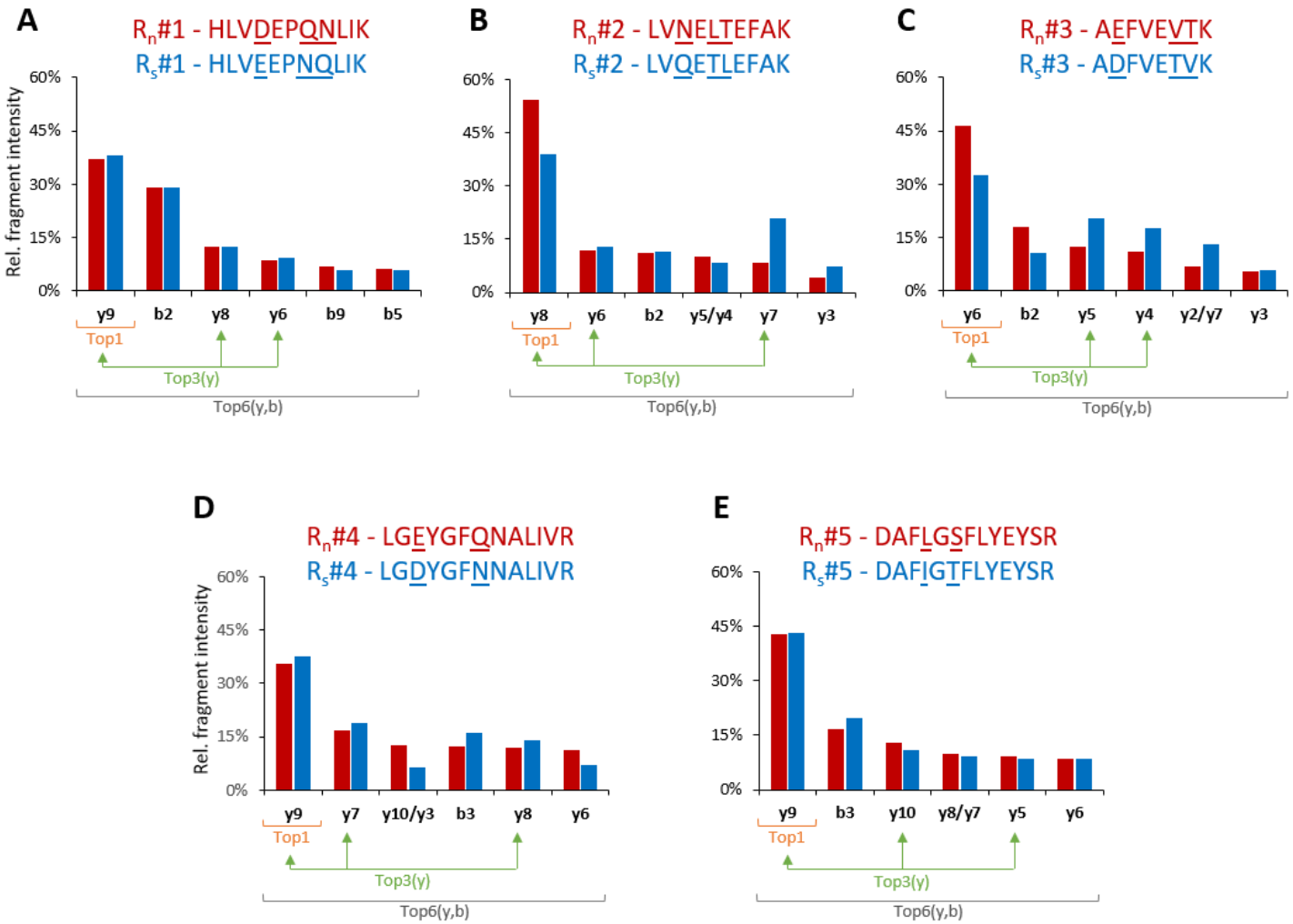


Figure 5. Relative abundance of the most intense six fragment ions (within y- and b-ions) from the native (BSA/light – red bars) and scrambled (CP/heavy – blue bars) R-peptides. In four of five pairs one fragment ion within the top 6 did not match between the native and scrambled forms and in this situation the non-matching fragments were presented together for better visibility of the fragmentation differences.

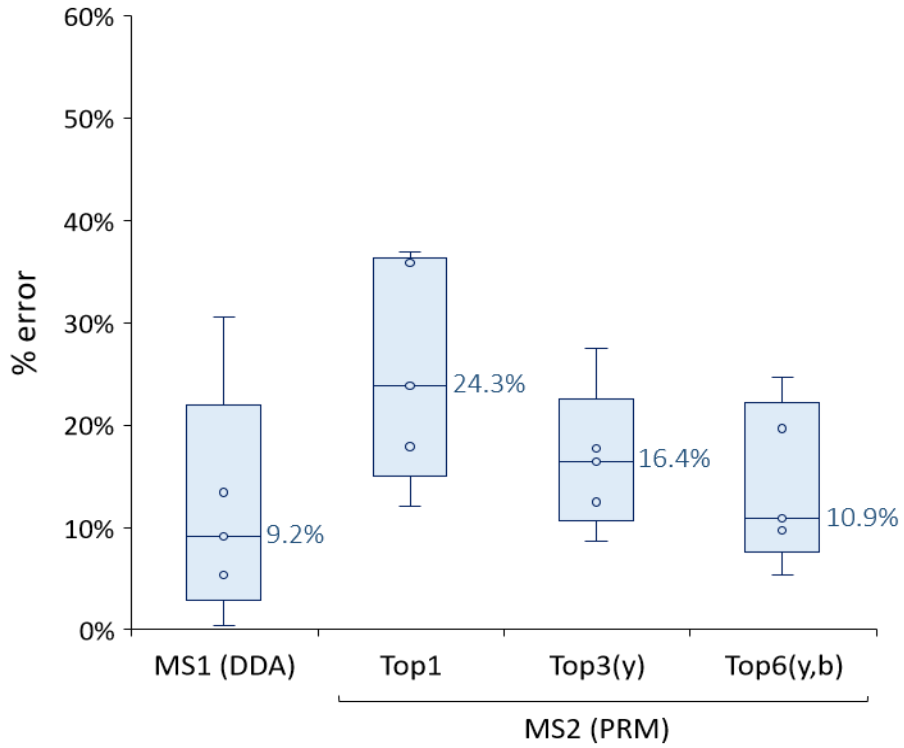


Figure 6. Relative error in the quantification of CPs by using R_s-peptides and either MS1 or MS2 (summed) intensities of different fragment ions. Top1 (the top fragment ion); Top3(y) (sum of the top three y-ions, matching between R_n and R_s); Top6(y,b) (sum of the top six fragments within y- and b-ions). Each box contains information from 5 R_s-peptides (detailed information is provided in Table S2 in SI), while the median values are given next to the boxes. Each box plot displays the median (line), the 25th and 75th percentiles (box), and the 5th and 95th percentiles (whiskers).

