ISSN 0006-2979, Biochemistry (Moscow), 2009, Vol. 74, No. 11, pp. 1239-1252. © Pleiades Publishing, Ltd., 2009.
Original Russian Text © M. A. Kovalyova, L. I. Kovalyov, I. Yu. Toropygin, S. V. Shigeev, A. V. Ivanov, S. S. Shishkin, 2009, published in Biokhimiya, 2009, Vol. 74, No. 11, pp. 1524-1538.

Proteomic Analysis of Human Skeletal Muscle (*m. vastus lateralis*) Proteins: Identification of 89 Gene Expression Products

M. A. Kovalyova[1], L. I. Kovalyov[1]*, I. Yu. Toropygin[2], S. V. Shigeev[3], A. V. Ivanov[1], and S. S. Shishkin[1]

[1]*Bach Institute of Biochemistry, Russian Academy of Sciences, Leninsky pr. 33, 119071 Moscow, Russia; fax: (495) 954-2733; E-mail: kovalyov@inbi.ras.ru*
[2]*Orekhovich Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Pogodinskaya ul. 10, 119121 Moscow, Russia; fax: (495) 245-0857*
[3]*People's Friendship University of Russia, ul. Miklukho-Maklaya 8, 117198 Moscow, Russia; fax: (495) 433-2794*

Received November 17, 2008
Revision received March 31, 2009

Abstract—Proteins from bioptates and autoptates of human skeletal muscle *m. vastus lateralis* were separated by O'Farrell two-dimensional gel electrophoresis (2DE). MALDI-TOF MS and MS/MS enabled identification of 89 protein spots as expression products of 55 genes. A modification of the O'Farrell's method including non-equilibrium electrophoresis in a pH gradient allowed detection – among major sarcomeric, mitochondrial, and cytosolic proteins – of several proteins, such as PDZ- and LIM domain-containing ones (pI > 8.70), fragments of known proteins, and a stable complex of heavy and light ferritin chains. The data underlie further studies of human skeletal muscle proteins in terms of molecular mechanisms of some physiological and pathological processes.

DOI: 10.1134/S0006297909110108

Key words: proteins of human skeletal muscle, proteomics

In contemporary proteomic investigations of tissues and biological fluids mainly associated with systemic analysis of blood plasma, liver, and brain proteins [1], the study on proteins from various muscular organs and tissues [2, 3] takes a distinct place. The muscular tissues and organs as objects for proteomic studies attract attention of investigators for a number of reasons. First, certain alterations of gene expression occurring in the course of development and functioning of muscular tissues and organs should be reflected in so-called protein profiles that proteomic analysis can precisely characterize [2, 4, 5]. Second, a significant number of muscular tissue diseases is known to date, in which proteomic analysis opens the way to elucidation of their pathogenesis and also favors the determination of diagnostically important molecular markers [3, 4, 6]. Third, because of differences in muscle cell differentiation and peculiarities of muscular organ morphology, proteomic technologies are used for the search for tissue-specific proteins, among which new protein isoforms or proteins can be found, which represents a significant item for study and seems to be important for realization of human proteome programs [7].

Data of protein studies in bioptates of the human skeletal muscle *m. vastus lateralis* are presented in this study, which continues the cycle of Russian national proteomic studies on human skeletal muscle proteins [2, 8-10] and broadens the earlier created database on these proteins.

MATERIALS AND METHODS

Preparation of biopsy samples of *m. vastus lateralis*. Samples (n = 23) for this proteomic study were kindly provided from the Institute of Medical and Biological Problems of the Russian Academy of Sciences (IMBP

Abbreviations: a.a., amino acid residues; FCC, ferritin chain complex; 2DE, O'Farrell two-dimensional electrophoresis; IEF, isoelectric focusing; MALDI-TOF MS, matrix-assisted laser desorption/ionization–time of flight mass spectrometry.

* To whom correspondence should be addressed.

RAS) within the framework of this joint study. The samples were obtained by needle biopsy described in detail earlier [11] from healthy volunteers of 20-24-year-old who gave informed consent. The samples were frozen immediately after the sampling and stored in liquid nitrogen. To prepare protein extract, a sample taken after biopsy (5 mg) was homogenized for 3-5 min in 200 μl of 9 M urea solution containing 5% β-mercaptoethanol, 2% Triton X-100, and 2% ampholytes, pH 3.5-10 (Sigma, USA). The homogenate was clarified by centrifugation followed by immediate fractionation of 100 μl of the extract by O'Farrell two-dimensional electrophoresis (2DE).

Preparation of samples taken after autopsy. Samples of various muscular organs ($n = 34$ – *m. vastus lateralis*, $n = 90$ – myocardium, and $n = 5$ – uterine myometrium) from persons who died in accidents, as well as brain cortex samples ($n = 5$) were obtained from the Bureau of Forensic Medicine of the Department of Public Health, Moscow. All except three samples were characterized by minimum autolysis duration, no more than 12 h. The postmortem interval was about one day for three samples. The postmortem samples were homogenized and extracts prepared by a method described earlier [10].

Proteins from prostate samples and human cell culture obtained as described earlier [12, 13] were used for comparative analysis.

Two-dimensional electrophoresis by O'Farrell method. The first-direction fractioning was isoelectrofocusing (IEF) in glass tubes (2.4 × 180 mm, instead of formerly used 3.5 × 120 mm) filled with 4% polyacrylamide gel prepared in 9 M urea solution containing 2% Triton X-100 and 2% ampholyte mixture. Ampholytes of 5-7, 5-8, and 3.5-10 pH ranges taken at 4 : 1 ratio were used in all the main experiments; for elucidation of precise distribution of proteins in marginal zones, ampholytes 4-6 or 7-9 pH ranges instead of 5-7 pH range were taken at the same ratio in separate experiments. The protein extracts (100-150 μl) were applied on the "acidic" edge of the gel, and IEF was carried out using a BioRad (USA) Model 175 electrophoretic cell until 2400 V/h was achieved. Then the polyacrylamide gel columns with the proteins separated by IEF were applied as a start zone for separation in the second direction, for which slab electrophoresis in polyacrylamide gel (200 × 200 × 1 mm) was used with a linear 7.5-20% gradient of acrylamide concentration in the presence of 0.1% SDS using a Helicon (Russia) vertical electrophoretic cell. A well was formed for protein marker application at the edge of each gel slab. Other details of the used 2DE modification were described earlier [10].

For protein visualization, the polyacrylamide gel slabs were stained with Coomassie Blue R-250 and then with silver nitrate according to the methods described in [14] modified by the addition of 0.8% acetic acid to sodium thiosulfate. The stained gels were documented by scanning on an Epson Expression 1680 scanner, and densitometry was carried out using the Melanie software (GeneBio, Switzerland) according to the producer's protocol.

Molecular masses (M) of the fractionated proteins were determined by their electrophoretic mobility in the second direction as compared with protein markers from standard heart muscle lysate, as well as with several previously identified reference proteins, as described earlier [8-10]. The results of the determinations were verified by a calibration curve created using a set of highly purified recombinant proteins with M ranging within 10-200 kDa (Fermentas, Lithuania) [15]. Isoelectric points (pI) of fractionated proteins were determined from their electrophoretic location in the second direction, as described earlier [8-10], taking into account the known localization of identified reference proteins. Theoretical values of M and pI were calculated using the "pI/Mw" software in open access at the ExPASy Proteomics Server (http://cn.expasy.org) and data on amino acid sequences of corresponding proteins in Swiss-Prot database taking into account the evidence for post-translational removal of signal sequences.

