Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography B, 817 (2005) 225–230
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

Database independent detection of isotopically labeled MS/MS spectrum peptide pairs

Frank Potthast[a,*,1], Jiri Ocenasek[b,1], Dorothea Rutishauser[a], Martin Pelikan[c], Ralph Schlapbach[a]

[a] *Functional Genomics Center Zurich, Uni/ETH Zurich, Winterthurer Strasse 190, 8057 Zurich, Switzerland*
[b] *Computational Laboratory, ETH Zurich, Hirschengraben 84, 8092 Zurich, Switzerland*
[c] *Department of Mathematics and Computer Science, University of Missouri at St. Louis, 320 CCB 8001 Natural Bridge Road, St. Louis, MO 63121, USA*

Received 1 April 2004; accepted 7 December 2004
Available online 30 December 2004

Abstract

Mass spectrometry data generated in differential profiling of complex protein samples are classically exploited using database searches. In addition, quantitative profiling is performed by various methods, one of them using isotopically coded affinity tags, where one typically uses a light and a heavy tag. Here, we present a new algorithm, ICATcher, which detects pairs of light/heavy peptide MS/MS spectra independent of sequence databases. The method can be used for de novo sequencing and detection of posttranslational modifications. ICATcher is distributed as open source software.

Keywords: Proteomics; ICATcher; ICAT; Peptide sequencing; PTM detection; Differential analysis; Chromatography

1. Introduction

Differential profiling of protein content is a central task of proteomics. Being the large scale study of protein content, it originated from two-dimensional gel electrophoresis almost 30 years ago [1,2]. A non-gel approach is liquid chromatography with isotopic labeling of the peptides [3,4]. Today, several methods for isotopic labeling exist, one of them using cleavable ICAT reagents [4,5]; in this approach one adds a chemical label to all cysteine residues. There are two types of tags, ICAT light and ICAT heavy, which differ by nine neutrons in composition; the monoisotopic heavy tag containing nine ^{13}C atoms. Relative quantification can be achieved by comparing the MS profiles of peptides following multidimensional chromatographic separation [6,7]. The approach includes the identification of the peptides by comparing the MS/MS spectra with protein databases [8–10]. If a peptide is identified it can be assigned a differential quantification by comparing the MS spectra [6,7]. However, identification might fail [11–13] when the protein is not in the database or the search settings are not appropriate; this may, for example, occur if the peptide is posttranslationally modified in an unexpected or unconsidered way. Search methods for trying to avoid such problems arising from posttranslational modifications exist [14,15], but depend on the availability of sequence databases.

In this paper, we present the ICATcher algorithm which detects heavy/light pairs of MS/MS spectra independent of sequence databases. Given a mass shift Δ and a pair of MS/MS spectra, it addresses the question whether or not the two peptide spectrum pairs are related by a modification of weight Δ. Methods for the special case $\Delta = 0$ Da are used for comparing theoretic spectra against real MS/MS spectra [8]; here, the $\Delta \neq 0$ case is addressed. We focus on the cleavable ICAT

* Corresponding author. Tel.: +41 1 6353905; fax: +41 1 6353922.
E-mail address: frank@fgcz.ethz.ch (F. Potthast).
[1] Equally contributing authors.

1570-0232/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2004.12.009

light/heavy pairs having a mass difference of $\Delta = 9.03$ Da. The quality of the algorithm is evaluated by a comparison with Sequest [8,16,17], and by a statistical approach. Note that the Sequest evaluation needs a sequence database while our approach does not. A distinct advantage of our algorithm is that we do not make assumptions about the fragmentation pattern, we just compare pairs of MS/MS spectra; in this way, we automatically deal with internal and rearranged peptide fragments in a correct way. We show the usability of ICATcher for the more general Δ-case and present a Bayesian model for the estimation of selectivity and sensitivity. An organization of classified pairs into links, clusters, and hyperclusters is also presented and discussed.

The method has several applications: first, it can be used for de novo sequencing [9,11,13,18–21], because it detects pairs of light/heavy ICAT spectra and thereby labels the MS/MS fragments containing cysteine. When extended to the more general Δ-case, the method can obviously be used for the detection of posttranslational modifications, probably limited to modifications not having a strong influence on the overall fragmentation and ionization behavior. The ICATcher software is distributed as open source [22] under the GNU General Public License version 2 [23].