Protein	UniProt accession	Q-peptides sequences	Intra-protein peptide concordance (median CV [%])	Concentration range (median) [ng/mL]	Lit. reported concentration range (method and ref.) [ng/mL]
APOE	P02649	SELEEQLTPVAEETR LGPLVEQGR QWAGLVEK LAVYQAGAR	7.5	922.1–2075.6 (1807.8)	900–3000 (SRM)[39] 2964–3112 (PRM)[40] 7000–12000 (ELISA)[51] 4000–13000 (ELISA)[50]
APOD	P05090	NILTSNNIDVK NPNLPPETVDSLK VLNQELR WYEIEK	5.7	1060.4–1327.3 (1256.6)	3000–12 000 (ELISA)[52]
AACT	P01011	ITLLSALVETR NLAVSQVVHK AVLDVFEEGTEASAATAVK ADLSGITGAR	10.1	540.9–1127.3 (724.0)	126–2954 (PRM)[41] 1400–3900 (ELISA)[44]
ZAG	P25311	EIPAWVPFDPAAQITK WEAEPVYVQR	9.0	85.2–113.2 (96.3)	n/a
SCG1	P05060	NYPSLELDK NYLNYGEEGAPGK WQQQGDLQDTK	12.4	388.6–942.3 (850.1)	300–900 (SRM)[42]
LRG1	P02750	VAAGAFQGLR GQTLLAVAK	14.0	27.1–114.3 (68.0)	25–350 (ELISA)[45] 90–800 (ELISA)[46]
CH3L1	P36222	FPLTNAIK ILGQQVPYATK VTIDSSYDIAK GNQWVGYDDQESVK	5.8	25.7–53.8 (35.2)	35–254 (SRM)[43] 70–160 (ELISA)[43] 29–182 (ELISA)[47] 30–350 (ELISA)[48]
CNTN1	Q12860	FIPLIPIPER ASPFPVYK	12.4	148.3–268.6 (238.9)	20–300 (ELISA)[44]
VGF	O15240	NSEPQDEGELFQGVDPR THLGEALAPLSK	30.9	60.6–175.9 (111.0)	n/a
NPTX1	Q15818	FQLTFPLR TPAAETLSQLGQTLQSLK	5.6	37.2–86.0 (72.0)	n/a

Table 2. Proteins concentrations (ranges and median) obtained from eleven CSF samples from patients with multiple sclerosis, together with UniProt accessions, Q-peptides used for quantification, intra-protein concordance (median CV from all samples) as well as the CSF concentrations reported for those proteins in literature by using both mass spectrometric (either PRM or SRM) and ELISA methods. Proteins: APOE (apolipoprotein E), APOD (apolipoprotein D), AACT (Alpha-1-antichymotrypsin), ZAG (Zinc-alpha-2-glycoprotein), SCG1 (Secretogranin-1/Chromogranin B), LRG1 (Leucine-rich alpha-2-glycoprotein 1), CH3L1 (Chitinase-3-like protein), CNTN1 (Contactin-1), VGF (Neurosecretory protein VGF), NPTX1 (Neuronal pentraxin-1).

Supporting Information

FastCAT accelerates absolute quantification of proteins by using multiple short non-purified chimeric standards

Ignacy Rzagalinski[1], Aliona Bogdanova[1], Bharath Kumar Raghuraman[1], Eric R. Geertsma[1], Lena Hersemann[1], Tjalf Ziemssen[2], Andrej Shevchenko[1*]

[1]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany
[2]Center of Clinical Neuroscience, Department of Neurology, University Hospital Carl Gustav Carus, Technical University of Dresden, 01307 Dresden, Germany

Table of Content

Figure S1 Scheme of chimeric protein (CP)
Figure S2 *E. coli* background in non-purified CP
Figure S3 Amino acid sequences of CP05/CP06
Figure S4 Truncation patterns for CP05/CP06
Figure S5 Tandem mass spectra for R_n/R_s-peptides
Figure S6 Relative R_n/R_s-peptides abundances for CP05/CP06 vs. BSA
Table S1 *E. coli* background proteins used for comparison in Figure S2
Table S2 Error in quantification of CPs by R_s-peptides
Table S3 Protein concentrations in CSF from individual patients
Table S4 Amino acid sequences for all used CPs

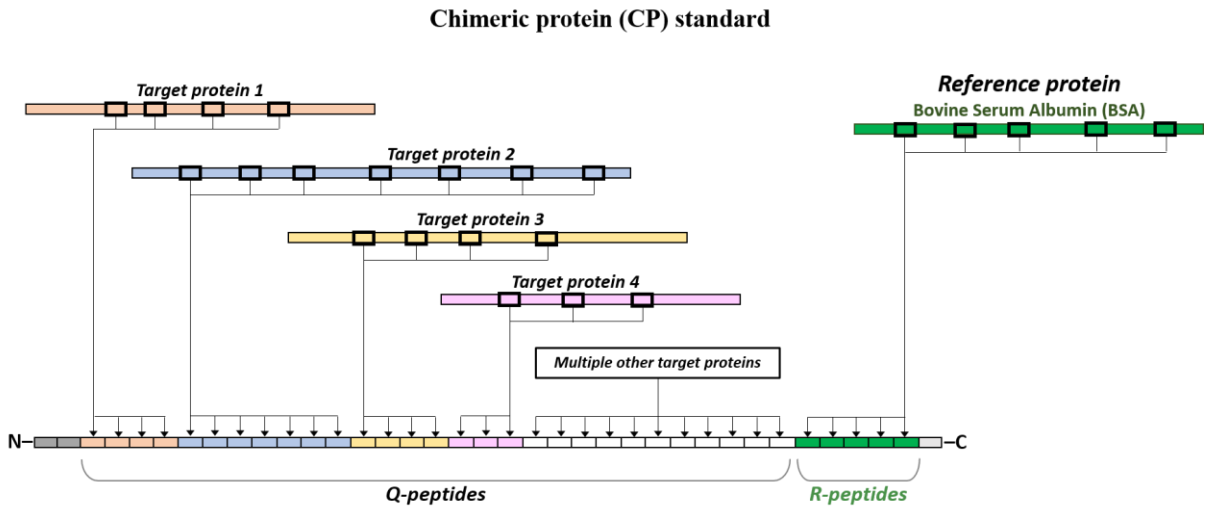


Figure S1. Scheme of chimeric protein (CP) standard.