Protein identification by mass spectrometry. Isolation of protein fractions from polyacrylamide gel slabs, hydrolysis with trypsin, and peptide extraction for protein identification by matrix assisted laser desorption/ionization time-of-flight mass-spectrometry (MALDI-TOF MS) were carried out according with published protocols [16] with some modifications [17]. A sample (0.5 μl) was mixed on the target with an equal volume of 20% acetonitrile containing 0.1% trifluoroacetic acid and 20 mg/ml of 2,5-dihydroxybenzoic acid (Sigma) and air-dried. Mass spectra were recorded on a Reflex III MALDI-TOF mass spectrometer (Bruker, USA) equipped with a UV-laser (336 nm) in the mode of positively charged ions with masses ranging within 500-8000 Da. When MS/MS analysis was carried out, the mass spectra of fragments were recorded in tandem mode of the mass spectrometer for detection of positively charged ions. Ion fragmentation was induced by injection of helium into the area of the track launch (initial region) of the charged ion free drift (at the pressure of the inert gas of $2 \cdot 10^{-7}$ Pa). The error of fragment mass determinations was below 0.05%. Only signals from C-terminal fragments split at a peptide bond (y-ions) were represented on the mass spectrum. Proteins were identified using Mascot software, Peptide Fingerprint option (Matrix Science, USA).

RESULTS AND DISCUSSION

One substantial difficulty experienced in proteomics studies of human cells and tissues is a considerable difference in representation of distinct proteins in the material under study. This difference can exceed seven orders of

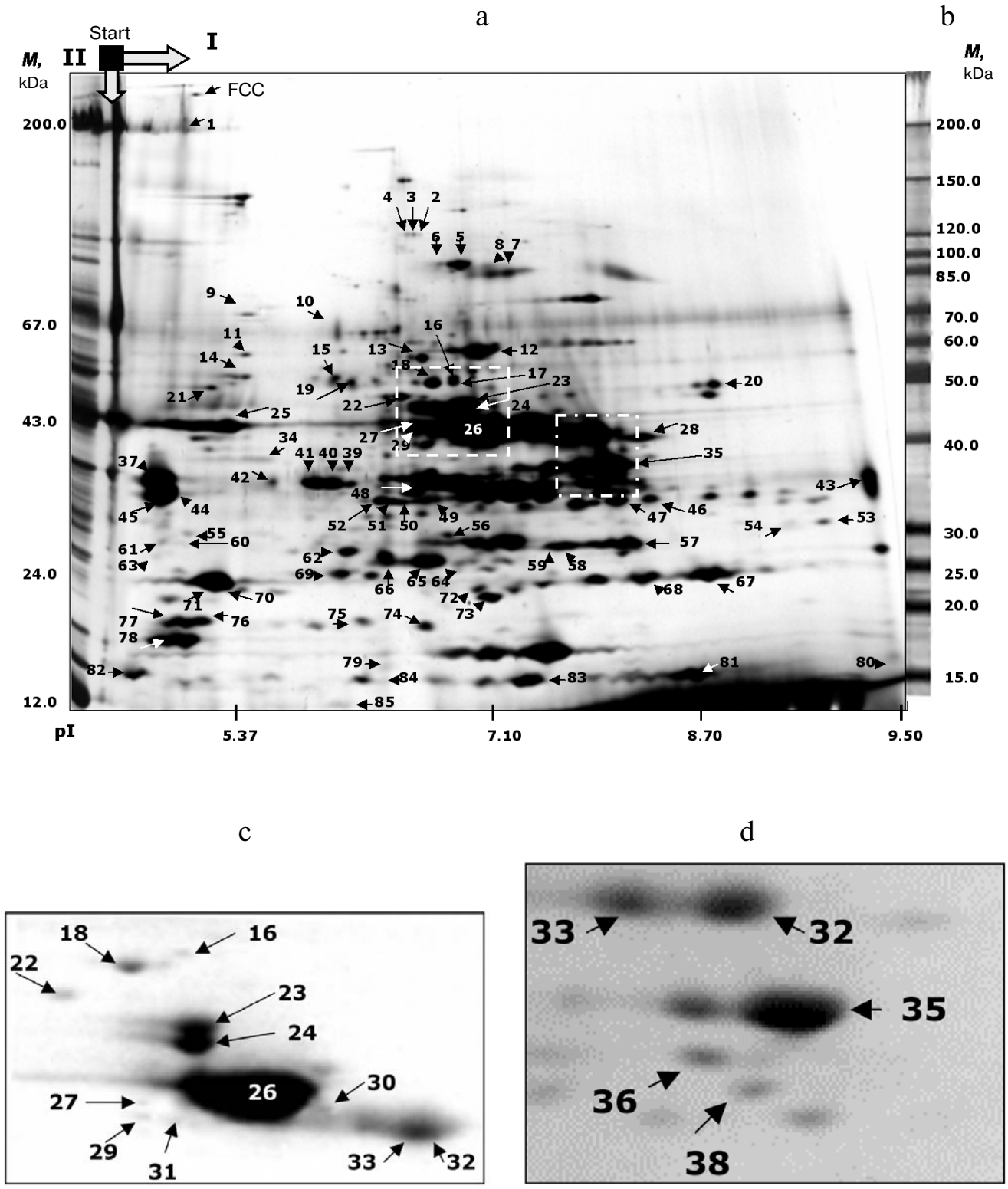


Fig. 1. Electrophoretic fractionation of proteins from bioptate of human *m. vastus lateralis*. Proteins identified by mass spectrometry are designated by arrows. a) 2DE pattern (70 μg protein), staining with silver nitrate; b) distribution of molecular mass markers (set of highly purified recombinant proteins with *M* ranging within 10-200 kDa (Fermentas)); c, d) fragments of 2DE gel (at low protein load) with centered fractions 26 and 35 identified as creatine phosphokinase and glyceraldehyde-3-phosphate dehydrogenase (see Table 1), respectively. FCC, ferritin chain complex.

Table 1. Major proteins and their isoforms including distinct electrophoretic variants (EV) found in human *m. vastus lateralis* samples taken at biopsy and identified in 2D electrophoretic gels by mass spectrometry

