Research Article **Open Access**

Proteomic Profiling of High-grade Glioblastoma Using Virtual-experimental 2DE

Stanislav N Naryzhny[1,2]*, Maria A Maynskova[1], Victor G Zgoda[1], Natalia L Ronzhina[2], Svetlana E Novikova[1], Natalia V Belyakova[2], Olga A Kleyst[2], Olga K Legina[2], Rimma A Pantina[2] and Michael V Filatov[2]

[1]Institute of Biomedical Chemistry, Pogodinskaya 10, Moscow, 119121, Russia
[2]BP Konstantinov Petersburg Nuclear Physics Institute, National Research Center "Kurchatov Institute", Leningrad Region, Gatchina, 188300, Russia

Abstract

Identification and quantitative analysis of different proteoforms (protein species) presented in a cell line generated from high grade glioblastoma was performed using two-dimensional electrophoresis (2DE), mass spectrometry (ESI LC-MS/MS), and immunodetection. A 2DE protein map containing an extensive data set comprising 937 spots with 1542 unique protein identifications (proteoforms) of 600 genes was obtained. Additionally, another set of experiments was performed where 16012 proteoforms coded by 4050 genes were identified by MS/MS according to their position in 96 gel sections (pixels). A special attention has been paid to the proteins that are the potential biomarkers of glioblastoma. The list of these biomarkers was compiled from literature. Next, we generated the graphs with theoretical and experimental information about proteoforms coded by the same gene. Such a virtual-experimental representation allowed better visualization of the state of these gene products. Many proteins, potential biomarkers of glioblastoma as well, are characterized by high numbers of protein species. We assume that these species could be a potential source of highly specific biomarkers of glioblastoma.

Keywords: Proteome; Proteoforms; Two-dimensional electrophoresis; Mass spectrometry; Biomarkers; Glioblastoma

Abbreviations: 2DE: Two-dimensional Electrophoresis; ESI LC-MS/MS: Liquid Chromatography-Electrospray Ionization-Tandem Mass Spectrometry; HCD: Higher Energy Collisional Dissociation; ABC: Ammonium Bicarbonate; CAN: Acetonitrile; PTM: Post-Translation Modification; emPAI: Exponential Modified Form of Protein Abundance Index

Introduction

High grade glioma (glioblastoma multiform, WHO grade IV) is the most common brain tumor with a very poor prognosis. Its malignancy makes glioblastoma the fourth biggest cause of cancer death [1,2]. The average survival period for this disease is one year, pointing the need to find ways of early diagnosis and treatment. Today, the main methods of diagnosis of this disease are computed tomography and brain biopsy. In clinic, only one brain tumor biomarker (MGMT, methyl-guanine DNA methyl transferase) is used [3]. Actually, MGMT is not a marker of the disease but a marker of epigenetic status of the patient. Accordingly, depending of this status DNA alkylating antineoplastic treatment is appointed. Therefore, there is an urgent need to develop new, noninvasive methods for early diagnosis. Qualitative and quantitative changes in the detected protein groups could serve as a good indicator of this cancer. Therefore, the systems biology approaches, such as proteomics, transcriptomics, and metabolomics, based on the use of high throughput techniques are very promising. They could afford the developing of the detection of cancerous and precancerous states based on not just one but a group of markers [4,5].

Indeed, recently, genomic and proteomic high-throughput technologies have provided molecular profiling of human gliomas [4-7]. These studies have generated a number of candidate diagnostic, prognostic and prediction of response to therapy markers. The list of these potential glioblastoma markers described in different publications is rather long. It indicates that the situation here is not very simple, even considering the publications, where possible molecular signatures associated with this tumor have been described [4,8,9]. Before adapting these biomarkers to the clinic, it is necessary

to address many challenges hindering an in-depth, nonbiased profiling of the glioblastoma proteome [10]. As glioblastoma cells are very heterogeneous, identifying of biological signatures for each subtype through protein biomarker profiling is therefore a high priority. Such a profiling not only may give a clue regarding tumor classification but may identify clinical biomarkers or targets for the development of clinical treatments [7,8].

In our previous experiments, searching for protein signatures and biomarkers for gliomas we used cell lines to obtain proteomics information specific to this disease [4]. We tested normal and cancer cell lines including six primary glioblastoma cell lines generated in our department. We found that proteome profiles in normal and glioblastoma cell lines are very similar but levels of several proteins have prominent differences between norm and cancer [4]. Importantly, our results confirm the information about some proteins as the potential cancer markers, common and associated with glioblastoma [4]. They also indicate the possibility of using PCNA and p53 proteins as biomarkers of glioblastoma assays [4]. Moreover, we have shown extremely higher heterogeneity of p53 in glioblastomas compare to normal cells, which itself requires a special attention and analysis.

In our present work, we decided to explore these aspects of proteoform heterogeneity, especially in potential cancer biomarkers. We have obtained more data using 2DE separation followed by imaging, immunochemistry, and mass-spectrometry. It may be relevant to mention here that in general our 2DE pictures look very

similar to the 2DE picture of proteins from the human glioma cell line U87MG [7]. We generated the 2DE map where 937 protein spots are annotated with 1542 proteoforms. Additionally, 16012 proteoforms coded by 4050 genes were identified according to their position in 96 gel sections (the so called "pixels-picking" approach). This approach allowed us to obtain a "panoramic" view of proteoforms corresponded to each of these genes (the so-called "one gene proteomes"). In our mind, this is a most comprehensive map of glioblastoma cell proteins for today.

Experimental Section

Chemicals and materials

All reagents used, unless other manufacturer is specified, were obtained from "Sigma-Aldrich" (USA). The remaining reagents were obtained from the following companies: Pierce: dithiotriitol (DTT), protease inhibitor cocktail; GE Healthcare: IPG DryStrip (gel strips), IPG-buffers, DryStrip-coating liquid, Coomassie R350; Promega: Trypsin Gold; Bio-Rad: molecular weight markers for protein electrophoresis; Biolot: RPMI-1640 medium and DMEM for cell growth, fetal calf serum; Orange Scientific: Carrel culture flasks.