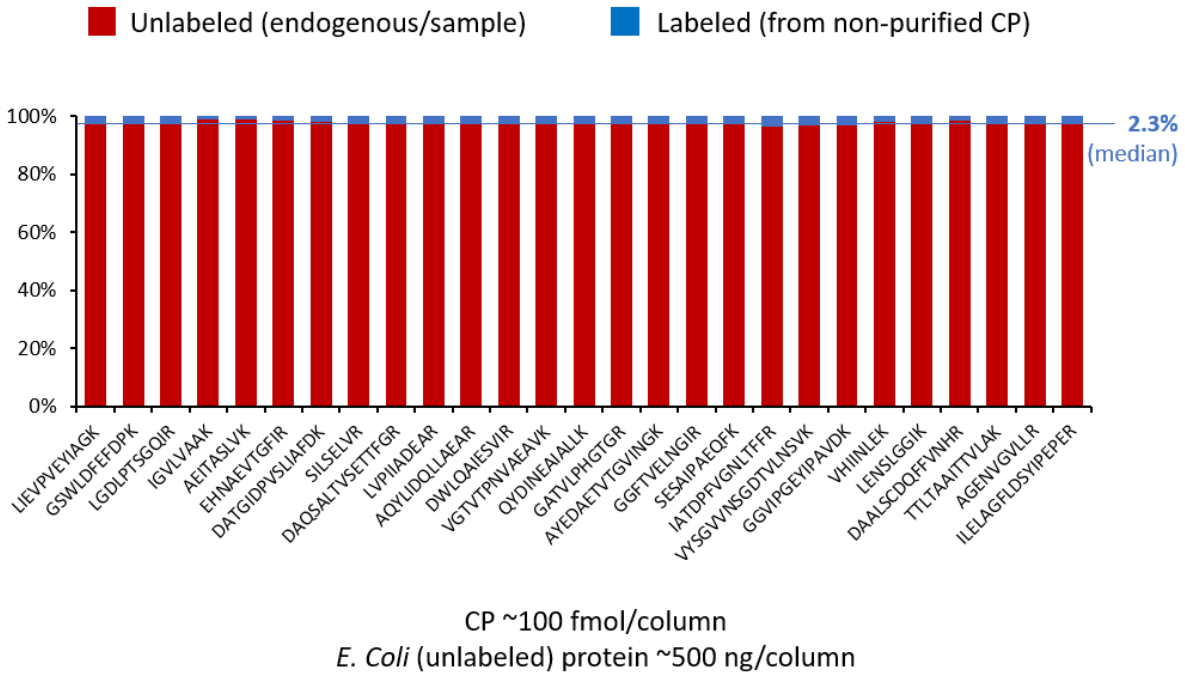


Figure S2. Contribution of an extra background from the expression host (*E. coli*) introduced with the non-purified CP standard (labeled, blue bars) in comparison with original E. coli sample (unlabeled/endogenous, red bars), shown for the typical working conditions with CP at approximately 100 fmol on column and total protein load on column ca. 500 ng. The contribution of the extra background (labeled) is just ~2 % of the total pool. Importantly, nine highly abundant E. coli proteins (mainly ribosomal and elongation factor proteins) were compared, which were represented by at least three peptides yielding consistent L/H ratios (see Table S1). Of note, only the peptides observed with one predominant charge state, free of missed cleavages as well as free of cysteine and methionine residues were selected. Moreover, CP labeling efficiency was above 99 % so virtually all unlabeled peptides originated from the original *E. coli* sample.

A CP05

MGSAWSHPQFEKGGGSGGGSGGSAWSHPQFEKLEVLFQGPAAAKVFADYEEYVKDFYELEPHKV
AAAFPGDVDRGLAGVENVTELKHLVDEPQNLIKLVQETLEFAKSELEEQLTPVAEETRLGPLVEQGR
AATVGSLAGQPLQERQWAGLVEKLAVYQAGARNILTSNNIDVKIPTTFENGRNPNLPPETVDSLKV
LNQELRWYEIEKAEFVEVTKITLLSALVETRNLAVSQVVHKAVLDVFEEGTEASAATAVKADLSGITG
ARHVEDVPAFQALGSLNDLQFFRYSLTYIYTGLSKEIPAWVPFDPAAQITKQKWEAEPVYVQRNYPS
LELDKNYLNYGEEGAPGKGEAGAPGEEDIQGPTKWQQQGDLQDTKLGEYGFQNALIVRDAFIGTF
LYEYSRGSGHHHHHH

Apolipoprotein E (APOE) – Apolipoprotein D (APOD) – Alpha-1-antichymotrypsin (AACT) – Zinc-alpha-2-glycoprotein (ZAG) – Secretogranin-1 (SCG1)
BSA (R-)peptides: native and scrambled

B CP06

MGSAWSHPQFEKGGGSGGGSGGSAWSHPQFEKLEVLFQGPAAAKVFADYEEYVKDFYELEPHKV
AAAFPGDVDRGLAGVENVTELKHLVEEPNQLIKLVNELTEFAKVAAGAFQGLRGQTLLAVAKTLDLG
ENQLETLPPDLLRENQLEVLEVSWLHGLKFPLTNAIKILGQQVPYATKTLLSVGGWNFGSQRVTIDS
SYDIAKGNQWVGYDDQESVKADFVETVKFIPLIPIPERASPFPVYKTDGAAPNVAPSDVGGGGGR
VQVTSQEYSARNSEPQDEGELFQGVDPRTHLGEALAPLSKLADLASDLLLQYLLQGGARVGEEDEE
AAEAEAEAEEAERFQLTFPLRTPAAETLSQLGQTLQSLKALSGNVIAWAESHIEIYGGATKLGDYGFN
NALIVRDAFLGSFLYEYSRGSGHHHHHH

Leucine-rich alpha-2-glycoprotein (A2GL) – Chitinase-3-like protein (CH3L1) – Contactin-1 (CNTN1) – Neurosecretory protein VGF (VGF) – Neuronal pentraxin-1 (NPTX1)
BSA (R-)peptides: native and scrambled

Figure S3. Amino acid sequences of the CP05 and CP06 constructs used for quantification of 10 neurological protein markers in CSF samples.

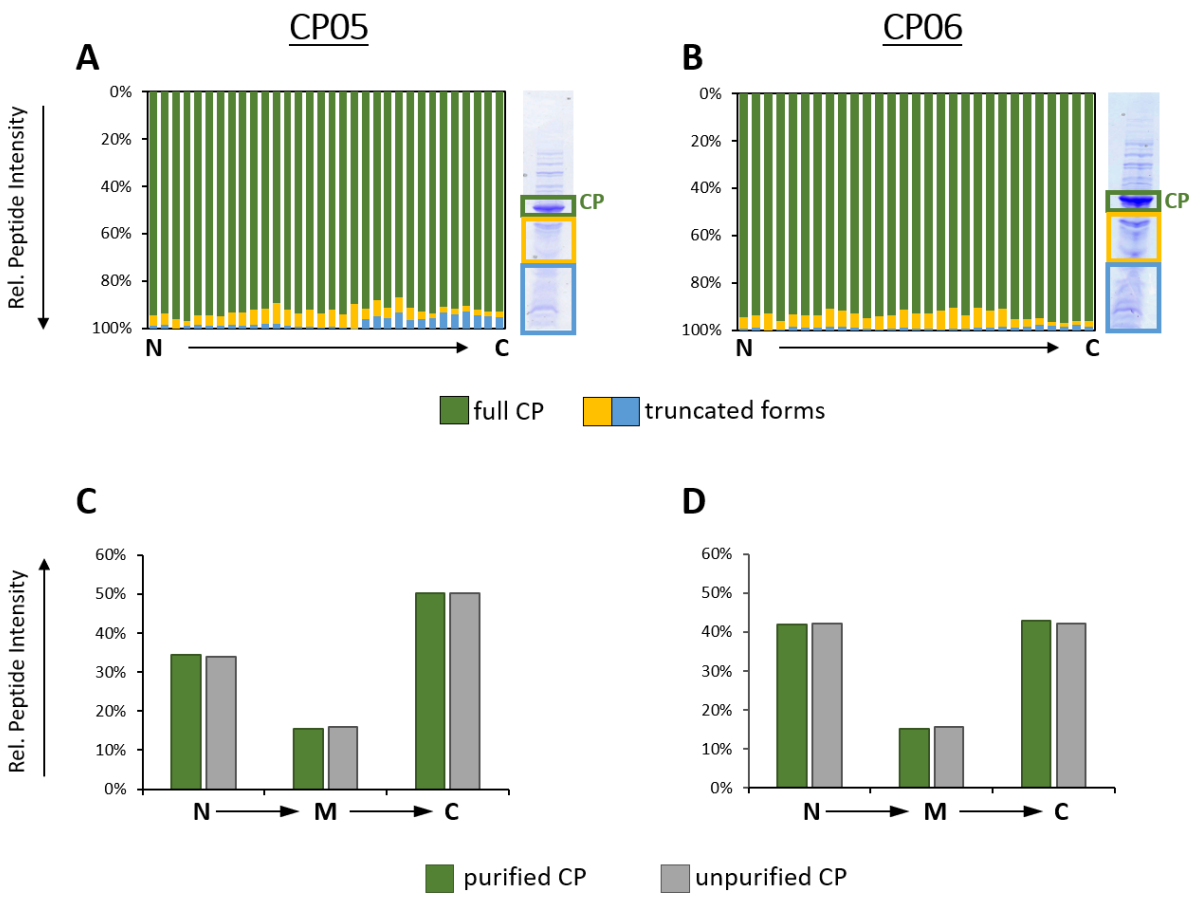
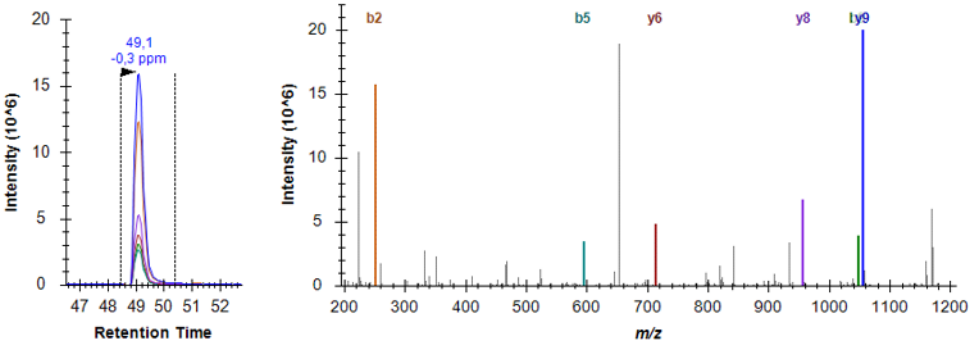


