Hiding behind Hydrophobicity

TRANSMEMBRANE SEGMENTS IN MASS SPECTROMETRY* [S]

Received for publication, May 26, 2004, and in revised form, September 22, 2004 Published, JBC Papers in Press, September 27, 2004, DOI 10.1074/jbc.M405875200

Lutz A. Eichacker†‡, Bernhard Granvogl§, Oliver Mirus‡, Bernd Christian Müller¶, Christian Miess¶, and Enrico Schleiff¶

From the 1Department für Biologie I, Ludwig-Maximilians Universität München, Menzinger Strasse 67, 80368 München, Germany and ¶Hoffmann-La Roche, Roche Centre for Medical Genomics, CH-4070 Basel, Switzerland

Proteomics of membrane proteins is essential for the understanding of cellular function. However, mass spectrometric analysis of membrane proteomes has been less successful than the proteomic determination of soluble proteins. To elucidate the mystery of transmembrane proteins in mass spectrometry, we present a detailed statistical analysis of experimental data derived from chloroplast membranes. This approach was further accomplished by the analysis of the Arabidopsis thaliana proteome after in silico digestion. We demonstrate that both the length and the hydrophobicity of the proteolytic fragments containing transmembrane segments are major determinants for detection by mass spectrometry. Based on a comparative analysis, we discuss possibilities to overcome the problem and provide possible protocols to shift the hydrophobicity of transmembrane segment-containing peptides to facilitate their detection.

Understanding the protein networks within a living cell is an ultimate goal in current biology. To convert this massive amount of information from gene expression into protein networks, tools were developed in the post-genomic era. Proteomic approaches integrating data from many organisms (see, for example, Refs. 1 and 2) as well as all compartments (see, for example, Refs. 3–8) are an essential tool for the understanding of the intracellular network of protein action. The advantages of such global approaches are also seen for medical applications (see, for example, Refs. 9–12). A further use of this technology is the analysis of post-translational modifications (e.g., Refs. 13 and 14). And last but not least, large-scale analysis of protein complexes further complements the picture of cellular action (Ref. 15 and references therein).

For most proteomic approaches, gel-based two-dimensional systems have been coupled with mass spectrometry, which is one of the most powerful tools for protein identification (16–18). The combination of isoelectric focusing or blue native PAGE as the first dimension followed by SDS-PAGE as the second dimension is a widespread technique for the effective resolution of large numbers of soluble and membrane proteins (5, 19–21). Proteins with a grand average hydropathy (GRAVY) score (22) above 0.5 (23–26) and even membrane protein complexes (27, 28) can be separated by blue native PAGE in contrast to isoelectric focusing. Fast and unequivocal identification of the separated proteins is possible by analyzing internal peptide sequences. In contrast to conventional protein sequencing procedures, mass spectrometry allows the rapid and automated analysis of many proteins in a short time frame. Even though the advantages of this technology are a major breakthrough for the high throughput analysis of proteomic compositions, the technology is also limited to a certain extent. For example, identification of proteins of the same size but different abundance raises difficulties. Furthermore, the procedure of peptide production by protease in gel digestion can reveal major losses, and therefore, vanishing of the signal will be observed. A problem always present is the detectability of certain classes of peptides. One set of these is the class of transmembrane segments. The inability of detection was even used to analyze the topology of certain protein classes (29). It was determined that only 10% of all peptides from transmembrane proteins analyzed represent at least portions of the transmembrane segments (29). Interestingly, some groups recently succeeded in detecting and sequencing transmembrane regions from GalP (30) and bacteriorhodopsin (31, 32) in model studies. In contrast to previous reports, high amounts of protein were used in these experiments, and samples were digested in solution without electrophoretic separation of the proteins. However, highly purified membrane fractions are often not as concentrated as in such model studies. Since major interest has been drawn to the identification of protein networks and protein complexes, which have to be isolated and concentrated beforehand, the proteomic analysis of membranes and membrane complexes suffers from the reduced ability to detect membrane proteins due to the reduced amount of analyzable peptides. Furthermore, not all expected proteins will be detected in one biological sample since the expression of some membrane proteins is specific for certain tissues or developmental stages of the individual.

We have now investigated several physicochemical parameters of peptides containing transmembrane segments by experimental and bioinformatic tools. For this, we have analyzed the thylakoid membrane by mass spectrometry and the Arabidopsis thaliana proteome by bioinformatic analysis. We specifically address the molecular forces intervening with detection of peptides from transmembrane segments by mass spectrometry. Based on the results, we propose that methionine modification

* The abbreviations used are: GRAVY, grand average hydropathy; BisTris, 2-[bis(2-hydroxyethyl)aminol]-2-(hydroxymethyl)propane-1,3-diol; MS, mass spectrometric.
Experimental data for outer envelope proteins were extracted from previous work (33). The topological models were constructed as described (34). The complete A. thaliana proteome was extracted from ftp://ftp.trpis.gsf.de/cress/arabiprot. The sequences were analyzed by TMHMM 2.0 (www.cbs.dtu.dk/services/TMHMM/) (35) to select proteins with proposed transmembrane segments.

