Improving Automatic Peptide Mass Fingerprint

Protein Identification by Combining Many Peak Sets

*Thorsteinn Rögnvaldsson[*α], Jari Häkkinen[β], Claes Lindberg[γ], György Marko-Varga[γ], Frank Potthast[δ], and Jim Samuelsson[ε]*

[α]School of Information Science, Computer and Electrical Eng., Halmstad University, Box 823, SE-301 18 Halmstad, Sweden
[β]Department of Theoretical Physics, Lund University, Sölvegatan 14A, SE-223 62 Lund, Sweden
[γ]Molecular Sciences, AstraZeneca R&D Lund, SE-221 87 Lund, Sweden
[δ]Functional Genomics Center Zürich, Winterthurerstr. 190 Y32 H52, CH-8057 Zürich, Switzerland
[ε]Genedata GmbH, Lena-Christ-Str. 50, D-82152 Martinsried, Germany

Abstract

An automated peak picking strategy is presented where several peak sets with different signal-to-noise levels are combined to form a more reliable statement on the protein identity. The strategy is compared against both manual peak picking and industry standard automated peak picking on a set of mass spectra obtained after tryptic in gel digestion of 2D-gel samples from human fetal fibroblasts. The set of spectra contain samples ranging from strong to weak spectra, and the proposed multiple-scale method is shown to be much better on weak spectra than the industry standard method and a human operator, and equal in performance to these on strong and medium strong spectra. It is also demonstrated that peak sets selected by a human operator display a considerable variability and that it is impossible to speak of a single "true" peak set for a given spectrum. The described multiple-scale strategy both avoids time-consuming parameter tuning and exceeds the human operator in protein identification efficiency. The strategy therefore promises reliable automated user-independent protein identification using peptide mass fingerprints.

---

[*] Corresponding author, email: denni@ide.hh.se, phone +46-35 16 74 77, fax: +46-35 12 03 48

1   Introduction

Mass spectrometry combined with database searching is today the preferred method for protein identification. The standard experiment is to separate the proteins, e.g. by 2D gel electrophoresis, digest the proteins with a highly specific enzyme, measure the masses of the peptide fragments with a mass spectrometer (typically a MALDI-TOF mass spectrometer) and then compare the peptide monoisotopic masses with expected monoisotopic masses from a database (protein or DNA database). This approach is known as peptide mass fingerprinting (PMF) [1-4]. Several algorithms exist for comparing the observed monoisotopic masses with expected peptide masses from a database [5-10] but most common today seem to be MASCOT [6], ProFound [7] and MSFit [8].

The success of a PMF experiment depends on several factors: The noise level in the spectrum, the mass accuracy, the amount and purity of the sample, the number of proteins in the sample, possible post-translational modifications, algorithm accuracy, and (to a considerable extent) operator skill and experience. A human operator is generally better than software algorithms at judging what is an interesting low intensity monoisotopic peak, as opposed to a noise peak, and what is a reasonable identity given mass deviations and sequence coverage. Automated mono-isotopic peak detection algorithms, e.g. [11-14], work excellently on strong spectra but their performance is often insufficient on weak spectra. This is unsatisfying since strong spectra are often produced by high abundant proteins, which typically represent the bulk of the cell activity. Low abundance proteins, on the other hand, produce weak spectra and often represent the interesting unknown processes. The typical situation in a laboratory today is to have human operators double-check the software results and the protein identities [15]. These operators also spend considerable time varying the parameters for the peak detection software, searching for "optimal" settings of parameters that yield the best identification results. An often overlooked but fundamental problem in this process is that a human operator is a subjective expert who makes judgments; there will always be borderline cases where the judgment can go either way. Two different operators will not pick identical peak sets, unless they are asked to pick only very few and strong peaks, and one operator can be more successful than another at identifying proteins from the same spectra. Also, a human operator's opinion on which peaks to pick may change with time, as demonstrated in this paper.

This paper describes a different approach. We abandon the idea of an optimal set of parameters for a peak picking algorithm and accept that monoisotopic peak picking is by nature a statistical process. It then follows that protein identification should be based on several peak sets, which represent different parameter settings for the peak picking algorithm, rather than a single peak set. We therefore combine information extracted at different signal-to-noise ratio levels in the spectrum into a composite judgment about the protein identity. In this way we achieve both high sensitivity on weak spectra, even better than the human operator, and avoid the dependence on the human operator. The proposed method is compared with manual peak picking, software-supported manual peak picking (the industrial standard) and two fully automated peak detection strategies that do not combine peak sets.