Fraction	**Proteins** [references for earlier identified ones]*	Experimental *M*/p*I* ratios	Coverage, %**	Accession numbers in NCBI and Swiss-Prot databases (gene name)***	Theoretical *M*/p*I* ratios (after removal of leader peptides)
1	2	3	4	5	6
1	**myosin heavy chain, isoform β**	200.0/5.00	15	115496169, NP_000248, 160760 (*MYH7*)	223.0/5.63
5	**aconitase 2**, **mitochondrial** [10]	89.0/7.04	48	20072188, AAH26196, 100850 (*ACO2*)	84.1/7.20
6	aconitase 2, mitochondrial (EV)	89.0/7.03	28	20072188, AAH26196, 100850 (*ACO2*)	84.1/7.20
7	**LIM domain-binding protein 3, isoform 1**	79.5/7.20	17	45592959, NP_009009, 605906, O75112-5 (*LDB3*)	77.1/8.47
8	LIM domain-binding protein 3, isoform 1 (EV)	79.5/7.12	17	45592959, NP_009009, 605906, O75112-5 (*LDB3*)	77.1/8.47
12	**pyruvate kinase**, **muscle isoform M1**, **spliced**	57.6/7.10	31	33286422, NP_872271, 179050, P14618-2 (*PKM2*)	57.9/7.59
14	**desmin** [34]	52.0/5.47	74	55749932, NP_001918, 125660, P17661 (*DES*)	53.4/5.21
22	**α-enolase** (enolase 1)	47.3/6.58	86	4503571, NP_001419, P06733 (*ENO1*)	47.1/7.0
23	**β-enolase (enolase 3**, **muscle-specific)**, **"heavy" isoform** [19]	45.2/7.04	37	16878083, AAH17249, 131370, P13929 (*ENO3*)	47.0/7.58
24	β-enolase (enolase 3, muscle-specific), "light" isoform [19]	44.7/7.04	62	16878083, AAH17249, 131370, P13929 (*ENO3*)	47.0/7.58
25	**α1-actin** (skeletal muscle) [10]	43.1/5.37	42	4501881, NP_001091, 102610, P68133 (*ACTA1*)	41.9/5.23
26	**creatine kinase M-type** (muscle type subunit) [10]	43.0/7.08	26	30582425, AAP35439, 123310, P06732 (*CKM*)	43.1/6.77
32	**aldolase A**, **muscle type**	39.7/8.30	46	49168540, CAG38765, 103850, P04075 (*ALDOA*)	39.3/8.39
33	aldolase A (EV)	39.7/8.00	48	49168540, CAG38765, 103850, P04075 (*ALDOA*)	39.3/8.39
35	**glyceraldehyde-3-phosphate dehydrogenase** [10]	37.0/8.35	37	31645, CAA25833, 138400, P04406 (*GAPDH*)	36.1/8.57
37	**tropomyosin 2** (β, isoform 1) [9]	36.2/4.85	34	42476296, NP_003280, 190990, P07951 (*TPM2*)	32.8/4.66
39	troponin T1, isoform 1, slow skeletal muscle (EV)	35.5/5.95	22	15305458, XP_048167, 191041, P13805 (*TNNT1*)	32.9/5.86 (isoform 1 by Swiss-Prot)
40	troponin T1, slow skeletal muscle (EV)	35.5/5.90	24	15305458, XP_048167, 191041, P13805 (*TNNT1*)	32.9/5.86 (isoform 1 by Swiss-Prot)
41	**troponin T1**, **slow skeletal muscle**	35.5/5.85	25	15305458, XP_048167, 191041, P13805 (*TNNT1*)	32.9/5.86 (isoform 1 by Swiss-Prot)

Table 1 (Contd.)

1	2	3	4	5	6
44	**tropomyosin 3** (α, isoform 3; γ-tropomyosin)	35.0/4.94	36	136085, CAH71264, 191030, P06753 (*TPM3*)	32.8/4.68
45	**tropomyosin 1** (α, isoform 1) [9]	34.7/4.92	64	63252898, NP_001018005, 191010, P09493 (*TPM1*)	32.7/4.69
47	**myozenin 1** (calsarcin 2)	34.0/8.50	62	10864053, NP_067068, 605603, Q9NP98 (*MYOZ1*)	31.7/8.86
49	troponin T1, isoform 2, slow skeletal muscle (EV)	33.0/6.85	28	15305458, XP_048167, 191041, P13805 (*TNNT1*)	32.9/5.86 (isoform 1 by Swiss-Prot)
50	troponin T1, isoform 2, slow skeletal muscle (EV)	33.0/6.75	23	15305458, XP_048167, 191041, P13805 (*TNNT1*)	32.9/5.86 (isoform 1 by Swiss-Prot)
51	troponin T1, isoform 2, slow skeletal muscle (EV)	33.0/6.55	24	187173288, NP_003274, 191041, P13805 (*TNNT1*)	32.9/5.86 (by NP_003274)
52	troponin T1, isoform 2, slow skeletal muscle (EV)	33.0/6.50	22	187173288, NP_003274, 191041, P13805 (*TNNT1*)	32.9/5.86 (by NP_003274)
57	**troponin T3**, **fast skeletal muscle**	28.5/8.40	33	49456809, CAG46725, P45378, (*TNNT3*)	29.7/8.40 (isoform 7)
58	**phosphoglycerate mutase 2**, **muscle**	28.8/7.50	39	17464807, XP_011580, 261670, P15259 (*PGAM2*)	28.6/9.00 (by P15259)
59	phosphoglycerate mutase 2, muscle (EV)	28.8/7.30	53	387016, AAA60073, 261670, P15259 (*PGAM2*)	28.5/8.77 (by AAA60073)
63	**tropomyosin 3, isoform 2** (cytoskeletal isoform TM30-NM)	26.0/4.80	18	24119203, NP_705935, 191030, P06753-2, (*TPM3*)	29.0/4.75 (by P06753-2)
64	**triosephosphate isomerase 1** [10]	25.0/6.80	54	4507645, NP_000356, 190540, P60174, (*TPI1*)	26.5/6.51
65	triosephosphate isomerase 1 (EV)	25.0/6.70	58	4507645, NP_000356, 190540, P60174, (*TPI1*)	26.5/6.51
67	**troponin I, isoform 1**, **fast-twitch skeletal muscle**	22.0/8.80	50	4507621, NP_003273, 191043, P48788 (*TNNI2*)	21.2/8.88
68	troponin I, isoform 1, fast-twitch skeletal muscle (EV)	22.0/8.48	56	4507621, NP_003273, 191043, P48788 (*TNNI2*)	21.2/8.88
69	**myosin light chain 1**, **slow-twitch muscle A; non-muscle myosin alkali light chain 6B**	22.0/6.15	54	4505303, NP_002466, 609930, P14649 (*MYL6B*)	22.7/5.56
70	**myosin alkali light chain 3, slow skeletal ventricular** [10]	21.9/5.14	49	4557777, NP_000249, 160790, P08590 (*MYL3*)	21.9/5.03
71	**myosin light chain 1**, **fast skeletal muscle isoform** [20]	21.9/5.05	66	127128, X05450, 160780, P05976 (*MYL1*)	21.0/4.97
73	**αB-crystallin** [10]	20.8/7.09	55	4503057, NP_001876, 123590, P02511 (*CRYAB*)	20.2/6.76

Table 1 (Contd.)

1	2	3	4	5	6
76	**myosin**, **light chain 2**, **slow**, **regulatory** [34]	18.5/5.00	94	21411329, AAH31006, 160781, P10916 (*MYL2*)	18.8/4.92
77	myosin, light chain 2, slow, regulatory (EV) [10]	18.5/4.95	41	21411329, AAH31006, 160781, P10916 (*MYL2*)	18.8/4.92
78	**myosin**, **light chain 2**, **fast**	18.0/4.97	60	28372499, NP_037424, Q96A32, (*MYLPF*)	18.8/4.91 (reported in 1981 and 2003)
82	**myosin light chain 1**, **fast skeletal**, **alkali isoform 3f** (produced by alternative splicing)	14.8/4.95	40	17986275, NP_524146, 160780, P06741 (MYL1)	16.6/4.62

* Major protein isoforms are bold-faced – here and in Tables 2 and 3.
** Overall length of identified tryptic peptides relative to that of protein amino acid sequence – here and in Tables 2 and 3.
*** Reference numbers are given in the order: NCBI (GenBank, Nucleotide or Protein, OMIM), Swiss-Prot – here and in Tables 2 and 3.

magnitude [1, 2, 18]. Taking this problem into account, several samples differing 10 times in amount from the same specimen were analyzed in parallel for effective solution by 2DE and subsequent protein identification by mass spectrometry. Besides, two subsequent staining steps were used for protein detection: the first (with Coomassie Blue R-250) has relatively low sensitivity, whereas the second (with silver nitrate) provides two orders of magnitude higher sensitivity.

The typical pattern of electrophoretic separation of proteins from human *m. vastus lateralis* bioptate at medium protein load (70 μg) is shown in Fig. 1a. The protein fractions analyzed by mass spectrometry are shown in Fig. 1 as consecutive numbers from higher *M* and p*I* values. The *M* and p*I* values given in Tables 1-3 are the arithmetic means of at least four samples at mean deviation below ±3% and ±2% from magnitudes of *M* and p*I*, respectively. The data of fractionation of the recombinant protein set used as *M* markers are shown in Fig. 1b for comparison. Since the gel at medium protein load proved to contain some overloaded regions scarcely analyzed (hatched and dot-and-dash lined rectangles in Fig. 1a), the corresponding regions on the gel are given additionally (Fig. 1, c and d) at low protein load.

