Identification of membrane proteins by tandem mass spectrometry of protein ions

Joe Carroll, Matthew C. Altman, Ian M. Fearnley, and John E. Walker*

Medical Research Council Dunn Human Nutrition Unit, Hills Road, Cambridge CB2 0XY, United Kingdom

Contributed by John E. Walker, July 20, 2007 (sent for review June 7, 2007)

The most common way of identifying proteins in proteomic analyses is to use short segments of sequence (''tags'') determined by mass spectrometric analysis of proteolytic fragments. The approach is effective with globular proteins and with membrane proteins with significant polar segments between membrane-spanning α-helices, but it is ineffective with other hydrophobic proteins where protease cleavage sites are either infrequent or absent. By developing methods to purify hydrophobic proteins in organic solvents and by fragmenting ions of these proteins by collision induced dissociation with argon, we have shown that partial sequences of many membrane proteins can be deduced easily by manual inspection. The spectra from small proteolipids (1–4 transmembrane α-helices) are dominated usually by fragment ions arising from internal amide cleavages, from which internal sequences can be obtained, whereas the spectra from larger membrane proteins (5–18 transmembrane α-helices) often contain fragment ions from N- and/or C-terminal parts yielding sequences in those regions. With these techniques, we have, for example, identified an abundant protein of unknown function from inner membranes of mitochondria that to our knowledge has escaped detection in proteomic studies, and we have produced sequences from 10 of 13 proteins encoded in mitochondrial DNA. They include the ND6 subunit of complex I, the last of its 45 subunits to be analyzed. The procedures have the potential to be developed further, for example by using newly introduced methods for protein ion dissociation to induce fragmentation of internal regions of large membrane proteins, which may remain partially folded in the gas phase.

mitochondria | complex I | ND6 subunit | membrane proteome | proteolipids

One-third of proteins encoded in genomes are hydrophobic membrane proteins, and their analysis by mass spectrometric methods is more difficult than the analysis of hydrophilic proteins. Many of the difficulties arise from their hydrophobicity and from the associated lack of procedures for purifying membrane proteins in a suitable form for mass spectrometric analysis. Usually, detergents are used to extract and purify membrane proteins, but they are incompatible with the protein ionization methods used in MS, and so they have to be removed from the purified protein (often by chromatography in organic solvents) before analysis can be undertaken (1–3). An alternative approach is to extract and fractionate the membrane proteins directly in organic solvents (3–5).

The measurement of the mass of a protein allows the presence but not the location of any posttranslational modifications to be detected, but it does not provide a reliable means of identifying a protein in sequence data-bases, as required in proteomic analyses. Usually, short segments of protein sequence (sequence tags), obtained by mass spectrometric analysis of proteolytic peptides, are used for this purpose (6). This method provides a highly effective means of generating partial sequences from globular proteins, and from membrane spanning proteins with significant polar segments between the helices and at their termini, but it is often ineffective with the more hydrophobic membrane proteins, where protease cleavage sites can be either rare or be completely absent. Another approach known as ''top down'' sequencing, that has been applied to globular proteins, is to fragment intact protein ions by collision induced dissociation (CID) with argon, and to deduce partial sequences from the fragment ion spectra (7–9). These spectra are complex, and often their interpretation requires the use of high resolution instruments, such as ion cyclotron resonance and linear ion trap mass spectrometers with Fourier transform capabilities, to separate the fragment ion isotopes and to determine the number of associated charges from the isotope spacings (9, 10). The expense of this instrumentation limits access to these techniques.

The top-down sequencing approach has been applied also to a few small membrane proteins in the molecular mass range 3–8 kDa, namely subunits VIIc and VIII of bovine mitochondrial cytochrome oxidase (4), subunit c of ATP synthase (11–13), and subunits of cytochrome b_6f from chloroplasts and cyanobacteria (3). The tandem mass spectra were produced by fragmentation of protein ions by CID in instruments with quadrupole or time-of-flight analyzers. They were rather simple and could be interpreted manually. However, because of the limited access to a range of membrane proteins purified in a form that is compatible with the generation of intact ion spectra, until now, no systematic study of tandem mass analysis of membrane proteins has been carried out. As described here, we have improved the procedures for solubilizing and fractionating hydrophobic membrane proteins by chromatography. We have extended the measurement of the molecular masses of membrane proteins by fragmenting the intact protein ions by CID. We have shown that, in contrast to the complex fragment ion spectra that arise from CID of globular proteins (7), the fragment ion spectra of many membrane proteins, containing from 1 to 18 transmembrane α-helices (TMHs), are relatively simple and are readily interpretable by manual inspection. Thus, they provide suitable tags for identifying unknown membrane proteins.

Results and Discussion

Extraction, Fractionation, and Characterization of Membrane Proteins. Membrane proteins were extracted from bovine heart mitochondria with solvents based on 2-propanol:acetonitrile and chloroform:methanol at neutral and acidic pH values. The latter solvent was also used to extract proteins from samples of purified complexes I and III and cytochrome oxidase. Significant amounts of subunit ND6 were found only in chloroform:methanol extracts at acidic pH. The extracts of mitochondria were fractionated by gel filtration and hydrophilic interaction chro-

Author contributions: J.E.W. designed research; J.C., M.C.A., and I.M.F. performed research; J.C., M.C.A., I.M.F., and J.E.W. analyzed data; and J.E.W. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Abbreviations: CID, collision induced dissociation; Q-TOF, quadrupole-time of flight; TMH, transmembrane α-helix.

