Research Article

The proteome survey of an electricity-generating organ (*Torpedo californica* electric organ)

Javad Nazarian[1], Yetrib Hathout[1], Akos Vertes[2] and Eric P. Hoffman[1]

[1] Research Center for Genetic Medicine, Children's National Medical Center, Washington, DC, USA
[2] Department of Chemistry, George Washington University, Washington, DC, USA

Torpedo californica is a species in class *Chondrichthyes*. Electric rays have evolved the electric organ, which is similar to the mammalian neuromuscular junction (NMJ). Here, we took a combined cDNA sequencing and proteomic approach to define the molecular constituents of the *T. californica* electric organ. For soluble proteins, 2-DE was used and 224 protein spots were mapped. Plasma membrane fractions were analyzed using the shotgun approach (LC-MS/MS). A *Torpedo* cDNA library was constructed and 607 cDNA clones were sequenced. Identification of electric organ proteins was done using cross-species comparisons, and a custom database was constructed from cDNA translations. We unambiguously identified 121 proteins and transcripts, 103 of which were novel additions to the existing databases of *Torpedo* fish. Fifteen proteins of known function, but not previously associated with either the electroplaque or NMJ, were present at high abundance. These included the heat shock and oxidative stress proteins, annexin V (calelectrin), and plectin 1. Most interesting were the unambiguous matches to 11 human ORFs of unknown function, including four potential RNA splicing proteins, a vacuolar sorting protein, and a tetraspanin containing protein. This analysis identified proteins that may play a role in the higher vertebrate neuromuscular junction or other electrical synapses.

Received: September 11, 2006
Revised: November 2, 2006
Accepted: November 11, 2006

Keywords:
Calelectrin / Electric organ / Neuromuscular Junction / *Torpedo californica*

1 Introduction

The electric organ of *Torpedo californica* and other electric rays are modified muscle fibers (electroplaques) primarily used to shock and stun prey by producing hundreds of volts of electricity [1]. Hexagonally shaped electroplaques are stacked in the pelvis area of *Torpedo* and form synapses on one side with innervating motorneurons [2–4]. The post-synaptic membrane of the electroplaque shows a very high density of acetylcholine receptors (AChRs). Neuronal acetylcholine (ACh), released in the synaptic cleft, binds to post-synaptic AChRs and facilitates a cascade of ion exchange across the polarized membrane. Proteins involved in this polarized high volume ion flux are very similar between the *Torpedo* electroplaque and the neuromuscular junction (NMJ). The NMJ is an electrical synapse, where nerve cells communicate with and activate muscle cells. Hence, the *Torpedo* electric organ has been extensively studied as a model for understanding the mammalian synapse and the NMJ [5–8]. The large size of the electric organ (about 800 g *per* fish) compared to the small size of the NMJ (30–50 μm of membrane specialization *per* myofiber) makes it a relatively abundant source of synaptic proteins.

During the 1970s and 1980s, initial studies using the *T. californica* electric organ led to the first biochemical identification, purification, and visualization of the key transmembrane ion channel, the AChR [9–11]. Intensive work on the

Correspondence: Dr. Eric P. Hoffman, Research Center for Genetic Medicine, Children's National Medical Center, 111 Michigan Avenue NW, Washington, DC 20010, USA
E-mail: EHoffman@cnmcresearch.org
Fax: +1-202-884-6014

Abbreviations: AChE, acetylcholinesterase; **AchR,** acetylcholine receptor; **LCM,** laser capture microscopy; **LIT,** linear IT; **NMJ,** neuromuscular junction; **SFRS5,** Arginine/serine-rich 5; **SMD3,** small nuclear ribonucleoprotein polypeptide D3; **Smpx,** small muscle protein X chromosome

© 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Torpedo electric organ in the 1980s and 1990s led to the identification of other major molecular players of the synapse, such as agrin and acetylcholinesterase (AChE) [7, 12–15]. However, with the emergence of high-throughput genomics, the requirement for large amounts of tissues for biochemical approaches was no longer limiting, and research on the *Torpedo* waned.

We have recently reported laser capture microscopy (LCM) and mRNA profiling studies of the vertebrate NMJ [16]. These experiments are technically challenging and limited to the study of mRNAs that are highly enriched in the nuclear domains underlying the NMJs. As such, they do not define protein modifications, or subcellular locations of NMJ protein constituents. We also found it difficult to validate candidate NMJ mRNAs and proteins, as there are no alternative data sources for the study of NMJs, such as EST libraries or proteomics sources. Finally, mRNA profiling using LCM is likely to be a relatively insensitive tool for the identification of protein components of the NMJ. For example, two proteins/complexes have been hypothesized to exist based on observed signaling cascades (MASC, RATL), yet the identity of these proteins or corresponding genes remains unknown [17–19]. MS has recently proven to be a powerful technique for identifying molecular players and events at the NMJ [20–22]. To increase the understanding of the mammalian NMJ, we felt that it was important to develop robust data sources of both mRNA (cDNA) sequence, and proteomic data of the *Torpedo* electric organ.

Here, we describe a three-pronged approach to *Torpedo* electric organ molecular characterization. Firstly, we created a cDNA library of the electric organ, and described the initial sequence data of 607 clones. This sequence data was used to provide initial assessments of transcript abundance, and to produce a peptide map and MS/MS spectral matching database for protein ID. Secondly, we defined a S100 cytosolic 2-D map with protein ID of the electric organ using a 2-DE MS-based approach. Finally, we studied the membrane fraction of the electric organ using the shotgun approach (in-solution digestion followed by LC-MS/MS analysis). We show that about 25% of proteins can be identified through cross-species mapping of MS data. This preliminary study also triples the number of *Torpedo* transcripts and proteins identified to date.

2 Materials and methods

2.1 Cell fractionation and protein extraction

Frozen *T. californica* electric organ (150 mg) was pulverized in liquid nitrogen and processed for cytosolic and membrane protein extraction using standard sucrose gradient fractionation techniques [23]. Aliquots were taken from the supernatant and protein concentration was measured using a BioRad protein assay reagent, following the manufacturer's instructions (BioRad, Hercules, CA). Protein extracts were then stored at –80°C until analysis.