Isolation of Protein Complexes Present in Native Membranes and Stromata

Native membrane and stromal protein complexes were prepared directly after isolation of chloroplasts from barley (Hordeum vulgare L. var. Steffi) as described (36). Chloroplasts were lysed at 4 °C in TMK buffer (10 mM Tris-HCl (pH 7.0), and 0.5 mM EDTA. Complexes were solubilized by incubation 25 mM Tris at 50 °C. After electrophoresis, the resolved proteins were visualized by staining with Coomassie Brilliant Blue R-250 (Serva, Heidelberg, Germany), and protein spots of interest were excised with a clean scalpel and stored in Eppendorf tubes at -80 °C.

Two-dimensional Gel Electrophoresis

First Dimension—Sample solutions of membrane proteins were applied to 10–15% blue native/PAGE systems (21).

Second Dimension—Individual lanes of the blue native polyacrylamide gel were cut and incubated for 15 min at room temperature in 2% (v/v) SDS, 66 mM sodium carbonate, and 2% (v/v) β-mercaptoethanol at 50 °C (37). Lanes were positioned on top of the second dimension gel (12.5%, containing 4 µm urea) and overlaid with a low melting agarose solution containing 0.5% (w/v) agarose, 192 µg/ml glycerine, 0.1% SDS, and 25 mM Tris at 50 °C. After electrophoresis, the resolved proteins were visualized by staining with Coomassie Brilliant Blue R-250 (Serva, Heidelberg, Germany), and protein spots of interest were excised with a clean scalpel and stored in Eppendorf tubes at ~20 °C. Standard protein solution (Mark 12, broad-range protein standard, Novex) was provided on a piece of filter paper next to the first dimension lane.

Mass Spectrometry

In gel digestion of proteins before subjection to mass spectrometry was carried out using the method described previously (38). Mass spectrometric analyses were carried out on a Micromass Q-TOF-I hybrid mass spectrometer equipped with an orthogonal nano-electrospray ionization source operating in the positive electrospray ionization mode. Samples were loaded into medium nano-electrospray capillaries (Pro-tana, Odense, Denmark) and positioned in the source. For high external calibration, a standard solution of 0.2% (v/v) formic acid (HFO) in 50% methanol was used. The peak at m/z 421.7589, which resulted from self-digestion of trypsin, was used for internal calibration of mass spectrometric (MS) spectra and for verification of successful digestion.

To obtain optimized spectral data, a capillary voltage of 750–900 V and a cone voltage of 45 V were used. For scanning peptide MS spectra, the quadrupole was set to the appropriate mass range of m/z 350–1600, where peptides were expected. After data collection, 25 scans were averaged at a rate of 2.0 s/scan. Instrument operation, data acquisition, and analysis were performed using MassLynx/Biolynx 3.5 and 4.0 software. For peptide detection coverage analysis, the following thylakoid proteins were analyzed in detail: Ps a K, Ps a G, Ps a H, Ps a L, Lh b C, Ps b, Ps b A, Ps b B, Ps b C, Ps b D, Ps b E, and Ps b F.

Calculation of Hydrophobicity

For calculation of the hydrophobicity, several scales were used, viz. the hydrophobicity scale of Kyte and Doolittle (22), the scale defined by Eisenberg et al. (39), and the scale of Black and Mould (40). To calculate the mean hydrophobicity of a pool of peptides or for a protein, all hydrophobicity values of the pool or of the protein were accumulated and finally normalized by the amount of amino acids in the pool or protein. For calculation of the central region (see Figs. 5 and 6) for comparison with the calculation performed by Eisenberg et al. (39), the hydrophobicity was calculated based on the Eisenberg scale in a window of 11 amino acids; values were normalized; and the highest value was assigned to the peptide. For the region bearing the highest hydrophobicity, the hydrophobic moment was calculated as follows (Equation 1),

\[
\mu_H = \left( \frac{\sum_{n=1}^{N} H \sin(\delta)}{N} + \frac{\sum_{n=1}^{N} H \cos(\delta)}{N} \right)^{1/2}
\]

where \(\mu_H\) is the hydrophobic moment, \(H\) is the hydrophobicity of the amino acid number \(n\), \(N\) is the window size analyzed (here, 11 amino acids), and \(\delta\) is the angle of the turn between two amino acids. For classical helices as discussed here, the angle was set to 100.

Calculation of the β-Barrel Score

The β-barrel score was calculated as described previously (34, 41). For calculation of the exact β-barrel score of the peptide pool selected, the bilayer score (41) was used. The score for each peptide was calculated for the case that the first amino acid would face the bilayer and for the case that the first residue would face the interior. The larger value was added to the exact β-score of the pool. Finally, the derived sum was divided by the number of amino acids in the pool.