2. Experimental

This article presents an algorithm for detecting pairs of MS/MS spectra being ICAT light/heavy pairs. In this paragraph, we shortly describe the experimental origin of the data used for evaluating the algorithm. For detailed information on performing ICAT experiments, have a look at, e.g. [4,5]. All used datasets are available on the ICATcher website [22].

2.1. Description of the experimental procedures

The study was done on human colon tumor cell lines HCT116 (deficient in wild-type hMLH1) and SW620. Equal amounts of soluble protein fraction from HCT116 and SW620 were reduced and labeled with the light and heavy cleavable ICAT reagent (Applied Biosystems, Foster City, CA, USA), respectively. Both processed samples were then combined and digested with Trypsin (Sequencing Grade Modified Trypsin, Promega, Madison, WI, USA) at 37 °C for 15 h. To reduce the complexity of the sample, the digest was fractionated into 24 fractions using Ethan HPLC system (Amersham Biosciences AB, Uppsala, Sweden) with a strong cation exchange chromatography column SCX Polysulfoethyl (PolyLC, Columbia, MD, USA). Fractions 1–4 and 21–24 were pooled and the resulting 18 fractions processed with an Avidin affinity column (ICAT Cartridge Avidin, Applied Biosystems, Foster City, CA, USA) to extract the ICAT labeled peptides. The acid cleavage of the biotin tag was performed as suggested by the manufacturer. All steps were performed according to the manufacturer's instructions. Sep-Pack columns (Vac C18 1cc, 50 mg, Waters, Milford, MA, USA) were then used for further cleanup of the affinity-purified fraction. One third of each fraction was used for LC–MS/MS analyses.

2.2. Capillary chromatography and mass spectrometry

Aliquots of 5 μl were loaded onto an analytical reversed-phase capillary column (Magic C18, 75 μm × 8 cm; 200 Å, TipiTips-ED, Switzerland) using a fully automated nanoflow LC-system consisting of a PAL auto sampler (CTC Analytics AG, Zwingen, Switzerland) and binary Rheos 2000 pump (Flux Instruments, Basel, Switzerland). All LC–MS/MS runs of ICAT labeled peptide fractions were performed using a 160 min binary gradient using solvent A (5% acetonitrile, 0.2% formic acid) and B (80% acetonitrile, 0.2% formic acid). Peptides were eluted with a linear gradient from 0 to 50% solvent B in 122 min and 50–100% B in 8 min, followed by 100% B for 5 min and 100% A for 22 min to equilibrate. Average flow at the tip was ∼0.25 μl/min after splitting.

The LC system was directly coupled to a Thermo Finnigan LCQ Deca ion trap mass spectrometer (San Jose, CA, USA) equipped with a nano-spray ionization source. Each MS full scan was followed by three MS/MS spectra of the three most intense peaks. Dynamic exclusion was enabled using the following conditions: repeat count, 2; repeat duration, 1 min; exclusion duration, 4 min; exclusion width of 3 Da and a reject mass list with the following masses: 304.00, 371.00, 391.00, 445.00, 1522.00, 1622.00, 1722.00, 1822.00.

2.3. Chemicals and reagents

Polyimide-coated fused-silica capillaries (75 μm i.d.) were from BGB (Analytik AG, Böckten, Switzerland) and Magic C18 resins (5 μm, 200 Å pore) were from Michrom BioResources (Auburn, CA, USA). Formic Acid was from Sigma–Aldrich (Steinheim, Germany), HPLC-grade acetonitrile from Scharlau (Scharlau Chemie S.A., Barcelona, Spain) and water from a Milli-Q Synthesis A10 (Millipore, Bedford, MA, USA).

3. Methods

In this section, we describe the algorithmic comparison of two MS/MS spectra, i.e. how it is decided whether or not they are related by a mass shift Δ. The algorithm is operated as a two-filter process: first, all spectra which do not have appropriate precursor mass differences are discarded. Second, the ICATcher algorithm, described below, is applied as a second quality filter. The section is concluded with an abstract organization of a large set of MS/MS spectra into links, clusters and hyperclusters and a description of a Bayesian statistical framework to estimate the sensitivity and selectivity of the method.