2.2 Analysis of the S100 cytosolic fraction

Two hundred micrograms of the (S100) cytosolic protein extract was processed for 2-DE/MS analysis. After desalting the protein sample against 10 mM Tris-HCl at pH 7 using p6 Bio-Spin column, the protein solution was dried by centrifugation under vacuum. Rehydration buffer (180 μL) containing 7 M urea, 2 M thiourea, 2% CHAPS, 50 mM DTT, and 0.5% ampholite at pH 3–10 was added to solubilize and denature the proteins. The first dimension electrofocusing was performed on IPG strips (17 cm, pH 3–10) using a BioRad electrofocusing chamber (BioRad) operated as follows: 12 h rehydration, 250 V for 15 min, 1000 V for 1 h, and 10 000 V for 4 h. The second dimension SDS-PAGE was performed on 18 × 18 cm Tris-HCl polyacrylamide (8–16%) precast gels (BioRad). Protein spots were fixed and visualized using Bio-Safe Coomassie Stain (BioRad) and the gels were scanned on GS800 densitometer (BioRad).

Protein spots were excised from the gel with the tip of a clean polypropylene pipette, transferred into a microcentrifuge tube containing 100 μL of deionized water, and processed for in-gel tryptic digestion as described previously [24]. The peptides were extracted, dried by vacuum centrifugation, redissolved in 10 μL of 0.1% TFA, and desalted using C18 ZipTip micropipette tips (Millipore, Bedford, MA) following the manufacturer's User Guide. The peptides were eluted from the ZipTip in 10 μL of ACN/0.1% TFA (70:30 v/v).

Peptide solution (0.3 μL) was mixed with 0.3 μL of matrix solution (50 mM CHCA in ACN/0.1% TFA (70:30 v/v)) and spotted on the MALDI plate. MS and MS/MS analyses were performed on a 4700 ABI TOF-TOF mass spectrometer (Applied Biosystems, Foster City, CA) equipped with Nd:YAG 200 Hz laser. A mixture of standard peptides was used to externally calibrate the instrument. Protein identification was carried out using the GPS explorer software (Applied Biosystems, Framingham, MA). Both MS and MS/MS data were combined to search nonredundant NCBI databases (all taxonomy included).

2.3 Analysis of the membrane fraction

The membrane fraction obtained above was resuspended at 1 mg/mL in 200 mM Na_2CO_3, pH 11, incubated on ice for 1 h, and titrated with 10% formic acid to obtain a pH of seven. The sample was then adjusted to 8 M urea and reduced and alkylated. Briefly, the extract was reduced with 10 mM DTT for 30 min at 37°C and alkylated by iodoacetamide (100 mM) for 1 h in dark at room temperature. The denatured proteins were then desalted using a biospin column and digested with trypsin (enzyme/protein ratio = 1:50 w/w) overnight at 37°C.

Aliquots of membrane protein digest were analyzed by LC-MS/MS using an LC-Packing system (DIONEX UltiMate™ Capillary/Nano LC System, Dionex, Sunnyvale, CA)

© 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

connected to a Linear IT (LTQ) mass spectrometer (Thermo Electron, San Jose, CA). Each sample was injected *via* an autosampler and loaded onto a C18 trap column (300 μm × 1 mm, LC Packings) for 6 minutes at a flow rate of 10 μL/min. The sample was subsequently separated by a C18 RP column (75 μm × 15 cm, Vydac, Columbia, Maryland) at a flow rate of 300 nL/min. The mobile phases consisted of water with 0.1% formic acid (A) and 90% ACN with 0.1% formic acid (B), respectively. A 90-min linear gradient from 5 to 50% B was typically employed. After LC separation, the sample was introduced into the mass spectrometer *via* a 10-μm silica tip (New Objective, Ringoes, NJ) adapted to a nanoelectrospray source (Thermo Electron). The spray voltage was set at 2.0 kV and the heated capillary at 160°C. The LTQ was operated in data-dependent mode in which one cycle of experiments consisted of one full-MS survey and subsequently three sequential pairs of intercalated zoom scans and MS/MS experiments. The targeted ion counts in the IT during full-MS, zoom scan, and MS/MS were 30 000, 3000 and, 10 000, respectively. Peptides were fragmented in the LIT using CID with the collision gas (helium) pressure set at 1.3 millitorrs and the normalized collision energy value set at 35%. Protein searches were performed with the BioWorks 3.2 software (Thermo Electron). Briefly, each file was searched against NCBInr and Swiss-Prot databases (all taxonomy included) using MASCOT search engine set with the following possible protein modifications: 16 Da shift for oxidized Met and 57 for Cys residue (peptide ion score $\geq$ 30). The same file was also searched against our in-house *Torpedo* library using the SEQUEST search engine. The acceptance criteria for peptide identification were set as follows: a DeltaCn (ΔCn)>0.1, a variable threshold of *Xcorr versus* charge state (*Xcorr* = 1.9 for $z = 1$, *Xcorr* = 2.5 for $z = 2$, and *Xcorr* = 3.5 for $z \geq 3$) and a peptide probability based score with a p value <0.001. The rate of false positive identification was less than 1.6% when a database search was performed against reversed FASTA formats of human and mice NCBInr and human Swiss-Prot databases using the same filtration criteria.

2.4 Construction of *Torpedo* electric organ cDNA library

T. californica electric organ was ground in liquid nitrogen and total RNA was isolated. Seven milligrams of total RNA was sent to Amplicon Express (Pullman, WA) for the construction of a cDNA library (5.16×10^6 cfu/mL primary colonies). Briefly, pBluescript II KS(+) vector was used with *E. coli* (EP-Max 10B) as the host strain. Cloning sites used for insertion were *Eco*RV (GGAATTC/GATATCAAG) and *Not*I (site of poly A tail [CCGCGGTGG/CGGCCGCT]). Gridded colonies (onto 384-well plates) were stored in SOC medium containing 15% glycerol and kept at −80°C.