Calculation of Fragment Coverage and Probability Values

For analysis of the digestion frequency, the sequences were split C-termically after arginine and lysine, and the corresponding fragments were sorted according to their amino acid length. To analyze the detection probability as shown in Fig. 7C, all peptides analyzed in this work were sorted according to their hydrophobicity or their amino acid length, and the fraction of detected peptides for each parameter in a certain window was calculated. To determine the probability for detection by the two parameters hydrophobicity and length, the probability for detection at a given peptide size \((P)\) and at a given hydrophobicity \((P)\) was used to calculate the detection probability by Equation 2.

\[
P = P_1 P_2 (P_1 + P_2 - 0.5)
\]

RESULTS AND DISCUSSION

Two-dimensional Protein Separation of Thylakoid Membranes—Transmembrane proteins from thylakoid membrane complexes were separated and identified in a two-dimensional blue native/SDS-PAGE system (21), which was optimized for separation of hydrophobic reaction center proteins from photosystems I and II (37). In agreement with Devree et al. (26), we confirmed that separation and MS sequencing of membrane proteins with GRAVY scores above 0.5 are possible after separation of protein complexes by two-dimensional native SDS-PAGE. Highly hydrophobic proteins belonging to this group are, for example, cytochrome b f gene products PetB (GRAVY score of 0.54) and PetD (GRAVY score of 0.56), reaction center II core protein PsbF (GRAVY score of 0.71), and the CF 0 -ATPase subcomplex unit AtpH (GRAVY score of 1.03). This indicates that the two-dimensional native/SDS-PAGE system is particularly well suited for separation of hydrophobic membrane proteins even of low molecular mass.

While analyzing the membrane proteome of chloroplasts, we recognized that several expected membrane proteins were not detectable, although they were described to be associated with thylakoid membrane complexes. Although low molecular mass proteins such as PsbF (4 kDa) were identified, proteins of photosystems I and II <10 kDa are often hardly detected by analysis of internal peptides. Because membrane proteins may be detected by means of their hydrophilic loop regions and because “short” peptides are detected more frequently than “long” peptides, extreme hydrophobicity and the lack of internal tryptic cleavage sites may have prevented these proteins from being detected.

Transmembrane Segments from Thylakoid Membrane Proteins Escape Detection by Electrospray Ionization MS—A detailed analysis of the thylakoid proteome revealed that especially membrane-inserted fragments could not be detected by
mass spectrometry. To gain an understanding of this phenomenon, we analyzed the observed fragments of helical membrane proteins of the thylakoid membrane and β-barrel membrane proteins of the outer envelope of chloroplasts (33, 34).

Tryptic digestion of proteins leads to peptides of different lengths corresponding to the distribution of tryptic cleavage sites (arginine and lysine). A number of these peptides could be detected by electrospray ionization MS (MS spectra). The quality of a spectrum depends on the amount of peptide extracted from the gel, which could be partially influenced by increasing the time of measurement. However, a relevant part of peptides of thylakoid proteins predicted from in silico analysis was not detected when spectra from tryptic digests were analyzed. In our MS spectra, peptides usually emerged between m/z ~300 and 1600. Therefore, large peptides >1.4 kDa have to be multiply charged to fit in this m/z range. However, peptides found by electrospray ionization MS normally carry charges from +2 to +4 anyway. As described previously (30, 42), peptides >4 kDa were not detected and sequenced. To decide whether a peptide signal was present or not, peak intensity had to be >50 counts/25 averaged spectra.

Doubly charged peptides <0.6 kDa (m/z 300, approximately five amino acids) were never detected in our MS spectra. A possible explanation could be charge repulsion within these peptides. For peptides longer than 15 amino acids, detectability was decreased with increased amino acid number and hydrophobicity. However, especially these long and hydrophobic peptides were of interest because tryptic digestions of α-helical membrane proteins yields peptides containing at least 20 amino acids.

About 35% of all peptides should contain helical transmembrane segments (Fig. 1A, bar 3). However, when the theoretical statistical occurrence of helical transmembrane segments (Fig. 1A, bar 3) was compared with the observed occurrence of such peptides of ~4% (bar 4), it became obvious that the detection of helical transmembrane regions is a problem. In contrast, peptides with transmembrane segments from β-barrel proteins were detected in the proposed statistical frequency (Fig. 1A, bars 1 and 2). We can therefore conclude that not the membrane insertion per se results in the loss of detection of such fragments.