3.1. Algorithm for computing the similarity between two spectra related by a mass shift

Given a large set of MS/MS spectra, the task is to find those pairs being related by a mass shift Δ, here with a focus of the $\Delta = 9.03$ Da shift induced by the ICAT light/heavy tags.

As mentioned above that the algorithm is operated as a two-filter process; in this paragraph we describe the second filter, the ICATcher score, which is supposed to discard the false positives remaining after the first filter. The inputs to ICATcher are two peaklists (for example, .dta files) extracted from the MS/MS raw-spectra. Extraction of peak coordinates from the raw data is done via external software shipped with the spectrometer. Output is a distance in the interval [0,1] with zero indicating a pair and one discarding the pair.

The list of n_A peaks of spectra A is defined by the array of masses M^A and the array of intensities I^A. To compute the distance between spectra A and B, first both spectra are normalized to compensate for different intensities. In praxis, this step is needed to process proteins with different levels of expression:

$$\boldsymbol{I}'^{A} = \frac{\boldsymbol{I}^{A}}{\sum_{j=1}^{n_A} \boldsymbol{I}_j^{A}} \qquad \boldsymbol{I}'^{B} = \frac{\boldsymbol{I}^{B}}{\sum_{j=1}^{n_B} \boldsymbol{I}_j^{B}} \tag{1}$$

Then, for each jth peak the algorithm computes the array of improvements after shift:

$$\mathbf{impr}_j = \left| \boldsymbol{I}'^{A}_j - \boldsymbol{I}'^{B}_{\text{nearest}(M_j^A)} \right| - \left| \boldsymbol{I}'^{A}_j - \boldsymbol{I}'^{B}_{\text{nearest}(M_j^A+\Delta)} \right| \quad j = 1, \ldots, n_A \tag{2}$$

where nearest(m) denotes

$$\text{nearest}(m) = arg \min_{k=1,\ldots,n_B} \left| \boldsymbol{M}_k^{B} - m \right| \tag{3}$$

and Δ is the mass difference between peptides A, B; Δ being positive when A is lighter than B, otherwise Δ has a negative sign.

Assuming that both b- and y-ions are contained in the spectrum B, about half of the peaks in B should be shifted back by $-\Delta$ to obtain a new spectrum $\boldsymbol{I}''^{B}$ closer to A. Thus, the **impr** array is sorted in descending order and the peaks corresponding to values in the first-half rank are shifted:

$$\boldsymbol{I}''^{B}_j = \begin{cases} \boldsymbol{I}^{B}_{\text{nearest}(M_j^A+\Delta)} & \mathbf{impr}_j \geq \text{median}(\mathbf{impr}) \\ \boldsymbol{I}^{B}_{\text{nearest}(M_j^A)} & \mathbf{impr}_j < \text{median}(\mathbf{impr}) \end{cases} \quad j = 1, \ldots, n_A \tag{4}$$

The ions containing ICAT-labeled cysteine typically have $\mathbf{impr}_j \gg \text{median}(\mathbf{impr})$, whereas the remaining ions have have $\mathbf{impr}_j \ll \text{median}(\mathbf{impr})$. The intermediate values in the **impr** array correspond usually to the peaks caused by contamination or noise, so the exact choice of the rank for shifting is not critical.

Finally, the evaluation of distance between A and B is based on the computation of dot product between $\boldsymbol{I}'^{A}$ and $\boldsymbol{I}''^{B}$:

$$\text{dist}(A, B) = 1 - \frac{\boldsymbol{I}'^{A} \cdot \boldsymbol{I}''^{B}}{\left|\boldsymbol{I}'^{A}\right| \cdot \left|\boldsymbol{I}''^{B}\right|} \tag{5}$$

This distance measurement returns a value between 0 and 1. The complete algorithm is run as a two-step filter process. First filter: only pairs of spectra that are of precursor mass difference $\Delta \pm 1$ Da from each other pass. This mass difference should generally be set to the precursor selection tolerance of the used instrument. Second filter: for those pairs passing the first filter, we compute the distance as described above, and declare pairs with a distance smaller than 0.2 to be true light/heavy pairs. A motivation for this cutoff value is given in the results section below. Depending on the exact task in mind, an operator might use a different cutoff, but she will always treat sensitivity against error rate.

