

From Relative to Absolute Quantification of Tryptic Peptides with Tandem Mass Tags: Application to Cerebrospinal Fluid

Loïc Dayon^{§a}, Natacha Turck^a, Alexander Scherl^a, Denis F. Hochstrasser^{ab}, Pierre R. Burkhard^c, and Jean-Charles Sanchez^{*a}

[§]SCS Mettler-Toledo Award Winner (Oral Presentation)

Abstract: Quantification is a major task in proteomics. Among the different analytical strategies to enable peptide and protein quantification, tagging with isotopic labels has emerged as a practical, versatile, and efficient alternative. In particular, isobaric labels, such as TMT or iTRAQ, are now widely employed to make relative comparison of the protein amounts in separate biological samples with tandem mass spectrometry (MS/MS). We used herein a shotgun proteomic approach based on labelling with tandem mass tags (TMTs) for the relative quantification of proteins, and the absolute quantification of their tryptic peptides in human cerebrospinal fluid (CSF). First, the comparison of *ante*- and *post-mortem* CSF samples was carried out for the discovery of protein marker candidates of brain-damage disorders. Second, tryptic peptides representative of these candidates were measured in CSF using reporter-ion calibration curves. These works highlighted the advantages and limitations of such strategies for quantification purposes in proteomics.

Keywords: Brain disorders · Cerebrospinal fluid · Mass spectrometry · Proteomics · Quantification

1. Introduction

In proteomics, the discovery of potential protein biomarkers of diseases is mainly carried out by comparison of samples from patients exhibiting the pathology under study with those from control individuals. Two-dimensional gel electrophoresis is a possible method to compare several protein samples, either through parallel or concomitant analyses after differential fluorescent labelling of each sample (*i.e.* fluorescence difference gel electrophoresis (DIGE)^[1]). Because of the central role of mass spectrometry (MS) for the identification of proteins, the use of mass labels (also called mass tags) has rapidly emerged as a rational alternative while

supporting the dissemination of gel-free proteomic methods. In this field, technologies like the isotope-coded affinity tagged (ICAT),^[2] and more recently tandem mass tags (TMTs),^[3] have opened the way to the quantitative comparison of human samples with shotgun proteomic methods. The use of ‘light’ and ‘heavy’ isotopes is the common denominator of these mass tagging

techniques, as it allows the easy differentiation of the tagged protein or peptide samples with MS or tandem MS (MS/MS). In particular, the TMT approach relies on the design of isobaric labels, which release a characteristic mass signature when fragmented by collision-induced dissociation (CID), well-known as reporter-ion (Fig. 1). Basically, TMT^[4] or isobaric tags for rela-