We have mostly used the Pepex peak picking software in our experiments because it was by design very simple to script and automate for the multiple peak sets in this study[1]. The results do not, however, depend on the specific peak picking software used and similar effects would probably have been observed with any other high quality peak picking software.

---

[1] Fully functional trial versions of the Pepex software are available from Halmstad University http://www.hh.se/staff/bioinf

We demonstrate our method on spectra that were acquired during an investigation of an airway wall remodeling process using human fibroblasts that had been treated with inflammatory cytokines. These samples had been metabolically labeled with S-35 to allow detection of newly synthesized proteins that were believed to also be low abundance proteins.

The data generation and analysis presented in this paper is outlined in Figure 1.
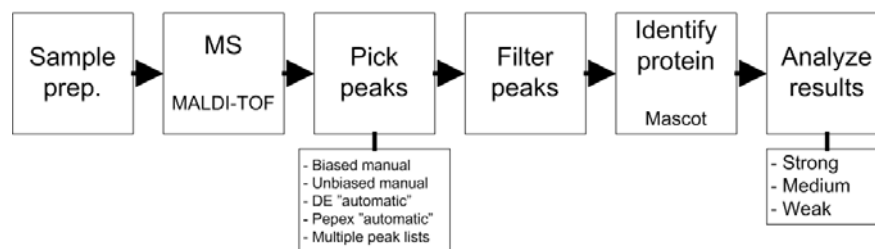
Figure 1 Overview of the experiments described in the paper. A typical peptide mass fingerprinting project was studied and our proposed multiple peak set method was compared to both automated and manual strategies for selecting the mono-isotopic peak lists. The final protein identification results were then analyzed for strong, medium, and weak spectra to see the benefits of each peak detection strategy for these different spectrum classes.

2   Experimental details

2.1   Sample preparation and instrumentation

Human fibroblasts were derived from the lung cell line HFL-1, obtained from American Tissue Culture Collection (ATCC). The samples were plated at 2.5×10^5 cells per well in 6-well plates and grown to confluence for one week. S-35 metabolic labeling was performed using 100 μCi S35-methionine per well in the presence of 10ng/mL of TGF-β, or vehicle, and incubated for 20 h. The cells were then lysed in 500 μL of 8M urea and 2% CHAPS. The amount of radioactivity in each gel spot was extremely low and did not impose any hazards or safety regulations related to the MS analysis. The small amount of S-35 in the labeled proteins could not be detected by MS and the mass shift of methionine containing peptides was insignificant.

Protein separation was done using 2D-gel electrophoresis with non-linear Immobiline strips with a pI window of 3-10. The strip was, after isoelectric focusing, treated with dithiothreitol and then with iodoacetamide to reduce and alkylate the proteins before running the second electrophoresis step. Gel pieces (2-3 mm in diameter) were cut out, washed with 25 mM ammonium bicarbonate and acetonitrile, and treated with porcine trypsin at 37 C overnight. Everything was completed according to standard protocols in the literature [16,17].

Mass spectra were acquired on a Voyager DE-Pro MALDI-TOF instrument (Applied Biosystems) operated in reflector mode using α-cyano-4-hydroxycinnamic acid matrix. Spectra were internally calibrated using the trypsin autodigest peaks at m/z 842.51 and m/z 2211.10 as reference masses.

2.2   Spectrum grouping

A set of 38 spectra was carefully selected to represent different protein sample amounts. The spectra were divided into three groups, designated "strong", "medium", and "weak", based on the intensity of the "unknown peptide peaks" relative to the trypsin autodigestion peak at m/z 2211. Spectra where more than 20% of the peaks (remaining after background peaks had been removed) had a higher intensity than the peak at m/z 2211 were classified as "strong" (9 spectra); "Medium" spectra were those where fewer than 20%, but more than 10%, of the peaks had a higher intensity

than the m/z 2211 peak (11 spectra); "Weak" spectra were those where fewer than 10% of the peaks had a higher intensity than the m/z 2211 peak (18 spectra).