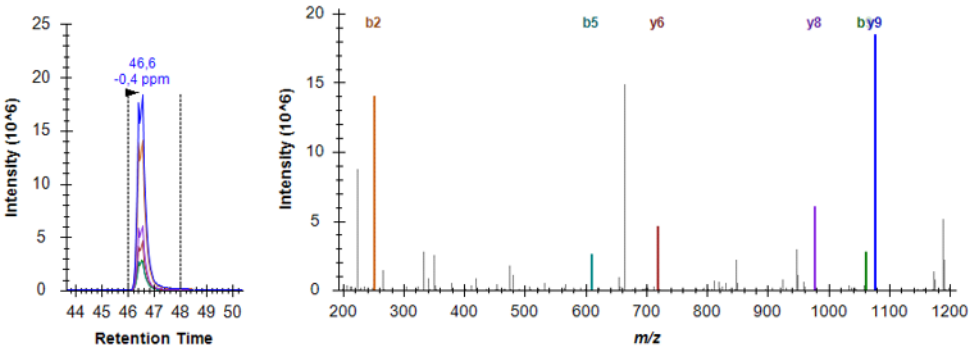
Figure S4. Truncation patterns for CP05 (A and C) and CP06 (B and D) standards.

A

HLVDEPQNLIK

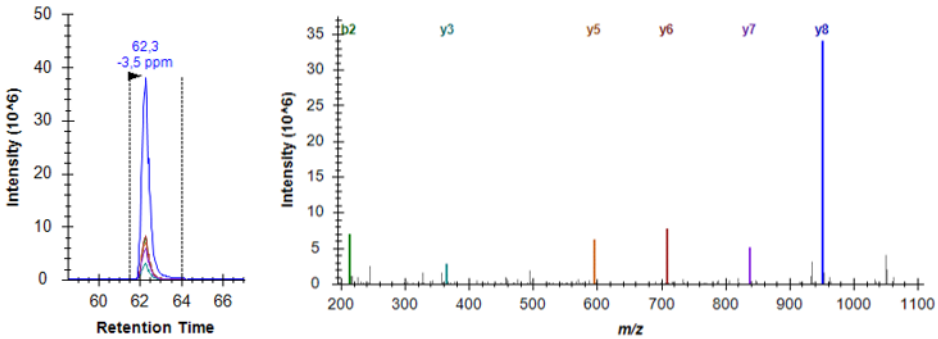


HLVEEPNQLIK

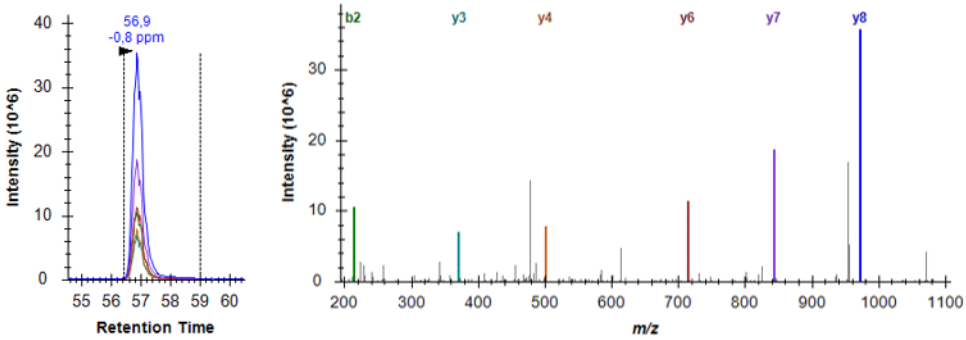


B

LVNELTEFAK

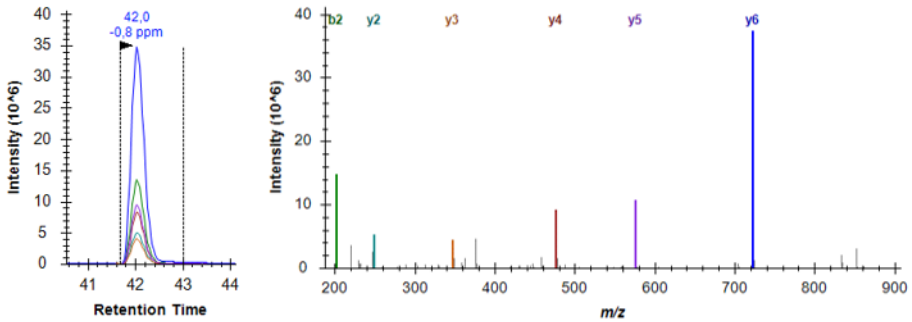


LVQETLEFAK

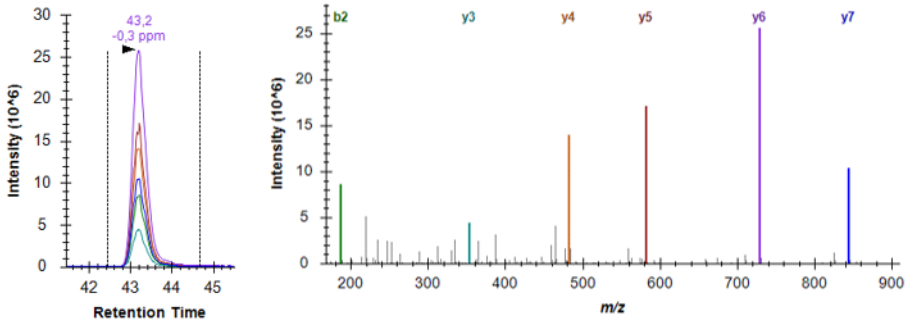


C

AEFVEVTK

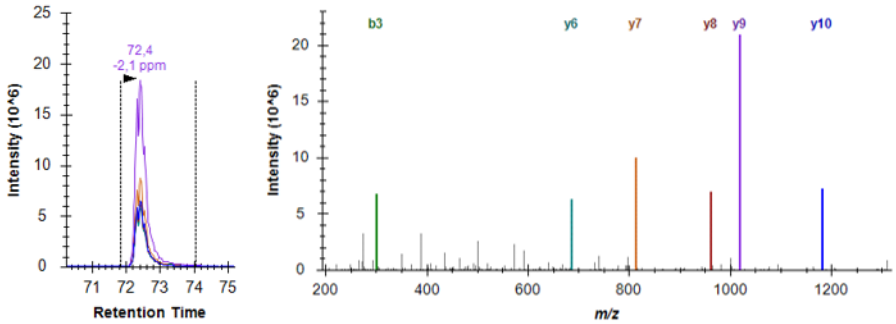


ADFVETVK

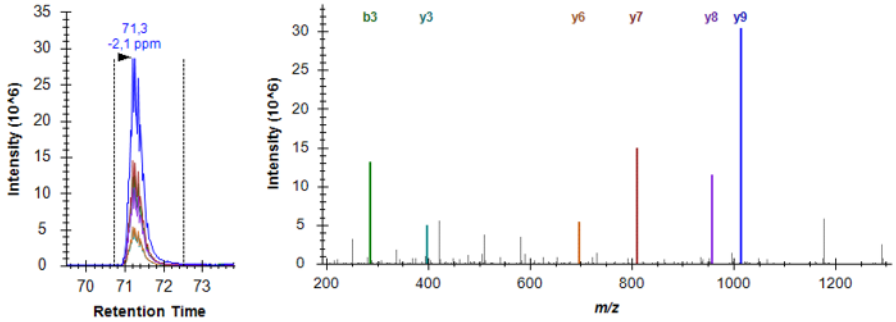


D

LGEYGFQNALIVR



LGDYGFNNALIVR



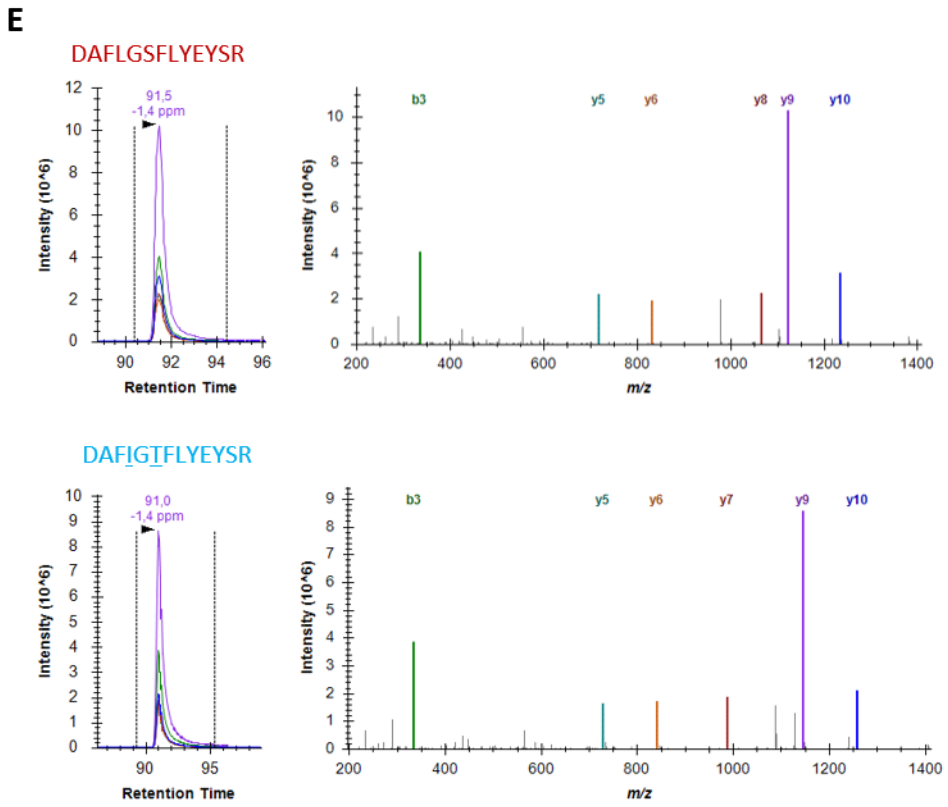