Cell culture and culture conditions

Human glioblastoma cells (a primary line L of glial tumor origin, developed in the laboratory of cell biology, PNPI) were cultured in DMEM or RPMI-1640 medium containing 5% fetal calf serum in 5% CO_2 at 37°C without antibiotics [4,11].

Sample preparation and two-dimensional electrophoresis (2DE)

All samples were prepared as described previously [12,13]. Cells (~10^7) containing 2 mg of protein, were treated with 100 μl of lysis buffer (7 M urea, 2 M thiourea, 4% CHAPS, 1% dithiothreitol (DTT), 2% ampholytes, pH 3-10, protease inhibitor mixture). Protein concentration in the sample was determined by the method of Bradford [14]. Proteins were separated by isoelectric focusing (IEF) using ImmobilineDryStrip pH 3-11 (7 or 18 cm) (GE Healthcare) following the manufacturer's protocol. The samples in the lysis buffer were mixed with rehydrating buffer (7 M urea, 2 M thiourea, 2% CHAPS, 0.3% DTT, 0.5% IPG buffer, pH 4-7 or 3-11 NL, 0.001% bromphenol blue) in the final volume of 130 μl (500 μg of protein) for 7 cm strip or 300 μl (800 μg of protein) for 18 cm strips. Strips were passively rehydrated for 4 h at 4°C. IEF was performed on the 3100 OFF GEL Fractionator (Agilent technologies) which was programmed as follows: For 7 cm strips-6 kV and 20 kVh, 14 h; for 18 cm strips-10 kV and 60 kVh, 14 h; temperature 20°C and maintained at the voltage 500 V. After IEF, the strips were soaked 10 min in the equilibration solution (50 mM Tris, pH 6,8, 6 M urea, 2% sodium dodecyl sulfate (SDS) and 30% glycerol) with 1% DTT. This process was followed by 10-min incubation in the equilibration solution containing 5% (w/v) iodacetamide. The strips were placed on the top of the 11% polyacrylamide gel of the second direction and sealed with a hot solution of 1 ml of 0.5% agarose in electrode buffer (25 mM Tris, pH 8.3, 200 mM glycine, and 0.1% SDS) and electrophoresed in the second direction under denaturing conditions using the system Hoefer miniVE (gel size 80×90×1mm, GE Healthcare) or Ettan™ DALTsix (240 × 200 × 1 mm, GE Healthcare). Electrophoresis was carried out at room temperature at 3 Watts constant for one gel [13,15].

Immunoblotting (Western Blot)

Transfer of proteins from the gel onto PVDF membrane (Amersham Hybond-P, GE Healthcare) was carried out using "semi-dry" process for 1 h at 15V by placing the gel and membrane between two sheets of extra thick blot paper (Bio-Rad), impregnated with a transfer buffer (48 mM Tris, 39 mM glycine, 0,037% SDS, 20% ethanol). After transfer, the membrane was treated according to protocol of Blue Dry Western, i.e., at first it was stained with 0.1% solution of Coomassie R350 and then dried and treated with antibodies [13]. We used the following antibodies with 1/1000 dilution: mouse monoclonal anti-p53 (Ab-6, DO-1, NeoMarkers), rabbit polyclonal anti-PKM2 SAB4200105-25 μL (Sigma), rabbit polyclonal Anti-cofilin (ab42824, Abcam), mouse monoclonal anti-PCNA (clone PC10, Abcam). Secondary goat anti-mouse or anti-rabbit immunoglobulin G labeled with horseradish peroxidase (PerkinElmer) was used at concentration of 0.5 μg/ml. The reaction was developed using ECL (Western Lightning Ultra, PerkinElmer) and X-ray film (Amersham Hyperfilm ECL) upon exposure from 10 seconds to 30 minutes. The gels, membranes, and films were scanned and analyzed using 2D Platinum (GE Healthcare).

Mass spectrometry

All procedures were performed according to the previously described protocol [16,17]. After separation by 2DE and staining with Coomassie R350, the gel pieces with a diameter of 1.5 mm corresponding to a protein spots were excised using Spot picker ("GE Healthcare", USA) and partially destained by a 15-minute incubation in 500 μl of 50% aqueous acetonitrile (ACN) with 25 mM ammonium bicarbonate. Further, the gel pieces were dehydrated by a 10-minute incubation in 200 μl of 100% ACN. ACN was removed, and the gel was dried for 20 minutes in a centrifugal evaporator Speed Vac. The dried gel pieces were rehydrated for 25 min on ice in a 12 μl solution of 25 mM ammonium bicarbonate (ABC) containing trypsin ("Trypsin Gold", 10 μg/ml) and proteolysis of the protein was performed by incubation at least 4 h at 37°C. Tryptic peptides were eluted from the gel with extraction solution (5% (v/v) ACN, 5% (v/v) formic acid) and dried in a vacuum centrifuge. For ESI LC-MS/MS analysis, peptides were dissolved in 5% (v/v) formic acid. Using an Agilent HPLC system 1100 Series (Agilent Technologies), 4 μg of peptides in 5% formic acid were injected onto a trap column Zorbax 300SB–C18, 5 × 0.3 mm (Agilent Technologies). After washing with 5% acetonitrile containing 0.1% formic acid, peptides were resolved on a 150 mm × 75 μm Zorbax 300SB-C18 reverse phase analytical column (Agilent Technologies) by a 30-min organic gradient of 5-60% ACN, 0.1% formic acid with a flow rate of 300 nL/min. Peptides were then ionized by nano-electrospray ionization at 2.0 kV using a fused silica emitter with an internal diameter of 8 μm (New Objective). MS/MS analysis was carried out in duplicate on an Orbitrap Q-Exactive (Thermo Scientific). Mass spectra were acquired in the positive ion mode. High resolution data was acquired in the Orbitrap analyzer with a resolution of 30 000 (m/z 400) for MS and 7500 (m/z 400) for MS/MS scans. Survey MS scan was followed by MS/MS spectra of five the most abundant precursors. For peptide fragmentation, higher energy collisional dissociation (HCD) was set to 35 eV, the signal threshold was set to 5000 for an isolation window of 2 m/z, and the first mass of HCD spectra was set to 100 m/z. Fragmented precursors were dynamically excluded from targeting for 90 s. Singly charged ions and ion with unassigned charge state were excluded from triggering MS/MS scans. The automatic gain control target value was regulated at $1 × 10^6$ with a maximum injection time of 100 ms and at $1 × 10^5$ with a maximum injection time of 250 ms for MS and MS/MS scans, respectively. The data were searched by Mascot "2.4.1" search engine (www.matrixscience.com) using the following parameters: Enzyme=Trypsin (allowing for cleavage before proline); Maximum missed cleavages=2; Fixed modifications=Carbamidometh ylation of cysteine; Variable modifications=Oxidation of methionine;