The results of identification of several "major" protein fractions, among which were myosin light and heavy chain isoforms, actin, tropomyosins, desmin, muscle creatine phosphokinase, enolase isoforms, and some other known muscle proteins, particularly some enzymes involved in glycolysis, are summarized in Table 1. Reference numbers of corresponding records in several NCBI databases and Swiss-Prot are assigned to each identified fraction in Table 1 (as well as in Tables 2 and 3) taking into account the regular changes in numbers after editions, with especially frequent changes in GenBank.

The data of analysis by MALDI-TOF (Fig. 2a) and tandem mass spectrometry (Fig. 2b) of fraction 22 that we characterized as α-enolase, the *ENO1* gene product, is given in Table 4 and in Fig. 2 as an example of identification. Two neighboring fractions, 23 and 24 (Fig. 1c), proved to be enolase isoenzymes as well, but they are allelic variants of β-enolase, the *ENO3* gene product. They differ not only in electrophoretic mobility but also in the N71S amino acid substitution [19]. The data suggest that biochemical polymorphism of enolase in human skeletal muscles results from both polygeny and polyallelism.

The data suggest typicality of biochemical polymorphism for many identified proteins, which manifests as several fractions with similar electrophoretic and structural properties. The causes are known of the observed polymorphism of some proteins (polygeny, polyallelism, and alternative splicing), which is reflected in corresponding records in NCBI and Swiss-Prot databases (Table 1). Allelic polymorphism of certain human muscle proteins revealed by 2DE and other proteomic techniques was described in our several previous publications [2, 20, 21].

Common features of protein distribution patterns on 2DE gels coincided with the data obtained in earlier studies of postmortem human striped muscles, including identification of distinct muscle proteins by immunoblotting and other methods [8-10]. The properties of the identified polymorphic proteins are consistent with the data of other studies. For instance, a comparison of myosin light chain isoform pattern (69-71, 76-78, 82; Fig. 1a) with the data given in [22] demonstrated coincidence of the main features of fraction distributions, suggesting the presence of both slow and fast fiber types in the muscle samples taken at biopsy. The polymorphism of many human myosin light chain isoforms due to polyge-

Table 2. Proteins characterized by high p*I* values and their electrophoretic variants (EV) found in human *m. vastus lateralis* samples taken at biopsy and identified in 2D electrophoretic gels by mass spectrometry

Fraction	Proteins and their EV (aliases)	Experimental *M*/p*I* ratios	Coverage, %	Accession numbers in NCBI and Swiss-Prot databases (gene name)	Theoretical *M*/p*I* ratios (after removal of leader peptides)
20	**mitochondrial trifunctional protein, β-subunit** (hydroxyacyl dehydrogenase subunit B)	49.0/9.00	13	4504327, NP_000174, 143450, P55084 (*HADHB*)	47.6/9.30
28	**glutamic-oxaloacetic transaminase 2, mitochondrial (**aspartate aminotransferase 2**)**	41.5/8.70	44	73486658, NP_002071, 138150, P00505 (*GOT2*)	44.6/8.98
43	**skeletal muscle LIM-protein FHL1**	35.4/9.50	17	2833246, U60115, 300163, Q13642 (*FHL1*, *SLIM1*)	36.1/9.25 (isoform 2)
53	**PDZ and LIM domain protein 5, cardiac, skeletal muscle**	31.0/9.30	59 (by 17389 778)	по 17389778, AAH17902 (*N*-terminal sequence encoded by known gene), 605904, Q96HC4 (*ENH*, *PDLIM5*)	by 17389778 - 23.8/9.93 (possibly produced by alternative splicing; the complete protein contains 595 a.a)
54	PDZ and LIM domain protein 5, cardiac, skeletal muscle (EV)	30.5/9.20	74 (by 17389 778)	the same	the same
67	**troponin I isoform 1, fast-twitch skeletal muscle**	22.0/8.80	50	4507621, NP_003273, 191043, P48788 (*TNNI2*)	21.2/8.88
81	α-globin	15.0/8.70	42	56967331, 1Y0D_A, 141800, P69905 (*HBA1*)	15.2/8.72
80	**cytochrome *c***	15.0/9.50	35	15929398, AAH15130, 123970, P99999 (*CYCS*)	11.8/9.63

ny (and of fast type myosin 1 light chain due to alternative splicing) was described already in the 1980s [23]; however, the *MYLPF* gene encoding the fast type myosin 2 light chain (regulatory) was identified only recently [24].

However, in many cases, when sets of several fractions with similar electrophoretic and structural properties were revealed (for instance, 2-4, 5-6, 7-8, 32-33, 39-41, and 49-52; Fig. 1 and Table 1), posttranslational modifications and/or other mechanisms can only be suggested as a cause of biochemical polymorphism of corresponding muscle proteins. The most abundant fraction was taken as the main form in each set, and other identified isoforms were regarded as electrophoretic variants (EV). For example, in Fig. 1 fraction 5 is assigned to aconitase main isoform and fraction 6 is its electrophoretic variant. The *M* values of the main isoforms of identified proteins obtained from their electrophoretic mobility were identical or very similar to the calculated ones in virtually all cases, and some underestimated experimental p*I* values of several proteins are apparently due to the fact that these proteins, although successfully fractionated in the pH gradient under the conditions of the modified 2DE protocol, did not achieve positions corresponding to their p*I* values (Tables 1-3).

Among the latter, fractions 7 and 8 are of particular interest: they were identified as EVs of isoform 1 of LIM domain-binding protein 3 (or ZASP Cypher). The proteins of this family, which are implicated in formation of Z-line structures and revealed during genomic studies [25], could only be detected earlier in proteomic studies employing chromatography methods for protein fractionation [26].

Figure 1 shows that, unlike previous studies [8-10], the modification of 2DE used in this work enables resolu-

Table 3. Proteins from various compartments and their electrophoretic variants found in human *m. vastus lateralis* samples taken at biopsy and identified in 2D electrophoretic gels by mass spectrometry