Figure S5. R-peptides tandem mass spectra for their native (BSA, unlabeled) and scrambled forms (CP05/06, labeled) together with their chromatographic peaks: HLVDEPQNLIK (nat) and HLVEEPNQLIK (scr) (**A**), LVNELTEFAK (nat) and LVQETLEFAK (scr) (**B**), AEFVEVTK (nat) and ADFVETVK (scr) (**C**), LGEYGFQNALIVR (nat) and LGDYGFNNALIVR (scr) (**D**), DAFLGSFLYEYSR (nat) and DAFIGTFLYEYSR (scr) (**E**).

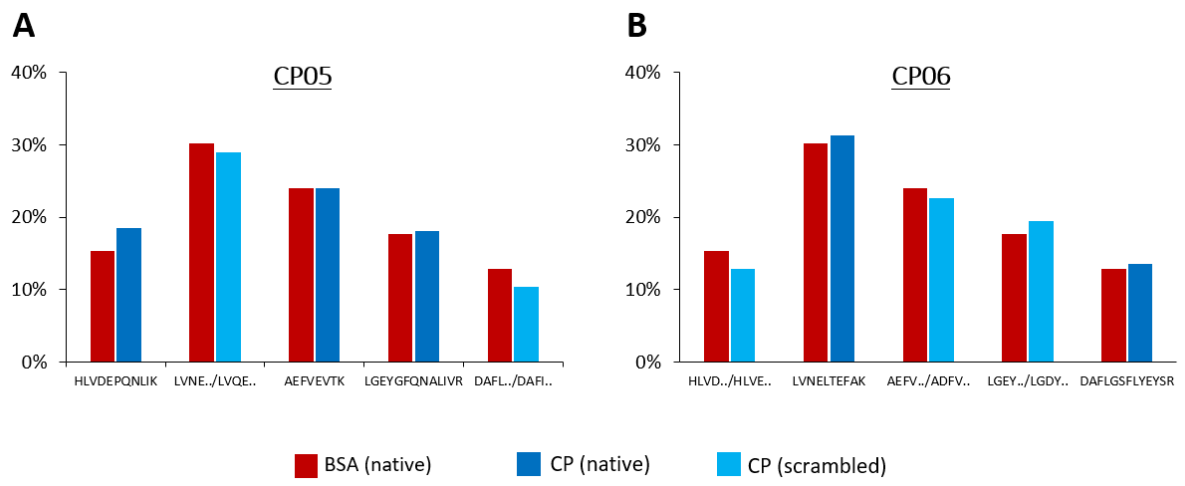


Figure S6. Matching R-peptides relative peak areas for BSA (native) and both CP05 (**A**) and CP06 (**B**) obtained using Top6(y,b) fragment ions. Differences were <5 % for all pairs.