*To whom correspondence should be addressed. E-mail: walker@mrc-dunn.cam.ac.uk.

This article contains supporting information online at www.pnas.org/cgi/content/full/0706817104/DC1.

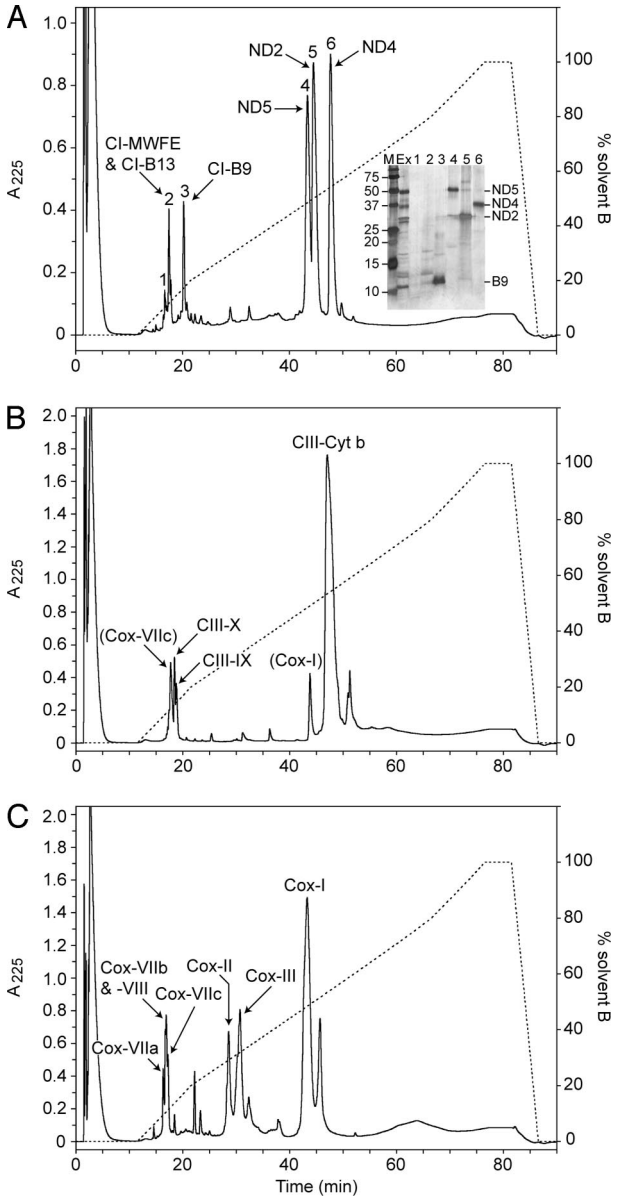

Fig. 1. Reverse-phase chromatography of proteolipids extracted from electron transfer complexes from bovine mitochondria. Shown are extracts of complexes I (*A*) and III (*B*) and cytochrome oxidase (*C*). Solid line, absorbance of the eluate at 225 nm; dashed line, the solvent gradient. The inset in *A* shows the SDS/PAGE analysis of fractions. Proteins were identified by measurement of their molecular masses, by tandem MS of intact protein ions and/or of proteolytic fragments.

matography, and those of complexes I and III and cytochrome oxidase were fractionated by chromatography on a macroporous reverse-phase column (Fig. 1). The resolution and recovery of hydrophobic membrane proteins achieved with the reverse-phase column were especially striking. Clearly, this method has the potential to be developed further as a general method for purifying membrane proteins for MS.

The propanol:acetonitrile extract of mitochondria contained 33 proteins (5), and the chloroform:methanol extract at least 20 proteins. Many of them were identified by tandem mass analysis of proteolytic peptides, and, as described below, most of them were also identified by tandem MS of intact protein ions. Their molecular masses were measured [supporting information (SI) Table 3]. The most significant result here is the measurement of the molecular mass of the ND6 subunit of complex I (SI Fig. 4). This subunit is the last protein encoded in mitochondrial DNA to have its molecular mass measured. It is the only protein encoded on the H (heavy)-strand of human (and bovine) mitochondrial DNA, and its gene overlaps partially the gene for ND5 encoded on the L (light)-strand (14). The measurement of the mass of ND6 (measured value 19,106.7; calculated value 19,077.7) shows that, in common with other proteins encoded in the mitochondrial genome, the protein is α-N-formylated and that otherwise it is not modified covalently (5). Thus, the chemical analysis of the 45 subunits of bovine complex I is now complete (15). None of its subunits is modified covalently by an unknown redox factor (stable under the experimental conditions used) that could participate in the energy transducing mechanism of the enzyme (16).

Fragment Ion Spectra. On the basis of the tandem mass spectra of intact protein ions, membrane proteins fall into two categories. In the first category, the fragment ion spectra were dominated by multiply charged series of fragment ions arising from internal regions of the protein, and they often contained other singly charged series arising from one or both terminal regions. In the second category, the fragment ions giving sequence information were usually singly charged series arising from N- and/or C-terminal regions. To a first approximation, the first category is composed of membrane proteins with masses up to ≈12 kDa with one to four TMHs (the "small proteolipids"), and the second category contained membrane proteins with masses >15 kDa with 4–18 TMHs (referred to as "proteins with multi-TMHs").