Fraction	Proteins [references for earlier identified ones]	Experimental *M*/p*I* ratios	Coverage, %	Accession numbers in NCBI and Swiss-Prot databases (gene name)	Theoretical *M*/p*I* ratios (after removal of leader peptides)
1	2	3	4	5	6
			Mitochondrial proteins		
11	**chaperonin**, **mitochondrial** (60-kDa heat shock protein isoform 1, GroEL)	58.5/5.51	29	41399285, NP_955472, 118190, P10809 (*HSPD1*)	57.9/5.24
13	**dihydrolipoamide dehydrogenase**, **mitochondrial**	55.0/6.80	32	51095146, EAL24389, 238331, P09622 (*DLD*)	50.1/6.35
15	**3-oxoacid CoA transferase 1**, **mitochondrial** (succinyl-CoA:3-ketoacid-coenzyme A transferase)	51.8/5.95	21	4557817, NP_000427, 601424, P55809 (*OXCT1*)	52.0/6.00
17	**ATP synthase**, **mitochondrial**, **α-subunit**, **isoform 1** [10]	51.6/7.00	42	30583257, AAP35873, 164360, P25705 (*ATP5A1*)	55.2/8.28 (by P25705)
18	ATP synthase, mitochondrial, α-subunit, isoform 1 (EV)	51.6/6.82	28	15030240, AAH11384,164360, P25705 (*ATP5A1*)	55.2/7.81 (by AAH11384)
19	3-oxoacid CoA transferase (EV)	51.5/6.19	26	4557817, NP_000427, 601424, P55809 (*OXCT1*)	52.0/6.00
21	**ATP synthase**, **mitochondrial**, **β-subunit**	49.0/5.15	64	32189394, NP_001677, 102910 (*ATP5B*)	51.7/5.00
27	**citrate synthase**, **mitochondrial (isoform short**, **CS protein)**	42.5/6.56	45	48257138, AAH00105, 118950, O75390 (*CS*)	45.6/6.49
30	**creatine kinase**, **mitochondrial 2 (sarcomeric)**	41.0/7.25	46	20810521, AAH29140, 123295, P17540 (*CKMT2*)	43.3/7.28
31	**3-hydroxyisobutyryl-CoA hydrolase**, **mitochondrial**	40.0/6.95	25	146324905, NP_055177, 610690 Q6NVY1 (*HIBCH*)	39.5/6.30
38	**voltage-dependent anion channel protein 1** (porin isoform 1) [34]	35.0/8.30	77	6063691, CAB58127, 604492, P21796 (*VDAC1*)	30.7/8.63
72	**superoxide dismutase**, **mitochondrial** [10]	21.8/7.08	42	134665, X59445, 147460 P04179, (*SOD2*)	22.1/6.58
85	**cytochrome *c* oxidase subunit VIb**	11.0/6.20	46	30582729, AAP35591, 124089, P14854, (*COX6B1*)	10.1/6.79
			Cytosolic proteins		
16	**UTP-glucose-1-phosphate uridylyltransferase**, **skeletal muscle**	51.7/7.01	29	2136353, S62599, 191760, Q16851 (UGP2)	56.8/8.15
29	**glutamic-oxaloacetic transaminase 1**, **soluble (aspartate aminotransferase 1)**	41.5/6.57	67	55663625, CAH73859, 138180, P17174 (GOT1)	46.1/6.57
36	**L-lactate dehydrogenase chain A**, **muscle**	36.6/8.20	35	13786849, 1I10_A, 150000, P00338 (LDHA)	36.6/8.46

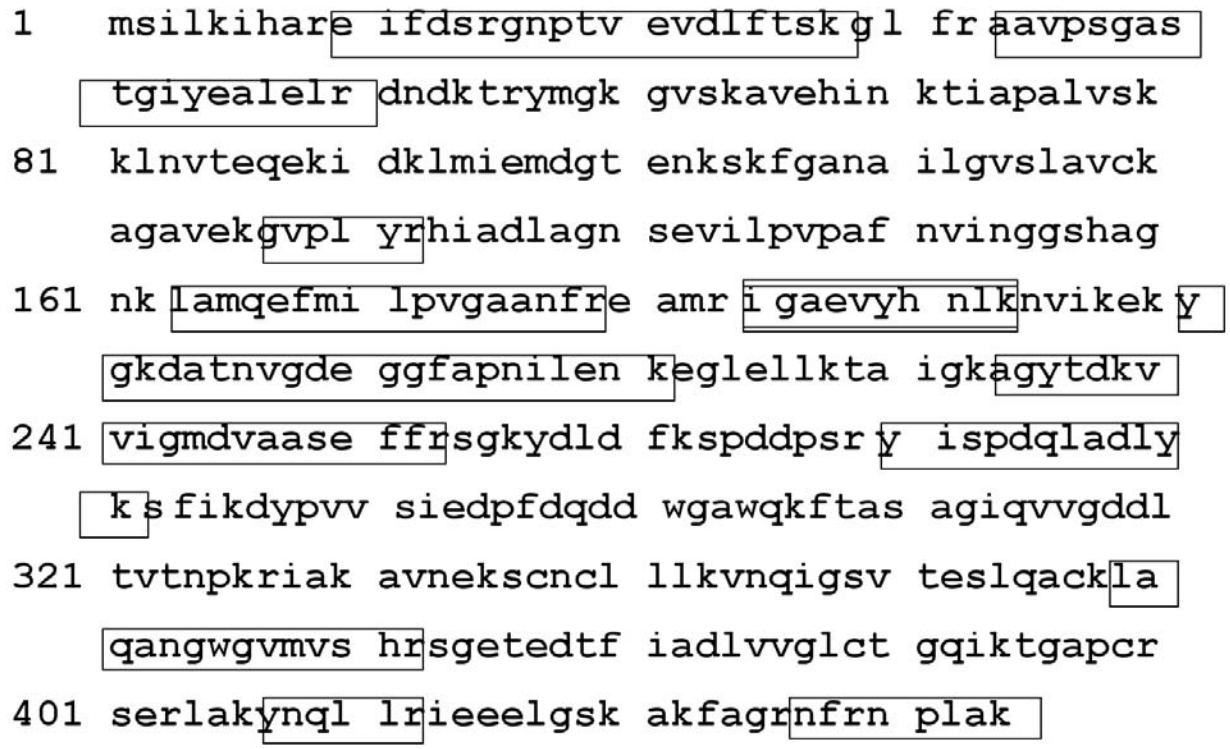
Table 3 (Contd.)

1	2	3	4	5	6
48	**malate dehydrogenase 1, cytosolic**	34.0/6.65	42	1255604, BAA09513, 154200, P40925 (MDH1)	36.3/6.89
56	**carbonic anhydrase I** [10]	28.7/6.90	52	4502517, NP_001729, 114800, P00915 (CA1)	28.7/6.63
61	**protein 14-3-3-Z-subtype (tyrosine 3/tryptophan 5 -monooxygenase activation protein, zeta polypeptide)**	27.8/4.85	27	4507953, NP_003397, 601288, P63104 (YWHAZ, 14-3-3-ZETA)	27.7/4.73
62	**heat shock 27-kDa protein β-1 isoform**	27.0/6.20	64	4504517, NM_001531, 602195, P04792 (HSPB1)	22.8/5.98 (by NM_001531)
66	**peroxiredoxin 6 (antioxidant protein 2; non-selenium glutathione peroxidase; acidic calcium-independent phospholipase A2, lysosomal)**	25.0/6.35	44	4758638, NP_004896, 602316, P30041 (PRDX6, PLA2)	24.9/6.02
74	Hsp27-like protein (EV)	18.5/6.89	45	123571, P04792, 602195 (HSPB1)	22.3/7.83
75	**heat shock protein, α-crystallin-related B6 (p20-like; Hsp20)**	18.5/6.26	42	6166215, NM_144617, 610695, O14558 (HSPB6)	17.1/5.95
79	**fatty acid-binding protein isoform 5**	15.5/6.50	54	4557581, NP_001435, 605168, Q01469 (FABP5)	15.0/6.82
84	**fatty acid-binding protein isoform 1**	14.5/6.33	54	4557577, NP_001434, 134650, P07148 (FABP1)	14.2/6.60
	Some other proteins				
10	**albumin** [10]	67.0/5.95	52	23243418, AAH36003, 103600, P02768 (ALB)	66.4/5.73 (isoform 1)
9	**heat shock 70 kDa protein isoform 5 (glucose-regulated protein, 78 kDa)**	72.0/5.10	37	18044381, AAH20235, 138120, P11021 (HSPA5)	72.3/5.07
2	α-collagen (EV) [10]	120.0/7.12	12	30041, CAA34683, 120160, P08123 (COL1A2)	123.9/9.08
3	**α-collagen isoform α-2(I)** [10]	120.0/7.10	12	30041, CAA34683, 120160, P08123 (COL1A2)	123.9/9.08
4	α-collagen (EV) [10]	120.0/7.08	14	30041, CAA34683, 120160, P08123 (COL1A2)	123.9/9.08
83	**β-globin**	14.7/7.23	95	4504349, NP_000509, 141900, P68871 (HBB)	15.9/6.81

tion of protein fractions with pI > 8.7. Some of these proteins were identified by mass spectrometry (Table 2). In particular, protein fraction 81 proved to be α-globin that, together with fractions 83 and 10 identified as β-globin and albumin, respectively (Table 3), apparently reflects the presence of traces of blood in the samples taken at biopsy. Besides α-globin, several well-known and functionally important proteins were revealed among the identified ones: cytochrome *c*, fast type skeletal muscle troponin I isoform, β-subunit of mitochondrial trifunctional protein (hydroxyacyl dehydrogenase, subunit B), and three protein fractions containing LIM- and PZD-domains (Fig. 1a and Table 2).