3.2. Abstract organization of results

One can describe the organization of involved entities on four levels of abstraction: (1) the spectra, (2) links between spectra having a mass distance Δ, (3) clusters of spectra linked by links of level 2, (4) hyperclusters made up of clusters of level 3.

The hypercluster concept is illustrated in Fig. 1 for two $\Delta = 9.03$ Da (ICAT) cluster and two $\Delta = 16$ Da (oxidation) clusters.

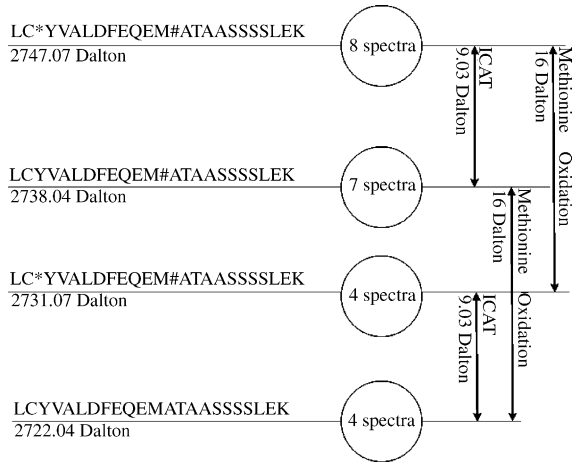

Fig. 1. Example of a hypercluster consisting of four clusters for a human gamma-actin peptide (Swissprot P02571). The peptide contains one cysteine which might be labeled with a light or a heavy ICAT tag, and one methionine which may be oxidized or not; i.e. there are four possible modification forms and therefore four possible clusters in total. Each of the four levels in the graph presents one modification form of the peptide. The light/heavy forms of ICAT-labeled cysteine are denoted with C/C*, respectively. Oxidized methionine is denoted with M#. In this example, all peptide forms have been detected several times, e.g. there were eight spectra measured on the C*M# precursor ion of this peptide. Indicated masses are average masses as calculated by Sequest.

3.3. Bayes rule for interpreting the ICATcher score

We currently use an ICATcher cutoff of 0.2 to differentiate between true and false positives. However, more precise ranking can be based on the estimation of the probability that a given ICATcher score really identifies a light/heavy pair. Let s denote the score, i the fact that the pair is an ICAT pair, f the fact that the pair passes the first filter and b the background of Bayesian inference. Then the probability that given score s is a true light/heavy pair can be stated using Bayes rule [24] as

$$p(i|s,\ f,\ b) = \frac{p(s|i,\ f,\ b)p(i|\ f,\ b)}{p(s|\ f,\ b)} \propto \frac{p(s|i,\ f,\ b)}{p(s|\ f,\ b)}$$

where $p(s|i,f,b)$ is the distribution of ICATcher scores (ICSs) among the ICAT pairs surviving the prefiltering, $p(i|f,b)$ is the proportion of ICAT pairs among all peptide pairs surviving the prefiltering, and $p(s|f,b)$ is the distribution of ICSs among all peptide pairs passing the first filter. Let us assume that the Sequest-identified pairs (described in Section 4 below) are an unbiased random sample out of all peptide pairs occurring in the data. Then, the distribution $p(s|i,f,b)$ can be estimated by taking the distribution of ICSs among the Sequest-identified and prefiltered ICAT pairs and renormalizing it. The distribution $p(s|f,b)$ is known by computing the ICSs of all prefiltered pairs. The probability $p(i|f,b)$ can be considered as a scaling factor. It is independent from s and thus does not affect the desired ordering of peptide pairs.

The ICATcher project is currently in the stage of analyzing the results from many datasets and approximating above distributions; they can be acquired only from the mass spectra of digested known proteins (since the Sequest identification is performed using a sequence database), but given our limited experience, we believe that they do not differ tremendously among different datasets and can be fixed as a part of ICATcher software for use on unknown proteins (work in progress).