The multiply charged ion series from the small proteolipids could be interpreted readily by manual inspection without consideration of isotope resolution and with no mathematical deconvolution or transformation of spectra on to a true mass scale. For example, a sequence of 14 aa was determined from a series of quadruply charged b type ions obtained by tandem MS of an intact protein ion of bovine subunit c of the F-ATPase with *m/z* 1,531 with five associated protons (see Table 1 and Fig. 2*A*), and other sequences were assigned with less certainty by examination of ions of lower abundance. The sequence was deduced from the mass differences between fragments (often adjacent peaks) in the same ion series, as in peptide analysis. Hence, the mass differences between the first four fragment ions were 32.8, 14.4, and 28.4, respectively. These masses were multiplied by successively higher integers until the differences corresponded to recognizable amino acid residue masses (131.2, 57.6, and 113.6 corresponding to Met, Gly, and Ile/Leu residues, respectively). Thus, in this example, the mutiplication factor was four. Other spectra from small proteolipids are shown in Fig. 2, and in SI Figs. 5–12, and the sequences deduced from them are summarized in Table 1. They include the c-subunits from *Escherichia coli* F-ATPases, the *E. coli* transport protein EmrE (17), and bovine phospholamban (see SI Figs. 8, 10, and 12), which has been found previously in mitochondria as well as in the sarcoplasmic reticulum (5, 18, 19).

One protein, PL-5283 (proteolipid with a mass of 5,283 Da) is novel. It copurified with subunit IX of complex III (mass 7,997 Da; see SI Fig. 13). An abundant protein ion with *m/z* 1,321.6 and four associated protons in its electrospray spectrum gave rise to a fragment ion spectrum containing a series of triply charged y ions y33–y44(3^+) in the *m/z* range 1,250–1,650 interpreted as the sequence FJJGFTJGNVV (Fig. 2*B*) where J is Ile or Leu. A partially overlapping sequence (MJ)(Q/K)FJJGFTJ was obtained from a series of singly charged ions in the *m/z* range 100–1,200. The sequence MJ came from a b2 fragment ion with *m/z* 245 (Fig. 2*B* and Table 1). A second protein ion with *m/z* 1,057.5 and five associated protons produced the related sequence GNVVGMYJAQNYDJ from the triply charged ions

BIOCHEMISTRY

Table 1. Sequences of small proteolipids identified by tandem MS of protein ions

			Ions			
			Precursor		Fragment	
Protein	TMH	Residues	*z*	*m/z*	(*z*)	Sequence (residues)
Cox-VIII	1	46	5^+	993.2	y17-y27 (2^+) b19-b28 (3^+)	SVTFLSFLLP (20–29)
PL-5283	1	47	4^+	1321.6	b2-b10 (1^+) y33-y44 (3^+)	(ML)QFLLGFTL (1–10) FLLGFTLGNVV (4–14)
PL-5283	1	47	5^+	1057.5	y23-y37 (3^+)	GNVVGMYLAQNYDI (11–24)
Cox-VIIc	1	47	5^+	1089.3	y12-y24 (2^+)	MMTLFFGSGFAA (24–35)
Plamban	1	52	4^+	1531.2	b30-b51 (4^+)	LFINFCLILICLLLICIIVML (31–51)
A6L	1	66	6^+	1328.4	b2-b10 (1^+) y52-y63 (5^+)	(f-MP)QLDTSTWLTMIL (1–14)
FA-c	2	75	5^+	1531.4	b59-b73 (4^+)	MGLFCLMVAFLILF (60–73)
CIII-IX	1	78	9^+	889.7	b3-b7 (1^+)	VAAR (4–7)
FA-c	2	79	5^+	1657.9	b6-b13 (1^+)	DLLYMAA (7–13)
FA-c	2	79	6^+	1377.2	b12-b23 (2^+)	AAVMMGLAAIG (13–23)
CI-B9	1	83	8^+	1158.4	b22-b30 (3^+) y48-y54 (4^+) y52-y61 (5^+)	AIAGLAVILPTLS (23–35)
EmrE	4	110	7^+	1713.4	y30-y53 (3^+)	IAYAIWSGVGIVLISLLSWGFFG (58–80)

Cox, cytochrome c oxidase; PL, proteolipid; FA, F-ATPase; CI, complex I; Plamban, phospholamban; ND, NADH dehydrogenase.

y23-y37(3^+) (Table 1 and SI Fig. 6). This partial sequence is related to hypothetical proteins with similar masses from zebrafish (*Danio rerio*), rat (*Rattus norvegicus*), and chicken (*Gallus gallus*), and to a 47-aa protein, with a calculated molecular mass of 5,283.3 Da encoded by a bovine cDNA. Residues 1–24 of the bovine protein correspond to the deduced sequence (Table 1). Entries in human, dog and rhesus monkey data-bases, where this sequence forms the C-terminal region of larger hypothetical proteins, will need to be reassessed. The sequence of bovine PL-5283 is 70% identical to seven vertebrate proteins (SI Fig. 14) and is predicted to contain a single transmembrane span (residues 1–21). The distribution of charges suggests that its C terminus will be in the mitochondrial inter-membrane space, and although there are no clues about its exact function, the large number of charged amino acids in its C-terminal region is striking. It is likely that this protein has escaped detection previously because it does not yield proteolytic fragments that are amenable to mass analyses.