These three protein fractions attracted our special attention because we failed to find any report of similar proteins in human muscles using 2DE, as well as because two of them (53 and 54), as evidenced by mass spectrometry, contained peptides encoded by one earlier-

a



1 msilkihare ifdsrgnptv evdlftskgl fraavpsgas
tgiyealelr dndktrymgk gvskavehin ktiapalvsk
81 klnvteqeki dklmiemdgt enkskfgana ilgvslavck
agavekgvpl yrhiadlagn sevilpvpaf nvinggshag
161 nklamqefmi lpvgaanfre amrigaevyh nlknvikeky
gkdatnvgde ggfapnilen keglellkta igkagytdkv
241 vigmdvaase ffrsgkydld fkspddpsry ispdqladly
ksfikdypvv siedpfdqdd wgawqkftas agiqvvgddl
321 tvtnpkriak avnekscncl llkvnqigsv teslqackla
qangwgvmvs hrsgetedtf iadlvvglct gqiktgapcr
401 serlakynql lrieeelgsk akfagrnfrn plak

Fig. 2. Generalized data of mass-spectrometric identification of fraction 22. a) Distribution of tryptic peptides identified using Mascot software (Table 4) in amino acid sequence of α-enolase. Totally 152 a.a are found in peptides (marked with frames) of 434 a.a. of the amino acid sequence accessible in NCBI database, Acc. No. NP_001419), that is, 35% is covered.

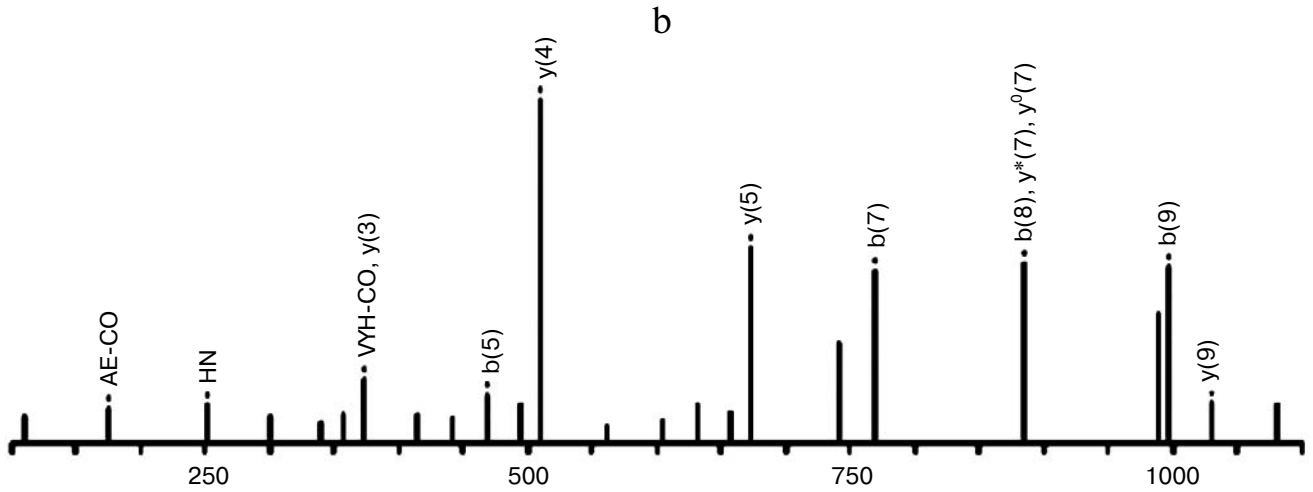


Fig. 2. Generalized data of mass-spectrometric identification of fraction 22. b) Results of fragmentation of one peptide (double frame in Fig. 2a) studied by tandem mass spectrometry (MALDI-TOF MS/MS).

revealed human transcript [27]. On the other hand, these peptides corresponded to the N-terminal sequence of the known PDZ- and LIM-domain 5-containing protein, which has, however, significantly larger *M* and is a product of the *PDLIM5* gene (synonym – *ENH*) (according to the Q96HC4 record in Swiss-Prot). Possibly, proteins 53 and 54 represent products of mRNA or a still uncharacterized gene, or products of mRNA formed by alternative splicing of the *ENH* gene product [27]. Besides, they can apparently be regarded as "fragments" of known protein (the presence of similar "fragments" has been observed in several studies [28, 29]).

Several "fragments" of known proteins were found in the present study. Some of them are shown on the Fig. 1a and others in Fig. 4 (see further). Fractions 34, 42, and 60 with *M* of 38.0, 33.2, and 27.8 kDa, respectively (Fig. 1a), were characterized by mass spectrometry as actin "fragments" (matches – 33, 31, and 21%). Such a decrease in *M* could be due to the shortening of polypeptide chains of the "fragments" in comparison with the full-length

Table 4. Identification of fraction 22 as enolase 1 by MALDI-TOF MS/MS using Mascot software

$M_{r\ obs}$, Da	$M_{r\ exp}$, Da	$M_{r\ calc}$, Da	Δ	Peptide position	Identified peptide
704.36	703.36	703.40	–0.05	127-132	GVPLYR
766.32	765.32	765.37	–0.05	10-15	EIFDSR
806.41	805.40	805.44	–0.04	407-412	YNQLLR
959.50	958.49	958.53	–0.04	427-434	NFRNPLAK
1143.60	1142.60	1142.61	–0.01	184-193	IGAEVYHNLK
1406.69	1405.68	1405.71	–0.03	16-28	GNPTVEVDLFTSK
1425.71	1424.70	1424.72	–0.02	270-281	YISPDQLADLYK
1541.68	1540.68	1540.76	–0.08	359-372	LAQANGWGVMVSHR
1556.67	1555.66	1555.77	–0.11	240-253	VVIGMDVAASEFFR
1804.87	1803.86	1803.94	–0.07	33-50	AAVPSGASTGIYEALELR
1939.82	1938.81	1938.97	–0.16	163-179	LAMQEFMILPVGAANFR
1960.78	1959.77	1959.92	–0.15	203-221	DATNVGDEGGFAPNILENK
2191.82	2190.81	2191.06	–0.25	234-253	AGYTDKVVIGMDVAASEFFR
2308.84	2307.83	2308.10	–0.26	200-221	YGKDATNVGDEGGFAPNILENK

Note: Result of identification: gi|4503571, molecular mass 47,566 Da – enolase 1 (phosphopyruvate hydratase, α-enolase), expressed in *Homo sapiens*.

actin by regions with average lengths of 35 to 140 a.a. Two similar "fragments" of α-actin (*M* 29.5 and 26.9 kDa) were described in study [28] of skeletal muscle proteins by 2DE fractionation.

In addition, a comparison between the data of mass determination of tryptic peptides of fraction 55 and its electrophoretic properties indicated that this fraction represents a "fragment" of slow type skeletal muscle troponin T1 isoform 3 (matches – 22%), because a concordance of *M*/p*I* values obtained from 2DE data (29.0/5.00) and calculated magnitudes (29.2/5.16) was observed in this case. Fraction 46 was identified in a similar way as a "fragment" of glyceraldehyde-3-phosphate dehydrogenase: matches – 37%, *M*/p*I* – 34.2/8.65 (experimental data) and 34.1/8.31 (calculated magnitudes).