3   Algorithmic details

3.1 User-dependent peak picking with single peak lists

Four user-dependent peak picking strategies were tried, two manual and two "automatic" (but manually tuned):

(a)   The first was an unbiased manual peak picking that was done without the aid of any peak-detection software, since peak-detection software will bias the decisions of a human operator. The spectra were plotted with a graphical display tool and two operators scanned through the spectra and marked the peaks that corresponded to monoisotopic peaks. The two operators spent between 20-45 minutes on each spectrum, listing the peaks that they visually deemed to be mono-isotopic peptide peaks.

(b)   The second was a biased manual peak picking done with the help of the Voyager 5 Data Explorer software. A third human operator tried different software parameter settings for each spectrum and selected an optimal set of parameters for each spectrum. This peak picking was considered typical for a pharmaceutical lab. The operator selected the peak lists by varying the absolute intensity parameters for each spectrum and adjusting the parameters to allow detection of small but significant peaks while excluding noise and background peaks. Additional mass ranges were inserted with different parameter settings when needed.

(c)   The third method was an industry standard "automated" peak picking where a human operator selected a single compromise set of parameters for the Voyager 5 Data Explorer software for all 38 spectra (the same human operator as in the biased manual peak picking described above). The operator tried several parameter settings with the Data Explorer software and the best overall results were achieved using an absolute intensity peak threshold. The subjectively best (compared to the biased manual peak picking) compromise performance was achieved using the following settings for all spectra: Use advance settings, set range 1 = m/z 750 – 1100 and minimum intensity = 500, set range 2 = m/z 1100 – 3500 and minimum intensity = 200.

(d)   The fourth method was a Pepex "automated" peak picking. A single set of Pepex parameters was chosen that matched the unbiased peak picking, (a) above, as well as possible. The subjectively, as judged by a human operator, best compromise performance was achieved using a relative peak intensity threshold, a signal/noise threshold of 1.5, and a noise density of 0.0005.

These four user-dependent peak picking methods were selected to provide relevant comparisons for the user-independent peak picking. The two manual methods, (a) and (b), represent how well a human operator can do when being very thorough (the unbiased peak picking) and when working quickly with the aid of a peak picking software (the biased peak picking). The two "automated" methods, (c) and (d), represent how good a specific peak picking tool can be at automatic peak picking when tuned by human operators. It is important to note, however, that both human operators who tuned the peak picking software were aware of a "true" answer (the biased and the unbiased manually picked peak lists) and the results with these automatically picked peak lists are therefore likely to be somewhat positively biased.

3.2 Identifying and removing contamination peaks

Contamination peaks were identified and removed before the peak lists were submitted to Mascot [6] for protein identification. All gel pieces included in this study were produced at the same time and it was therefore reasonable to expect similar contamination in essentially all samples. The contamination peaks were defined as those peaks that occurred unreasonably often in the spectra, using the algorithm described by Levander et al. [18]: Peaks with signal/noise ratios above 1.5 were picked from all 38 spectra and it was noted how often an individual peak occurred. For instance, if a peak occurred 38 times then it occurred 38/38 = 100% of the time; this was the case for, e.g., the trypsin 842.5 Da peak. The number of occurrences for each peak was then compared to the chance probability for a peak with the observed mass to occur several times in the experiment, based on the distribution of peptide masses in protein databases. If the observed number of observations exceeded the chance probability by more than two standard deviations, then the observed peak was labeled as a background peak. This resulted in 37 background peaks, many of which matched to (methylated) peptides from pig trypsin and human keratin.

3.3 Database protein scoring

All user-dependent peak sets were scored against SwissProt, version October 31 2001, using the Mascot tool from Matrix Science [6]. The settings for the Mascot searches were: Mass tolerance = 50 ppm, missed cleavages = 1, fixed modifications = carbamidomethylation of cysteine, variable modifications = methionine oxidation, species = all.