Examples of the second category of proteins with multi-TMHs, where sequence information was deduced manually almost exclusively from fragment ion spectra arising from N- and/or C-terminal regions of the proteins, are summarized in Table 2; spectra are shown in Fig. 3 and SI Figs. 15–22. They include the c-subunit from the canine V-ATPase (20) (four TMHs), where sequence was deduced from an ion with nine associated protons, the a-subunit of the F-ATPase (predicted to contain five to six TMHs); the ND1, ND2, and ND5 subunits of complex I (8, 8, and 18 predicted TMHs, respectively); and the ND6 subunit (four or five predicted TMHs) from the same complex. Three hydrophobic proteins that were examined gave no interpretable fragment ion spectra; possible explanations are discussed below. They are subunit Cox-I and Cox-III (12 and 7 TMHs, respectively) of the cytochrome oxidase complex and subunit ND4 (12 or 13 predicted TMHs) from complex I. They are three of the 13 proteins encoded in mitochondrial DNA; the 10 remaining proteins (Cox-II, cytochrome *b*, ND1–3, ND4L, ND5, ND6, and F-ATPase a and A6L subunits) all gave rise to partial sequences (see Tables 1 and 2, Figs. 2*C* and 3 *A* and *B*, and SI Figs. 15 *A* and *B*, 17, 18, and 20–22).

Not all of the membrane proteins that were examined conform exactly to the two categories. For example, subunits ND4L and ND3 of mitochondrial complex I (measured molecular masses of 10,825 and 13,082, respectively) are predicted to have two and three TMHs, respectively, but their fragmentation ion spectra were derived from small singly charged fragment ions. It is possible that slight variation in experimental conditions could produce multiply charged fragmentation spectra.

These observations correlate with the tandem MS data obtained on membrane proteins with one or two TMHs from chloroplasts and cyanobacteria (3, 13, 21), where extensive sequences were obtained from spectra containing multiple and also singly charged fragment ions.

Spectral Characteristics. The relative simplicity of the fragmentation patterns and the ease of interpretation of fragment ion spectra of intact membrane proteins are a consequence of their sequences, and particularly of the distribution of charged and hydrophilic residues. Fragmentation of peptides in the gas phase in CID is assisted by protonation of amide bonds. Hence, cleavages are charge directed, and peptide bond cleavage occurs preferentially at amino acid residues with higher proton affinities, especially to the N-terminal side of arginine and proline (22, 23). Because the TMHs of most membrane proteins contain few charged residues, a random pattern of amide bond cleavage by CID is favoured that is less influenced by differences in proton affinities. Thus, cleavage of the amide backbone was observed frequently in hydrophobic regions, and often the patterns of multiply charged fragment ions used for assigning sequences were derived from them also. This effect is illustrated by the tandem mass spectra of EmrE (SI Fig. 12), where the sequence of residues 58–80 is defined by an extensive series of 24 y^{3+} fragment ions from the TMH extending from residues 61–79 (24). This sequence is bounded by Arg-82 and Pro-55, and no fragment ion was observed from peptide bonds C-terminal to Arg-82.

In the analysis of ions from larger membrane proteins with multiple TMHs, the abundant products of the fragmentation were singly charged ions that defined either one or both ends of the protein (examples are provided by subunits ND1, ND2, ND3, ND4L, ND5, and ND6 of complex I and subunit a of ATP synthase). These terminal regions were generally hydrophobic, and only the C-terminal parts of subunits ND3 and ND6 of