4. Results and discussions

Our LCQ run has 4518 scans yielding 3966 MS/MS spectra in total, from which 1804 are doubly charged. As in ref. [11], we only look at the doubly charged spectra. Because our pair comparison has a light > heavy direction, we have to check for a total of 3,250,809 pairs. A 3705 of 3,250,809 pairs pass the first filter and are assigned an ICATcher score (ICS) as described in the methods section. Of these 3705 pairs, 284 pairs also pass the second filter having an ICS below 0.2. Processing time is roughly 10 min on a Pentium IV processor.

The 3705/284 pairs passing the first/second filter, respectively, all have a scan number distance and an ICS. Observe that scan number distance corresponds to elution time difference. In Fig. 2, we show the histograms of the first/second filter pairs with respect to scan number distance (Fig. 2A) and with respect to ICATcher score (Fig. 2B).

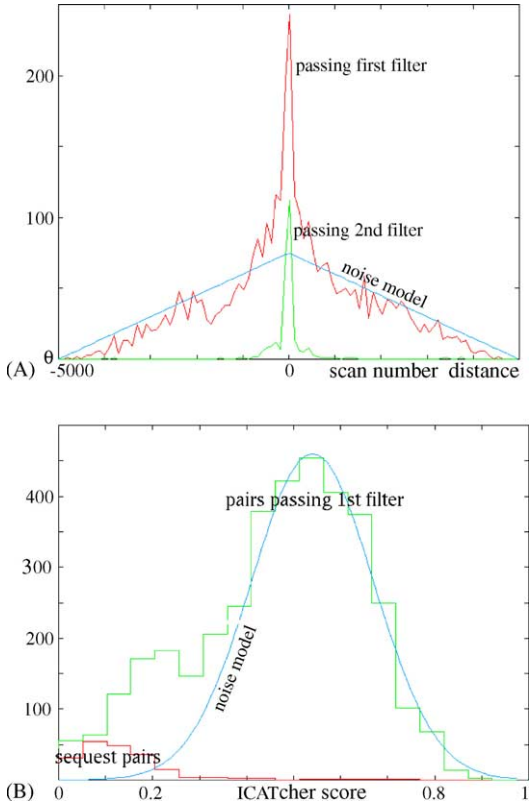

Fig. 2. Quality control of the classification algorithm. (A) Scan-number distance (corresponding to elution time difference) histograms (bin size 100) of all light/heavy pairs passing the first and those passing the second filter. The triangle indicates the noise model. It is clearly visible that pairs passing the second filter have a strong tendency to co-elute, which is in line with experimental knowledge of the ICAT reagent. The total distribution is close to the expected random distribution plus the second filter ICATcher signal, which is a footprint of a good classification. (B) Histogram of ICATcher scores (bin size 0.05). The two step-histograms are the distribution of all ICATcher scores passing the first filter, and (smaller one) the distribution of scores for all correct pairs identified by Sequest and Peptide Prophet. The continuous curve is a scaled Gaussian model fit for those peptide links which pass the first filter but are false positives. Parameters: $\sigma \approx 0.13$, maximum at $\mu \approx 0.54$. Only pairs with an ICATcher score below 0.2 are classified as true $\Delta = 9.03$ light/heavy pairs.

If one would pick 3705 pairs randomly, one would expect on average the triangular scan-number-distance distribution of Fig. 2A. It is easy to see that the first-filter passers follow this random distribution well except the subset also passing the second filter and thereby being classified as correct pairs. Another quality indicator for the method is that the second filter passers are strongly peaked around zero scan number distance which is a sign of co-elution. This is well known [4] and not used in the ICATcher score. Therefore, Fig. 2A is both a confirmation of the algorithmic quality as well as a confirmation for co-elution of the cleavable ICAT reagents.