3.4 Combining results from multiple peak lists

There is no such thing as a "true" peak list: Different human experts pick different peak lists and automatic peak detection tools pick different peaks depending on the parameter settings. It should therefore be a good idea to combine results from several users or from several parameter settings; effectively integrating over parameter values so that the result becomes user-independent (i.e. no human operator is needed to tune the software). To test this idea, we varied the spectrum noise level estimation and the signal-to-noise ratio of accepted peaks. In the Pepex software this was done by varying the relative noise density parameter with the two values {0.0005, 0.001} and the signal-to-noise cut off with the seven values {1, 1.5, 2, 2.5, 3, 3.5, 4}, respectively, yielding a total of 2×7 = 14 peak sets for each spectrum. The number of peaks in the peak sets then varied between 0 and 450 (after contamination peaks had been removed). These peak sets were then submitted individually to the Mascot protein identification tool and combined using a logistic perceptron [19]. The combined protein identification success was then compared to the user-dependent peak sets described above.

The logistic perceptron is a single layer neural network with a logistic output unit, i.e. it is a function that takes a number of input variables $x_1, x_2, ..., x_m$, and outputs a number y between zero and one. The logistic perceptron's functional form is

$$y(\mathbf{x}) = \left[1 + \exp\left(v_0 + v_1 x_1 + v_2 x_2 + ... + v_m x_m\right)\right]^{-1}$$

where $v_0, v_1, v_2, ..., v_m$ are free parameters.

The logistic perceptron was trained to recognize a correct match and output a high value (i.e. close to one) for a correct protein identity and a low value (i.e. close to zero) for an incorrect protein identity. This was done is the following way: A human expert listed all the correct protein identities (this is described later in this paper) and all other reported protein identities were labeled incorrect. The desired output for correct (true) protein identities was set to one and the desired output for

incorrect (false) protein identities was set to zero. The perceptron then received information about the size of the database protein, the Mascot scores for the 14 peak sets, and the query lengths for the 14 peak sets (a total of 29 inputs). That is, the input variables where: x_1 = size of matched database protein, x_{k+1} = the Mascot score for peak list number k, and x_{k+15} = the number of peaks in peak list number k. The protein identities reported by Mascot on 37 of the 38 spectra were then used to train the perceptron, meaning that the parameters v_0, v_1, v_2, ..., v_m were tuned so that the sum square error

$$E = \sum_n \left[y(\mathbf{x}(n)) - d(n) \right]^2$$

was minimized ($d(n)$ denotes the desired output value, zero or one, for inputs $\mathbf{x}(n)$). The resilient propagation method [20] with 300 weight updates was used for the minimization; this method is similar to the more commonly known gradient descent method. The trained perceptron was then used to categorize the reported protein identities on the 38th spectrum, which it had not seen during training, as being either correct or incorrect. This procedure was then repeated 38 times so that out-of-sample judgments could be made by the perceptron on all the 38 spectra (this procedure is often referred to as leave-one-out cross-validation).

The described logistic perceptron method corresponds to a pattern recognition approach to the protein identification problem. The logistic perceptron is trained to recognize the characteristics for a correct protein identity when several different peak sets are used. The logistic perceptron is similar, but not identical, to the classical logistic regression model [21], and the logistic perceptron's output can be interpreted as an estimate of the posterior probability $p(\text{correct} \mid \mathbf{x})$. That is, the probability that the reported protein identity is correct, given the observation $\mathbf{x}$. This output can be further processed; one can require the estimated probability to be higher than a given threshold, e.g., 0.7. A threshold of 0.5 was used in the work described here.

4   Results

4.1   Similarities and differences between user-dependent peak lists

Some general observations, illustrated in Figure 2, could be made for manual and "automatic" peak sets. First of all, the human operator was an inconsistent peak picker and the unbiased manual peak sets did not agree completely with the biased peak picking. Furthermore, there were significant differences between the peak sets that were picked "automatically".

The peaks picked by the operator by manually tuning Data Explorer (biased peak picking) were compared to peak sets picked by the same individual two years earlier on the same spectra, also using Data Explorer. The new peak lists were on average about twice as big as the old ones but the smaller peak lists were not complete subsets of the larger peak lists. On average, only 56% of the new peaks matched the old peaks for the same spectrum and 92% of the smaller peak sets matched the larger peak sets for the same spectra. Thus, the operator disagreed with himself on at least one out of every ten peaks even though the same peak detection software was used on both occasions. This was because the operator had changed his opinion on how peaks should be picked. The operator had prioritized avoiding false positive peaks in the two year old lists, whereas priority was given to avoiding false negatives in the new peak lists; a "false positive" peak denotes a peak that is not in the spectrum but that is nevertheless picked by the software, and a "false negative" peak is a peak that is in the spectrum but that is missed by the software. Both priorities are reasonable and the observed difference between the peak lists illustrates how subjective manual peak picking is.