2.5 Sequencing of 607 clones and establishment of an in-house proteome database

Torpedo clones were sequenced from the 5′ end using the T3 primer (5′ATTAACCCTCACTAAAGGGA 3′). Clones were sequenced either by plasmid purification (79 clones by Polymorphic, Alameda, CA) or by direct PCR sequencing (528 clones by Amplicon Express). cDNA sequences were then saved in FASTA format and imported into BioWorks software. The software then converted these sequences into six reading frames, *in silico* trypsin digested and searched using the same parameters as above (alkylation of Cys residues (+57 Da) and possible oxidation of Met residues (+16 Da)). The resulting file was indexed as a local library searchable by peptide mass spectra.

3 Results

3.1 Construction and partial sequencing of a cDNA library from *Torpedo* electric organ

We produced a cDNA library of *T. californica* electric organs consisting of 5.16×10^6 primary colony forming units (cfu/mL) in pBluescript II KS+. The cloning protocol enriched for 5′ ends of all transcripts. Ninety nine percent of clones contained inserts, and the average insert size was 1.3 ± 0.86 kbp. Colonies (37 056) were plated and 607 clones were picked, DNA was purified and sequenced from the 5′ end. The average sequence length was 700 bp, while the average clone insert size was 1.3 kb. All sequences were deposited in NCBI GenBank (GenBank PubMed sequence IDs are available as Supplementary Table 1).

Homology searches (BLAST) within the 607 clones showed 50.8% to be singletons (represented only once in the group sequenced). However, nonoverlapping clones of the same sequence could not be ruled out. We then queried matches against *T. californica* sequences previously deposited in GenBank. The NCBI database currently contains 38 nucleotide and 140 protein sequences for *T. californica*. All of these (mainly redundant) sequences represent only 36 distinct transcripts. We identified 11 of the 36 by cDNA sequence: eight that have been previously characterized as components of the mammalian NMJ (Table 1A), and three that had been sequenced but not associated previously with the NMJ (seven) or electric organ (four) (Table 1B). These 11 previously characterized transcripts were inclusive of the most abundant cDNAs in our library, including creatine kinase (13/607 sequenced clones), AChE (eight clones), and α-Na^+/K^+ ATPase (eight clones).

The 607 cDNA sequences were then tested for homology (BLAST) against all GenBank NCBInr sequences. For our homology search, we ensured that our sequences had orthologs in human and/or mice and that the homology was at least 55%. This provided positive identification of 18 addi-

© 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 1. Known *Torpedo* genes/proteins. (A) Known *Torpedo*, NMJ associated, (B) known *Torpedo*

(A)

Gene symbol	Swiss-Prot ID	gi ID	cDNA clones #	MS/MS peptide number identified			
				Membrane LTQ			S100 2-DE
				Indexed	Swiss-Prot	NCBI	
AChE	P04058	224895	8		7	4	
Rapsyn	P09108	131129	4	1	6	2	
α-AchR	P02710	223528	2	2	7	4	
γ-AchR	P02714	223654	4	1			
β-AchR	P02712	223613	2				
δ-AchR	P02718	223614	1				
Presynaptic Ca^{2+} channel	P56101	213227	2	1			
VAT-1	P19333	137463			7	6	12
α-Na^+/K^+ ATPase	P05025	114388	8	4	28	22	
Chloride channel protein	P35522	544028			10	1	
SITS-binding protein	P19965	134751			7		
Vimentin	P23729	64402			4	2	15
Synuclein	P37379	419969			2	2	

(B)

Gene symbol	Swiss-Prot ID	gi ID	cDNA clones #	MS/MS peptide number identified			
				Membrane LTQ			S100 2-DE
				Indexed	Swiss-Prot	NCBI	
Calmodulin	P62151	49037467	2	1	3	2	4
Creatine Kinase	P04414	125309	13	10	9	8	26
Type III intermediate filament	P23729	124212			4		
α-MHC class II chain protein	AAD54120	5852593	1				
Synuclein	P37379	419969			2	1	

tional *Torpedo* genes through cross-species homology, including six that have been characterized as components of mammalian NMJs (Table 2A), and 12 that have been characterized as mammalian proteins but have not yet been described as a component of the NMJ (Table 2B).

To identify potential novel mammalian NMJ proteins, we queried for unambiguous matches of *Torpedo* cDNAs to human ORFs, corresponding to poorly characterized or uncharacterized (hypothetical) genes. We identified ten human ORFs matching *Torpedo* electric organ cDNA, with a range of 77–98% amino acid homology over regions of 64 to 137 residues (Table 3). These included putative proteins with RNA metabolism protein motifs (small nuclear ribonucleoprotein polypeptide D3 (SMD3), RPR38, FLJ37697), a vacuolar sorting protein motif (VPS8), and a tetraspanin motif (CD9).

We translated the 607 cDNA sequences into six reading frames, and developed both a predicted trypsin peptide map and MS/MS spectral database using BioWorks 3.2 software. The in-house database was then searched using BioWorks software (see Section 2) where known and novel *Torpedo* proteins were identified.

3.2 Proteome survey of the soluble fraction of the *Torpedo* electric organ

Soluble proteins (S100) were isolated from *T. Californica* electric organ tissue, and processed for 2-DE and MS analysis, as described in Section 2 (Fig. 1). A total of 224 protein spots were excised, digested by trypsin, and the resulting peptides analyzed by MALDI-TOF using both MS and MS/MS mode. Database searches were performed by combining both mass fingerprinting data and MS/MS data for each spot. We used MASCOT to search the nonredundant (NCBInr) database (all taxonomies included), and the following stringent criteria (peptide count >3, protein score >65, and peptide ion score >40) were used to retain or reject the identified proteins. Among the 224 excised spots, we confidently iden-

© 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 2. Novel *Torpedo* genes/proteins. (A) Known NMJ proteins previously unknown in *Torpedo*, (B) novel gene/protein of the *Torpedo* electric organ

(A)

Gene symbol	Protein ID/GI	cDNA clones #	MS/MS peptide number identified	
			Membrane LTQ	S100 2-DE
Desmin	P17661			6
β-Synatogamin	NP_795905		3	
Cathepsin D	NP_034113	1		
Calcineurin binding protein 1 (Cabin1)	AAH57551	1		
Intersectin 1	NP_034717	1		
MEF2C	AAH57650	3		
Integral membrane protein 2B (ITM2B)	NP_032436	1		
Apoliprotein A-I binding protein (AIBP)	NP_000473	1		