We subsequently analyzed the proteolytic fragments of the thylakoid proteins in more detail. All peptides were divided in two pools with regard to length. The cutoff was defined by 20 amino acids, a typical length of a transmembrane domain. We first analyzed all peptides shorter than 21 amino acids (Fig. 1B, bars 1–3). We observed an experimental coverage of ~50% of all peptides when no dissection regarding the localization of the peptide was performed (Fig. 1B, bar 1). In contrast, when only peptides comprising soluble regions of the proteins were analyzed, 60% of all peptides predicted were experimentally observed as well (Fig. 1B, bar 2). For peptides containing membrane regions, the overall detection rate was ~35% (Fig. 1B, bar 3) and therefore significantly smaller than the detection rate for soluble peptides. We next analyzed all peptides longer than 20 amino acids. Interestingly, the overall detection rate for such peptides was decreased to ~15% (Fig. 1B, bar 4), and the soluble phase peptides (bar 5) were significantly reduced to ~20% as well. However, almost no peptides of membrane proteins were detected (Fig. 1B, bar 6).

We then wanted to understand why certain peptides of transmembrane proteins were detectable and others not. One physical parameter defining a transmembrane region is hydrophobicity. Therefore, the average hydrophobicity of the pool of detected (Fig. 1C, bars 2 and 4) and non-detected (bars 1 and 3) peptides was compared. Interestingly, the pool of detected peptides showed an average hydrophobicity of 0.2 independent of the cutoff length, whereas the range of the hydrophobicity determined for the detected peptides covered a range from −2 to 0.5, with two exceptions where the peptides revealed hydrophobicity values of 2.0 and 1.1. In contrast, the non-detected peptides revealed a high hydrophobicity (Fig. 1C, bars 1 and 3). Furthermore, peptides not detected and longer than 21 amino acids exceeded the hydrophobicity index of shorter peptides of the same pool by a factor of ~4. These two observations point to the interpretation that the hydrophobicity and the length of peptides are critical determinants for detection by experimental means.

Comparison of the Thylakoid and the A. thaliana Proteome—From the analysis of the experimental data obtained from sequencing of the thylakoid membrane, it was determined that longer and more hydrophobic peptides are not as frequently detected compared with shorter and more hydrophilic peptides (Fig. 1B). We therefore addressed the specific role of peptide length in the increase in the hydrophobicity index. To answer this question, the proteome of the thylakoids was digested in silico at trypsin-specific cleavage sites, and the hydrophobicity and the number of fragments were calculated for each pool of
peptides of a defined length (Fig. 2). Strikingly, the pools of peptides shorter than 20 amino acids showed a hydrophilic behavior in general, whereas the pools of longer peptides revealed a general hydrophobic character. To exclude that this observation is limited to the proteins of the thylakoid membrane, we analyzed the entire A. thaliana proteome. As expected, the number of peptides in a certain pool size decreased exponentially with the length of the peptides (Fig. 3A). Next, we analyzed the β-barrel probability and the hydrophobicity of each peptide pool (Fig. 3, B and C). The probability for a membrane-inserted β-sheet was found to be very low. Only the pools with a peptide size above 50 amino acids in length revealed a score above zero. For comparison, a cutoff score of 0.25 was chosen for selection as a membrane-inserted β-sheet (34, 41). However, it is worthwhile to mention that the exact β-strand score of the pools increased almost linearly with the peptide size of the pools (Fig. 3B). This might suggest that the portion of peptides derived from membrane-inserted β-barrels increases in the pools of peptide of increasing size. As seen before for the thylakoid proteins, the average hydrophobicity of the A. thaliana proteome fragments increased almost linearly in the peptide size range of up to 30 amino acids (Fig. 3C). The pools of peptides longer than 20 amino acids showed a mean hydrophobic character, whereas the pools of peptides shorter than 21 amino acids revealed an overall hydrophilic character. Therefore, the observation made for the thylakoid proteins can be generalized for the whole proteome. Peptides longer than 20 amino acids are more hydrophobic in general, indicating that hydrophobicity is a major determinant for detection by mass spectrometry.

From the results observed, one question has to be raised. If longer peptides are more hydrophobic (Figs. 2 and 3C) and if more hydrophobic segments are not as frequently detected (Fig. 1C), do transmembrane proteins have a higher portion of long peptides compared with soluble proteins when trypsinized? To answer this question, we analyzed the peptide frequency of four protein pools, viz. the complete A. thaliana proteome, proteins with a β-barrel score above 1, proteins with a mean hydrophobicity above 0.25 (according to the Kyte and Doolittle scale (22)), and all proteins with at least one transmembrane region identified or predicted by TMHMM (Fig. 4). The first pool was chosen as a control. The second pool contains proteins with a putative membrane β-barrel fold (34, 41). This pool was analyzed to compare the transmembrane β-barrel and α-helical proteins (fourth pool) since analysis of the experimental data revealed a drastic difference in the coverage rate (Fig. 1A). The third pool was chosen since a mean hydrophobicity of 0.25 and larger accounts for very hydrophobic proteins for the following reasons.