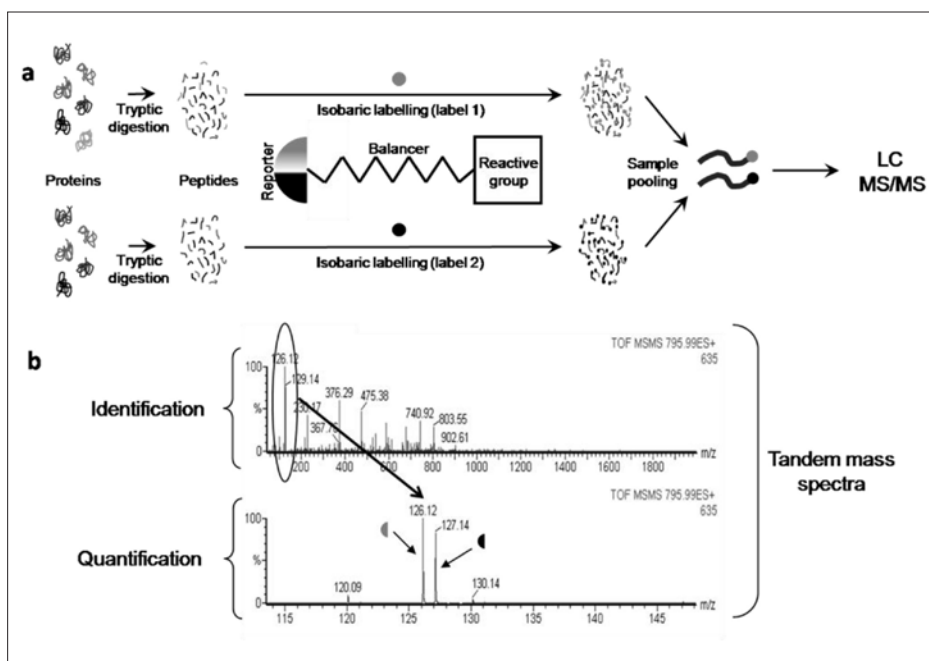


Fig. 1. Schematic representation of quantification with isobaric tandem mass tags (TMTs).

*Correspondence: Dr. J.-C. Sanchez^a

Tel.: +41 22 379 54 86

Fax: +41 22 379 55 05

E-mail: jean-charles.sanchez@unige.ch

^aBiomedical Proteomics Group

University of Geneva

CH-1211 Geneva 4

^bClinical Proteomics Group

Geneva University Hospitals

CH-1211 Geneva 14

^cDepartment of Neurology

Geneva University Hospitals and Faculty of Medicine

CH-1211 Geneva 14

tive and absolute quantification (iTRAQ)^[5] enable relative quantification of protein samples by comparison of reporter-ion abundances obtained after CID of the labelled tryptic peptides.

We applied such an approach to identify potential protein markers related to brain-damage disorders. *Post-mortem* cerebrospinal fluid (CSF) was taken as a model of massive brain injury based on the assumption that the mechanisms that follow death might share some similarities with those related to brain disorders.^[6] Using the TMT technology and MS/MS, comparison of *post-* and *ante-mortem* CSF provided several relevant proteins with increased levels in the CSF of deceased patients. Some of these concentration differences were validated with enzyme-linked immunosorbent assays (ELISAs).

To provide reliable markers, the discovery phase is followed by verification and validation phases in increasingly large cohorts of patients. In this context, absolute quantification of the biomarker candidates is required to assess their diagnosis and prognosis performances. Several studies using antibody-based assays have been carried out in our group to stage Human African trypanosomiasis patients,^[7] predict outcome following aneurismal subarachnoid hemorrhage,^[8] and manage stroke^[9] with proteins found initially in *post-mortem* CSF.^[4,6,10] Because the costly and time-consuming immunoassay development excludes systematic verification of all marker candidates found in the CSF discovery phase, TMT-based assays for the absolute quantification of CSF tryptic peptides with MS/MS were developed, and evaluated to help the early assessment of these markers.

2. Results and Discussion

2.1 Relative Quantification of CSF Proteins

TMTs comprise a set of structurally identical tags, which label peptides on free amino-terminus and epsilon-amino functions of lysine residues. Their chemical structure enables the incorporation of 'heavy' isotopes (¹³C or ¹⁵N) in the reporter group, and/or 'heavy' isotopes (¹³C or ¹⁵N) in the balancer group to provide tags with the same total mass. After activation with CID and subsequent MS/MS, the tags give reporter-ions at different masses depending on the number of 'heavy' isotopes in the reporter group. The so-called sixplex TMTs^[4] allow the comparison of up to six different extracts while the duplex TMTs^[11] are suited for two samples.

The general analytical strategy is depicted in Fig. 1. Basically, different protein extracts are digested with trypsin, and the resulting peptides are differentially labelled with different isobaric versions of the TMTs (Fig. 1a). Samples are then pooled. Labelled peptides originating from both extracts present the same physicochemical properties such as isoelectric point (pI), hydrophobicity, and mass. They can nevertheless be distinguished under CID (Fig. 1b). Their fragmentation pattern is identical, and allows their sequence identification through matching with theoretical tandem mass spectra computed from protein databases. But, the release of the reporter-ions from the TMTs varies for peptides coming from the different extracts with mass differences due to the 'heavy' isotopes. By comparing the abundances of the reporter-ions, it is possible to relatively quantify the identified peptides, and by extension their parent proteins in the compared extracts.