The above-described "fragments" of known proteins were represented in virtually all gels after 2DE of extracts from samples taken at biopsy, whereas another pattern was observed for two "fragments" of αB-crystallin (Fig. 3), one of them (F1-cr) being revealed in 10 volunteers (Fig. 3b), and the second (F2-cr) in only two persons. However, both "fragments" were revealed in analysis of most postmortem skeletal muscle samples, with remarkable increase in amount of F2-cr in samples with postmortem interval of about one day (Fig. 3c). Thus, the revealed αB-crystallin "fragments" apparently represent the proteolytic products of this protein and can serve as objective signs of autolysis. Still another protein fraction was identified as αB-crystallin with 57% matches (Fig. 3, EV-cr); it did not differ in its *M* from the main fraction (73), but it was characterized by lower p*I* value (6.72) and was less abundant. The EV-cr fraction was assigned to an electrophoretic variant of αB-crystallin, which is possibly its phosphorylated form [30].

In general, the data suggest that the "fragments" of known proteins revealed under proteomic studies can be of different origin. Elucidation of the causes is interesting for adequate estimation of roles of the observed proteins in healthy volunteers and patients with various pathologies.

The used modification of 2DE has broadened not only possibilities for proteome analysis of muscle proteins due to inclusion of several fractions in the p*I* range 8.7-9.5, but also revealed several protein fractions with *M*

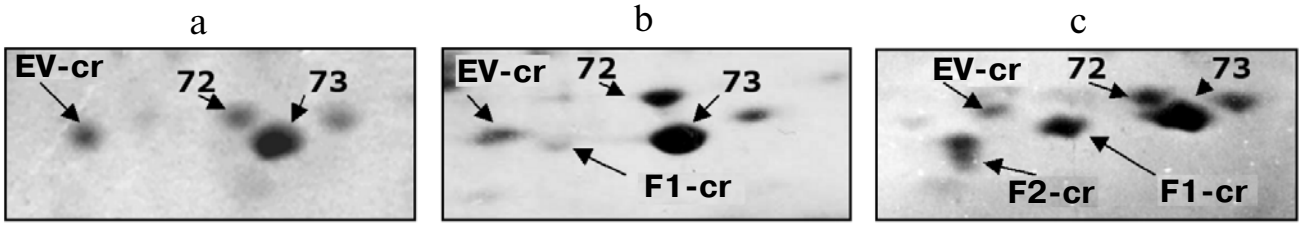


Fig. 3. Fragments of 2DE gels with identified αB-crystallin main fraction (73) obtained from *m. vastus lateralis* samples taken at biopsy (a) and after autopsy with postmortem interval below 12 h (b) and above 12 h (c). Proteins identified by mass spectrometry as crystallin EV (EV-cr) and crystallin fragments (F1-cr and F2-cr) are indicated by arrows.

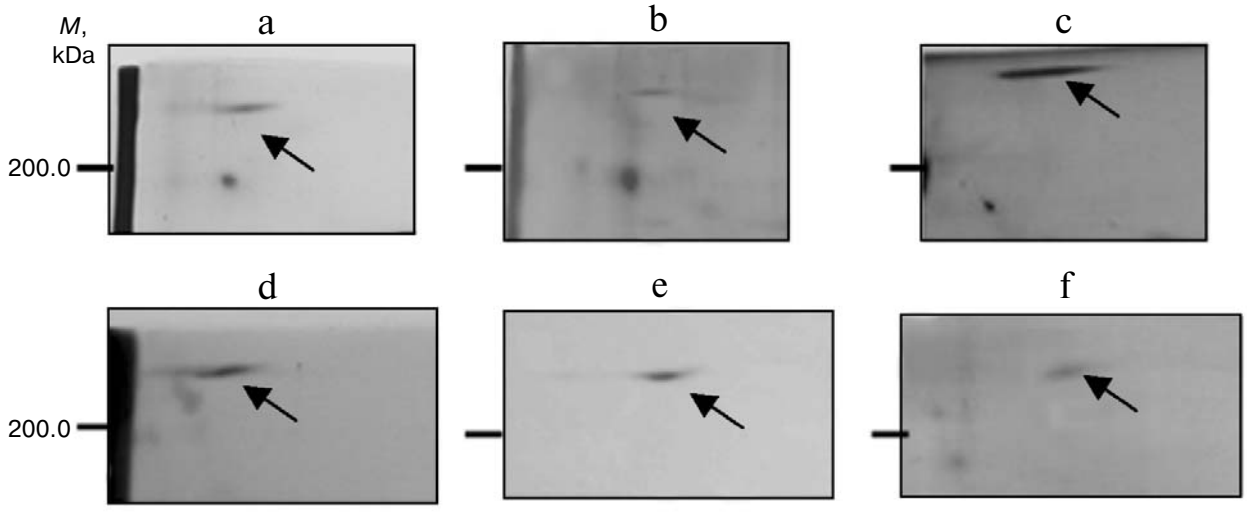


Fig. 4. Fragments of 2DE gels with identified ferritin chain complex (arrow) obtained in analysis of proteins from cultivated myoblasts (a), fibroblasts (b), prostate samples taken at biopsy (c), DU145 cell line (d), brain samples taken after autopsy (e), and lymphocytes (f).

above 200 kDa (Fig. 1a). To judge from mass-spectrometric analysis, one of them contained tryptic peptides corresponding to ferritin heavy and light chains (*FTH1* and *FTL* genes, sequence coverage 41 and 40%, respectively). Since *M* of ferritin heavy and light chains are evaluated as 26.2 and 16.4 kDa, respectively, one can admit that the given protein fraction represents a very tight complex of ferritin chains, and given this fact it is designated as the FCC abbreviation in Fig. 1a.

Ferritin isoforms have been studied in muscle since the 1980s. In parallel, a distinct stability of oligomeric ferritin complexes has been demonstrated [31, 32], but the complexes we have revealed are extremely tight: they do not dissociate onto monomers even under the action of high urea concentration and in the presence of detergents used in 2DE of proteins. Apparently, this stability is only possible if covalent bonds are formed between the ferritin chains.

We found FCCs that are similar in their properties (including mass-spectrometry results) in all other studied muscular tissues and organs. Moreover, complexes like these were also detected in cultured myoblasts and fibroblasts, in samples of prostate tissues, in cultures of epithelial prostate cancer line DU145 cells, and in lymphocytes (Fig. 4). Hence, one can suppose that the FCCs revealed by the proteome studies and represented in various human cells are associated with features of ferritin metabolism. During the intracellular catabolism of ferritin, some portion of it becomes detectable in lysosomes (intralysosomal ferritin) [33], and FCCs might represent either this intralysosomal fraction, which is sustained as a stable depot for iron ions, or an intermediate of ferritin complex catabolism.

A series of known mitochondrial, cytoplasmic, and some other proteins identified by mass spectrometry is shown in Fig. 1a, in line with above described "major" muscle proteins and several protein fractions unusual for routine 2DE analysis (the ones with pI > 8.70, "fragments" of known proteins, and FCC) (Table 3).

In total, we have identified 89 protein fractions in 2DE of samples taken at biopsy of human *m. vastus lateralis* using mass spectrometry (1-85 and FCC in Figs. 1-3; F1-cr, F2-cr, and EV-cr in Fig. 3), the vast majority of which remained unidentified earlier. Complete coincidence of the data was observed for fractions previously identified by other methods [8-10] (Table 1). In the aggregate, 89 identified protein fractions proved to be the expression products of 55 genes; 53 genes are presented in Tables 1-3, and the two others are the ferritin heavy and light chains described above.