Eisenberg et al. (39) analyzed a large pool of membrane proteins with regard to their hydrophobicity using the Eisenberg scale. They observed that all helices had an average hydrophobicity above 0.29 on their scale. Analyzing the hydrophobicity values of the Eisenberg scale (39) and the Kyte and Doolittle scale (22) revealed an almost linear relation between the two scales (Supplemental Fig. 1A). Taking the linear relation into account, a value of 0.29 on the Eisenberg scale would correspond to 0.5 on the Kyte and Doolittle scale. Also taking into account that membrane proteins expose soluble loops, the overall hydrophobicity of a membrane protein should be lower than this value. We therefore chose a cutoff value of 0.25 according to the Kyte and Doolittle scale since only six amino acids of the standard 20 amino acids account for a hydrophobicity value above 0.25 (Supplemental Fig. 1A). The amino acid sequences of the selected proteins should be dominated by such amino acids, which are known to be present in transmembrane regions (43). Only 6% of all protein sequences were selected by the defined cutoff value (Supplemental Fig. 1B). In contrast, 17.8% of all proteins of the A. thaliana proteome contain at least one single transmembrane domain as determined by TMHMM analysis of all A. thaliana sequences. These sequences were selected in the fourth pool to directly analyze all sequences with putative transmembrane regions.

We then analyzed the peptide frequency after in silico digestion of the four pools of sequences. The distribution of peptides derived from putative β-barrel proteins did not significantly differ compared with complete proteome (Fig. 4, A and B). In contrast, for the third and fourth pools, we observed a significant enrichment of peptides of 20–35 amino acids (Fig. 4, A and B). Such a result might have been expected for the pool of hydrophobic sequences since arginine and lysine (representing the cleavage sites for digestion by trypsin) are the most hydrophilic amino acids and are possibly less abundant in proteins with a high average hydrophobicity. Strikingly, however, is the result with proteins containing putative transmembrane regions. 50% of all protein sequences in the pool have only a single transmembrane domain, and only 35% have more than two predicted transmembrane regions as determined by TMHMM (data not shown). The average hydrophobicity of the complete protein pool containing putative transmembrane domains was below zero (according to the Kyte and Doolittle scale (22)), suggesting the presence of large hydrophilic domains. In addition, the arginine and lysine frequency in thylakoid membrane proteins was not significantly lower compared with the
complete A. thaliana protein pool (data not shown). Still, a similar cluster of longer peptides as for the hydrophobic proteins was observed. This suggests that those transmembrane domains are usually covered by long peptides. Furthermore, the maximum peptide length was found to be in the range of 28 amino acids, but the peak expanded from 20 to 40 amino acids. Therefore, the length of the transmembrane segments derived by tryptic digestion could be one reason for the low abundance in the MS experiment as discussed above (Fig. 1B). This result can be further extended to other proteomes. While analyzing all sequences from either Escherichia coli (Fig. 4C) or Homo sapiens (Fig. 4D), we observed a similar profile as for A. thaliana (Fig. 4B). We are therefore convinced that the observations for A. thaliana can be generalized.

How Can the Detection Problem Be Solved?—We have shown that two phenomena account for the lower detectability of transmembrane segments. An increased peptide size reduced the probability for detection by experimental means (Fig. 1B), and the frequency of enlarged peptides was increased in proteins with putative membrane localization (Fig. 4, A and B). However, not only the length but also the hydrophobicity of the segment (Fig. 1C) accounted for the low detectability of the transmembrane regions since the longer peptides from soluble proteins remained detectable by mass spectrometry to a higher extent (Fig. 1B). In general, the increase in the peptide length was shown to correlate with an increase in the mean hydrophobicity of the peptides (Figs. 2, 3C, and 6). Therefore, we concluded that a reduction of the hydrophobicity and/or length of the transmembrane segment-containing peptides would allow their detection.

To analyze and prove the last idea, we went back to the experimentally derived data set. We analyzed the length of the peptides as well as their mean hydrophobicity and hydrophobic moment. To do so, we used the Eisenberg scale to compare the features of the peptides with the previously reported characteristics of transmembrane segments (39). As discussed above, most peptides not detected were of high hydrophobicity (Figs. 1C and 5C). The maximum of the gaussian distribution of the hydrophobicity for peptides not detected was 0.59 (Fig. 5C, dashed line), whereas the center of the hydrophobicity of the detected peptides was 0.15 (solid line). This clearly supports the idea that strong hydrophobicity contradicts detectability. It is worthwhile mentioning that a small cluster of highly hydrophilic peptides were not detected as well. (15% of all nondetected peptides had a peak of hydrophobicity of about –0.33.) Analyzing the peptide length supported the observation outlined in Fig. 1B since mainly short sequences were detected, with a size distribution peaking at 10 amino acids (Fig. 5B). For peptides not detected by MS analysis, two peaks were obtained. Whereas the first peak (at approximately seven amino acids) covered only 10% of the complete area, the main peak had a wide distribution with a center at ~30 amino acids, confirming
the above-mentioned idea that length is a major determinant for detectability.