(B)

Gene symbol	Protein ID/GI	cDNA clones #	MS/MS peptide number identified	
			Membrane LTQ	S100 2-DE
Annexin V	P08758	2	7 (indexed)	
Heat shock protein 90	P07901	1	3	16
Carbonic anhydrase	P00915	1	1	
Phosphoglycerate mutase 2	P15259	1	1	
Phospholipase A1	Q86WX6	1	1	
Phosphoglycerate kinase	P51903		2	10
Tubulin β-1 chain	Q9YHC3	1	7	23
Tubulin β-5 chain	P09653	1	3	
Adenylate kinase	P00568		1	8
Glyceraldehyde-3-phosphate dehydrogenase	P04406		2	6
Ubiquitin	Q05550	1		6
Thioredoxin peroxidase	Q63716	1		6
6-Phosphofructokinase	P08237	1		9
γ-Enolase	O57391	3	1 (indexed)	7
Plectin 1	Q6S383	1		39

tified 58 protein spots corresponding to 33 distinct proteins (Supplementary Table 2).

3.3 Proteome survey of the membrane fraction of the *Torpedo* electric organ

Membrane proteins were digested by trypsin and analyzed by LC-MS (Fig. 2, panel A) and MS/MS (Fig. 2, panel B) as described in Section 2. Proteins were identified against NCBInr (all taxonomies included) as well as our in-house constructed *Torpedo* cDNA database using the SEQUEST search engine. This approach generated matches to 78 human or mouse proteins (Supplementary Table 3). Indexing to the local cDNA translational data identified ten proteins with eight of these overlapping with those identified *via* the NCBInr database searches.

3.4 Analyses of identified proteins and cDNA sequences

In order to provide validation of the MS/MS generated protein IDs, we first inspected each list for known *Torpedo* proteins. We found 18 out of 36 known *T. californica* proteins currently available in the public database. Most of these proteins were identified by multiple methods (cDNA, MALDI MS/MS, or LC-MS/MS) and are either known NMJ-associated proteins (Table 1A) or non-NMJ genes/proteins (Table 1B).

We then searched our data for known NMJ (Table 2A) and non-NMJ (Table 2B) proteins which were not previously identified in *Torpedo*. Shotgun membrane protein IDs with low peptide counts ($n = 1$) were retained if visual inspection of MS/MS data confirmed the peptide sequence found in the

© 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 3. cDNA clones from *Torpedo* electric organ as potential components of the NMJ

Torpedo		Human					
cDNA Clone ID	Number of amino acids translated	Gene ID	Type	Potential function	Amino acid		Locus
					Homology	Number	
TFI_004_D8.T3	138	SMD3	SM, SML motif	Pre-mRNA splicing	97%	138	chr22:23, 283, 643-23, 297, 942
TFI_005_H4.T3	99	PRP38	PRP motif	pre-mRNA splicing	96%	63	chr1:52, 650, 837-52, 654, 947
TFI_1_E3.T3	178	FLJ37697	RRM motif	RNA recognition	94%	108	chr6:17, 389, 792-17, 391, 140
TFI_2_G2.T3	145	VPS8	Vps	Vacuolar sorting	84%	107	chr3:186, 194, 465-186, 196, 936
TFI_004_E2.T3	124	CD9	Tetraspanin Motif	Ca^{2+} signaling/ adhesion	76%	100	chr12:6, 212, 058-6, 212, 144
TFII_1_A9.T3	81	SMPX	ORF	Unknown	77%	57	chrX:73,310,450-73,310,635
Contig 0228	184		Motif	Unknown (DUF1308)	92%	137	chr7:42, 915, 828-42, 916, 238
TFI_005_B12.T3	173		Motif	Unknown (DUF766)	89%	155	chr14:59, 031, 490-59, 035, 378
TFI_003_B3.T3	209		ORF	Unknown	81%	167	chr15:32, 167, 576-32, 181, 209
TFI_2_B1.T3	64		ORF	Unknown	82%	64	chr19:3, 562, 924-3, 570, 149

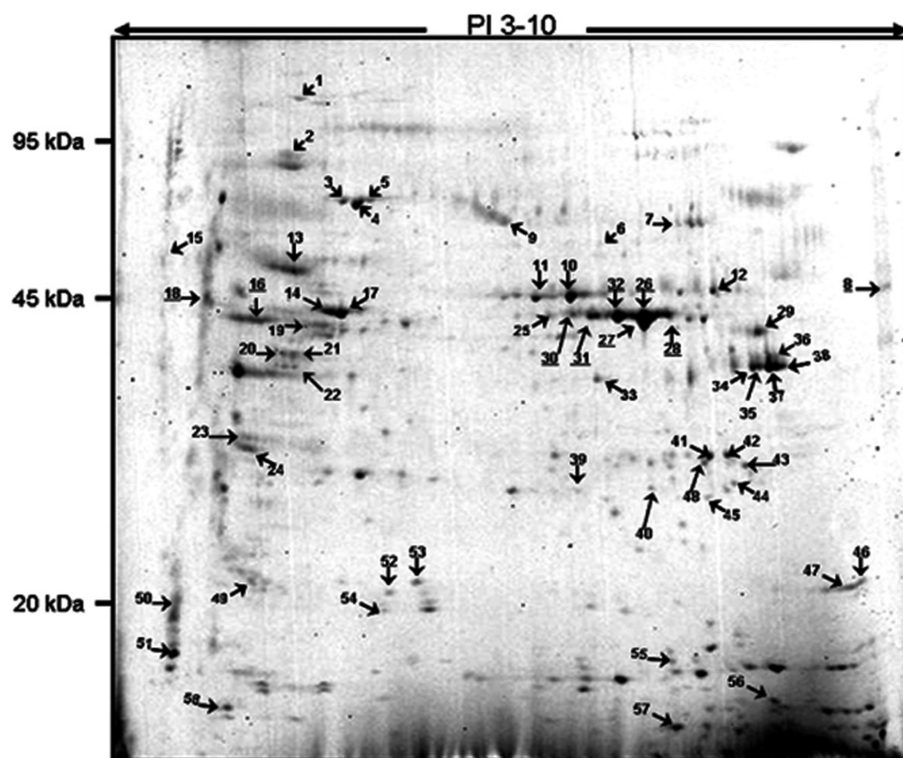

Figure 1. Cytosolic (S100) 2-DE of *Torpedo* electric organ with spots identified by MALDI-TOF-MS/MS. Protein spots (224) were isolated, subjected to in-gel trypsin digestion, and analyzed by MALDI-TOF and MS/MS analyses. Shown are the 58 spots positively identified by a combination of cross-species MS/MS spectral matching, *Torpedo* cDNA matching, and/or LC/MS matching. The 58 spots corresponded to 33 distinct proteins. Multiple isoforms of creatine kinase were identified, and these are indicated by underlined numbers (8, 16, 18, 26–28, 30–32).