Phosphorylation of serine and threonine, acetylation of lysine; Precursor mass tolerance=20 ppm; Product mass tolerance=0.01 Da. NeXtProt database (October 2014) was used as a protein sequence database. For FDR assessment, a separate decoy database was generated from the protein sequence database. False-positive rate of 1% was allowed for protein identification. These parameters have previously been shown to be enough for identifying true positive matches [18]. Exponentially modified PAI (emPAI), the exponential form of protein abundance index (PAI) defined as the number of identified peptides divided by the number of theoretically observable tryptic peptides for each protein, was used to estimate protein abundance [19].

Results and Discussion

2DE map of proteins from glioblastoma cells

We have performed high resolution two-dimensional electrophoresis (2DE) of proteins extracted from glioblastoma cells followed by mass spectrometry or immunodetection (Western blot). Using a similar approach, we have previously tested several glioblastoma cell lines generated in our department [4]. This time, using large gels (24×20 cm), more protein (800 µg), and ESI LC-MS/MS mass spectrometry instead of MALDI-TOF MS we have managed to obtain much more detailed 2DE map. Initially, we obtained the 2DE map using the classical approach-the gel was stained, protein spots were cut out, and proteins were analyzed by ESI LC-MS/MS (Supplementary Figure S1 and Supplementary Table S1). We detected more than 1000 protein spots, which were excised from the gel using SpotPicker (GE Healthcare). The gel plugs were treated for mass spectrometry according to Materials and Methods. The peptides were analyzed using LC-ESI-MS/MS mass spectrometer (Orbitrap Q-exactive). The database (NeXtProt) search by Mascot allowed identifying with high score (detection of at least two specific sequences) 1542 protein species (proteoforms) that are products of 600 genes and localized in 937 spots (numbers of spots are shown in Supplementary Figure S1(B). A complete list of identified proteoforms is represented in Supplementary Table S1. This table contains theoretical and experimental information about these proteoforms and data about spots (spot relative abundance, %V), where these proteoforms were detected. As the spot intensity is proportional to its protein amount, the gel densitometry can be used for protein quantitation. The parameter %V can be used also for the proteoform abundance, but only in case, where a spot contains only a single proteoform. The situation is more complicated if several different proteoforms are detected in a single protein spot. The names of proteoforms are shown in annotations in Supplementary Figure S1(A). Additional information about each spot number (Spot ID), the protein name (Protein), the protein UniProt number and the gene name (Gene) is presented in Supplementary Table S1. Also experimental and theoretical physicochemical parameters are shown. Experimental parameters are the coordinates (pI, Mw) of the protein spot in 2DE map. Theoretical parameters are pI and Mw of the canonical sequence of the protein taken from the NextProt database. The relative abundance (%V) of each spot is based on its cumulative staining intensity. Some proteoforms of different origin (genes) have very similar physicochemical parameters (pI/Mw), so we detected them in the same spots (Supplementary Table S1). Roughly, it is possible to estimate contribution from each proteoform by emPAI parameter, which is calculated based on number of peaks in MS spectra. The names of these proteoforms are presented in annotations in Supplementary Figure S1(A). Many proteins were identified in several spots. For instance, actin (ACTB_HUMAN) was found in 53 spots, pyruvate kinase PKM isoform M2 (KPYM_HUMAN)-in 30 spots, ATP synthase subunit alpha, mitochondrial (ATPA_HUMAN)-in 11 (Supplementary Table S1). The most probable causes of such a variety are alternative splicing and multiple posttranslational modifications. Specific and nonspecific proteolytic degradation should also be considered.

Biomarkers and virtual-experimental 2DE

In this study we applied the so-called "virtual-experimental 2DE", which was recently introduced by us for proteomics study [16,17]. We have paid special attention to the genes that are the potential source of biomarkers for glioblastoma and which products were detected in our experiments. The list of these biomarkers was compiled according to our study and literature data (Table 1). Next, we generated the graphs (Figure 1 and Supplementary Figure S2), where each graph represents the distribution of the proteoforms coded by the same gene (Supplementary Table S1). Virtual-experimental representation of the data obtained (Supplementary Figure S2) gives us idea about profiles of proteoforms coded by the same gene. Usually the most abundant proteoform is a form, which parameters are close to the theoretical parameters of master protein coded by canonical sequence. The less abundant proteoforms are most likely the product of splicing or post-translation modifications.