A shotgun proteomic workflow was developed for the analysis of the pooled TMT-labelled peptides from CSF with strong cation-exchange (SCX) liquid chromatography (LC) or isoelectric focussing (IEF) off-gel electrophoresis (OGE) as first dimension separation before reversed-phase (RP) MS/MS. *Ante-* and *post-mortem* CSF samples were compared using such approaches with sixplex TMTs,^[4] and duplex TMTs (unpublished work). Glutathione S-transferase P (GSTP1), protein DJ-1 (PARK7), and glial fibrillary acidic protein (GFAP) are a few examples of proteins which were found to be increased in the CSF of deceased patients with off-line RP-LC matrix-assisted laser desorption ionization (MALDI) tandem time-of-flight (TOF/TOF) MS. Fig. 2 provides an example of tandem mass spectrum zoomed on the sixplex TMT reporter-ion region that clearly evidenced a tryptic peptide with increased concentration in *post-mortem* CSF with respect to *ante-mortem* CSF.

The relative protein concentration differences found in CSF with TMTs were

checked with ELISA in a few cases.^[4] For instance, PARK7 was found with *post-mortem* CSF/*ante-mortem* CSF ratios of 55 and 58 respectively with TMTs and ELISA. As well, GFAP ratios were 8 and 10 with both techniques. First, these results validated the isobaric tagging approach with TMT for relative quantification of proteins. Second, they supported the protein list found in these studies. Several of these proteins were previously described to be related to brain disorders.^[4]

2.2 Absolute Quantification of CSF Tryptic Peptides

Isobaric TMTs were evaluated for the absolute concentration measurements of tryptic peptides of CSF. These developments aimed at providing a tool to select which proteins might be worth being validated by preliminary multiplexed verification with MS/MS of their tryptic peptides in a small number of patients.

Sixplex TMT reagents were used to label CSF, and four identical mixtures of synthetic proteotypic peptides (representative of a panel of protein marker candidates), which were prepared at four known concentrations. Labelled samples were mixed, and the pooled peptide sample was subjected to OGE (Fig. 3). Efficient fractionation of the peptides helped to identify lower abundance proteins, and reduce accidental CID events for both consistent identification and accurate quantification. The analyses of two identical CSF samples, separated into 24 off-gel fractions, and analyzed with RP-LC MALDI TOF/TOF MS (Fig. 3), demonstrated high resolution of the IEF separation. MALDI MS/MS showed 79% of the peptides in a unique fraction, and 43% in the same off-gel fraction between both samples. When considering the 24 fractions together, 54% of the peptides were commonly identified in sample 1 and sample 2. This result was very close to the percentage obtained for individual fractions, supporting the reproducibility of the IEF. The fraction pH

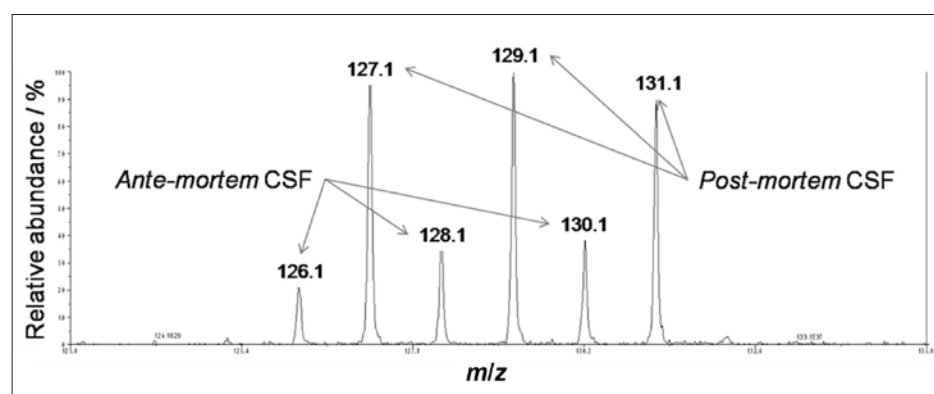


Fig. 2. Example of tandem mass spectrum (zoomed on the reporter-ion region).