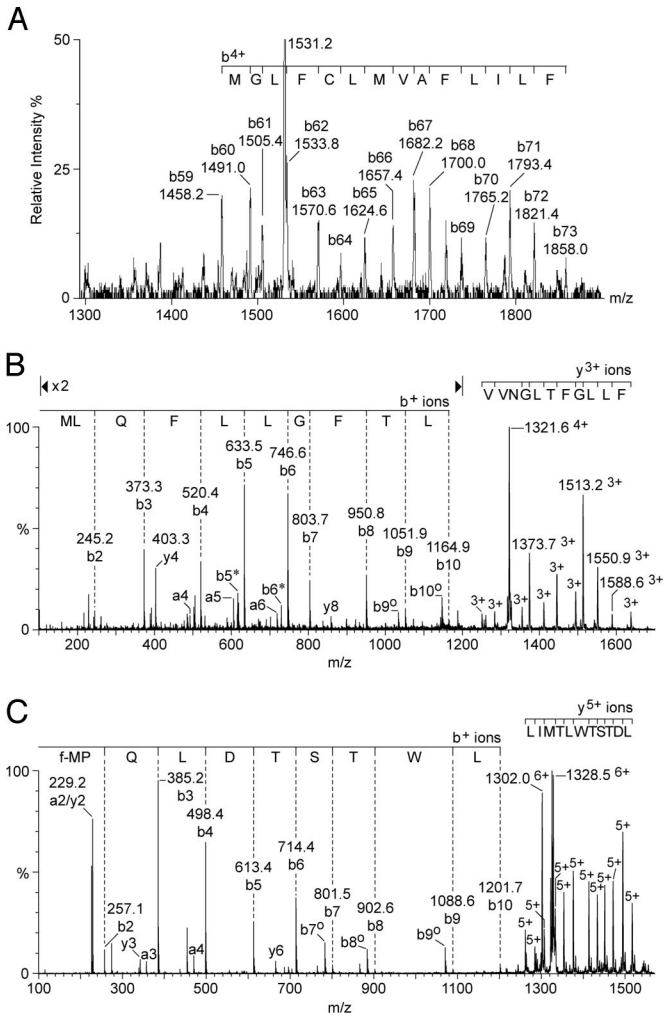

Fig. 2. Tandem MS analyses of protein ions from small proteolipids. Shown are fragment ion spectra from the c subunit of bovine F-ATPase (*A*), bovine mitochondrial PL-5283 (*B*), and the A6L subunit of the F-ATPase (*C*). For further details, see *SI Text*.

complex I contain a basic residue. Despite being restricted to the terminal regions, the sequences provide sufficient information to allow the protein to be identified unambiguously with their associated molecular masses. No significant interpretable fragment ion spectra were obtained from subunit ND4 of complex I or from the cytochrome oxidase subunits Cox-I and Cox-III. The presence of basic amino acids and proline residues in the terminal regions of ND4 and Cox-I (but not in Cox-III) may provide an explanation. No significant data were obtained from the family of mitochondrial transporter proteins, exemplified by the adenine nucleotide carrier (see SI Fig. 23). They have six TMHs, but they are interspersed with extensive hydrophilic and charged sequences (25, 26).

Perspectives for Improving Tandem MS Analysis of Membrane Proteins. The current study has been made possible by our development of methods for the extraction and purification of a wide range of membrane proteins in organic solvents, especially the development of improved methods for normal phase chromatography on polyhydroxyethylaspartamide, and reverse phase chromatography on macroporous reverse-phase columns (27). The latter method is a high resolution method of high potential that allows large hydrophobic proteins to be recovered pure in high yield.

Because of the lack of access to a representative range of suitably purified membrane proteins, so far there have been almost no studies of the impact of different forms of instrumentation on top-down sequencing of membrane proteins. In preliminary comparisons of the same protein ions from subunit c of mitochondrial F-ATPase and subunits Cox-VIIc and Cox-VIII, obtained by low energy CID in triple quadrupole and quadrupole-time of flight (Q-TOF) instruments, we have observed few significant differences in prominent ions or in abundance of fragment ions that are useful for sequence analysis. Although the superior fragment ion resolution of the Q-TOF instrument has enabled quadruply and quintuply charged ions to be distinguished, this feature was not important for any of the sequence assignments. However, a 19^+ ion from bacteriorhodopsin has been analyzed by CID in both linear ion trap with Fourier transform capability and Q-TOF instruments. Although, the baseline resolution of fragment ions was superior in the Fourier transform instrument, no sequence was deduced manually from the spectrum (28), whereas it was possible to deduce the sequence of residues 11–19 in the fragment ion spectrum obtained with the Q-TOF instrument (SI Fig. 19).

The lack of interpretable sequence data from the internal regions of large hydrophobic proteins may be a consequence of membrane proteins having a greater tendency than globular proteins to remain folded in the gas phase (28). CID depends on protonation of amide groups, and so folded hydrophobic regions will incorporate fewer charges during electrospray ionization and fragmentation will be more likely to occur in the less well folded N-and C-terminal regions, as observed. Nonetheless, fragmentation of internal hydrophobic regions was observed in some membrane proteins, for example, the tandem MS data from bacteriorhodopsin contained a doubly charged series of y ions corresponding to residues 11–19 (WLALGTALM; see Table 2), and that of cytochrome *b* contained a series of triply charged b ions corresponding to residues 19–23 (Table 2). Also, the fragmentation of Cox-II (26 kDa, two TMHs) by CID cleaves the bond Leu-68-Pro-69 and singly charged b ion series arise from both the N terminus of the new fragment and the intact molecule. A similar effect has been noted in the SGDH subunit (17 kDa, one TMH) of complex I (15).

It is possible that more extensive unfolding of membrane proteins may be achieved in mass spectrometers equipped with ion traps or ion cyclotrons, where ions can be confined, thereby increasing the opportunity for a protein to unfold. Confinement increases the likelihood of fragment ions forming and of ions being activated and fragmented. However, access to these instruments is likely to remain a difficulty. Another foreseeable source of improvement in tandem MS of membrane proteins may come from recently introduced methods of protein ion dissociation, such as electron capture dissociation and electron transfer dissociation, where peptide chain cleavages are sequence independent and the pattern of fragment ions is less influenced by proton affinities (23, 29). These techniques are likely to give increased random fragmentation of membrane proteins, producing interpretable tandem mass spectra that cover more of the sequence of the protein than at present (30). These benefits have been realized already with some hydrophilic proteins (31, 32).