© 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

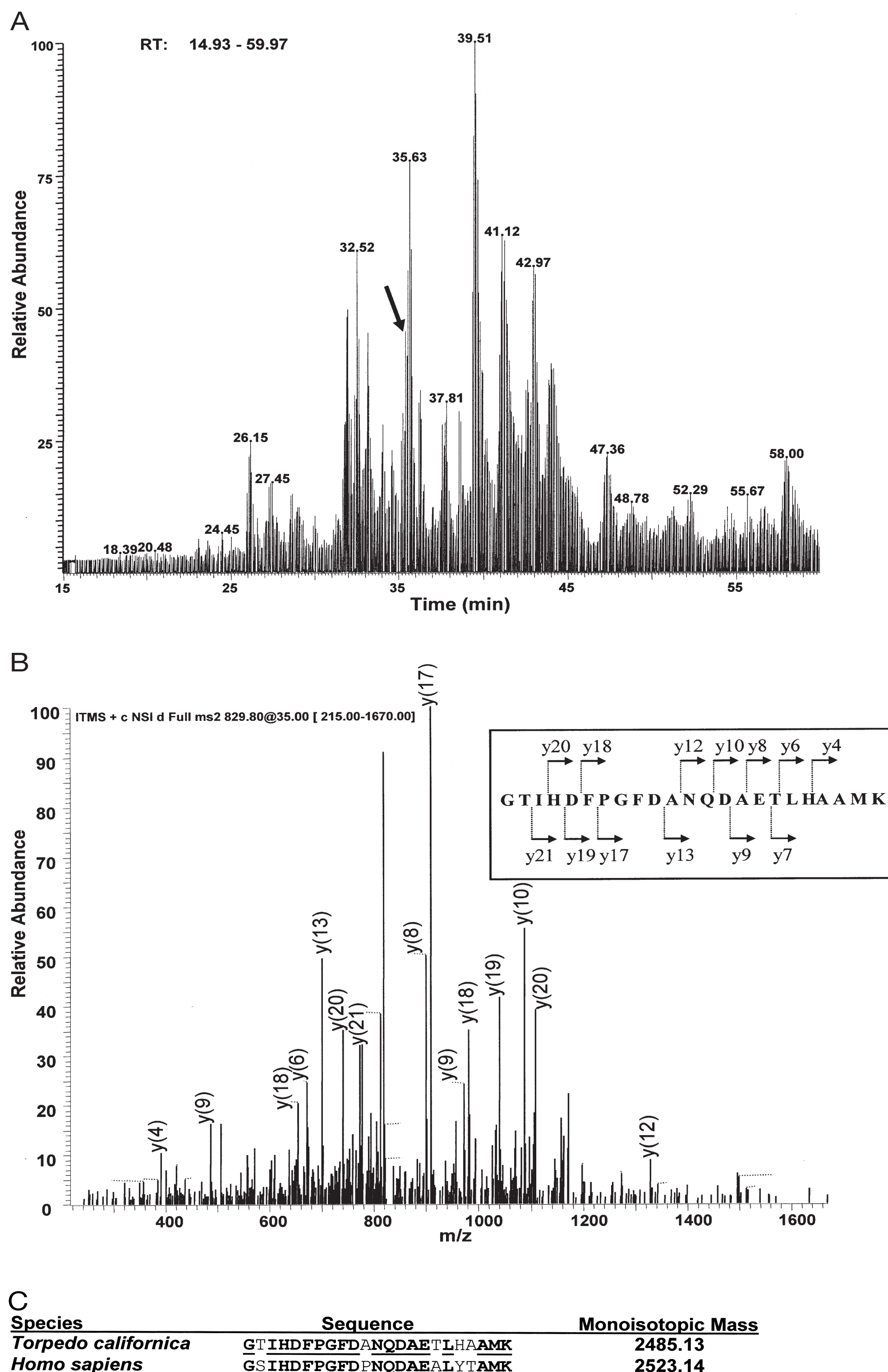

Figure 2. Shotgun proteomic analyses of membrane proteins by LC-MS and MS/MS using a ThermoFinnigan LIT quadropole (LTQ) mass spectrometer. (A) shows the PMF of *T. californica* membrane proteins by LC-MS. The arrow indicates the peak selected for MS/MS analysis. (B) shows peak 1107.70 which was selected, fragmented, and sequenced by MS/MS, where *y* ions were positively identified by matching to the custom electric organ cDNA translated database. Inset shows the peptide sequence with assigned fragment ions. (C) shows the tryptic peptide of *Torpedo* annexin (calelectrin) and its human ortholog. Although this peptide is 78% similar to its human ortholog the difference in its monoisotopic molecular mass will hinder its identification from the database. (Similar amino acids are in bold and underlined).

© 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

database searches, and if they were identified by more than one method (cDNA, MALDI MS/MS, or LC-MS/MS). We found eight known NMJ transcripts using multiple methods (Table 2A). Our search for non-NMJ transcripts that were identified by at least two methods yielded a total of 14 transcripts (Table 2B). The remainder of the identified proteins are listed in Supplementary Tables 2 (S100) and 3 (membrane) with their peptide count, protein score, or percent protein coverage.