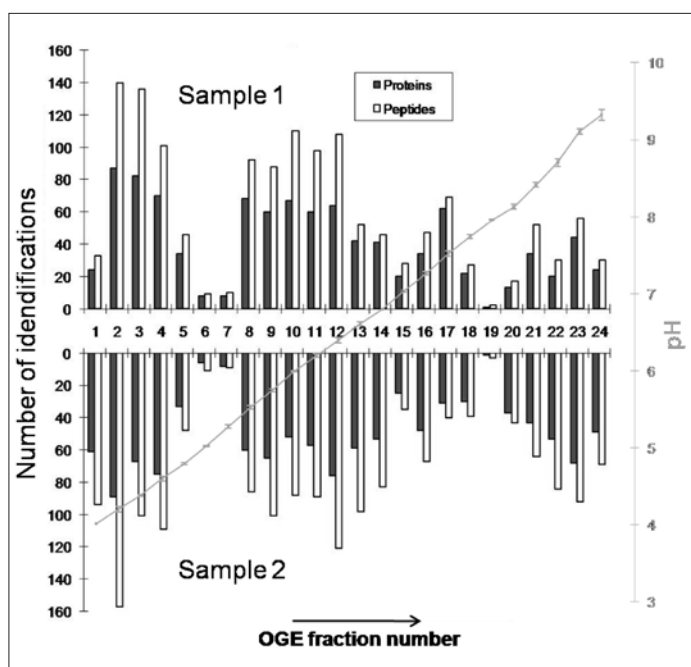


Fig. 3. Number of peptides and proteins identified in each OGE fraction for two identical CSF samples. Measured pH in each fraction.

new tool for the multiplexed early evaluation of peptide markers in complex protein samples.

3. Conclusions

Isobaric tagging with TMT was demonstrated to be a very useful technique for the differential quantitative analysis of peptides, and proteins in complex biological samples, and in particular CSF. In a model of massive brain injury, increased protein levels in *post-mortem* CSF with respect to *ante-mortem* CSF were found. These increases were confirmed with transversal methods. Proof-of-principle experiments for the absolute quantification of tryptic peptides with TMT reporter-ion calibration curves were shown using MS/MS platforms usually employed for biomarker discovery.

4. Experimental

4.1 Materials

β -Lactoglobulin (LACB) from bovine milk (~90%), iodoacetamide (IAA, $\geq 99\%$), tris(2-carboxyethyl) phosphine hydrochloride (TCEP) 0.5 M, potassium phosphate monobasic (99%), and α -cyano-4-hydroxycinnamic acid were purchased from Sigma. Triethylammonium hydrogen carbonate buffer (TEAB) 1 M pH = 8.5, sodium dodecyl sulphate (SDS, $\geq 98\%$), ortho-phosphoric acid (85%), and trifluoroacetic acid (TFA, $\geq 99.5\%$) were from Fluka. Formic acid (FA, 99%) was from Biosolve. Hydroxylamine solution 50 wt.% in H_2O (99.999%) was from Aldrich. Hydrochloric acid (25%), potassium chloride (> 99.5%), and ammonium dihydrogen phosphate were from Merck. Water for chromatography LiChrosolv[®], and acetonitrile Chromasolv[®] for HPLC ($\geq 99.9\%$) were respectively from Merck and Sigma-Aldrich. The Tandem Mass Tags[®] (TMT[®]) were provided by Proteome Sciences. Sequencing grade modified trypsin was from Promega. Synthetic peptides were prepared by Proteome Sciences, and received as lyophilized powder.

4.2 CSF Samples

Post-mortem CSF samples were collected by ventricular puncture at autopsy. Control *ante-mortem* CSF samples were collected by routine diagnostic lumbar puncture from living healthy patients. Clinical data of deceased and living patients have been previously reported.^[6] Each patient or patient's relatives gave informed consent prior to enrolment. The local institutional ethical committee board approved the clinical protocol.