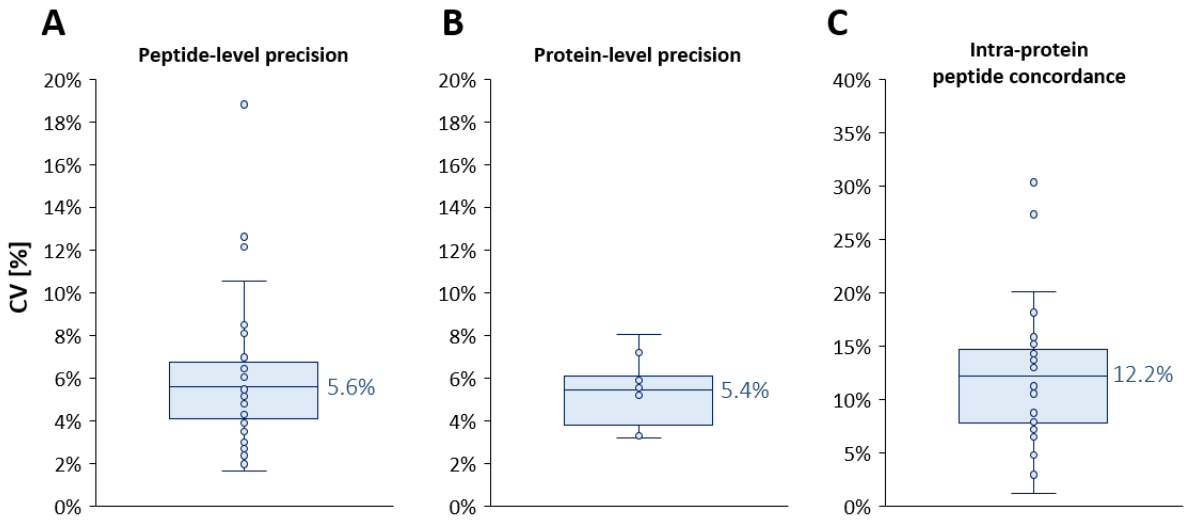


Figure S7. Precision of FastCAT method obtained from the PRM analysis of the pooled CSF sample. Panels A and B show peptide-level and protein-level precision, respectively. The graphs are created with data acquired for all investigated proteins (10 target proteins and two CPs) based on three sample preparation replicates and two LC injections each. Panel C shows intra-protein peptide concordance obtained from the same data set. Median CV values are given on the graphs. Each box plot displays the median (line), the 25th and 75th percentiles (box), and the 5th and 95th percentiles (whiskers).

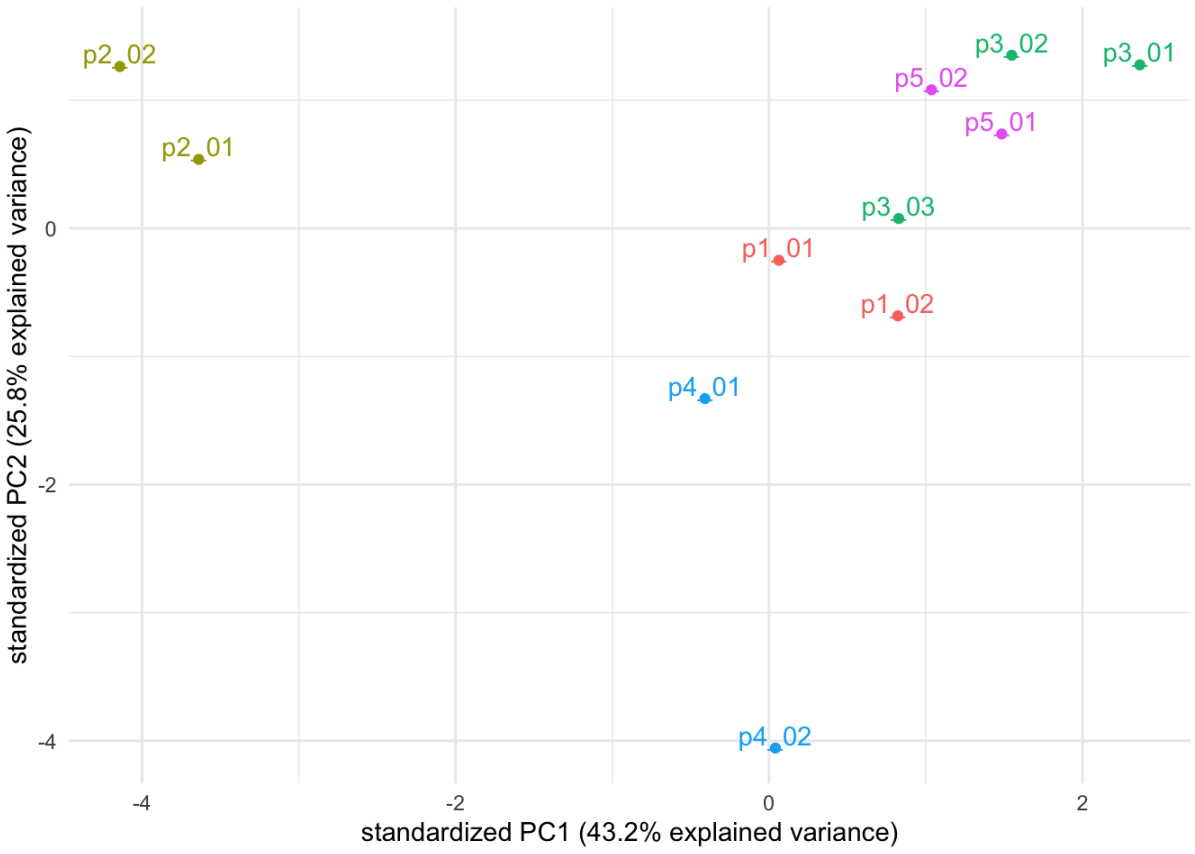


Figure S8. Principal component analysis (PCA) graph for all quantitative data (from Table S4) obtained for patients 1-5 (p1-p5) with two (or 3 for p3) times of sample collection.

Table S1. List of the representative *E. coli* proteins (selected from those of the highest abundance) used for an estimation of the amount of an extra E. coli background introduced with non-purified CP standard, together with their UniProt accession numbers, as well as the peptides (three per protein) with their retention times and charge states. Peptides were selected to be free of missed cleavages, not containing cysteine and methionine and represented by the clearly dominant one charge state.

Protein	UniProt Accession	Peptide	RT	Charge state
DNA-directed RNA polymerase	P0A8V2	LIEVPVEYIAGK	76.7	2
		GSWLDFEFDPK	87.1	2
		LGDLPTSGQIR	57.5	2
Elongation factor Ts	P0A6P1	IGVLVAAK	53.2	2
		AEITASLVK	55.2	2
		EHNAEVTGFIR	53.7	3
50S ribosomal protein L4	P60723	DATGIDPVSLIAFDK	88.9	2
		SILSELVR	76.7	2
		DAQSALTVSETTFGR	69.9	2
Pyruvate dehydrogenase E1	P0AFG8	LVPIIADEAR	64.4	2
		AQYLIDQLLAEAR	89.5	2
		DWLQAIESVIR	105.7	2
50S ribosomal protein L1	P0A7L0	VGTVTPNVAEAVK	56.0	2
		QYDINEAIALLK	85.9	2
		GATVLPHGTGR	41.1	2
30S ribosomal protein S1	P0AG67	AYEDAETVTGVINGK	64.3	2
		GGFTVELNGIR	70.9	2
		SESAIPAEQFK	55.2	2
Elongation factor G	P0A6M8	IATDPFVGNLTFFR	89.9	2
		VYSGVVNSGDTVLNSVK	67.9	2
		GGVIPGEYIPAVDK	70.7	2
30S ribosomal protein S2	P0A7V0	VHIINLEK	51.9	2
		LENSLGGIK	51.7	2
		DAALSCDQFFVNHR	66.1	3
Elongation factor Tu1	P0CE47	TTLTAAITTVLAK	82.1	2
		AGENVGVLLR	61.4	2
		ILELAGFLDSYIPEPER	98.4	2

Table S2. Relative error of the CP05/06 quantification based on the scrambled R-peptides and using either MS1-derived (DDA) peptides peak areas or different MS2 fragment ions (PRM).

Native form	**Scrambled form in CP**	**Δ RT [min]**	**CP with scrambled form**	**MS1-based (DDA)**			**MS2-based (PRM)**						
				CP [fmol] by nat	CP [fmol] by scr	% error	CP [fmol] by nat	CP [fmol] by scr					
								Top1	% error	Top3(y)	% error	Top6(y,b)	% error
HLVDEPQNLIK	HLVEEPNQLIK	2.5	CP06	275	290	5.4%	294	241	17.9%	241	17.7%	236	19.7%
LVNELTEFAK	LVQETLEFAK	5.4	CP05	270	295	9.0%	280	180	35.8%	245	12.5%	250	10.9%
AEFVEVTK	ADFVETVK	1.2	CP06	275	274	0.4%	294	185	36.9%	268	8.7%	265	9.7%
LGEYGFQNALIVR	LGDYGFNNALIVR	1.1	CP06	275	312	13.4%	294	329	12.1%	341	16.4%	310	5.4%
DAFLGSFLYEYSR	DAFIGTFLYEYSR	0.5	CP05	270	188	31.0%	280	214	23.9%	203	27.4%	211	24.7%

Table S3. Protein concentrations (ng/mL) detected in 11 CSF samples from 5 individual patients with multiple sclerosis. CSF was collected two times for four patients and three times for the patient 3. Patient age at 1st collection is given in years (y), while the time difference (Δt) to the 2nd (and 3rd for patient 3) puncture is given in months (m).

Protein	Concentration in CSF [ng/mL]										
	Patient 1		Patient 2		Patient 3			Patient 4		Patient 5	
	time 1	time 2	time 1	time 2	time 1	time 2	time 3	time 1	time 2	time 1	time 2
	age, 43 y	Δt, 40 m	age, 61y	Δt, 17 m	age, 36	Δt, 10 m	Δt, 16 m	age, 30 y	Δt, 32 m	age, 30 y	Δt, 26 m
APOE	1489.28	1604.15	1106.19	922.07	1926.15	1807.81	1772.67	1854.96	2075.59	1996.34	1811.10
APOD	1197.68	1307.78	1185.15	1203.23	1284.59	1281.67	1297.31	1083.58	1060.41	1246.55	1327.29
AACT	759.84	684.61	765.58	723.96	580.30	564.96	770.17	810.35	1127.28	540.90	588.94
ZAG	100.05	113.18	102.01	96.32	85.98	85.89	87.58	102.12	108.05	85.19	85.76
SCG1	636.56	684.09	451.77	388.55	942.32	885.69	869.22	788.86	862.85	852.69	850.08
LRG1	67.98	71.42	77.14	80.84	50.02	47.78	114.29	38.45	92.91	27.13	33.82
CH3L1	51.77	35.09	35.16	33.37	53.82	49.85	44.63	38.90	27.48	27.90	25.74
CNTN1	238.90	256.07	159.70	148.28	229.68	230.81	229.77	241.20	268.62	257.04	246.15
VGF	139.34	163.29	62.64	60.65	175.94	133.49	141.99	73.35	98.14	110.97	99.89
NPTX1	75.67	86.02	42.50	37.18	77.01	72.52	71.98	64.28	81.85	70.87	68.53

Table S4. All chimeric protein (CP) standards used in the study, together with the information on their molecular weights, labeling efficiencies as well as amino acid sequences.

Chimeric Protein	Molecular weight	Labeling efficiency	Amino acid sequence
CP01	264.8 kDa	>99 %	MGSAWSHPQFEKGGGSGGGSGGSAWSHPQFEKLEVLFQGPAAAKVFADYEEYVKDFYELEPHKVAAAFPGD VDRGLAGVENVTELKGSTLSAQQGSQFKNEPVSYILELINGRNIGDSYVAAKLDNGYNHLIEVVRLVLGIPTYG RFSPLVASNERLTEAEGSSLYIGGRSTDAEEDPQVIKFADQDNDLVNLRLSNFVDTTVAWLNYRFFVSVTRVQL TDLNRDVFIPDVFNNYKILAIFPFPGPSQYINVVPYLKVPIWPYTLYFRGTFFVSHKGHEISVFSLTHKAPLGLHF HASWLKIINNPEATQRAVYWVEHVSRSHYHVGSALAKAEQNGYGVTVHYEELSSAKGSAYAHAENTLRSEGD LTQITPPSALGVQVRAEVEAVQIIAESLKAANQLTDRPTIINVAISPSSDRGHGAGGADVLTYKDLNQGNSYVQD KSLVFVDNHDNQRSEYTGLGAITEFRVVEFLDHLIDLGVAGFRIGALDTSRLIELLVRLADLLLQEERVNGEPVD VGAFFIRITPDVVGAIVQEATTYFRATQLSEQLGSELTPKLLAADYADGVSQPRILSWNAVNLYGLKASEQPGL TAIHTAFLRLLPAQYEDGISAPRGIATNDVGIAKDTFSFLTGSGRAINEGGFQSLPQKFGLNEGSEPQAYGIGLKI TTHYTLNPRAILVYLVEKITPGWEENWAGALDVKINPQHSIPTLVDNGFTIWESRALGLEFNKINLDPYALYVIV EKTNFDDSALFYASGESLKVLTTGGDAGSNRLGSDVQPPGRVFLQLLENQRNLSQTFGNIWRVFDLPDFLTWR GSTVQFAARDGAISHVVFKYFNAYLSDRFSEPYDSTLSDVIKNVANLEATKDADAIVQQTLAKGTQAEIVVAD VTKVGDVTEVAEAVAFLASSKSVLLTLRIDALEGVYKALNLDFDYKAITFPLFWENKLTLYGIDGSPPVRELPY YEEANGSRAVGVELNKAPADPEAFKAIQVYLVEKLLNLQAGEHLKPEFLKINPQHTIPTLVDNGFALWESRGL VVSGTRSLSDVSLTGRGNSYLILPPPRISGPSNHVTVVRNAAFSGDSYVSHRSVVNVGLGQALRSTQEEVDHIRS AVHDVEVFLKVSTAIDAISVSGRVALQIQDVATSSRLEEISLRLLELHDNRLFNNFDVLRLAVLDLSHNRASLSG IQSHAFKTFFDGNPIHTLRDFGVELEDLQITRSTISSTTTVRTSTTSSLTGNPRTSVATVAGGAVVGATKLLVPGS SSTTTTSSLRLLPDTTDEQLLTSALEEKSTGNIFAAKVFAVNSAGRIFADNVYGRGVAEDFAPSFVKHDGGSPIT GYIIEKHYPNPAVRVVGSEADTGRDAPTTESYLRSTVPFATSESNRLVGGNDLDTPLIITRIAELEGLGSRIQYLE DSTRIGEAISQSSIRFITQIVDVTKAEDAGVYLVVARTYYDFGFVALDIIKVWQVSVSAHALLWDLNDGKDVLS VAFSADNRLWDLAAGKLNNDLIARFQEALAGLSKLAAHDALGGAAKSAGSGVSTTAIEKIVQVQIDDVGKSEL DVFSDWLQVARFSQSDFGLDQGETLLRESLLEITIYHQKGAAYQEAPVDAEVAVTPKISAFGLFTYSVFSIIAGS LKVFANPVQLEFYGFVPWSIIQRQVGASSGSSGAVSKIVQIQQVGPAHLNVKVQDNSVLVEGNHEERVLPAIEA FQKFSEATLDEIIRFHGVALAFNALDSKENAILTDIWNITPFKVIVDILLKTILVDLQVGKLAELHAASVVAKEAG LEIELAPKAGFAGDDAPRSYELPDGQVITIGNERVAPEEHPVLLTEAPLNPKGYSFTTTAEREITSLAPSTIKEITA LAPSTIKQEYDESGPGIVHRQEYDESGPSIVHRSIELLEKLQASGDFPTTKEQNSPIYISRATVAAFAASEGHAHPR YQVIVKSSDAQSQATASEAESKALSASSLIALSSRDPVLTAFQLSWELKDDYGITEDDIIEVRATAWVIVHRVEL ELWLKVVLPDLAVLRALGIEEQSGGAARVDEYGFFLYWKYYLSDLARIEQADYLPTEQDILRVPTTGILEYPFD LDGIVFRTIITYPWFQNSSVILFLNKAGIAVEGDIKDAFGIIVSYAVKDTALASTTLIASQDARDFLLSPGELELEV TLDKVFGQLATTYRLQYAPLNRISLEVTLDRLLQPAPGTIEKFYPEDVAEELIQDITLRIGFPWSEIRTTHTAGFL ANDRLFYLQVKAAIELNRGDNILLTKVPPEVPPSKLVDFIINRSLVGVQKQIDVETVADRYIQATDLSEASRNAE YDQFLEDIQKATAWVIVHRALSASSLIALSSRHLVDEPQNLIKYLYEIARQTALVELLKLGEYGFQNALIVRDAF LGSFLYEYSRLVNELTEFAKGSGHHHHHH