"Pixel-picking approach" in creation of 2DE map

It is worth noting that the sensitivity of spot detection is the major bottleneck of this analysis. Actually, there is a greater quantity of proteoforms than those 1542 distributed around the gel and located in more than 937 spots (Supplementary Figure S1). For instance, if we stain this gel with silver instead of Coomassie we will detect at least 4000 spots (data not shown). It came to our attention that the least abundant proteoforms were detected often just by chance as they have pI/Mw similar to the most abundant proteoform. Definitely, much more minor proteoforms exist, but they didn't come into analysis just because of low staining sensitivity. Since the bottleneck is the sensitivity of spot detection, we decided to analyze the whole gel. In this case we would be able to analyze all of the proteins presented in the gel no matter if they are stained or not. To accomplish this, we have divided the gel into 96 sections ("pixels") (as in 96-well plates), identified as 1-12 along the Mw dimension and A-H along the pI dimension (Figure 2). Exactly the same approach we have recently applied to analyze the proteins from HepG2 cells [16,17]. Based on the position of the stained protein spots that were identified earlier (Supplementary Figure S1), the gel was calibrated, and the coordinates of borders for each section were determined. For instance, the section C5, where a most abundant spot corresponds to actin (ACTB_HUMAN) has coordinates: pI 5.11-5.80/Mw 40000-52000. All these sections were cut and treated with trypsin according to protocol for mass spectrometry. The tryptic peptides were analyzed by ESI LC-MS/MS (see Materials and Methods) using an Orbitrap Q-Exactive mass spectrometer (Thermo Scientific). Finally, protein identification and relative quantification (emPAI) were performed using Mascot "2.4.1". Up to 500 unique proteins were identified in each section by Mascot search with high score (at least two specific peptides for polypeptide). Herewith, the same unique protein presented in different gel sections was considered to be in different proteoforms. Using this approach, we detected 16012 proteoforms coded by 4050 genes (Supplementary Table S2). It means that in average 4 proteoforms were found for each gene. 3D graphs for each gene representing its proteoforms distribution around 2DE gel were built. The examples (ANXA1, TPIS, ENOA, and KPYM) of this representation are shown in Figure 3. More data (graphs for 35 genes representing biomarkers from the Table 1) are shown in Supplementary Figure S3.

№	Protein	UniProt	Protein name	pI/Mw[1]	Spot % V[2]	emPAI[3]	Reference
	PCNA_HUMAN	P12004	PCNA	4.57/28769	0.127	12.45	[4,9,20]
	COF1_HUMAN	P23528	Cofilin-1	8.22/18502	0.571	111.96	[4]
	KPYM_HUMAN	P14618	Pyruvate kinase PKM	7.96/57936	2.89	56.30	[4,21,22]
	ANXA1_HUMAN	P04083	Annexin A1	6.57/38714	0.406	24.48	[4,23]
	ANXA2_HUMAN	P07355	Annexin A2	7.57/38604	2.018	77.3	[4,20,23,24]
	TPIS_HUMAN	P60174	Triosephosphate isomerase	5.65/30791	0.504	137.58	[4,23,25]
	NPM_HUMAN	P06748	Nucleophosmin	4.64/32575	0.077	7.4	[4,7]
	VIME_HUMAN	P08670	Vimentin	5.05/53651	2.513	518.49	[4,32]
	TERA_HUMAN	P55072	Transitional endoplasmic reticulum ATPase	5.14/89321	0.322	17.65	[4,21,23,24,26]
	ENOA_HUMAN	P06733	Alpha-enolase	7.01/47168	1.390	76.21	[4]
	PRDX1_HUMAN	Q06830	Peroxiredoxin-1	8.27/22110	0.241	140.54	[4,10,21,22]
	PRDX6_HUMAN	P30041	Peroxiredoxin-6	6.00/25034	0.263	92.72	[4,27]
	SYAC_HUMAN	P49588	Alanine--tRNA ligase, cytoplasmic	5.34/106810	0.063	4.48	[4,7,27,28]
	TCTP_HUMAN	P13693	Translationally-controlled tumor protein	4.84/19595	0.237	152.61	[4]
	HSPB1_HUMAN	P04792	Heat shock protein beta-1	5.98/22782	0.076	13.00	[4,7,32]
	GSTP1_HUMAN	P09211	Glutathione S-transferase P	5.43/23355	0.100	23.17	[4,7]
	CH60_HUMAN	P10809	60 kDa heat shock protein, mitochondrial	5.70/61054	0.559	181.26	[4,8,21,31]
	ACTB_HUMAN	P60709	Actin, cytoplasmic 1	5.29/41736	3.06	102.96	[4,32]
	PHB_HUMAN	P35232	Prohibitin	5.57/29804	0.128	7.54	[4,21]
	SODM_HUMAN	P04179	Superoxide dismutase [Mn], mitochondrial	8.35/24722	0.108	40.31	[10,29,32]
	MOES_HUMAN	P26038	Moesin	6.08/67820	0.822	11.93	[4,7,20]
	IDHC_HUMAN	O75874	Isocitrate dehydrogenase [NADP] cytoplasmic isoform	6.53/46659	0.076	3.68	[8,9,10]
	IDHP_HUMAN	P48735	Isocitrate dehydrogenase [NADP], mitochondrial isoform	8.88/50909	N/A	2.11	[8,9]
	FSCN1_HUMAN	q16658	Fascin	6.84/54530	0.080	2.29	[8,9]
	EZRI_HUMAN	p15311	Ezrin	5.94/69413	0.067	1.43	[7,9]
	CNN3_HUMAN	Q15417-1	Calponin-3	5.69/36413	0.162	6.56	[7,9]
	CNN2_HUMAN	Q99439-1	Calponin-2	6.94/33697	N/A	3.92	[7,9]
	CAPG_HUMAN	P40121	Macrophage-capping protein	5.82/38499	0.046	1.54	[7,9]
	DPYL3_HUMAN	Q14195-2	Dihydropyrimidinase-related protein 3	6.04/61963	0.120	6.84	[7,9]
	DPYL2_HUMAN	Q16555	Dihydropyrimidinase-related protein 2	5.95/62293	0.120	11.11	[7,9,32]
	GFAP_HUMAN	p14136	Glial fibrillary acidic protein isoform Iso 1	5.42/49880	N/A	0.16	[7,9,20,32]
	MK01_HUMAN	P28482	Mitogen-activated protein kinase	6.50 /41389	0.043	1.18	[7,10,20,21,24]
	P53_HUMAN	P04637	Cellular tumor antigen p53	6.33/43653	N/A	N/A	[4,8]
	HIF1A	Q16665	Hypoxia-inducible factor 1-alpha	5.17/92670	N/A	N/A	[30]
	EGFR	P00533	Epidermal growth factor receptor	6.26/134277	N/A	N/A	[8]
	PTEN	P60484	Phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN	5.94/47166	N/A	N/A	[8]
	MDM2	Q00987	E3 ubiquitin-protein ligase Mdm2	4.6/55233:	N/A	N/A	[8]
	MDM4	O15151	Protein Mdm4	4.85/54864	N/A	N/A	[8]
	RB	P06400	Retinoblastoma-associated protein	8.13/106159	N/A	N/A	[8]
	VEGFA	P15692	Vascular endothelial growth factor A	9.21/27042	N/A	N/A	[8]
	MGMT	P16455	Methylated-DNA--protein-cysteine methyltransferase	8.28/21646	N/A	N/A	[8]
	STMN1	P16949	Stathmin	5.76/17303	0.036	29.90	[9]
	NDKA	P15531	Nucleoside diphosphate kinase A	5.81/17149	0.168	164.31	[9]
	SORCN	p30626	Sorcin	5.32/21676	N/A	12,55	[7]
	S10AB	p31949	Protein S100-A11	6.56/11740	N/A	17.32	[7]
	CRYAB	P02511	Alpha-crystallin B chain	6.76/20159	N/A	N/A	[7]
	PEA15	q15121	Astrocytic phosphoprotein PEA-15	4.93/15040	N/A	5.38	[7]
	FABP7	o15540	Fatty acid-binding protein, brain	5.40/14889	N/A	N/A	[7]
	FABP5	q01469	Fatty acid-binding protein, epidermal	6.59/15164	N/A	120.00	[7]
	GATM	P50440	Glycine amidinotransferase, mitochondrial	8.26/48455	N/A	N/A	[7]
	PAI1	p05121	Plasminogen activator inhibitor 1	6.68/45060	N/A	0.49	[7]
	SPRC	p09486	SPARC	4.73/34632	0.267	9.74	[7]
	LMNB1	P20700	Lamin B1	5.11/66408	0.032	26.72	[7]