The unbiased manual peak lists contained on average about twice as many peaks as the biased manual peak lists. On average 43% of the unbiased manual peaks agreed with the biased peaks, and 81% of the biased peaks agreed with the unbiased peaks. Thus, the persons selecting peaks in an unbiased way disagreed with the operator who used Data Explorer on at least two in every ten peaks.

The peak sets picked "automatically" by the Voyager 5 Data Explorer software were on average about half the size of the biased manual peak lists selected by the operator. Almost all peaks picked "automatically" by Data Explorer agreed with the operator, which could be explained by the fact that the operator used Data Explorer to select peaks. It is clear that the operator preferred a conservative parameter setting for Data Explorer, leading to small peak sets. The parameter setting for Pepex was, in comparison, much less conservative and the peak lists were about 50% larger than the peak lists selected by the operator but smaller than the fully manual peak sets. On average, about 80% of the operator's peaks agreed with the Pepex peaks. The situation is illustrated with Venn diagrams in Figure 2.
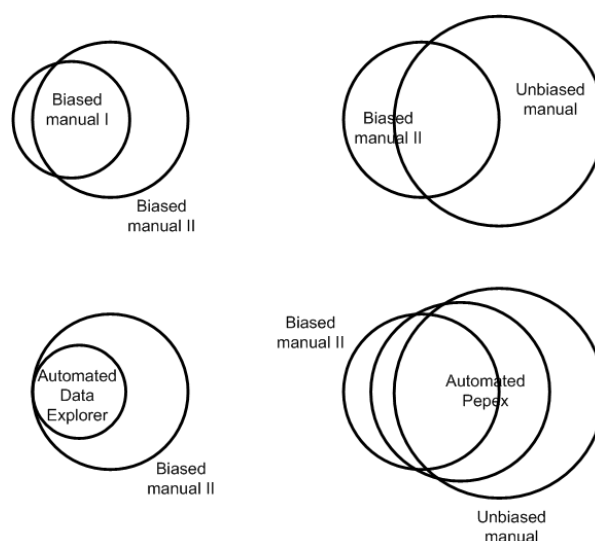

Figure 2 Venn diagrams illustrating the agreement and differences between the different peak sets. The "biased manual I" and "biased manual II" refers to the two different occasions when the biased manual peak picking was done; "biased manual I" is the older of the peak sets. The peak sets are explained in the text.

The subjectivity and uncertainty of peak picking is further illustrated in Figure 3, where the different peak sets are compared to a biased and an unbiased peak set using the efficiency and purity statistics. These are defined as *efficiency = (number of agreeing peaks)/(number of peaks picked manually)* and purity = *(number of agreeing peaks)/(number of peaks picked by automated software)*. If the software finds all the peaks that were picked manually, then the efficiency equals one. If the software pick no peaks that disagree with the manual peak list, then the purity equals one. In data mining literature, efficiency is often referred to as "recall" and purity as "precision". In medical literature, efficiency is called "sensitivity" and purity is called "positive predictive value".

Figure 3 (lower part) illustrates how the operator has the freedom to choose anything from a small peak set that is pure but inefficient to a peak set that is impure but efficient. However, there is no rule that tells which peak set is better than the others and peak picking is therefore a process where one should use more than one parameter setting to get stable and reliable results. The top left panel in Figure 3 shows the number of correctly (i.e. agreeing with the manually picked peaks) and incorrectly (i.e. not agreeing with the manually picked peaks) picked peaks in a single spectrum as a function of a user specified signal-to-noise threshold. A high signal-to-noise threshold means making few "mistakes" but also missing potentially informative peaks. The top right panel in Figure 3 shows

the corresponding purity and efficiency numbers (the unbiased manual peak list contained 87 peaks in this case). The lower left panel in Figure 3 shows the (efficiency, purity) pairs when a correct peak is defined by the unbiased manual peak set. The lower right panel in Figure 3 shows the (efficiency, purity) pairs plotted in an efficiency-purity graph when a correct peak is defined by the biased manual peak set. The filled square marks the particular peak set that was used in the automated Pepex peak picking. The fixed parameters for Pepex were selected to be as close (on average for all spectra) as possible to the upper right corner when comparing against the unbiased manual peak set.