CP02	79.3 kDa	>99 %	MGSAWSHPQFEKGGGSGGGSGGSAWSHPQFEKLEVLFQGPAAAKVFADYEEYVKDFYELEPHKVAAAFPGD VDRGLAGVENVTELKLGDYGFNNALIVRDAFIGTFLYEYSRAAAYVLQETPVVNALVDENEIVYRSDGEVEIA SEKVIDTAYEIIKGIILIGEGIGNAEEQAAEFLKVQNDDSIVFFDYRVLEATLAQDFSKLVTWYDNEFGYTNRYA GEDAAAGAAETLFVAKAEIEAVQIIAETLKVVEFLEHILDLGVAGFRIVQVNLDDVGKVGDVTEVAEAAVFLA TSKAGLAIEGDIKANEILSDIWNITPFKGIAEDFAPTFVKIADLEGIYKALQLDFEYKALVFETWQGPLEVRWVAI DGEQYGEGSSRLIDDNVANALKVVDHAYEAVVIGAGGAGLRYLYDVARFHGATSINLVGDLDTVTNPKGTVA HDGDYLIVAKTVEADAAHGSVTRTIEADAAHGSVTRVVELITYIATKYAVFDTGSRVTENVLAFIYKQLLFSAG AELNKLDLGTVVSPVSGPKLGANTLLELVIFGRAFGGNTQDFGRVFQFLEASAGSKLNADTSLFILASKGQETST QPIATIFAWSRAGQSHLGLPIFGSAVEAKGVEPSHAISGARALIANGTGPYFYLPKNTVIASGGYGRAAAAQINY IRSGNVVPGYHGAVLRVALLGAGAGIGNPLGLLLKVPQVILAVGLPARHLVEEPNQLIKLVNELTEFAKGSGHH HHHH
CP03	42.3 kDa	>99 %	MGSAWSHPQFEKGGGSGGGSGGSAWSHPQFEKLEVLFQGPAAAKVFADYEEYVKDFYELEPHKVAAAFPGD VDRGLAGVENVTELKIGEEYISDLDQLRVIFLENYRLLSYVDDEAFIRVLVDLERTLSDYNIQKESTLHLVLRTIT LEVEPSDTIENVKEGIPPDQQRIGDYAGIKVLGIDGGEGKEALDFFARVVGLSTLPEIYEKANELLINVKGVIFYE SHGKGNPTVEVELTTEKTFAEALRAADALLLKTAGIQIVADDLTVTNPKNVNDVIAPAFVKVNQIGTLSESIKA VDDFLISLDGTANKHLVDEPQNLIKYLYEIARQTALVELLKLGEYGFQNALIVRDAFLGSFLYEYSRLVNELTEF AKGSGHHHHHH
CP04	43.0 kDa	>99 %	MGSAWSHPQFEKGGGSGGGSGGSAWSHPQFEKLEVLFQGPAAAKVFADYEEYVKDFYELEPHKVAAAFPGD VDRGLAGVENVTELKHLVDEPQNLIKLVNELTEFAKIGEEYISDLDQLRVIFLENYRLLSYVDDEAFIRVLVDLE RTLSDYNIQKESTLHLVLRTITLEVEPSDTIENVKEGIPPDQQRYLYEIARLVTDLTKAEFVEVTKIGDYAGIKVL GIDGGEGKEALDFFARVVGLSTLPEIYEKANELLINVKGVIFYESHGKGNPTVEVELTTEKTFAEALRAADALLL KTAGIQIVADDLTVTNPKNVNDVIAPAFVKVNQIGTLSESIKAVDDFLISLDGTANKLGEYGFQNALIVRDAFLG SFLYEYSRGSGHHHHHH
CP05	45.4 kDa	>99 %	MGSAWSHPQFEKGGGSGGGSGGSAWSHPQFEKLEVLFQGPAAAKVFADYEEYVKDFYELEPHKVAAAFPGD VDRGLAGVENVTELKHLVDEPQNLIKLVQETLEFAKSELEEQLTPVAEETRLGPLVEQGRAATVGSLAGQPLQ ERQWAGLVEKLAVYQAGARNILTSNNIDVKIPTTFENGRNPNLPPETVDSLKVLNQELRWYEIEKAEFVEVTKI TLLSALVETRNLAVSQVVHKAVLDVFEEGTEASAATAVKADLSGITGARHVEDVPAFQALGSLNDLQFFRYSL TYIYTGLSKEIPAWVPFDPAAQITKQKWEAEPVYVQRNYPSLELDKNYLNYGEEGAPGKGEAGAPGEEDIQGP TKWQQQGDLQDTKLGEYGFQNALIVRDAFIGTFLYEYSRGSGHHHHHH
CP06	46.0 kDa	>99 %	MGSAWSHPQFEKGGGSGGGSGGSAWSHPQFEKLEVLFQGPAAAKVFADYEEYVKDFYELEPHKVAAAFPGD VDRGLAGVENVTELKHLVEEPNQLIKLVNELTEFAKVAAGAFQGLRGQTLLAVAKTLDLGENQLETLPPDLLR ENQLEVLEVSWLHGLKFPLTNAIKILGQQVPYATKTLLSVGGWNFGSQRVTIDSSYDIAKGNQWVGYDDQESV KADFVETVKFIPLIPIPERASPFPVYKTDGAAPNVAPSDVGGGGGRVQVTSQEYSARNSEPQDEGELFQGVDPRT HLGEALAPLSKLADLASDLLLQYLLQGGARVGEEDEEAAEAEAEAEEAERFQLTFPLRTPAAETLSQLGQTLQS LKALSGNVIAWAESHIEIYGGATKLGDYGFNNALIVRDAFLGSFLYEYSRGSGHHHHHH