Experimental Procedures

Preparation, Extraction, and Fractionation of Membrane Proteins. Lysosomal storage bodies associated with canine ceroid lipofuscinosis (provided by D. N. Palmer, Lincoln University, Canterbury, New Zealand), containing mainly subunit c of the V-ATPase (20), were extracted with chloroform:methanol (2:1, vol:vol). Subunit c from the *E. coli* ATP synthase (provided by

Table 2. Sequences of proteins with multi-TMHs determined by tandem MS of protein ions

			Ions			
			Precursor		Fragment	
Protein	TMH	Residues	*z*	*m/z*	(*z*)	Sequence (residues)
ND4L	2	98	6^+	1805.3	y2-y8 (1^+)	QNLNLL(QC) (91–98)
ND3	3	115	8^+	1636.5	b2-b9 (1^+)	(f-MN)LMLALLT (1–9)
					y2-y9 (1^+)	TQK(GL)EW(TE) (107–115)
VA-c	4	155	9^+	1683.0	y4-y15 (1^+)	LGLYGLIVALI (141–151)
ND6	4–5	175	11^+	1738.5	b2-b9 (1^+)	(f-MM)LYIVFIL (1–9)
					y2-y7 (1^+)	MEITR(GN) (169–175)
FA-a	5–6	226	13^+	1910.0	b2-b11 (1^+)	(f-MN)ENLFTSFIT (1–11)
					y3-y11 (1^+)	LLVSLYLH(DNT) (216–226)
Cox-II	2	227	15^+	1737.6	b2-b4 (1^+)	YP (3–4)
Cox-II		69–227	15^+	1737.6	b2-b9 (1^+)	(PA)IILILIA (69–77)
BR	7	247	19^+	1410.6	b10-b19 (2^+)	WLALGTALM (11–19)
CI-ND1	8	318	21^+	1701.1	b2-b8 (1^+)	(f-MF)MINILM (1–8)
CI-ND2	8	347	22^+	1786.5	b2-b11 (1^+)	(f-MN)PIIFIIILL (1–11)
CIII-Cyt-b	8	379	21^+	2030.6	b18-b23 (3^+)	IDLPA (19–23)
CI-ND5	18	606	40^+	1709.0	y3-y7 (1^+)	ILFN (600–603)

VA, V-type ATPase; BR, apo-bacteriorhodopsin. Other abbreviations are as defined in Table 1.

I. Arechaga, Medical Research Council (MRC) Dunn Human Nutrition Unit) obtained by over-expression in *E. coli* (33) was purified by hydroxylapatite chromatography in lauryldimethylamine oxide. Multidrug resistance protein E (EmrE) from *E. coli* (provided by C. G. Tate, MRC Laboratory of Molecular Biology, Cambridge, U.K.) was extracted from *E. coli* membranes with chloroform:methanol (2:1, vol:vol) (17) and purified by chromatography on hydroxylapatite and by gel filtration in the same solvent. Proteins were precipitated from chloroform:methanol extracts and from detergent solutions with 4 vol of diethyl ether at −20°C. For mass spectrometric analysis, protein precipitates of membrane proteins were dissolved in chloroform:methanol (2:1, vol:vol) and diluted into a solution of chloroform, methanol and 0.1% aqueous formic acid (4:4:1 by vol). Apo-bacteriorhodopsin was purified from *Halobacterium salinarium* (34, 35).

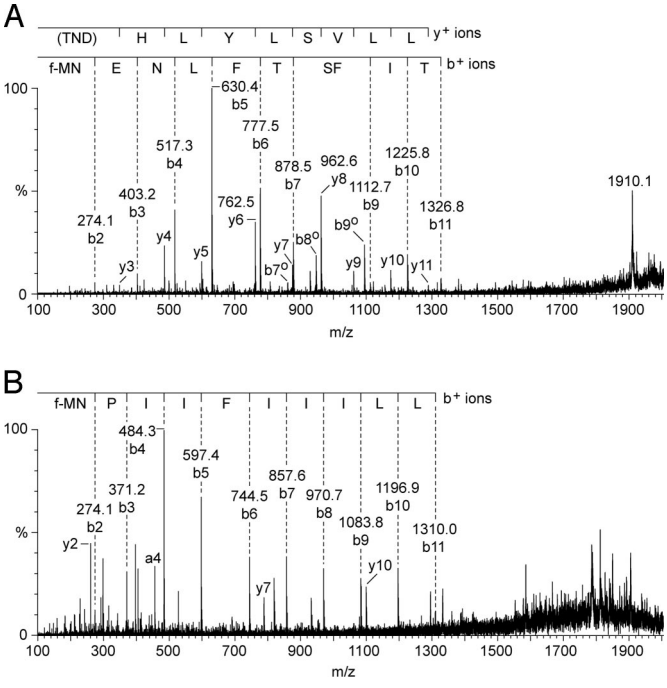

Fig. 3. Tandem mass analyses of protein ions from proteins with multi-TMH. (*A*) Fragment ion spectrum of a protein ion from subunit a of bovine F-ATPase with *m/z* 1,910.0 and 13 associated protons. The N-terminal sequence, (f-MN)ENLFT(SF)IT, is defined by the singly charged ions b2–b11 and the C-terminal region, LLVSLYLH(DNT), by the ions, y3–y11. Ions denoted b° correspond to the loss of water. (*B*) Fragment ion spectrum of protein ion with *m/z* 1,786.5 *m/z* and 22 associated protons from subunit ND2 of bovine complex I. The N-terminal sequence (f-MN)PIIFIIILL is defined by the singly charged ions, b2–b11.