It should be evident from Figure 3 that there exists no such thing as an objectively optimal set of parameters. The optimality of parameters depends very much on who defines the correct peaks.
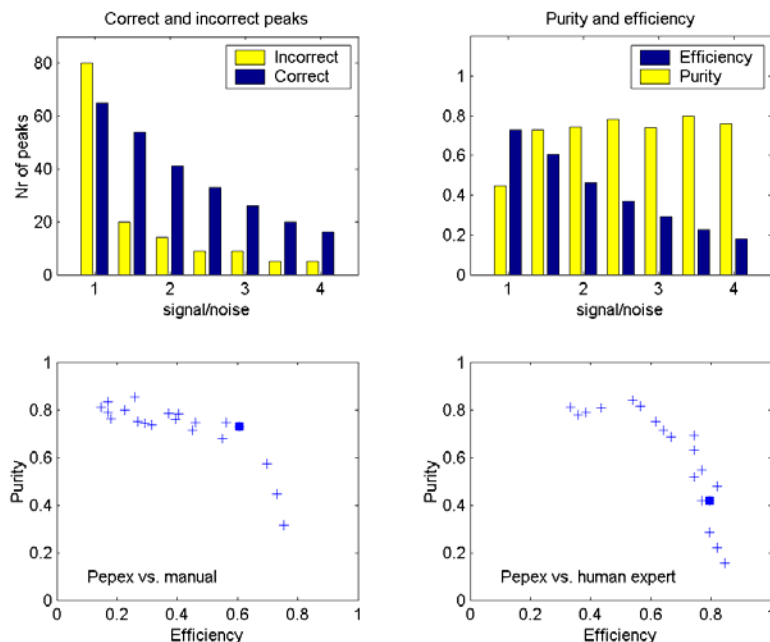
Figure 3 Illustration of the statistical nature of peak detection. The "manual" peak picking referred to is the unbiased manual peak list, and the "human expert" is the biased manual peak list. Each cross in the lower graphs corresponds to one peak set (one parameter setting for the peak picking algorithm), and the filled square marks the peak set achieved when the best compromise set of parameters were used for the Pepex peak picking algorithm. Each pair of purity/efficiency bars in the upper right graph corresponds to one cross in the lower left graph (the bar graph only shows some of the peak sets, in order to make the bar graph easier to read). More explanation is given in the text.

4.2   Protein identification efficiency

The number of correct and false protein identifications was compared for the different peak picking strategies and for the strategy where results were combined. A protein identity was deemed to be correct by looking at the size and pI of the protein compared to the spot position on the gel, the mass error of matched peaks versus theoretical peptide masses, the protein being of human origin, and the sequence coverage of the protein by the matching peptides. All these different aspects were judged subjectively by an experienced operator. A false hit was defined as a reported significant identity that deviated from one or more of the expected values mentioned above. Proteins with more than 1600 residues were ignored, because of the poor database statistics for large proteins, and obvious contamination proteins (keratins) were ignored.

The protein identification results are summarized in Figure 4. The best "automatic" result was achieved using the user-dependent Pepex peak set and the worst "automatic" result was achieved

with the user-dependent Data Explorer peak set. This does probably not reflect any quality difference between the peak picking algorithms. It is rather a consequence of how the parameters were set for the two tools. The difference in result illustrates how difficult/risky it is to manually tune peak picking software and select a single set of parameters.

The unbiased manual peak sets (a) produced two false hits, of which one was of human origin. The biased manual peak sets (b) produced one false hit, which was a protein of non-human origin. The user-dependent "automatic" Data Explorer peak sets (c) produced two false hits, none of human origin, and the user-dependent "automatic" Pepex peak sets (d) produced no false hit. These "automatic" results are likely to be overly optimistic since the parameters of the peak picking algorithms were set based on knowledge from the manual peak picking.