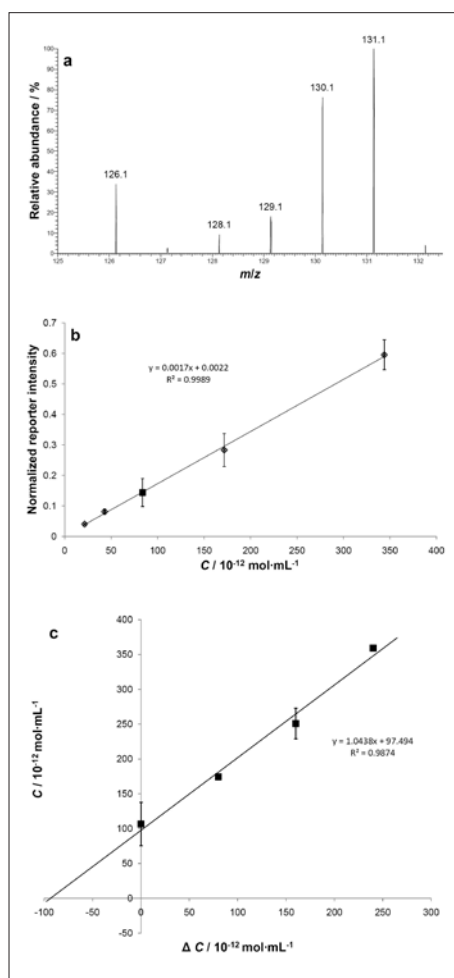


Fig. 4. Zoomed tandem mass spectrum of the reporter-ion region for ALPGQLKPFETLLSQNQQGK peptide (a). Determination of the concentration of ALPGQLKPFETLLSQNQQGK peptide (square) in CSF according to the synthetic standard analogues (diamonds) (b). Standard addition experiments for ALPGQLKPFETLLSQNQQGK peptide (c).

after IEF increased linearly from 4.0 to 9.4 with mean relative standard deviation (RSD) of 0.4%.

MS/MS was used to identify the peptides, and also obtained quantitative data. The spiking of the synthetic proteotypic peptides at four known concentrations (Fig. 4a) allowed four-point calibration curves to be established (Fig. 4b) to retrieve the concentrations (C) of the targeted peptides in CSF. For instance, ALPGQLKPFETLLSQNQQGK peptide of GSTP1, labelled with TMT with reporter-ion at $m/z = 126.1$ in CSF, could be quantified with respect to the reporter-ions at $m/z = 128.1$, 129.1, 130.1, and 131.1 released from the spiked standard synthetic peptides. The results of Fig. 4 were obtained with on-line RP-LC electrospray ionization (ESI) hybrid linear ion trap-orbitrap (LTQ-OT) MS. The quantitative values were noticed to occasionally vary from one peptide to another, even though belonging to the same protein. Quantitative differences and high RSDs might have resulted from different factors such as post-translational modifications, co-fragmentation of several peptides, tryptic digestion artefacts, peptide recovery, and stability.

To further characterize the quantification method, standard addition experiments were performed with incremental spiking of standard synthetic peptides in CSF to induce known concentration increases (ΔC). The results obtained for ALPGQLKPFETLLSQNQQGK of GSTP1 are displayed in Fig. 4c. They evidenced linear response, and accurate quantification. Absolute quantification using reporter-ion calibration curves can be thus achieved, and might be used as a

4.3 Sample Preparation

The CSF samples were dried, and then dissolved in TEAB 100 mM (adjusted to pH = 8 with diluted HCl) with SDS 0.01%, and 1 mM TCEP. The reduction was carried out at 60 °C for 1 h. Alkylation was performed with IAA 4 mM in the dark during 30 min. Trypsin was added (enzyme-to-protein ratio of 1:50 to 1:20), and the digestion was proceeded overnight at 37 °C. TMT labelling was achieved for 1 h, after addition of TMT reagent in CH₃CN. Hydroxylamine was added for 15 min reaction. The differentially TMT-labelled samples were pooled in a new tube, and dried. The pooled sample was desalted with Oasis® HLB 1 cc (30 mg) extraction cartridges from Waters.