4 Discussion

4.1 *T. californica* electric organ proteomics: Technical considerations

The *Torpedo* electric organ is a whole organ model for the minute NMJ in other animals. The components and function of the NMJ remain poorly understood, and our goal was to use proteomics to understand the components of the electric organ and, by inference, the NMJ. *T. californica*, like sharks and other rays in class *Chondrichthyes*, is evolutionarily distantly related to vertebrates and to the few fish for which protein/gene sequences are available (Fig. 3). Proteomics requires databases for spectral matching to determine protein ID, either using peptide fingerprints, or MS/MS spectra, and these databases are typically generated using genomic DNA and cDNA/mRNA sequence data. None of the *Chondrichthyes* species have yet been targeted for genomic sequencing. Thus, protein ID using peptide fingerprinting (2-D) or MS/MS spectra (2-D, LC-MS) was challenging, as expected. The 2-DE MS data of soluble proteins were more successful for protein identification: out of 224 spots, 58 (25%) could be assigned, with a peptide ion score better than 44 in a MASCOT search (sum of >3 peptides), using NCBInr databases (Supplementary Table 2), with robust ID of 33 distinct proteins. As anticipated, LC/MS shotgun analysis of membrane fractions was less successful using cross-species matching, where it was not possible to identify proteins using similarly stringent criteria. However, the limited number of *Torpedo* proteins previously characterized, have been predominantly abundant membrane proteins, and LC/MS showed robust ID of 18 proteins by intraspecies matching (of the total 36 distinct proteins previously reported).

To increase the sensitivity and specificity of the proteomics of the *Torpedo*, we constructed a cDNA library of mRNA from the electric organ, and created a pilot proteomics matching database of sequence data from an initial 607 cDNA clones, 50% of which were tentatively distinct non-overlapping sequences. This provided validation for ten LC/MS peptides and parent proteins. The cDNA sequencing also provided new ORFs that could be matched to DNA sequence data in human and mouse. Taking all data together, we report unambiguous identification of 121 *Torpedo* proteins, with 103 of these as new additions to the existing databases, along with 190 distinct electric organ cDNAs matching to human or mouse sequences.

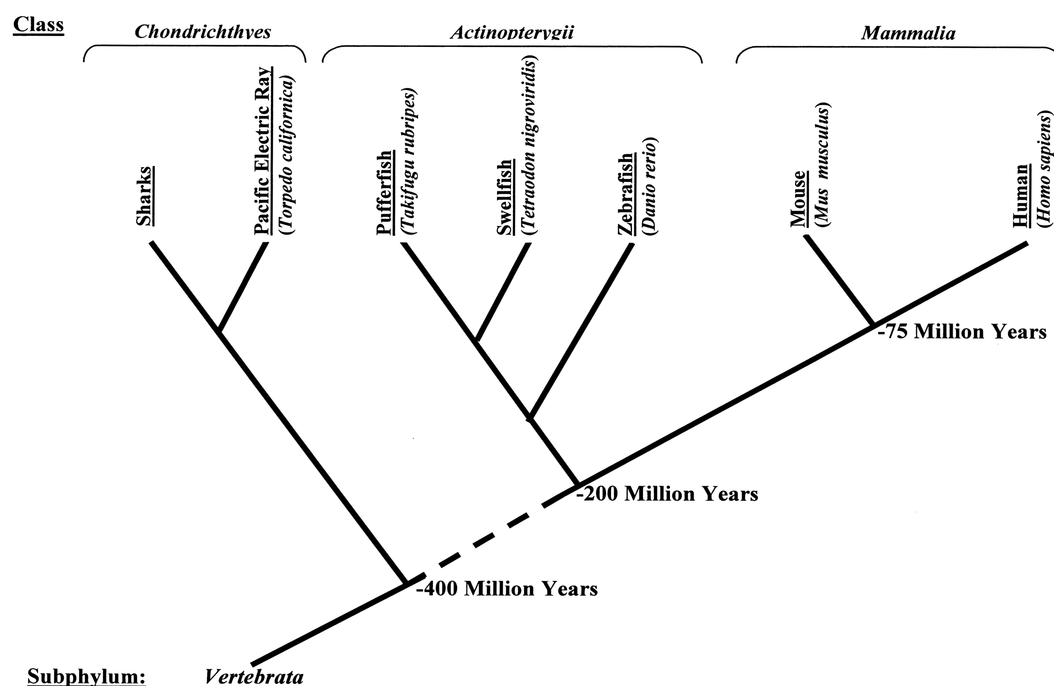

Figure 3. An evolutionary cladogram showing the relationship of *Torpedo* to other vertebrates. The common ancestor of shown species is subphylum *Vertebrata*. *T. californica* like sharks and other rays are more ancient species that fall under the *Chondrichthyes* class (cartillagenous fish which lack "true bones"). Zebrafish, puffer fish, and swell fish belong to the *Actinopterygii* (ray-finned fish) class. Human and mice belong to class *Mammalia* and are more closely related to *Actinoptergii* than *Chondrichtyes*.

© 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

4.2 Insights into the structure and function of the electroplaque and NMJ

The large electric organ of *Torpedo* is a traditional model for studying the small (30 μm) NMJ. The syncytial multinucleated myofiber contains nuclear domains within which gene expression and protein production is specialized and restricted by local nuclei, an example being the almost exclusive expression of AChRs by the nuclear domains directly beneath the NMJ [25–28]. This specialization restricts 95% of muscle AChR expression to the 0.1% of the muscle fiber surface that constitutes the NMJ [29]. Dissection and analyses of nuclear domains is hindered by their small size. The electric organ, however, is a large (~1 kg) modified muscle that has the basic structure of an electric synapse (Na^{2+}/K^{+} exchange across the membrane due to ACh binding) and has served as a primary source for identification of NMJ proteins [7, 12–15].

In the proteomic and cDNA effort described here, we identified 103 proteins and 190 cDNAs of proteins that are components of the electric organ, and potential components of the NMJ of other animals. Consistent with the validity of the electric organ as a model for NMJs, we identified eight additional known NMJ proteins in the *Torpedo* electric organ (Table 2A).

Fifteen proteins were identified by two or more of the three methods used (2-D MS/MS of soluble fractions; LC/MS of membrane fractions; cDNA sequencing) (Table 2B), and were considered potential novel components of the NMJ. Many of these were well-characterized stress- and energyrelated proteins, consistent with both the high-energy demands and electricity-induced stress that would be expected of the electric organ. We hypothesize that these same proteins are abundant at the NMJ. Plectin was identified by one cDNA and 39 peptides in soluble fractions. Plectin is a 500 kDa cytoskeletal protein that crosslinks the intermediate filament network, providing mechanical strength to cells. The very large size of plectin explains the many peptides identified.