The logistic perceptron (Pc) gave six false hits. One of the false hits (for a spectrum classified as "strong") was on an extremely short protein (26 residues) of non-human origin. In one case, the perceptron gave as top hit a protein that was not reported as a significant hit by Mascot when any of the other peak sets were used. This protein (human HSP 90) was, however, judged as correct by the experienced operator based on molecular weight, pI and peptide mass precision. In this particular case, the perceptron also reported three other proteins of non-human origin as hits (i.e. this spectrum accounted for half of the false hits produced by the perceptron). If the decision threshold for the perceptron was increased from 0.5 to 0.7 then the number of false hits decreased by four and the number of correctly identified proteins decreased by only one. The perceptron was thus more uncertain about false hits than about correct hits.

The sequence coverage varied between the peak sets. The average sequence coverage with the automatic Data Explorer peak set was 25%. The average sequence coverage with the automatic Pepex peak set was 32%. The largest sequence coverage was achieved with the logistic perceptron method; the average sequence coverage was here 45%, 1.5 times the average sequence coverage achieved by the human operator (30%). This is of course not surprising since the multiple peak lists method includes some very large peak lists.
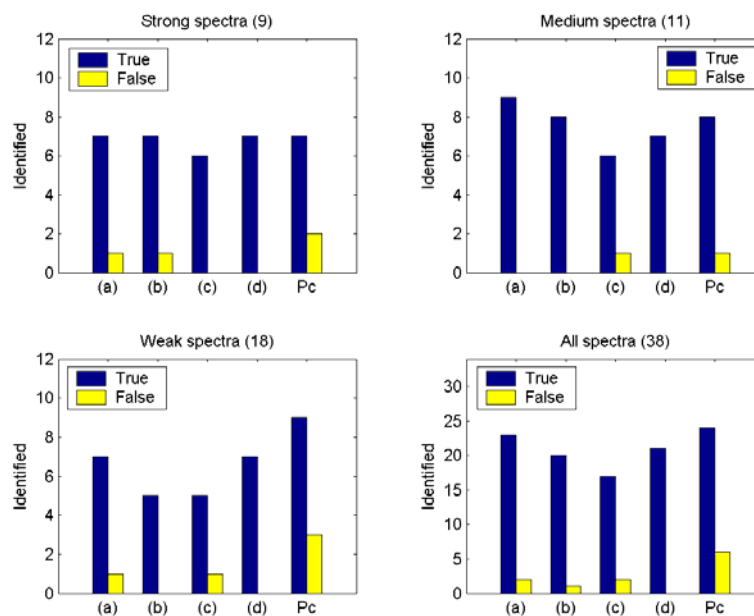
Figure 4 The protein identification results when using the different peak detection strategies and weighting strategies. The "Pc" label refers to the logistic perceptron approach where scores from several peak lists are combined. The other labels refer to the listing of different peak picking strategies in the text. The results are discussed in more detail in the text.

The time complexity varied much between methods. For the biased manual peak picking the operator spent about 10-15 minutes per spectrum. The unbiased manual peak picking took 20-45 minutes per spectrum. The automated Pepex runs took about 3-5 seconds per spectrum (excluding the time devoted to parameter tuning) and the automated Data Explorer took only a few seconds per spectrum (again excluding the time for parameter tuning). The Mascot runs took 5-10 seconds per peak list. A repeated search with 14 peak sets, using Pepex combined with Mascot, took about 2 minutes per spectrum, without the need for any manual parameter tuning.

5   Conclusions

We have suggested a method for combining results from several peak sets in peptide mass fingerprint experiments. The motivation being that one should not be satisfied with just a single parameter setting for a peak picking software but combine information from several peak sets that contain information about both strong and weak peaks in the spectrum. The suggested method provides completely automatic and reproducible peak picking and achieves at least as good protein identification as manual operation by a human operator using a commercial peak selection tool. Combining several peak sets in this way removes the arbitrariness of using a single parameter setting for a peak picking software. This is important since manual tuning of peak detection software turned out, in our experiments, to produce protein identification performance that fell anywhere between poor and excellent.