Thus, taking into account the previously reported data on identification of 60 protein fractions in proteomic studies of human muscle proteins [10, 13, 19, 34] and results of the present work, the total number of identified protein fractions has increased to 121, thus broadening and modernizing the earlier created national GEP-HM database on human muscle proteins [2, 10, 34]. In general, the results also including data on biochemical polymorphism of some muscle proteins form a background for further studies on human skeletal muscle proteins in terms of molecular mechanisms of various physiological and pathological processes.

REFERENCES

1. Hanash, S. (2004) *Mol. Cell Proteom.*, **3**, 298-301.
2. Shishkin, S. S., Kovalyov, L. I., and Kovalyova, M. A. (2004) *Biochemistry (Moscow)*, **69**, 1283-1298.
3. Cieniewski-Bernard, C., Acosta, A., Dubois, E., Lamblin, N., Beseme, O., Chwastyniak, M., Amouyel, P., Bauters, C., and Pinet, F. (2008) *Clin. Exp. Pharmacol. Physiol.*, **35**, 362-366.
4. Perry, S. V. (2001) *J. Muscle Res. Cell Motil.*, **22**, 5-49.
5. Guillet, C., Boirie, Y., and Walrand, S. (2004) *Curr. Opin. Clin. Nutr. Metab. Care*, **7**, 531-538.
6. Barboro, P., Rubagotti, A., Orecchia, P., Spina, B., Truini, M., Repaci, E., Carmignani, G., Romagnoli, A., Introini, C., Boccardo, F., Carnemolla, B., and Balbi, C. (2008) *Cell Oncol.*, **30**, 13-26.
7. Barry, S. P., Jayasinghe, S. N., Pericleous, C., Hubank, M., Latchman, D. S., and Stephanou, A. (2008) *Biotechnol. J.*, **3**, 530-535.
8. Kovalyova, M. A., Kovalyov, L. I., Khudaidatov, A. I., Efimochkin, A. S., and Shishkin, S. S. (1994) *Biochemistry (Moscow)*, **59**, 493-498.
9. Laptev, A. V., Shishkin, S. S., Egorov, Ts. A., Kovalyov, L. I., Tsvetkova, M. N., Galyuk, M. A., Musalyamov, A. Kh., and Efimochkin, A. S. (1994) *Mol. Biol. (Moscow)*, **28**, 52-58.
10. Kovalyov, L. I., Shishkin, S. S., Efimochkin, A. S., Kovalyova, M. A., Ershova, E. S., Egorov, T. A., and Musalyamov, A. K. (1995) *Electrophoresis*, **16**, 1160-1169.
11. Shenkman, B. S., Netreba, A. I., Litvinova, K. S., Tarakin, P. P., Vikhlyantsev, I. M., Podlubnaya, Z. A., Nemirovskaya, T. L., Kovalyova, M. A., Stekhanova, T. N., Popov, D. V., and Vinogradova, O. L. (2004) in *Medical and Biological Technologies for Enhancement of Capacity for Work under Physical Stress* [in Russian], Fon, Moscow, pp. 102-116.
12. Kovalyov, L. I., Shishkin, S. S., Khasigov, P. Z., Dzeranov, N. K., Kazachenko, A. V., Kovalyova, M. A., Toropygin, I. Yu., and Mamykina, S. V. (2006) *Appl. Biochem. Microbiol.*, **42**, 424-427.
13. Makarov, A. A., Kovalyov, L. I., Kovalyova, M. A., Toropygin, I. Yu., and Shishkin, S. S. (2009) *Rus. J. Dev. Biol.*, **39**, 83-89.
14. Blum, H., Beir, H., and Cross, H. G. (1987) *Electrophoresis*, **8**, 126-129.
15. Osterman, L. A. (1981) *Methods of Protein and Nucleic Acid Research*, *Vol. 1, Electrophoresis and Ultracentrifugation* [in Russian], Nauka, Moscow, pp. 56-58.
16. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) *Anal. Chem.*, **68**, 850-858.
17. Govorun, V. M., Moshkovskii, S. A., Tikhonova, O. V., Goufman, E. I., Serebryakova, M. V., Momynaliev, K. T., Lokhov, P. G., Khryapova, E. V., Kudryavtseva, L. V., Smirnova, O. V., Toropygin, I. Yu., Maksimov, B. I., and Archakov, A. I. (2003) *Biochemistry (Moscow)*, **68**, 42-49.

18. Nikolaev, E. N. (2007) in *Human Molecular Polymorphism* (Varfolomeev, S. D., ed.) [in Russian], Vol. 1, University of People's Friendship, Moscow, pp. 1-82.
19. Kovalyov, L. I., Kovalyova, M. A., Toropygin, I. Yu., Shigeev, S. V., and Shishkin, S. S. (2006) in *Medical and Biological Technologies for Enhancement of Capacity for Work under Physical Stress* [in Russian], Anita, Moscow, Vol. 2, pp. 52-62.
20. Laptev, A. V., Shishkin, S. S., Kovalyov, L. I., Galyuk, M. A., Musalyamov, A. Ch., and Egorov, Ts. A. (1993) *Biochem. Genet.*, **31**, 253-258.
21. Kovalyov, L. I., Kovalyova, M. A., Kovalyov, P. L., Serebryakova, M. V., Moshkovskii, S. A., and Shishkin, S. S. (2006) *Biochemistry (Moscow)*, **71**, 448-453.
22. Billeter, R., Heizmann, C. W., Howald, H., and Jenny, E. (1981) *Eur. J. Biochem.*, **116**, 389-395.
23. Schiaffino, S., and Reggiani, C. (1996) *Physiol. Rev.*, **76**, 371-423.
24. Sachdev, S., Raychowdhury, M. K., and Sarkar, S. (2003) *DNA Seq.*, **14**, 339-350.
25. Faulkner, G., Pallavicini, A., Formentin, E., Comelli, A., Ievolella, C., Trevisan, S., Bortoletto, G., Scannapieco, P., Salamon, M., Mouly, V., Valle, G., and Lanfranchi, G. (1999) *J. Cell Biol.*, **146**, 465-475.
26. Leonoudakis, D., Conti, L. R., Anderson, S., Radeke, C. M., McGuire, L. M., Adams, M. E., Froehner, S. C., Yates, J. R., 3rd, and Vandenberg, C. A. (2004) *J. Biol. Chem.*, **279**, 22331-22346.
27. Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., Wagner, L., Shenmen, C. M., Schuler, G. D., Altschul, S. F., et al. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 16899-16903.
28. Gelfi, C., de Palma, S., Cerretelli, P., Begum, Sh., and Wait, R. (2003) *Electrophoresis*, **24**, 286-295.
29. Vorum, H., Ostergaard, M., Rice, G. E., Honore, B., and Bek, T. (2007) *Proteome Sci.*, **5**, 5-14.
30. Den Engelsman, J., Keijsers, V., de Jong, W. W., and Boelns, W. C. (2003) *J. Biol. Chem.*, **278**, 4699-4704.
31. Linder, M. C., Nagel, G. M., Roboz, M., and Hungerford, D. M. (1981) *J. Biol. Chem.*, **256**, 9104-9110.
32. Chiaraluce, R., Consalvi, V., Cavallo, S., Ilari, A., Stefanini, S., and Chiancone, E. (2000) *Eur. J. Biochem.*, **267**, 5733-5741.
33. Kidane, T. Z., Sauble, E., and Linder, M. C. (2006) *Am. J. Physiol. Cell Physiol.*, **291**, C445-455.
34. Shishkin, S. S., Kovalyov, L. I., and Gromov, P. S. (2000) in *Multiformity of Modern Human Genetics* (Shishkin, S. S., ed.) [in Russian], Gilem, Moscow-Ufa, pp. 17-50.