For quality control purposes, we also analyzed all MS/MS spectra with Sequest [8]. The spectra were searched against a protein sequence database [25] consisting of homo sapiens proteins in SWISS-Prot and TrEMBL. Sequest search

parameters were as follows: peptide mass tolerance: 3 Da, MS/MS tolerance: 0.5 Da, maximum number of missed cleavages in a peptide = 2, enzyme = trypsin (cleaving at K,R). Cysteines were searched to be statically modified by the ICAT tag of weight 227.2613 Da and an extra weight for the heavy tag of 9.03 Da. Additionally, we looked for variable oxidations of methionine. The quality of MS/MS spectra to peptide assignments was estimated with peptide prophet [16,17]; 630 peptide assignments had a reliability of above 95%. For this 95% cutoff, the Peptide Prophet statistical model estimates the sensitivity to 92% and the error rate to below 1%. The 630 high-quality MS/MS-peptide assignments give rise to 288 doubly-charged ICAT light/heavy links. In general, the number of true pairs will be higher than the number of pairs detected correctly by Sequest.

Of the 288 "Sequest true" pairs, 224 (77%) passed the first filter. This loss of true positives is due to the fact that the precursor mass tolerance was set to 3.0 Da for the Sequest search; whereas ICATcher only allows a maximum tolerance of 1.0 Da. A 157 of the 224 filter passers (70%) also pass the second filter. In total, we detect 54.4% of the "Sequest true" light/heavy pairs.

Of the 284 "ICATcher true pairs", 157 (55.3%) are thus confirmed by Sequest. It is reasonable to expect that quite some of the 127 other pairs, not confirmed by Sequest, are true ICAT light/heavy pairs nevertheless; the quantitative aspect is speculative. A first guess might be that many of the 127 unexplained pairs are due to unspecific cleavage; thus we also did a Sequest search taking into account trypsin-unspecific cleavages. In this scenario, Peptide Prophet yields 620 peptide assignments with a probability above 95%. Of these 620 peptide assignments, eight have a single unspecific "cleavage"; six of them having a proline on the N-terminal edge. This gives some evidence that the tryptic digest was highly specific in our experimental setup. It has recently been claimed [26] that this might generally be the case for trypsin digests. Our eight "unspecific" assignments contain one light/heavy pair, belonging to the PCNA_HUMAN peptide I.PEQEYSCVVK.M. The corresponding spectrum pair has an ICS of 0.09 and was therefore correctly classified by ICATcher. Considering monomethylation of Arginine, Sequest identified one additional ICAT pair AVCMLSNTTAIAEAWAR with an ICS of 0.07. Dimethylation did not identify any additional pair. One hundred and twenty-five pairs thus remain to be explained. We lack software which could search through this set of 125 spectrum pairs in a systematic way for PTMs; it would be highly desirable to have a mechanism like that in place. Furthermore, we tried to perform some DeNovo sequencing on unidentified spectra, but without any obvious success. We believe the ICATcher > DeNovo approach would be feasible on high-accuracy instrument data like those obtained from a modern Q-Tof or equivalent. DeNovo sequencing software taking two spectra plus the additional information (ICAT pairs) into account is currently lacking as well.

Considering the ICS distribution, one might think that the distribution of "false first passers" can be fitted with a Gaussian distribution as indicated in Fig. 2B. This would suggest that most of the ICATcher declared pairs not confirmed by Sequest would be true pairs. Both selectivity and specificity of the algorithm could easily be improved by including the scan-number-distance into the ICATcher algorithm. We do not find this attractive because we plan to use the method also for light/heavy tags which do not co-elute.

A central issue in the classification of light/heavy pairs is the mass accuracy of the spectrometer, both as precursor mass selection and MS/MS mass accuracy are concerned. The instrument used for this study has a precursor mass accuracy of around 1 Da and an MS/MS mass accuracy of approximately 0.5 Da. It is reasonable to expect that the discriminative power of ICATcher will be even more distinct on instruments having better mass accuracies. In our lab, we can achieve 0.01 Da mass accuracy in MS/MS mode on a Q-Tof with internal calibration.

5. Conclusions

An algorithm, ICATcher, for the detection of MS/MS light/heavy spectrum pairs has been described and we have shown its quality by (1) comparing it with high-quality peptide identifications and (2) checking that the expected elution characteristics is in line with experimental knowledge. Despite the rather low mass accuracy of the instrument used, ICATcher is able to detect light/heavy pairs with good sensitivity. It seems reasonable to expect it to work even better on the new generation of high-accuracy MS/MS spectrometers.

This paper describes an algorithm establishing light/heavy links between peptide pairs with a focus on the 9.03 Da mass shift of the light/heavy cleavable ICAT reagent. We also presented an abstract framework organizing such links into clusters and hyperclusters and illustrated the concept with a cleavable ICAT/oxidation hypercluster detected with ICATcher, being perfectly confirmed by Sequest. The potential application of the algorithm for detection of many different PTMs is obvious; as is the usability of such information for de novo sequencing projects.

Our future work will focus on the usage of ICATcher to the identification of phosphorylated peptides as well as the generalization of the ICATcher method towards a universal PTM classification framework.

Acknowledgements

We would like to thank Giancarlo Mara from the Institute for Molecular Cancer Research Zurich for providing the mismatch repair proficient/deficient cell lines, Sabine Schrimpf for help with HPLC fractionation, and Erich Brunner for support with the ICAT technology.

References

[1] P.H. O′Farrell, J. Biol. Chem. 250 (1975) 4007.
[2] J. Klose, U. Kobalz, Electrophoresis 16 (1995) 103.
[3] A. Chevchenko, I. Chernushevich, K.G. Standing, B. Thompson, M. Wilm, M. Mann, Rapid Commun. Mass Spectrom. 11 (1997) 1015.
[4] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Nature Biotechnol. 17 (1999) 994.
[5] J. Li, H. Steen, S.P. Gygi, Mol. Cellular Proteom. 2–11 (2003) 1198.
[6] D.K. Han, J.K. Eng, H. Zhou, R. Aebersold, Nature Biotechnol. 19 (10) (2001) 946.
[7] X.J. Li, H. Zhang, J.A. Ranish, R. Aebersold, Anal. Chem. 75 (23) (2003) 6648.
[8] J.K. Eng, A.L. McCormack, J.R. Yates III, J. Am. Soc. Mass Spectrom. 5 (1994) 976.
[9] M. Mann, M. Wilm, Anal. Chem. 66 (1994) 4390.
[10] D.N. Perkins, D.J. Pappin, D.M. Creasy, J.S. Cottrell, Electrophoresis 20 (18) (1999) 3551.
[11] D.L. Tabb, A. Saraf, J.R. Yates III, Anal. Chem. 75 (23) (2003) 6415.
[12] P.A. Pevzner, Z. Mulyukov, V. Dančík, C.L. Tang, Genome Res. 11 (2001) 290.
[13] P.A. Pevzner, V. Dančík, C.L. Tang, J. Computat. Biol. 7 (6) (2000) 777.
[14] C.L. Gatlin, J.K. Eng, S.T. Cross, J.C. Detter, J.R. Yates III, Anal. Chem. 72 (4) (2000) 757.
[15] Error tolerant search in Mascot, see www.matrixscience.com.
[16] A. Keller, A.I. Nesvizhskii, E. Kolker, R. Aebersold, Anal. Chem. 74 (20) (2002) 5383.
[17] A.I. Nesvizhskii, A. Keller, E. Kolker, R. Aebersold, Anal. Chem. 75 (17) (2003) 4646.
[18] R.J. Johnson, K. Biemann, Biomed. Environ. Mass Spectrom. 18 (1989) 945.
[19] C. Bartels, Biomed. Environ. Mass Spectrom. 19 (1990) 363.
[20] V. Bafna, N. Edwards, Bioinformatics 17 (1) (2001) 13.
[21] M. Münchbach, M. Quadroni, G. Miotto, P. James, Anal. Chem. 72 (17) (2000) 4047.
[22] J. Ocenasek, M. Pelikan, www.proteios.org/icatcher.
[23] GNU, http://www.gnu.org/licenses/licenses.htmlGPL.
[24] E.T. Jaynes, Max. Entropy Bayesian Methods Appl. Stat. (1986) 1.
[25] B. Boeckmann, A. Bairoch, R. Apweiler, M.C. Blatter, A. Estreicher, E. Gasteiger, M.J. Martin, K. Michoud, C. O'Donovan, I. Phan, S. Pilbout, M. Schneider, Nucleic Acids Res. 31 (2003) 365.
[26] J.V. Olsen, S.E. Ong, M. Mann, Mol. Cellular Proteom. 6 (2004) 608.