The largest number of proteins was found using multiple peak sets combined with the logistic perceptron. This method even found proteins that had low Mascot score but where an experienced operator believed the identification to be correct (albeit at the price of three false positives); this is remarkable and shows that the logistic perceptron was able to make judgments that reflected the human operator. The overall best protein identification, maximizing correct hits and minimizing false hits, was achieved with the tuned Pepex algorithm. However, this required considerable time in the tuning process and the end result is biased; it is questionable whether this could be repeated at a different laboratory by other people since parameter tuning is so subjective. User-independent methods that employ several peak sets must therefore be considered better in the long run.

The annotation hit-rate was essentially equally good for all methods on high-abundant proteins or proteins with strong MALDI response (Figure 4) whereas the hit-rate on low-abundant protein spots, or proteins with weak MALDI response, from the 2D-gels was 50% at best and 25% at worst. The best yield on weak spectra was achieved using the logistic perceptron strategy with multiple peak sets and the worst yield was achieved by using a user-dependent strategy with a single peak set from software tuned by a human operator (Figure 4). This difference can be critical since valuable new biological and pharmacological information is likely to come from the group of proteins with weak MALDI response and/or low abundance.

The particular strengths of the proposed method are its objectivity, automation, and improved sensitivity for weak spectra. The method can be completely automated, it requires no subjective parameter tuning, and it is very successful at identifying proteins. These are all essential properties for a high-throughput protein identification method.

References

[1] D.J.C. Pappin, P. Hojrup and A.J. Bleasby *Curr. Opin. Biol.* **3** (1993) 327

[2] M. Mann, P. Hojrup and P. Roepstorff *Biol. Mass Spectrom.* **22** (1993) 338

[3] P. James, M. Quadroni, E. Carafoli and G. Gonnet *Biochem. Biophys. Res. Comm.* **195** (1993) 58

[4] W.J. Hentzel, T.M. Billeci, J.T. Stults, S.C. Wong, C. Grimley and C. Watanabe *Proc. Natl. Acad. Sci.* **90** (1993) 5011

[5] M.R. Wilkins, E. Gasteiger, C.H. Wheeler, I. Lindskog, J.C. Sanchez, A. Bairoch, R.D. Appel, M.J. Dunn and D.F. Hochstrasser *Electrophoresis* **19** (1998) 3199

[6] D.N. Perkins, D.J.C. Pappin, D.M. Creasy and J.S. Cottrell *Electrophoresis* **20** (1999) 3551

[7] W. Zhang and B.T. Chait *Anal. Chem.* **72** (2000) 2482

[8] K.R Clauser, P. Baker and A.L. Burlingame *Anal. Chem.* **71** (1999) 2871

[9] K.C. Parker *J. Am. Soc. Mass Spectrom.* **13** (2002) 22

[10] V. Egelhofer, J. Gobom, H. Seitz, P. Giavalisco, H. Lehrach and E. Nordhoff *Anal. Chem.* **74** (2002) 1760

[11] P. Berndt, U. Hobohm and H. Langen *Electrophoresis* **20** (1999) 3521

[12] R. Gras, M. Muller, E. Gasteiger, S. Gay, P.A. Binz, W. Bienvenut, C. Hoogland, J.C. Sanchez, A. Bairoch, D.F. Hochstrasser and R.D. Appel *Electrophoresis* **20** (1999) 3535

[13] E.J. Breen, F.G. Hopwood, K.L. Williams and M.R. Wilkins *Electrophoresis* **21** (2000) 2243

[14] H. Field, D. Fenyö and R.C. Beavis *Proteomics* **2** (2002) 36

[15] J.S. MacNeil *Genome Technology* **29** (2003) 31

[16] B. D. James (ed.) *Gel Electrophoresis of Proteins - A Practical Approach, 3rd ed.* Oxford University Press (1998)

[17] T. Berkelman and T. Stenstedt (eds.) *2-D Electrophoresis Using Immobilized pH Gradients - Principles and Methods* Amersham Biosciences Handbook (1998)

[18] F. Levander, T. Rögnvaldsson, J. Samuelsson and P. James *To appear in Proteomics* (2004)

[19] C.M. Bishop *Neural Networks for Pattern Recognition* Oxford Univ. Press, Oxford (1995)

[20] M. Riedmiller and H. Braun *Proc. IEEE Int. Conf. on Neural Networks (ICNN)* (1993) 586

[21] B. D. Ripley *Pattern Recognition and Neural Networks* Cambridge Univ. Press, Cambridge (1996)