Table 1: Virtual 2DE maps of proteoforms coded by the same genes are presented. The proteoforms detected in different spots but coded by the same genes, ANXA2 (ANXA2_HUMAN), ENO1 (ENOA_HUMAN), PKM (KPYM_HUMAN), or TPI1 (TPIS_HUMAN) (Supplementary Table S1), are shown. A round red marker indicates the position of the theoretical master proteoform (a form coded by the canonical sequence). More graphs can be found in Supplementary Figure S3.
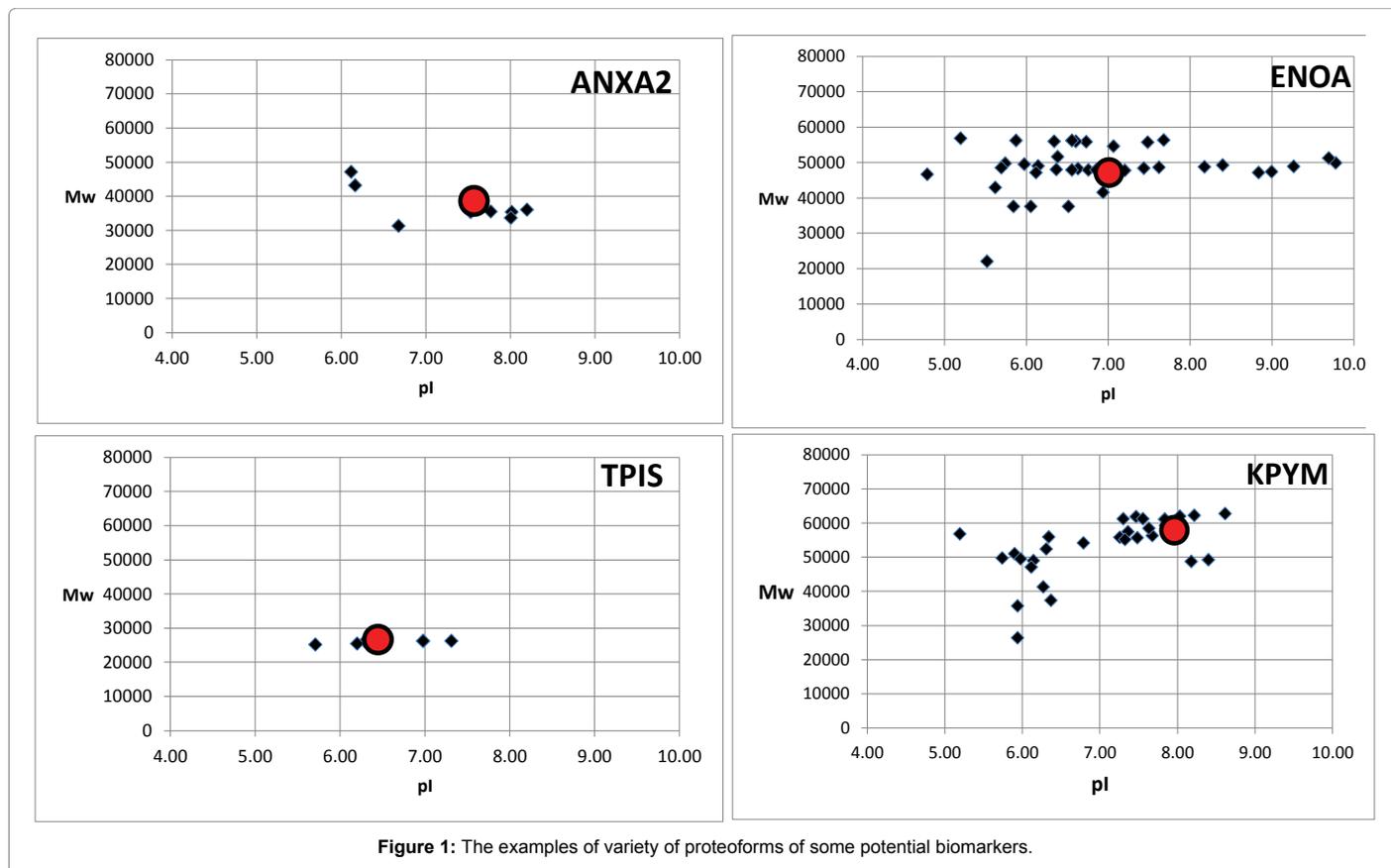
Figure 1: The examples of variety of proteoforms of some potential biomarkers.
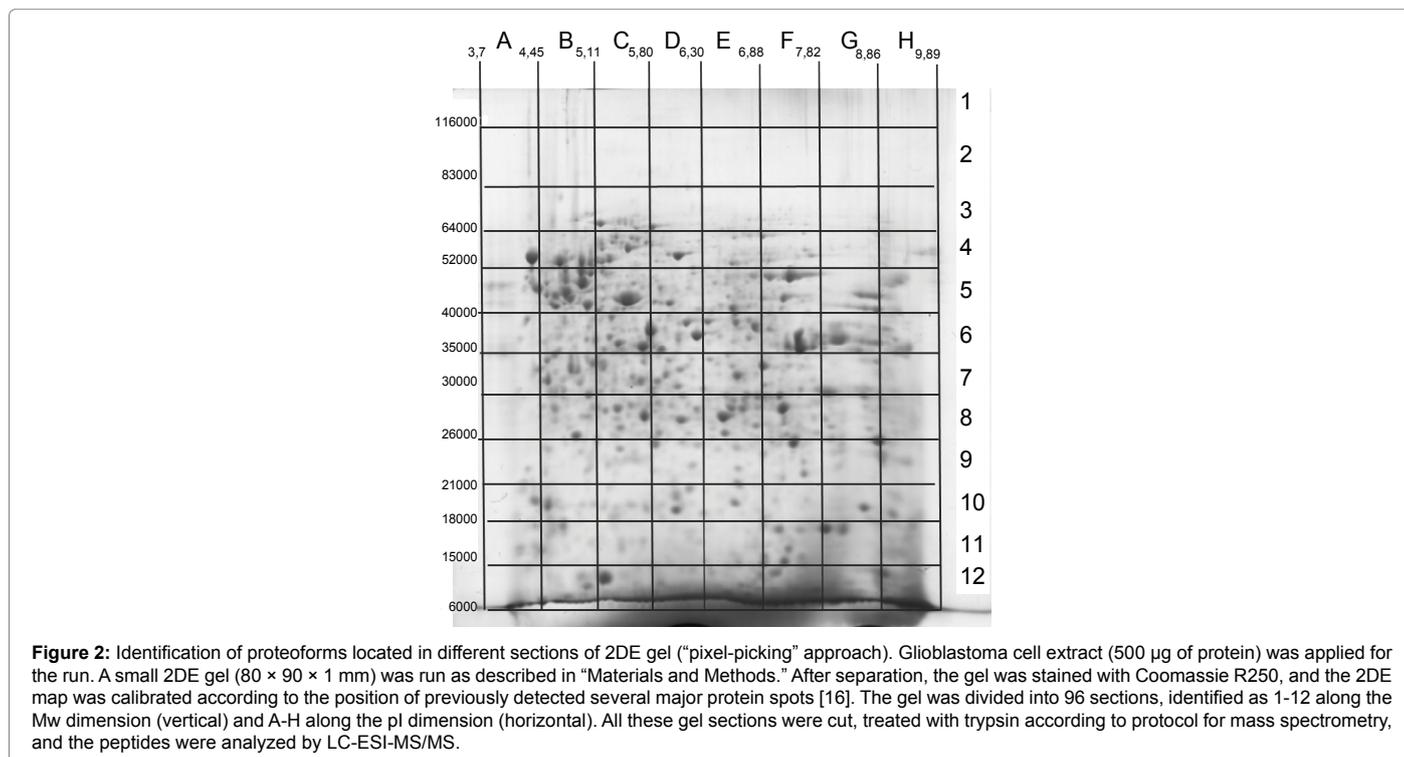
Figure 2: Identification of proteoforms located in different sections of 2DE gel ("pixel-picking" approach). Glioblastoma cell extract (500 µg of protein) was applied for the run. A small 2DE gel (80 × 90 × 1 mm) was run as described in "Materials and Methods." After separation, the gel was stained with Coomassie R250, and the 2DE map was calibrated according to the position of previously detected several major protein spots [16]. The gel was divided into 96 sections, identified as 1-12 along the Mw dimension (vertical) and A-H along the pI dimension (horizontal). All these gel sections were cut, treated with trypsin according to protocol for mass spectrometry, and the peptides were analyzed by LC-ESI-MS/MS.
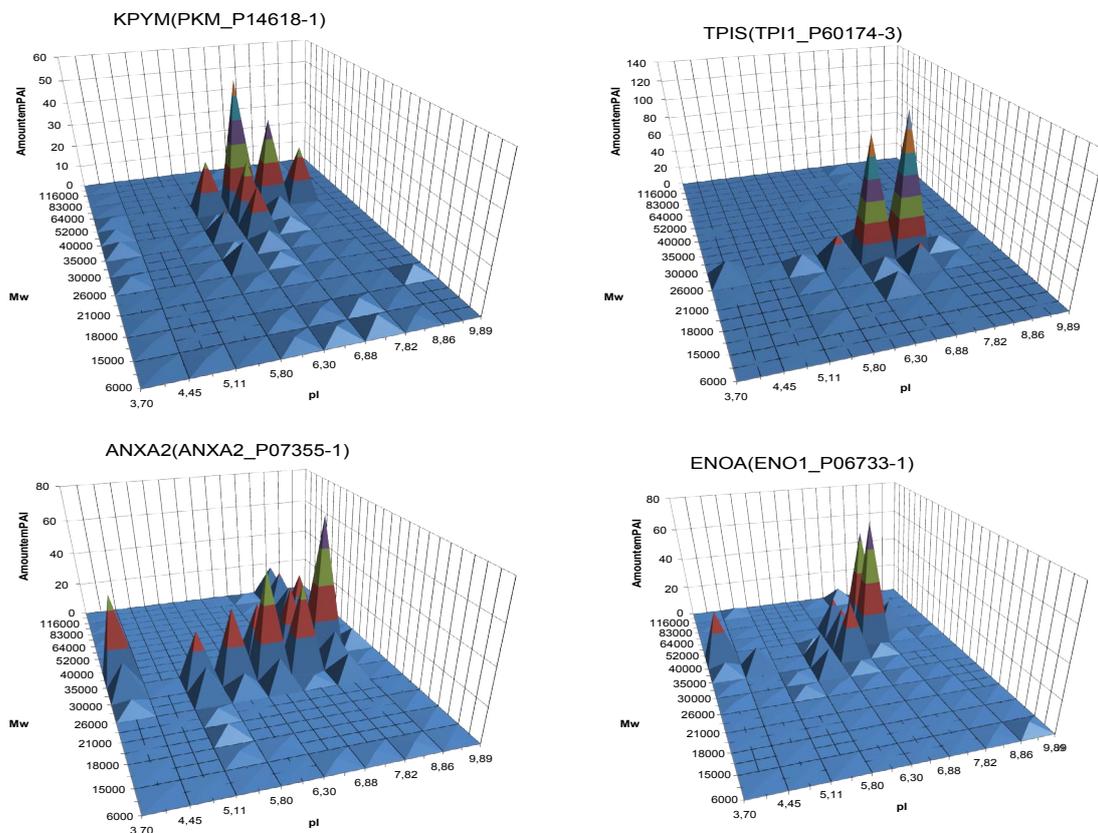
Figure 3: The examples of 3D graphs showing distribution of proteoforms between different sections of the 2DE map. The 3D graphs were generated from the data presented in Supplementary Table S2. A semi quantitative (estimated by emPAI) distribution of the same protein (gene) around the different gel sections was plotted. Distribution of proteoforms of ANXA2 (ANXA2_HUMAN), ENO1 (ENOA_HUMAN), PKM (KPYM_HUMAN), TPI1 (TPIS_HUMAN) is shown (Supplementary Table S2). More graphs can be found in Supplementary Figure S4.
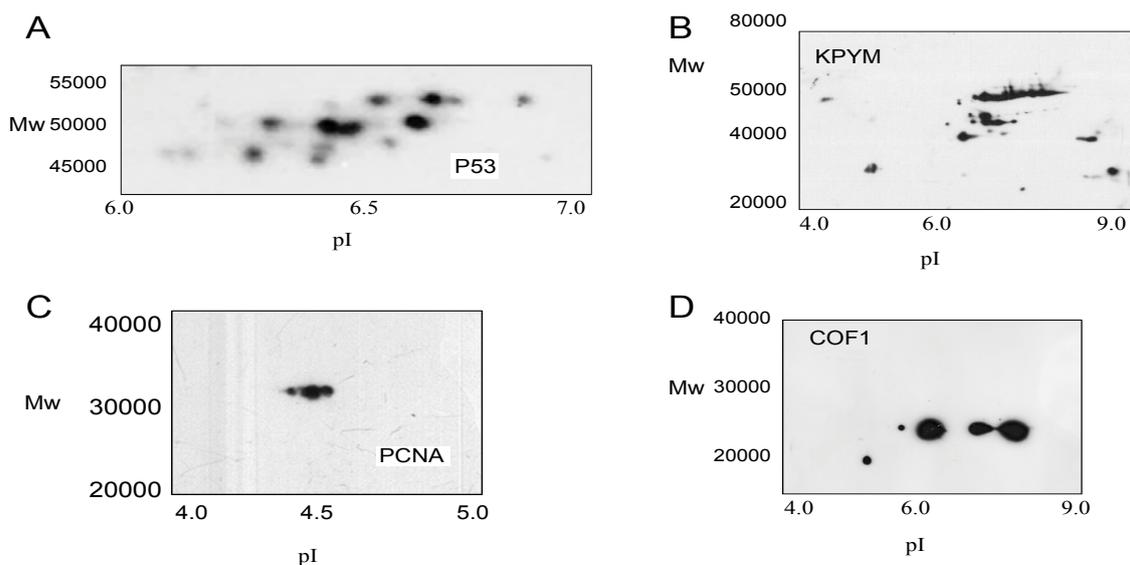
Figure 4: 2DE of glioblastoma cell extract with following Western-blot analysis. Glioblastoma cell extract (500 µg of protein) was run in a small 2DE gel (80 × 90 × 1 mm) as described in "Materials and Methods." After separation, proteins were transferred from the gel to the PVDF membrane that was treated according to the Blue Western blot protocol using available antibodies (see Materials and Methods). A− p53, B−KPYM, C−PCNA, D−COF1.

Western blot

In addition to detection by mass spectrometry we have applied Western blot for immunodetection of some proteins (Figure 4). It should be mentioned that we failed to detect protein p53 in the gel using mass-spectrometry (Supplementary Tables S1 and S2). But this protein seems to be a very promising biomarker of glioblastoma cells [4]. In our case, the levels of p53 proteoforms were not high enough for detection by mass-spectrometry. In opposite, Western blot nicely reveals multiple p53 proteoforms in glioblastoma cells (Figure 4). As far as the other tested proteins (KPYM, COF1, PCNA), Western blot analysis confirmed our mass spectrometry data (Figure 4).

Conclusion

A detailed 2DE map of proteoforms from the primary human glioblastoma cells was generated. The precise location and quantitation of 1542 proteoforms in 937 spots was performed. The less précised location (ΔMw=5000, ΔpI=0.8) and quantitation of 16012 proteoforms coded by 4050 genes was determined using our panoramic "pixel-picking" approach. The 2D and 3D graphs showing distribution of proteoforms of the same gene (for 4050 genes) in 2DE map were also built. We assume that due to a high variety of proteoforms, many potential glioblastoma biomarkers (KPYM, p53, ENOA, VIME, ACTB, CH60…) could be a source of highly specific biomarkers of glioblastoma. But this aspect warrants further investigation.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Acknowledgement

Conflict of Interest

The authors declare that there is no conflict of interest.

References

1. Furnari FB, Fenton T, Bachoo RM, Mukasa A, Stommel JM, et al. (2007) Malignant astrocytic glioma: genetics, biology, and paths to treatment. Genes Dev 21: 2683-2710.

2. Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, et al. (2007) The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol 114: 97-109.

3. Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, et al. (2005) MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med 352: 997-1003.

4. NaryzhnyÄ SN, Ronzhina NL, MaÄnskova MA, Beliakova NV, Pantina RA, et al. (2014) Development of barcode and proteome profiling of glioblastoma. Biomed Khim 60: 308-321.

5. Naryzhny SN, Ronzhina NL, Maynskova MA, Belyakova NV, Pantina RA, et al. (2014) Development of barcode and proteome profiling of glioblastoma. Biochemistry (Moscow) Supplement Series B: Biomedical Chemistry 8: 243-251.

6. Sallinen SL, Sallinen PK, Haapasalo HK, Helin HJ, Helén PT, et al. (2000) Identification of differentially expressed genes in human gliomas by DNA microarray and tissue chip techniques. Cancer Res 60: 6617-6622.

7. Zhang R, Tremblay TL, McDermid A, Thibault P, Stanimirovic D (2003) Identification of differentially expressed proteins in human glioblastoma cell lines and tumors. Glia 42: 194-208.

8. Cohen AL, Colman H (2015) Glioma biology and molecular markers. Cancer Treat Res 163: 15-30.

9. Hanash SM, Bobek MP, Rickman DS, Williams T, Rouillard JM, et al. (2002) Integrating cancer genomics and proteomics in the post-genome era. Proteomics 2: 69-75.

10. Kalinina J, Peng J, Ritchie JC, Van Meir EG (2011) Proteomics of gliomas: initial biomarker discovery and evolution of technology. Neuro Oncol 13: 926-942.

11. Shtam TA, Naryzhny SN, Landa SB, Burdakov VS, Artamonova TO, et al. (2012) Isolation and proteomic analysis of exosomes secreted by human cancer cells in vitro. Tsitologiya 54: 430-438.

12. Naryzhny SN, Lee H (2007) Characterization of proliferating cell nuclear antigen (PCNA) isoforms in normal and cancer cells: There is no cancer-associated form of PCNA. FEBS Lett 581: 4917-4920.

13. Naryzhny SN (2009) Blue Dry Western: simple, economic, informative, and fast way of immunodetection. Anal Biochem 392: 90-95.

14. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72: 248-254.

15. Naryzhny SN (1996) Upside-down stopped-flow electrofractionation of complex protein mixtures. Anal Biochem 238: 50-53.

16. Naryzhny SN, Maynskova MA, Zgoda VG, Ronzhina NL, Kleyst OA, et al. (2016) Virtual-Experimental 2DE approach in chromosome-centric human proteome project. J Proteome Res 15: 525-530.

17. Naryzhny SN, Zgoda VG, Maynskova MA, Novikova SE, Ronzhina NL, et al. (2016) Combination of virtual and experimental 2DE together with ESI LC-MS/MS gives a clearer view about proteomes of human cells and plasma. Electrophoresis 37: 302-309.

18. Larance M, Ahmad Y, Kirkwood KJ, Ly T, Lamond AI (2013) Global subcellular characterization of protein degradation using quantitative proteomics. Mol Cell Proteomics 12: 638-650.

19. Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, et al. (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 4: 1265-1272.

20. Koncarevic S, Urig S, Steiner K, Rahlfs S, Herold-Mende C, et al. (2009) Differential genomic and proteomic profiling of glioblastoma cells exposed to terpyridineplatinum(II) complexes. Free Radic Biol Med 46: 1096-1108.

21. Shim JH, Kim YS, Bahk YY (2006) Proteome profile changes that are differentially regulated by lipid and protein phosphatase activities of tumor suppressor PTEN in PTEN-expressing U-87 MG human glioblastoma cells. Proteomics 6: 81-93.

22. Bian XW, Xu JP, Ping YF, Wang Y, Chen JH, et al. (2008) Unique proteomic features induced by a potential antiglioma agent, Nordy (dl-nordihydroguaiaretic acid), in glioma cells. Proteomics 8: 484-494.

23. Vogel TW, Zhuang Z, Li J, Okamoto H, Furuta M, et al. (2005) Proteins and protein pattern differences between glioma cell lines and glioblastoma multiforme. Clin Cancer Res 11: 3624-3632.

24. Grzendowski M, Wolter M, Riemenschneider MJ, Knobbe CB, Schlegel U, et al. (2010) Differential proteome analysis of human gliomas stratified for loss of heterozygosity on chromosomal arms 1p and 19q. Neuro Oncol 12: 243-256.

25. An JH, Lee SY, Jeon JY, Cho KG, Kim SU, et al. (2009) Identification of gliotropic factors that induce human stem cell migration to malignant tumor. J Proteome Res 8: 2873-2881.

26. Trog D, Fountoulakis M, Friedlein A, Golubnitschaja O (2006) Is current therapy of malignant gliomas beneficial for patients? Proteomics evidence of shifts in glioma cells expression patterns under clinically relevant treatment conditions. Proteomics 6: 2924-2930.

27. Odreman F, Vindigni M, Gonzales ML, Niccolini B, Candiano G, et al. (2005) Proteomic studies on low- and high-grade human brain astrocytomas. J Proteome Res 4: 698-708.

28. Park CK, Kim JH, Moon MJ, Jung JH, Lim SY, et al. (2008) Investigation of molecular factors associated with malignant transformation of oligodendroglioma by proteomic study of a single case of rapid tumor progression. J Cancer Res Clin Oncol 134: 255-262.

29. Chumbalkar VC, Subhashini C, Dhople VM, Sundaram CS, Jagannadham MV, et al. (2005) Differential protein expression in human gliomas and molecular insights. Proteomics 5: 1167-1177.

30. Li Z, Bao S, Wu Q, Wang H, Eyler C, et al. (2009) Hypoxia-inducible factors regulate tumorigenic capacity of glioma stem cells. Cancer Cell 15: 501-513.

31. McNamara MG, Sahebjam S, Mason WP (2013) Emerging biomarkers in glioblastoma. Cancers (Basel) 5: 1103-1119.

32. Collet B, Guitton N, Saïkali S, Avril T, Pineau C, et al. (2011) Differential analysis of glioblastoma multiforme proteome by a 2D-DIGE approach. Proteome Sci 9: 16.