The physical features of the analyzed peptides can now be compared with the prediction rules for transmembrane segments as discussed previously (39). About 55% of the non-detected peptides fall into the region typical for transmembrane helices (Fig. 5A) as defined by Eisenberg et al. (39), and only 20% of all peptides are clearly assigned as soluble (Fig. 5A, to the left of the dashed line). The gaussian distribution of the hydrophobic moment of the non-detected peptides had a maximum at 0.03 (Fig. 5A, dashed line), meaning that hydrophobicity is equally distributed throughout the sequence, as the hydrophobic moment (Equation 1) is near zero. From this, it becomes obvious that a combined reduction of the hydrophobicity and an increase in the hydrophobic moment would change the physical properties of a peptide, where detection becomes more likely. But how can such a shift of the physical features be achieved?

One solution to this problem could be an oxidation of the hydrophobic transmembrane peptides. As Black and Mould (40) observed, oxidation of the amino acid methionine results in a decrease in hydrophobicity. When compared with the Eisenberg scale (39), a single oxidation of methionine would reduce the hydrophobicity from 0.64 to −0.76 (Fig. 6A). A double oxidation would decrease the hydrophobicity of the peptide to −2.0 (Fig. 6A). Such a drastic shift of one single amino acid may contribute to the physical properties of the hydrophobic transmembrane peptide in a way that the peptide becomes detectable.

To investigate this idea, the peptides were oxidized before MS analysis. This procedure revealed the detection of two previously unidentified sequences (Fig. 6, B and C). The hydrophobicity and hydrophobic moment of the two peptides were clearly shifted by oxidation (Fig. 6B, open circles). Furthermore, the oxidation placed the hydrophobicity value in a range where 80% of all analyzed peptides were detectable disregarding their length (Fig. 6B, dashed line). In contrast, before oxidation, the hydrophobicity of the two peptides was in a range where, under our experimental conditions, only 15–20% of all peptides could be found by MS analysis.

A second solution would be a reduction of the size of the fragments. One possibility is the subsequent cyanogen bromide cleavage of the peptides. Since both methods would target methionines in transmembrane regions, we analyzed the methionine content in different proteomes (Fig. 7A). We observed that the overall methionine content in eukaryotes was reduced in comparison with the analyzed prokaryotic proteome (Fig. 7A, compare bars 1, 4, and 7). However, when only proteins with putative transmembrane regions were analyzed, the methionine content was increased in sequences from E. coli and H. sapiens (Fig. 7A, bars 2 and 5), but not from the plant proteome (bar 8), compared with the corresponding content of the complete proteome. Analysis of the transmembrane segments then showed a significant increase in the methionine distribution in all three cases (Fig. 7A, bars 3, 7, and 9). Therefore, methionine is an obligate target to further optimize transmembrane peptide analysis. Supporting this notion, almost 60% of all predicted transmembrane segments in E. coli and almost 50% in the eukaryotic proteomes contain a methionine (Fig. 7B). We conclude that targeting methionine by chemical modification or cleavage should increase the detectability of transmembrane segments by mass spectrometry.

Indeed, when the transmembrane peptides not detected by mass spectrometry in our approach (Fig. 7C, shaded triangles) were subsequently cleaved in silico by cyanogen bromide (open circles), we determined that most of the peptides fall into the region with respect to amino acid length and hydrophobicity of the proteolytic fragment where detection of peptides becomes most likely according to our experience (0, 25, 50, and 75% detection probability boundaries) (Fig. 7C). In a subsequent exper-
FIG. 6. Hydrophobicity loss of oxidized methionine allows transmembrane peptide detection. A, correlation between the hydrophobicity \((H)\) scales of Eisenberg et al. (39) and Black and Mould (40). The positions of 20 amino acids (\(\text{aa}\)) with methionine (open circle) highlighted are shown. The proposed positions of the singly (Msx) and doubly (Mso) oxidized methionines (shaded squares) according to the linear relation are given. B, see the legend to Fig. 5A for details. The hydrophobicity and hydrophobic moment shift of two oxidized peptides are indicated by open circles connected by a solid line. The dashed line gives the percentage of non-detected peptides in relation to the hydrophobicity (right scale). M, methionine; Ms, monooxidized methionine; Mso, di-oxidized methionine. C, Coomassie Blue-stained spot of CB23 (LhbC) from barley was cut out of a second dimension SDS gel and digested with trypsin. A positive mode electrospray (Es+) ionization Q-TOF mass spectrum was recorded, which is enlarged here between \(m/z\) 995 and 1020. 25 single spectra were averaged for this experiment. The tryptic peptide appearing at \(m/z\) 1013.52 contains two mono-oxidized methionines. The mass without modification would be \(m/z\) 997.52. The peptide sequence is given at the top.

FIG. 7. Rule of methionine oxidation is applicable to all proteomes. A, the content of methionines as a percentage of the total amino acid (\(\text{aa}\)) content was analyzed for the E. coli proteome (bars 1–3), the H. sapiens proteome (bars 4–6), and the A. thaliana proteome (bars 7–9). For comparison, the methionine content of all sequences (all; bars 1, 4, and 7), of all sequences with putative transmembrane regions as predicted by TMHMM 2.0 (TM-Prot; bars 2, 5, and 8), and the predicted transmembrane region itself (TM; bars 3, 6, and 9) was analyzed. B, the number of predicted transmembrane regions containing at least a single methionine was compared with the number of transmembrane regions predicted. Shown is the percentage of transmembrane regions containing a methionine for E. coli (bar 1), H. sapiens (bar 2), and A. thaliana (bar 3). C, a pool of non-detected tryptic peptides (shaded triangles) was digested in silico by an algorithm mimicking cyanogen bromide cleavage (open circles). The probability borders for detectability of peptides were calculated as described in the legend to Fig. 5. Experimentally determined peptide fragments are highlighted (open squares). The probability borders for detectability of peptides were calculated as described under “Experimental Procedures,” and probability borders are shown. The numbers are the minimum percentage of detectability of the area surrounded. \((\text{HE11})_{\text{aa}}\), average hydrophobicity/amino acid using an 11-amino acid window.
short peptides containing transmembrane segments still soluble regions, the decrease in the peptide size cannot be problematic since those regions are present in long and tides containing membrane segments is an intrinsic experimental approach, we were able to detect 14 new materials. However, we have clearly demonstrated that transmembrane segment-containing fragments escape detection by mass spectrometry because of their size and their hydrophobicity (Fig. 5). In addition, longer fragments were found to be more hydrophobic than shorter fragments (Fig. 3); and in turn, (Fig. 5). In addition, longer fragments were found to be more hydrophobic than shorter fragments (Fig. 3); and in turn, transmembrane region-containing proteins revealed a higher yield of fragments longer than 20 amino acids compared with transmembrane region-containing fragments escape detection by mass spectrometry because of their size.

**Conclusion**—It is well known that membrane proteins are detected by means of their soluble regions and that short peptides are detected more frequently than long peptides (Fig. 1B) (44). However, we have clearly demonstrated that transmembrane segment-containing fragments escape detection by mass spectrometry because of their size and their hydrophobicity (Fig. 5). In addition, longer fragments were found to be more hydrophobic than shorter fragments (Fig. 3); and in turn, transmembrane region-containing proteins revealed a higher yield of fragments longer than 20 amino acids compared with soluble proteins (Fig. 4). Therefore, the low abundance of peptides containing membrane segments is an intrinsic experimental problem since those regions are present in long and hydrophobic peptides. In contrast to peptides representing soluble regions, the decrease in the peptide size cannot be proposed as a tool for an increase in peptide coverage since many short peptides containing transmembrane segments still escape detection (Fig. 5B). However, alteration of the hydrophobicity would clearly increase the probability of detection. One way to modify the hydrophobicity is the oxidation of such peptides (Fig. 6A) since methionines are highly enriched in transmembrane segments (Fig. 7, A and B). As demonstrated, such oxidation revealed the detection of two peptides not detected before (Fig. 6, B and C) since the hydrophobicity value was drastically altered. Furthermore, analysis of the second method of targeting methionines, viz. subsequent cyanoargent bromide digestion, revealed that the shorter and sometimes less hydrophobic peptides became detectable (Fig. 7C). We therefore propose that the development of techniques quantitatively targeting the methionines of peptides will reveal the detectability of transmembrane segment-containing fragments by mass spectrometry.

**REFERENCES**

1. Canovas, F. M., Dumas-Gaudot, E., Recorbet, G., Jorrin, J., Mock, H. P., and Rossignol, M. (2004) Proteomics 4, 285–298
2. Cash, P. (2003) Adv. Biochem. Eng. Biotechnol. 83, 93–115
3. Lister, R., Chew, O., Lee, M. N., Heazlewood, J. L., Clifton, R., Parker, K. L., Millar, A. H., and Whelan, J. (2004) Plant Physiol. 134, 777–789
4. Sickmann, A., Sickmann, A., Reinders, J., Wagner, Y., Joppich, C., Zahedi, R., Meyer, H. E., Schonfisch, B., Perschil, I., Chacinska, A., Giard, B., Rehling, P., Pfanner, N., and Meisinger, C. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 13207–13212
5. Ulmschneider, M. B., and Sansom, M. S. (2001) EMBO Rep. 3, 557–562
6. Eckart, K., Eichacker, L., Schleiff, E., Heins, L., and Soll, J. (2002) EMBO J. 21, 6136–6145
7. Embley, E., Langen, H., and Meyer, H. E. (2001) Protein Sci. 10, 595–606
8. Devree, E., Vanrobaey, S., Smet, J., Van Beeumen, J., and Van Coster, R. (2002) Electrophoresis 23, 2525–2533
9. Anderson, N. L., Matheson, A. D., and Steiner, S. (2000) Curr. Opin. Biotechnol. 11, 408–412
10. Aebersold, R., and Goodlett, D. R. (2001) Chem. Rev. 101, 269–295
11. Cash, P. (2003) Electrophoresis 24, 242–247
12. Venter, H., Ashcroft, A. E., Keen, J. N., Henderson, P. J., and Herbert, R. B. (2002) Biochem. J. 363, 243–252
13. Hixson, K. K., Rodriguez, N., Camp, D. G., III, Strittmatter, E. F., Lipton, M. S., and Smith, R. D. (2002) Electrophoresis 23, 3224–3232
14. Eichacker, L., and Schleiff, E., Heins, L., and Soll, J. (2002) EMBO Rep. 3, 557–562
15. Schleiff, E., Eichacker, L. A., Eckart, K., Becker, T., Mirus, O., Stahl, T., and Soll, J. (2003) Protein Sci. 12, 748–759
16. Rowley, A., Choudhary, J. S., Marzioch, M., Ward, M. A., Weir, M., Solari, R. C., and Blackstock, W. P. (2000) Methods 20, 383–397
17. Anderson, N. L., Matheson, A. D., and Steiner, S. (2000) Curr. Opin. Biotechnol. 11, 408–412
18. Aebersold, R., and Goodlett, D. R. (2001) Chem. Rev. 101, 269–295
19. Kashino, Y. (2003) J. Chromatogr. B 797, 191–216
20. van Wijk, K. J. (2001) Plant Physiol. 126, 501–508
21. Schagger, H., Cramer, W. A., and von Jagow, G. (1994) Anal. Biochem. 217, 220–230
22. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132
23. Santoni, V., Rabilloud, T., Deumus, P., Rouquier, D., Mansson, M., Kieffer, S., Garin, J., and Rossignol, M. (1999) Electrophoresis 20, 705–711
24. Santoni, V., Kieffer, S., Desclaux, D., Masson, F., and Rabilloud, T. (2000) Electrophoresis 21, 3329–3344
25. Hippler, M., Klein, J., Pink, A., Allinger, T., and Hoerth, P. (2001) Plant J. 28, 595–606
26. Devree, E., Vanrobaeys, S., Smet, J., Van Beeumen, J., and Van Coster, R. (2002) Electrophoresis 23, 2525–2533
27. Bucher, M., Decker, S., Hoermann, F., Soll, J., and Heins, L. (2002) EMBO J. 21, 6136–6145
28. Thylholm, E., Lindstrom, V., Tissie, C., Robinson, C., Schroeder, W. P., and Funk, C. (2002) FEBS Lett. 513, 217–222
29. Washburn, M. P., Wolters, D., and Yates, J. R., III (2001) Nat. Biotechnol. 19, 242–247
30. Venter, H., Ashcroft, A. E., Keen, J. N., Henderson, P. J., and Herbert, R. B. (2002) Biochem. J. 363, 243–252
31. Hixson, K. K., Rodriguez, N., Camp, D. G., III, Strittmatter, E. F., Lipton, M. S., and Smith, R. D. (2002) Electrophoresis 23, 3224–3232
32. Blonder, J., Conrads, T. P., Yu, L. R., Terunuma, A., Janini, G. M., Issaq, H. J., Vogel, J. C., and Veenastra, T. D. (2004) Proteomics 4, 31–45
33. Eckart, K., Eichacker, L., Schleiff, E., Heins, L., and Soll, J. (2002) EMBO Rep. 3, 557–562
34. Schleiff, E., Eichacker, L. A., Eckart, K., Becker, T., Mirus, O., Stahl, T., and Soll, J. (2003) Protein Sci. 12, 748–759
35. Krugh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) J. Mol. Biol. 309, 567–580
36. Eichacker, L. A., Hehrich, M., Rudiger, W., and Muller, B. (1996) J. Biol. Chem. 271, 32174–32179
37. Eichacker, L., and Muller, B. (April 21, 2001) German Patent 1001-19-597 DE
38. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Anal. Chem. 68, 850–858
39. Eisenberg, D., Schwarz, E., Komaromy, M., and Wall, R. (1984) J. Mol. Biol. 179, 125–142
40. Black, S. D., and Mould, D. R. (1991) Anal. Biochem. 193, 72–82
41. Wilm, M., and Mann, M. (1996) Anal. Chem. 68, 1–8
42. Wilm, M., and Mann, M. (1996) Anal. Chem. 68, 1–8
43. Ooms, M. B., and Somsen, M. S. (2001) Biochim. Biophys. Acta 1512, 1–14
44. Wu, C. C., and Yates, J. R., III (2003) Nat. Biotechnol. 21, 262–267