A novel finding in our study is the reporting of the cDNA sequence and matching proteomic data for Annexin V (seven peptides; two cDNAs) (Table 2B). Although the identified *Torpedo* sequence is highly homologous to its human ortholog (59%), the monoisotopic masses of some of its tryptic digested peptides greatly differ (Fig. 2, panel C), hindering its identification by cross-species comparison. Annexins in mouse and human include seven related genes that have been shown to perform a variety of functions [30, 31]. The entire annexin class of proteins was originally identified through purification and functional characterization of "calelectrin", an abundant 35 kDa membrane protein from *T. marmorata* electric organ [32]. Calelectrin was found to be able to aggregate vesicles in a calcium-dependent manner [33], and the aggregates formed calcium-selective channels in lipid bilayers. These data suggested that calelectrin was a voltage-sensitive ion channel; however, functional channels comprised of annexins have never been demonstrated *in vivo* [30]. The only amino acid or cDNA sequence data for *Torpedo* calelectrin was a 33 amino acid peptide, and this single peptide was used to show homology to the annexin family, and more specifically assign calelectrin as most highly related to human and mouse Annexin VI [34]. The extensive cDNA and peptide data presented in our current study show that "calelectrin" is closely related to Annexin V, and not Annexin VI (Fig. 4). Our study confirms earlier findings that Annexin V/calelectrin is a very abundant component of the electroplaque membrane, and further suggests that it may serve a role in NMJs.

We also identified a series of novel proteins and mRNA sequences that had not been previously seen in either the electric organ or NMJs. Arginine/serine-rich 5 (SFRS5) is a splicing factor for which we identified a cDNA clone. Little is known about the SFRS5 protein, other than that it is highly related to proteins known to be critical for 5′ recognition of splice sites and tissue remodeling (SFRS1). The NMJ is perhaps the most dramatic model for "nuclear domains", where specific nuclei in a multinucleated syncytial cell (myofiber) carry out distinct transcriptional programs. The myonuclei underlying the NMJ clearly have a distinct transcriptional profile; however, little is known regarding the molecular steps involved in the initiation and propagation of this. Neuregulin, for example, is a synaptic protein whose expression is important for the maintenance of a high density of AChRs at the NMJ [35]. Alternative splicing of neuregulin, specifically at the NMJ, plays an important role in its domain-specific expression as well as AChR clustering [36]. The high concentration of SFRS5 in the electric organ suggests that this uncharacterized RNA regulatory protein may have a role in this process.

Finally, we identified ten mRNA sequences (cDNAs) from *Torpedo* electric organ cDNA library that can be considered as candidate components of the NMJ (Table 3), with amino acid homology ranging from 76 to 97%. Five of these were ORFs with no evidence for function assigned to the human or mouse protein. Interestingly, three of the ten mapped to human proteins containing motifs involved in RNA metabolism (SMD3, PRP38, FLJ37697) (Table 3). One cDNA (TFI_005_H4.T3) was found to contain a pre-mRNA processing factor 38 (prp38) domain. PRP38, a small 28 kDa acidic protein initially identified in yeast, is involved in pre-mRNA splicing and maintenance [37]. We blasted our translated cDNA against the human genome and found it to be 96% homologous to the poorly characterized human PRPF38A cDNA (chr1: 52,643,936–52,644,106).

Clone (TFI_004_D8.T3) was found to have 97% homology at the mRNA level to the uncharacterized human SMD3 (accession number P62318). SMD3 function has been inferred through the presence of SM and SM-like (SML) protein domains. SM domain-containing proteins interact with RNA to form the core domain of the ribonucleoprotein particles involved in an array of RNA processes

© 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

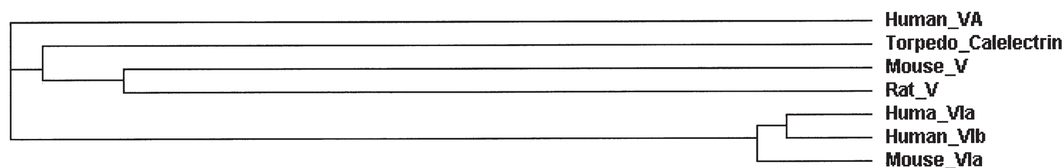

Figure 4. A cDNA homology cladogram showing that *Torpedo* calelectrin (annexin) is most closely related to human, mice, and rat annexin V. The ClustalW program was used to compare the nucleotide sequences of *Torpedo* annexin (calelectrin) to that of human, mouse, and rat annexin V and VI. *Torpedo* annexin is indeed annexin V; yet, being an evolutionarily ancient form of annexin, it is more distant to human than to rodent.

such as pre-mRNA splicing, telomere replication, and mRNA degradation [38]. Finally, clone (TFI_1_E3.T3) showed 94% sequence homology to human hypothetical protein (FLJ37697) on chromosome 6 (chr6: 17,389,792-17,391,140). The predicted protein encodes a full RNA recognition motif (RRM/RRM_1).

Another interesting ORF (clone TFII_1_A9.T3) codes for an 81 amino acid peptide that is 77% similar to an uncharacterized human protein named Stretch Regulated Skeletal Muscle Protein. The murine homolog of this protein is known as small muscle protein X chromosome (Smpx). Smpx is primarily expressed in cardiac and skeletal muscle and plays a possible role in skeletal muscle hypertrophy [39]. Exercise is known to increase the size of the NMJs and induce structural remodeling of the NMJ [40, 41]. Therefore, expression of Smpx could be a direct result of synaptic remodeling and architectural restructuring.

Overall, the survey of the proteome and transcriptome of *T. californica*, for which a genomic database does not currently exist, resulted in the identification of a total of 103 unique proteins and 190 cDNAs. While this study offers a list of potential molecular players of the NMJ, future studies should focus on validation and localization of *Torpedo* electric organ orthologs within the NMJ, and structural and functional characterization of these proteins.