4.4 Fractionation and Separation

The pooled TMT-labelled CSF sample was fractionated with SCX-LC or OGE. For SCX-LC, a 30 min gradient at 200 µL·min⁻¹ was used with a 2.1 × 100 mm, (5 µm, 200 Å) Polysulfoethyl A column (PolyLC), and an Alliance system (Waters). The gradient was run as already described.^[4] Fractions were collected every minute. The OGE separation was carried out overnight with the 3100 OFFGEL Fractionator (Agilent Technologies) using the 24 well format. Immobiline™ DryStryp pH 3-10, 24 cm, and IPG buffer pH 3-10 (GE Healthcare) were used. The fractions were collected, and their pH was measured (744 pH Meter, and Biotrode (Metrohm)). All fractions were dried, cleaned with C18 ultramicrospin column (Harvard Apparatus), and dried again.

RP-LC was performed using an Alliance system equipped with a flow splitter before MALDI MS or a NanoAcquity system (Waters) before ESI MS. In the off-line MALDI set-up, the fractions were separated with a home-packed 5 µm 200 Å Magic C18 AQ (Michrom) 0.1 × 100 mm column. The separation was run at 400 nL·min⁻¹ as already described.^[4] One minute fractions were deposited onto MALDI plates using a home-made LC-robot. In the on-line ESI set-up, peptides were trapped on a home-made 5 µm 200 Å Magic C18 AQ (Michrom) 0.1 × 20 mm pre-column, and separated on a home-made 5 µm 100 Å Magic C18 AQ (Michrom) 0.75 × 150 mm column comprising a gravity-pulled emitter. The analytical separation was run at a flow rate of 220 nL·min⁻¹ as already described.^[12]

4.5 Mass Spectrometry

MALDI TOF/TOF MS was performed with a 4800 Proteomics Analyzer (Applied Biosystems). The matrix (α -cyano-4-hydroxycinnamic acid in H₂O/CH₃CN/TFA 50%/50%/0.1% with 10 mM NH₄H₂PO₄) was added onto the plates.

All mass spectra were acquired in positive ionization mode with an *m/z* scan range of 800–4000. After selection of 20 precursors at the maximum, MS/MS experiments were done from the less to the most intense precursors with medium collision energy. ESI LTQ-OT MS was performed with a LTQ Orbitrap XL (Thermo Electron). For MS survey scans, the OT resolution was set to 60000, and the ion population was 5 × 10⁵ with an *m/z* window from 400 to 2000. Maximum of three precursors were selected for both CID in the LTQ and high-energy C-trap dissociation (HCD) with analysis in the OT.^[12] The normalized collision energies were 35% for CID, and 50% for HCD. For MS/MS in the LTQ, the ion population was 1 × 10⁴ (isolation width of 2 *m/z*), while for MS/MS detection in the OT, it was 2 × 10⁵ (isolation width of 4 *m/z*), with resolution of 7500, first mass at *m/z* = 100, and maximum injection time of 750 ms.

4.6 Identification and Quantification

Peak lists were generated from TOF/TOF and LTQ-OT raw data using respectively the 4000 Series Explorer software (Applied Biosystems), and an in-house written Perl script. CID and HCD data obtained with LTQ-OT were merged using a custom-made program.^[12] The *mgf* files were searched against UniProt-Swiss-Prot database using Phenyx (GeneBio). *Homo sapiens* taxonomy was specified for database search. Variable amino acid modifications were oxidized methionine. TMT-labelled peptide amino terminus, TMT-labelled lysine, and carbamidomethylation of cysteines were set as fixed modifications. Trypsin was selected as the enzyme with one potential missed cleavage. The peptide *p* value was 1 E-6 for searching TOF/TOF data and 1 E-3 for LTQ-OT data with respective parent ion tolerance of 1.1 Da, and 5 ppm. Protein and peptide scores were set to maintain the false positive discovery rate below 1%.

For quantification, the reporter-ion areas or intensities were corrected according to the isotopic purities of the reporter-ions. Because of the equal spiking of LACB inside the CSF samples, a normalization of the data could be achieved to reduce the manipulation bias. A normalization of the reporter intensities by the sum of all or a part of the reporter intensities was made in order to determine the relative abundance of each reporter.^[4,13]

Acknowledgements

Alexandre Hainard, Virginie Licker, Xavier Robin, Nadia Walter, Catherine Fouda, Dr. Yohann Couté, Dr. Olivier Lassout, Dr. Carla Pasquarello, Dr. Karsten Kuhn, Dr. Stefan Kienle, Dr. Peter Schulz-Knappe, and Dr. Ian

Pike are thanked for their help, and very fruitful discussions. The authors thank Proteome Sciences R&D GmbH & Co. KG for financial support.

Received: January 18, 2010

- [1] M. Unlu, M. E. Morgan, J. S. Minden, *Electrophoresis* **1997**, *18*, 2071.
- [2] S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, R. Aebersold, *Nat. Biotechnol.* **1999**, *17*, 994.
- [3] A. Thompson, J. Schafer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, T. Neumann, C. Hamon, *Anal. Chem.* **2003**, *75*, 1895.
- [4] L. Dayon, A. Hainard, V. Licker, N. Turck, K. Kuhn, D. F. Hochstrasser, P. R. Burkhard, J. C. Sanchez, *Anal. Chem.* **2008**, *80*, 2921.
- [5] P. L. Ross, Y. L. N. Huang, J. N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlett-Jones, F. He, A. Jacobson, D. J. Pappin, *Mol. Cell. Proteomics* **2004**, *3*, 1154.
- [6] P. Lescuyer, L. Allard, C. G. Zimmermann-Ivol, J. A. Burgess, S. Hughes-Frutiger, P. R. Burkhard, J. C. Sanchez, D. F. Hochstrasser, *Proteomics* **2004**, *4*, 2234.
- [7] A. Hainard, N. Tiberti, X. Robin, V. Lejon, D. M. Ngoyi, E. Matovu, J. C. Enyaru, C. Fouda, J. M. Ndung'u, F. Lisacek, M. Muller, N. Turck, J. C. Sanchez, *PLoS Negl. Trop. Dis.* **2009**, *3*, e459.
- [8] N. Turck, L. Vutskits, P. Sanchez-Pena, X. Robin, A. Hainard, M. Gex-Fabry, C. Fouda, H. Bassem, M. Mueller, F. Lisacek, L. Puybasset, J. C. Sanchez, *Intensive Care Med.* **2010**, *34*, 107.
- [9] N. Turck, P. R. Burkhard, X. Robin, A. Rosell, L. Dayon, N. Walter, C. Fouda, G. Wagner, M. Gex-Fabry, D. F. Hochstrasser, M. Mueller, J. Montaner, J. C. Sanchez, submitted for publication.
- [10] J. A. Burgess, P. Lescuyer, A. Hainard, P. R. Burkhard, N. Turck, P. Michel, J. S. Rossier, F. Reymond, D. F. Hochstrasser, J. C. Sanchez, *J. Proteome Res.* **2006**, *5*, 1674.
- [11] P. van Ulsen, K. Kuhn, T. Prinz, H. Legner, P. Schmid, C. Baumann, J. Tommassen, *Proteomics* **2009**, *9*, 1771.
- [12] L. Dayon, C. Pasquarello, C. Hoogland, J. C. Sanchez, A. Scherl, *J. Proteomics* **2010**, *73*, 769.
- [13] L. Dayon, N. Turck, S. Kienle, P. Schulz-Knappe, D. F. Hochstrasser, A. Scherl, J. C. Sanchez, *Anal. Chem.* **2010**, *82*, 848.