Bovine heart mitochondria were extracted with chloroform:methanol (2:1, vol:vol) (36). The proteins were precipitated with ether, redissolved in chloroform:methanol:water (46.75:45.75:7.5 by vol) containing 60 mM ammonium acetate, pH 7.0, and fractionated by gel filtration on Toyopearl HW-55 in the same solvent. The column fractions were extracted with water to remove salt, and the proteins were precipitated from the organic phase with ether. They were redissolved in chloroform:methanol (2:1, vol:vol) and introduced into the mass spectrometer in a solution of chloroform:methanol:1% formic acid (4:4:1 by vol). This procedure yielded samples of Cox-VIIc, Cox-VIII, CIII-IX, phospholamban and FA-c.

A6L and adenine nucleotide translocase were obtained similarly from bovine heart mitochondria extracted with chloroform:methanol:water (75:23:2 by vol) containing 20 mM ammonium formate, pH 3.7. These proteins were dissolved in formic acid (60%) and purified by reverse-phase HPLC on a PRP-3 column at 40°C (Hamilton, Reno, NV) in 0.1% trifluoroacetic acid with an acetonitrile gradient.

Bovine heart mitochondria and complex I were extracted with 9 vol of 2-propanol: acetonitrile:hexafluoro-isopropanol:water (70:25:0.56:4.44, by vol) containing 20 mM ammonium formate, pH 3.7 (5). The proteins were fractionated by HILIC on columns of polyhydroxyethyl-aspartamide with a gradient of decreasing organic solvent (5) yielding samples of PL-5283, CI-ND2, CI-ND3, CI-ND4L, CI-B9 and FA-a.

Bovine complexes I and III and cytochrome oxidase prepared by ion exchange chromatography (37) were extracted with 12 vol of chloroform:methanol:water (66.7:31.3:2 by vol) containing 20 mM ammonium formate, pH 3.7. Proteins were precipitated with ether, redissolved in the extraction solvent, and precipitated again with ether. The precipitate was redissolved in a mixture of formic acid:trifluoroethanol:hexafluoro-2-propanol:water (60:20:5:15, by vol) and fractionated by reverse-phase chromatography on an mRP column (75 × 2.1 mm i.d.; Agilent,

Cheadle, U.K.). The column was equilibrated in 0.1% trifluoroacetic acid containing 20% trifluoroethanol and 5% hexafluoro-2-propanol at 60°C and eluted with a gradient of isopropanol from 0–70% (vol:vol) and hexafluoro-2-propanol from 5–1% (vol:vol) at a flow rate of 0.1 ml/min. This procedure gave samples of CI-ND4, CI-ND5, CIII-Cytb, Cox-I, Cox-II, and Cox-III.

To obtain samples that contained the ND6 subunit of bovine complex I, the enzyme was extracted with chloroform:methanol:water (66.7:31.3:2 by vol) containing 20 mM ammonium formate, pH 3.7, as described above, and precipitated with ether. The ether precipitate was redissolved in 0.1 vol of extraction solvent, and centrifuged (16,000 × *g* for 10 min at 4°C). The solution contained complex I subunits ND1, ND2, ND3, ND4, ND5, ND6, and ND4L (SI Fig. 24).

SDS/PAGE Analysis of Proteins. Proteins in extracts and column fractions were examined by SDS/PAGE in 12–22% gradient gels. They were stained with Coomassie blue dye, and many of them were identified by mass spectrometric analyses of proteolytic digests (5). Primary chicken antibodies were raised against the synthetic peptide KDKEVEVVFEFNGLGD representing residues 109–124 of the ND6 subunit of complex I by Agrisera (Vännäs, Sweden). They were used to detect ND6 in extracts of bovine heart mitochondria by immunoblotting (5).

MS. Electrospray ionization-MS analyses were performed in positive ion mode with either a Sciex API III$^+$ triple quadrupole mass spectrometer (MDS Sciex, Concord, ON, Canada) equipped with a nanoelectrospray interface, or a Q-TOF1 Q-TOF mass spectrometer (Waters–Micromass, Manchester, U.K.) with a nanoflow electrospray interface (4, 5, 38). Protein ions were disrupted by CID with argon under similar conditions to those used for the tandem MS analysis of peptides (39). Tandem mass spectra were interpreted manually without, either knowledge of isotope spacing for determination of charges on ions, or deconvolution of raw spectral data. Isoleucine and leucine residues, denoted as J, were not distinguished by MS. Where protein sequences were derived from DNA sequences and there was no ambiguity, Ile and Leu were used. The number of charges on fragment ions was determined by arithmetic. Sequences obtained by tandem MS were compared with protein data-bases with BLAST (40), and a bovine expressed sequence tag data-base was screened with TBLASTN.

We thank D. N. Palmer, I. Arechaga, and C. G. Tate for providing samples of proteins; S. Palmer (MRC Dunn Human Nutrition Unit) for help with producing bacteriorhodopsin; and P. James (University of Lund, Lund, Sweden) for suggesting the use of the macroporous supports and for advice on the composition of solvents for reverse phase chromatography of membrane proteins.

1. Schindler PA, van Dorsselaer A, Falick AM (1993) *Anal Biochem* 213:256–263.
2. Musatov A, Robinson NC (1994) *Biochemistry* 33:10561–10567.
3. Whitelegge JP, Zhang H, Aguilera R, Taylor RM, Cramer WA (2002) *Mol Cell Proteomics* 1:816–827.
4. Fearnley IM, Walker JE (1996) *Biochem Soc Trans* 24:912–917.
5. Carroll J, Fearnley IM, Walker JE (2006) *Proc Natl Acad Sci USA* 103:16170–16175.
6. Mann M, Wilm MS (1994) *Anal Chem* 66:4390–4399.
7. Loo JA, Edmonds CG, Smith RD (1990) *Science* 248:201–204.
8. Ginter JM, Zhou F, Johnston MV (2004) *J Am Soc Mass Spectrom* 15:1478–1486.
9. Kelleher NL, Lin HY, Valaskovic GA, Aaserud DJ, Fridriksson EK, McLafferty FW (1999) *J Am Chem Soc* 121:806–812.
10. Reid GE, McLuckey SA (2002) *J Mass Spectrom* 37:663–675.
11. Buzy A, Ryan EM, Jennings KR, Palmer DN, Griffiths DE (1996) *Rapid Commun Mass Spectrom* 10:790–796.
12. Chen R, Fearnley IM, Palmer DN, Walker JE (2004) *J Biol Chem* 279:21883–21887.
13. Whitelegge JP (2005) *Trends Anal Chem* 24:576–582.
14. Anderson S, de Bruijn MHL, Coulson AR, Eperon IC, Sanger F, Young IG (1982) *J Mol Biol* 156:683–717.
15. Carroll J, Fearnley IM, Skehel JM, Shannon RJ, Hirst J, Walker JE (2006) *J Biol Chem* 281:32724–32727.
16. Friedrich T, Brors B, Hellwig P, Klintscher L, Rasmussen T, Scheide D, Schulte U, Mäntele W, Weiss H (2000) *Biochim Biophys Acta* 1459:305–309.
17. Yerushalmi K, Lebendiker M, Schuldiner S (1995) *J Biol Chem* 270:6856–6863.
18. Taylor SW, Fahy E, Zhang B, Glenn GM, Warnock DE, Wiley S, Murphy AN, Gaucher SP, Capaldi RA, Gibson BW, *et al.* (2003) *Nat Biotechnol* 21:281–286.
19. Forner F, Foster LJ, Campanaro S, Valle G, Mann M (2006) *Mol Cell Proteomics* 5:608–619.
20. Palmer DN, Jolly RD, van Mil HC, Tyynela J, Westlake VJ (1997) *Neuropediatrics* 28:45–48.
21. Whitelegge JP, Laganowsky A, Nishio J, Souda P, Aguilera R, Zhang H, Cramer WA (2006) *J Exp Bot* 57:1515–1522.
22. Wysocki VH, Tsaprailis G, Smith LL, Breci LA (2000) *J Mass Spectrom* 35:1399–1406.
23. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF (2004) *Proc Natl Acad Sci USA* 101:9528–9533.
24. Ubarretxena-Belandia I, Baldwin JM, Schuldiner S, Tate CG (2003) *EMBO J* 22:6175–6181.
25. Saraste M, Walker JE (1982) *FEBS Lett* 144:250–254.
26. Pebay-Peyroula E, Dahout-Gonzalez C, Kahn R, Trezeguet V, Lauquin GJ, Brandolin G (2003) *Nature* 426:39–44.
27. Martosella J, Zolotarjova N, Liu H, Moyer SC, Perkins PD, Boyes BE (2006) *J Proteome Res* 5:1301–1312.
28. Whitelegge JP, Halgand F, Souda P, Zabrouskov V (2006) *Expert Rev Proteomics* 3:585–596.
29. Zubarev RA, Kelleher NL, McLafferty FW (1998) *J Am Chem Soc* 120:3265–3266.
30. Zabrouskov V, Whitelegge JP (2007) *J Proteome Res* 6:2205–2210.
31. Mikesh LM, Ueberheide B, Chi A, Coon JJ, Syka JEP, Shabanowitz J, Hunt DF (2006) *Biochim Biophys Acta* 1764:1811–1822.
32. Han X, Jin M, Breuker K, McLafferty FW (2006) *Science* 314:109–112.
33. Arechaga I, Butler PJG, Walker JE (2002) *FEBS Lett* 515:189–193.
34. Oesterhelt D, Stoeckenius W (1974) *Methods Enzymol* 31:667–678.
35. Hufnagel P, Schweiger U, Eckerskorn C, Oesterhelt D (1996) *Anal Biochem* 243:46–54.
36. Fearnley IM, Walker JE (1987) *Biochemistry* 26:8247–8251.
37. Buchanan SK, Walker JE (1996) *Biochem J* 318:343–349.
38. Wilm MS, Mann M (1994) *Int J Mass Spectrom Ion Processes* 136:167–180.
39. Morris HR, Paxton T, Dell A, Langhorne J, Berg M, Bordoli RS, Hoyes J, Bateman RH (1996) *Rapid Commun Mass Spectrom* 10:889–896.
40. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) *J Mol Biol* 215:403–410.