This work was supported in part by the WM Keck Foundation and the Erynn Godla Family via the Juvenile ALS Foundation (www.jals.org). We would like to thank Dr. Zohra Olumee for her help and assistance in protein fractionation.

5 References

[1] Whittaker, V. P., Essman, W. B., Dowe, G. H., *Biochem. J.* 1972, *128*, 833–845.

[2] Kawashima, T. I. M., Sasaki, H., *Anat. Histol. Embryol.* 2004, *33*, 294–298.

[3] Sheridan, M. N., *J. Cell Biol.* 1965, *24*, 129–141.

[4] Koester, D. M., *Anat. Rec. A Discov. Mol. Cell Evol. Biol.* 2003, *273*, 648–662.

[5] Miller, C., *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 1982, *299*, 401–411.

[6] Sealock, R., *J. Neurosci.* 1982, *2*, 918–923.

[7] Nitkin, R. M., Smith, M. A., Magill, C., Fallon, J. R. *et al., J. Cell Biol.* 1987, *105*, 2471–2478.

[8] Carr, C., Fischbach, G. D., Cohen, J. B., *J. Cell Biol.* 1989, *109*, 1753–1764.

[9] Kistler, J., Stroud, R. M., *Proc. Natl. Acad. Sci. USA* 1981, *78*, 3678–3682.

[10] Kistler, J., Stroud, R. M., Klymkowsky, M. W., Lalancette, R. A. *et al., Biophys. J.* 1982, *37*, 371–383.

[11] Mitra, A. K., McCarthy, M. P., Stroud, R. M., *J. Cell Biol.* 1989, *109*, 755–774.

[12] Cohen, I., Rimer, M., Lomo, T., McMahan, U. J., *Mol. Cell. Neurosci.* 1997, *9*, 237–253.

[13] Gramolini, A. O., Burton, E. A., Tinsley, J. M., Ferns, M. J. *et al., J. Biol. Chem.* 1998, *273*, 736–743.

[14] Jones, G., Herczeg, A., Ruegg, M. A., Lichtsteiner, M. *et al., Proc. Natl. Acad. Sci. USA* 1996, *93*, 5985–5990.

[15] McMahan, U. J., *Cold Spring Harb. Symp. Quant. Biol.* 1990, *55*, 407–418.

[16] Nazarian, J., Bouri, K., Hoffman, E. P., *Physiol. Genomics* 2005, *21*, 70–80.

[17] Apel, E. D., Glass, D. J., Moscoso, L. M., Yancopoulos, G. D. *et al., Neuron* 1997, *18*, 623–635.

[18] Zhou, H., Glass, D. J., Yancopoulos, G. D., Sanes, J. R., *J. Cell Biol.* 1999, *146*, 1133–1146.

[19] Glass, D. J., Bowen, D. C., Stitt, T. N., Radziejewski, C. *et al., Cell* 1996, *85*, 513–523.

[20] Strochlic, L., Cartaud, A., Labas, V., Hoch, W. *et al., J. Cell Biol.* 2001, *153*, 1127–1132.

[21] Lukas, R. J., Tubbs, K. A., Krivoshein, A. V., Bieber, A. L. *et al., Anal. Biochem.* 2002, *301*, 175–188.

[22] Watty, A., Neubauer, G., Dreger, M., Zimmer, M. *et al., Proc. Natl. Acad. Sci. USA* 2000, *97*, 4585–4590.

[23] Hall, Z. W., *J. Neurobiol.* 1973, *4*, 343–361.

[24] Jensen, O. N., Wilm, M., Shevchenko, A., Mann, M., *Meth. Mol. Biol.* 1999, *112*, 513–530.

[25] Allen, D. L., Roy, R. R., Edgerton, V. R., *Muscle Nerve* 1999, *22*, 1350–1360.

[26] Apel, E. D., Lewis, R. M., Grady, R. M., Sanes, J. R., *J. Biol. Chem.* 2000, *275*, 31986–31995.

[27] Rossi, S. G., Vazquez, A. E., Rotundo, R. L., Allen, D. L. *et al., J. Neurosci.* 2000, *20*, 919–928.

[28] Fontaine, B., Sassoon, D., Buckingham, M., Changeux, J. P., *EMBO J.* 1988, *7*, 603–609.

[29] Tsim, K. W., Barnard, E. A., *Neurosignals* 2002, *11*, 58–64.

© 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

[30] Demange, P., Voges, D., Benz, J., Liemann, S. *et al.*, *Trends Biochem. Sci.* 1994, *19*, 272–276.

[31] Monastyrskaya, K., Babiychuk, E. B., Hostettler, A., Rescher, U., Draeger, A. *Cell Calcium* 2006 (In Press), doi: 10.1016/j.ceca.2006.06.008.

[32] Walker, J. H., *J. Neurochem.* 1982, *39*, 815–823.

[33] Sudhof, T. C., Walker, J. H., Obrocki, J., *EMBO J.* 1982, *1*, 1167–1170.

[34] Rodriguez-Garcia, M. I., Kozak, C. A., Morgan, R. O., Fernandez, M. P., *Genomics* 1996, *31*, 151–157.

[35] Sandrock, A. W., Jr., Dryer, S. E., Rosen, K. M., Gozani, S. N. *et al.*, *Science* 1997, *276*, 599–603.

[36] Ponomareva, O. N., Ma, H., Dakour, R., Raabe, T. D. *et al.*, *Neuroscience* 2005, *134*, 495–503.

[37] Blanton, S., Srinivasan, A., Rymond, B. C., *Mol. Cell Biol.* 1992, *12*, 3939–3947.

[38] Malatesta, M., Fakan, S., Fischer, U., *Exp. Cell Res.* 1999, *249*, 189–198.

[39] Kemp, T. J., Sadusky, T. J., Simon, M., Brown, R. *et al.*, *Genomics* 2001, *72*, 260–271.

[40] Deschenes, M. R., Judelson, D. A., Kraemer, W. J., Meskaitis, V. J. *et al.*, *Muscle Nerve* 2000, *23*, 1576–1581.

[41] Somasekhar, T., Nordlander, R. H., Reiser, P. J., *J. Neurocytol.* 1996, *25*, 315–331.

© 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim