Expression and Characterization of Human Fragile X Mental Retardation Protein Isoforms and Interacting Proteins in Human Cells

Proteomics Insights
Volume 10: 1–9

DOI: 10.1177/1178641818825268
SAGE

Jiang Zhang[1], Guangli Wang[2], Wei-Wu He[2], Molly Losh[3], Elizabeth Berry-Kravis[4] and William E Funk[1]

[1]Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. [2]OriGene Technology, Inc., Rockville, MD, USA. [3]Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL, USA. [4]Departments of Biochemistry, Neurological Sciences and Pediatrics, Rush University, Chicago, IL, USA.

ABSTRACT: Fragile X mental retardation protein is an mRNA-binding protein associated with phenotypic manifestations of fragile X syndrome, an X-linked disorder caused by mutation in the *FMR1* gene that is the most common inherited cause of intellectual disability. Despite the well-studied genetic mechanism of the disease, the proteoforms of fragile X mental retardation protein have not been thoroughly characterized. Here, we report the expression and mass spectrometric characterization of human fragile X mental retardation protein. *FMR1* cDNA clone was transfected into human HEK293 cells to express the full-length human fragile X mental retardation protein. Purified fragile X mental retardation protein was subjected to trypsin digestion and characterized by mass spectrometry. Results show 80.5% protein sequence coverage of fragile X mental retardation protein (Q06787, *FMR1*_HUMAN) including both the N- and C-terminal peptides, indicating successful expression of the full-length protein. Identified post-translational modifications include N-terminal acetylation, phosphorylation (Ser600), and methylation (Arg290, 471, and 474). In addition to the full-length fragile X mental retardation protein isoform (isoform 6), two endogenous fragile X mental retardation protein alternative splicing isoforms (isoforms 4 and 7), as well as fragile X mental retardation protein interacting proteins, were also identified in the co-purified samples, suggesting the interaction network of the human fragile X mental retardation protein. Quantification was performed at the peptide level, and this information provides important reference for the future development of a targeted assay for quantifying fragile X mental retardation protein in clinical samples. Collectively, this study provides the first comprehensive report of human fragile X mental retardation protein proteoforms and may help advance the mechanistic understanding of fragile X syndrome and related phenotypes associated with the *FMR1* mutation.

KEYWORDS: fragile X mental retardation protein, mass spectrometry, characterization, post-translational modifications, interacting proteins

RECEIVED: December 12, 2018. **ACCEPTED:** December 21, 2018.

TYPE: Original Research

FUNDING: Support for this research was provided through a pilot grant from the Northwestern Clinical and Translational Sciences Institute (grant no.: 8UL1TR000150), as well as through the National Institutes of Health (NIH) (grant no.: R01MH091131).

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: William E Funk, Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA. Email: w-funk@northwestern.edu
Jiang Zhang, Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA.
Email: jzhang@uwalumni.com

Introduction

Fragile X syndrome (FXS) is the most common inherited cause of intellectual disability affecting approximately 1:4000-5000[1] and also the most common known monogenic cause of autism spectrum disorder (ASD). Fragile X syndrome is caused by a cytosine-guanine-guanine (CGG) trinucleotide repeat expansion of over 200 in the 5′-untranslated region (5′-UTR) of the fragile X mental retardation 1 gene (*FMR1*). Expansions over 200 undergo hypermethylation, leading to transcriptional silencing of *FMR1* and the loss or reduction of expression of the fragile X mental retardation protein (FMRP).[2] FMRP is an important protein involved in the localization and translation of neuronal mRNAs, processes that are essential for neuroplasticity and the development of neurons. Therefore, loss of expression of FMRP causes significant changes in neuronal structural morphogenesis, synaptic differentiation and formation, synaptic plasticity, neurotransmission properties, and consequently a range of phenotypic and behavioral manifestations associated with FXS.

Despite advances in the understanding of the functions of *FMR1* and its pathogenic association with FXS, the detailed mechanism of FXS at the protein level is not completely understood. The *FMR1* gene encodes a total of 11 known FMRP isoforms in humans, as a result of alternative splicing. These FMRP isoforms share a highly conserved N-terminal block of ~400 residues and variable C-terminal sequences with varying mRNA-binding affinities. The structure of FMRP contains multiple conserved binding domains, including two ribonucleoprotein K homology domains (KH domains) in the central region and a cluster of arginine and glycine rich domains (RGG box) near the C-terminus. These binding domains form an extended network of interactions with mRNAs as well as nuclear and cytoplasmic protein partners. Along with FMRP alternative splicing isoforms, two closely related human paralogs, fragile X-related protein 1 (FXR1P) and fragile X-related protein 2 (FXR2P) share over 60% amino acid identity with FMRP and possess similar KH and RGG domains and collectively constitute the *FMR1* RNA-binding protein family.

Fragile X mental retardation protein is highly expressed in neurons in the brain, particularly in the hippocampus, nucleus basalis, and granular layer of the cerebellum, with a primary association with ribosomes in nuclei and cytoplasm of neurons.[2,3] FMRP has a diverse array of functions through different areas and development stages of the neuron and has been implicated to play important roles in nucleo-cytoplasmic shuttling of mRNA, dendritic mRNA localization and synaptic protein synthesis.[4–6] Fragile X mental retardation protein interacts with and regulates the translation of ~ 4%-6% of brain mRNAs,[1] and studies have shown that post-translational modifications (PTMs) on the binding domains, including arginine methylation on the RGG domain and phosphorylation on the conserved Ser residues, affect the ability of FMRP to bind to mRNAs and protein partners. Given the molecular complexity of FMRP, a thorough characterization of FMRP proteoforms is necessary to provide important mechanistic insight into the role of FMRP in the pathogenesis of FXS.

Current knowledge of FMRP has been gained primarily from studies using immunocytochemistry approaches. Despite the sensitivity of these approaches, antibody-based methods generally lack the specificity to distinguish highly similar protein homologs and sequence variants arising from polymorphisms and mutations, as well as various PTMs. In addition, immunochemical methods are low throughput techniques and usually require long experiment time and laborious procedures. Recently, mass spectrometry (MS) has emerged as a promising alternative for the direct analysis of proteins, offering the advantages of shorter analysis time and improved sensitivity and specificity. Mass spectrometry analyses are based on the direct measurement of analyte mass (i.e. *m/z*) and are therefore able to provide unambiguous identification of various protein isoforms. When integrated with high performance separation techniques, MS is especially well suited for the analysis of complex analyte mixtures in a so-called "shotgun" MS approach, affording greatly improved analytical throughput and sensitivity compared to conventional immunochemistry methods. However, thus far, MS analyses of human FMRP have been scarce. Evans et al[7] expressed three FMRP isoforms in *Escherichia coli* cell lines and used peptide mass fingerprinting (PMF) to confirm the protein identities. Although PMF has been widely used for protein identification in relatively simple mixtures, this method only provides mass information at the peptide level, and the lack of MS/MS information poses a challenge for peptides with similar masses and for identifying PTMs. In this work, we expressed the full-length human FMRP (UniProt ID Q06787, *FMR1*_HUMAN) in human HEK293 cells transfected with purified *FMR1* cDNA clone. The expressed FMRP was purified by immune-precipitation and subsequently digested and analyzed using a bottom-up MS approach. We showed successful expression and identification of the full-length human FMRP, as well as two endogenous FMRP isoforms and homolog proteins FXR1P and FXR2P co-purified from the human cells, suggesting the potential interaction networks of human FMRP. Peptide quantification was performed to identify the high abundance tryptic peptides, which provides an important basis for the development of future targeted MS methods for directly quantifying FMRP in human blood samples.

Experimental Section

Protein expression and purification

HEK293 cells growing in cell culture dishes or flasks were transfected with OriGene RC222699 TrueORFGold cDNA clone, encoding a full-length *FMR1* transcript variant ISO6 fused with C-terminal Myc-DDK tag. RC222699 plasmid DNA was amplified and purified from *E. coli* and was mixed with MegaTran Transfection Reagent (TT200002) in an optimized ratio for transient transfection. After 2-3 days of transfection, the transfected cells were harvested, lysed in modified radio immune precipitation assay (RIPA) buffer (25 mM Tris-HCl pH7.6, 150 mM NaCl, 1% NP-40, 1 mM ethylenediaminetetraacetic acid [EDTA], 1× proteinase inhibitor cocktail mix, 1 mM phenylmethylsulfonyl fluoride [PMSF] and 1 mM Na_3VO_4), and clarified through high-speed centrifugation.

Clarified cell lysate was incubated with OriGene anti-DDK 4C5 monoclonal antibody affinity resin at 4°C for 2 hours to overnight. The amount of resin added to protein lysate and incubation time was adjusted based on the number of transfected cells. Expressed FMRP was purified by OriGene 4C5 monoclonal antibody affinity resin specifically recognizing the DDK tag at the C-terminus of the expressed FMRP. After incubation, protein bound to the anti-DDK affinity resin was collected by centrifugation for 5-10 minutes. Collected resin was washed with lysate buffer two to three times to remove non-specific binding and was then transferred into disposable polypropylene purification columns. The packed resin was washed again, and FMRP was eluted with 0.1 M of glycine buffer, pH2.7 and was immediately neutralized with 1 M Tris to a final pH range of 7.2-7.3. Peak fractions of eluates containing the highest amount of human FMRP were pooled and concentrated through a disposable membrane ultrafiltration unit or Vivaspin 20 with MWCO 10 000 Da (GE Healthcare) to a final 1 mL volume. The purified protein was characterized by NuPAGE Novex Bis-Tris 4%-12% gel in MOPS running buffer, and the protein concentration was determined by bicinchoninic acid assay (BCA) protein Assay (Thermo Scientific Inc).

Reagents and sample preparation

Chemicals were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise noted. Sequencing-grade trypsin was purchased from Promega (Madison, WI). All solutions were prepared in Milli-Q water (Millipore Corporation, Billerica, MA).

For the protein digest, 20 µL FMRP solution (0.1 µg/µL) was subjected to denaturing condition at 60°C for 15 minutes and then tryptic digestion at 37°C overnight. No reduction was used since no disulfide bonds are present in FMRP. The resulting tryptic peptide solution was desalted by ZipTip and prepared in the aqueous buffer (5% acetonitrile (ACN), 0.1% formic acid).

LC-MS/MS of purified FMRP

The peptides were loaded directly onto a 15-cm, 75-µM reversed phase capillary column (ProteoPep™ II C18, 300 Å, 5 µm size, New Objective, Woburn, MA) and were separated using a 70-minute gradient from 5% acetonitrile to 100% acetonitrile on a Proxeon Easy n-LC II (Thermo Scientific, San Jose, CA). The peptides were directly eluted into a linear trap quadropole (LTQ) Orbitrap Velos mass spectrometer (Thermo Scientific) with electrospray ionization at 350 nl/min flow rate. The mass spectrometer was operated in data-dependent mode, and for each MS1 precursor ion scan, the 10 most intense ions were selected from fragmentation by collision-induced dissociation (CID). The other parameters for MS analysis were as follows: resolution of MS1 was set at 60 000, normalized collision energy 35%, activation time 10 ms, isolation width 1.5, and $\geqslant +4$ charge states were rejected.

Data analysis and protein identification

The data were processed using Proteome Discoverer (version 1.4., Thermo Scientific) and searched using Mascot search engine. The peptide MS/MS data were searched against a customized database containing all 11 human FMRP isoforms in the UniProt database concatenated with the SwissProt human protein database (version June 2014) to identify the proteins in the purified fraction. The search parameters were set as follows: (1) enzyme specificity: trypsin; (2) variable modification: methionine oxidation and N-terminal acetylation; (3) precursor mass tolerance was ±10 ppm; and (4) fragment ion mass tolerance was ±0.8 Da. All the spectra were searched against target/decoy databases, and results were used to estimate the q values in Percolator algorithm as embedded in Proteome discoverer 1.4. The peptide identification was considered valid at q value <0.1 and were grouped for protein inference to satisfy the rule of parsimony. Furthermore, in the final protein list, protein identification was considered only valid if supported by a minimum of one unique peptide.

Results and Discussion

FMRP expression and purification

The expressed and purified FMRP was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE; Figure 1). The main band of the SDS-PAGE result shows ~MW 73 000, consistent with the size of the full-length

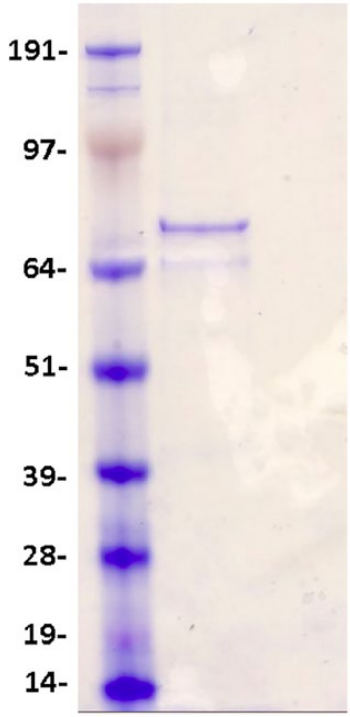

Figure 1. SDS-PAGE of the expressed human FMRP. The band at ~MW 73 000 corresponds to the expressed full-length FMRP. The band of lower intensity at ~MW 69 000 are FMRP isoforms co-purified from the human HEK 293 cells.

FMRP (*FMR1*_HUMAN, isoform 6) with the C-terminal expression tag. In addition, SDS-PAGE shows that the presence of a less-abundant protein band corresponding to ~MW 64 000-69 000, suggesting possible co-purified proteins may be interacting with the expressed FMRP or fragments of the purified protein. This result indicates the successful expression and purification of the full-length 73 kDa human FMRP from the human HEK293 cells.

Despite the successful expression of the full-length human FMRP, it is challenging to harvest the protein in high concentration and yield from the HEK293 cells. The highest protein concentration that can be obtained by enrichment is ~0.1 µg/µL, or ~1 µM. This suggests the low expression level of FMRP in HEK293 cells or precipitation of the protein at higher concentrations.

MS characterization of expressed human FMRP

Digested FMRP was analyzed by reversed phase HPLC tandem MS. Database search identified the human FMRP (Q06787, *FMR1*_HUMAN, isoform 6; the isoform 6 in the UniProt and other protein databases corresponds to the ISO1 of the cDNA transcript variant; notice the nomenclature difference between the UniProt protein database and GeneBank identifications) with sequence coverage of 80.5%, providing unambiguous identification of the FMRP protein with a false-positive rate (FDR) at 0.01 (99% confidence probability; Figure 2). The identified peptides include both the N- and C-terminal

Figure 2. Sequence coverage of the expressed full-length human FMRP (Q06787, *FMR1*_HUMAN, isoform 6). Underlines residues correspond to identified peptides by LC-MS/MS.
Bold font colored residues indicate residues with post-translational modifications (Ser phosphorylation and Arg methylation).

peptides, indicating the successful expression of the full-length FMRP (Figure 3). This result is consistent with the SDS-PAGE of the purified protein, supporting that the most abundant protein in the purified fraction is the full-length 632 residue FMRP isoform 6.

Given the potential roles of PTMs in the regulation of the FMRP function and mRNA/protein-binding properties, important relevant PTMs were examined for the peptide MS/MS data. Several PTMs including acetylation, Arg methylation, and Ser phosphorylation were identified. The N-terminal peptide of the FMRP (MEELVVEVR) was primarily found in the acetylated form (Figure 3A). Met1 acetylation is one of the most common co-translational covalent modifications of eukaryotic proteins, with an important role in the synthesis, stability, and localization of proteins.[8] About 85% of all human proteins are acetylated at their N-terminus,[9] and this is also observed in FMRP. In addition to N-terminal acetylation, multiple Met residues were found to be oxidized. The functional role of Met oxidation is generally believed to be a mechanism of antioxidation defense to protect other functionally more important residues from oxidative damage, as well as a regulatory component of cellular metabolism.[10,11] However, the specific role of Met oxidation in FMRP isoforms is yet unknown. Oxidized Met residues could be either a native PTM or an artifact of sample preparation, given the susceptibility of Met residues to oxidation.

MS/MS data identified Ser600 as a confident phosphorylation site (Figure 3B). Phosphorylation plays an important role in the dynamic regulation of FMRP function. For example, phosphorylation induces structural changes in the protein interaction domain due to modifications in hydrophobicity and charge state, which in turn affects its ability to bind mRNA and proteins. Previous reports have found phosphorylation at several sites within a highly conserved domain in vertebrate FMRP proteins as well as in lower order organisms. Ceman et al[12] used site-directed mutagenesis, phosphoamino acid analysis and precursor ion scan MS to study FMRP purified from murine brains and found that Ser499 is a primary phosphorylation site and possibly induces hierarchical phosphorylation of neighboring residues in the same domain, including Ser496, Thr501, and Ser503. However, the exact site of the phosphorylated residues could not be determined due to insufficient analytical sensitivity. Bartley et al[13] used total FMRP and pFMRP Ser499 antibodies and immunoblotting to study mice FMRP, and the results showed that mFMRP Ser499 phosphorylation is preserved in both heterozygous and a *Tsc1* knockout mice. Studies of *Drosophila* FMRP (d*FMR1*) also showed phosphorylation of Ser406, the conserved phosphorylation site corresponding to Ser500 in human FMRP.[14] These studies suggest a conserved Ser500 phosphorylation in human FMRP. However, phosphorylation of these sites was not observed in our MS data. Another possibility is the low stoichiometry of phosphorylation among these sites that is below the sensitivity limit of the method.

Arginine methylation is a functionally important PTM for FMRP in modulating its mRNA-binding activity. Several methylation sites including Arg290, 471 and 474 have been observed in our MS/MS data (Figure 3B and C). Methylation of Arg471 has been previously reported in mice.[15] However, Arg290 and Arg474 methylation have not been previously reported. The biological function of FMRP is also modulated by methylation of arginine residues of the RGG box domain.[16,17] Arginine methylation is important for protein-RNA interactions in general, and it has been shown that Arg544 is a key methylation site in human FMRP.[18] However, our experiments did not yield peptide information in the domain between Glu523-Arg572, primarily because of the high arginine content in the primary sequence of this domain, which upon trypsin digestion is cleaved into multiple short peptides that are not as efficiently detected by LC/MS analysis. As a result, it is possible that more methylation sites are present in this region that is not detected.

Identification of endogenous FMRP isoforms and interacting proteins

In addition to the full-length FMRP isoform, two additional FMRP proteoforms—isoform 4 (Q06787-5, *FMR1*_HUMAN) and isoform 7 (Q06787-7, *FMR1*_HUMAN)—were identified in our experiments. Figure 4A shows the sequence homology alignment of all three isoforms (identical sequences before Ser490 not shown). Fragile X mental retardation protein isoform 4 and isoform 7 share the same N-terminal sequence with the full-length isoform 6 but have C-terminal truncations of 25 and 17 residues relative to isoform 6, respectively. These two alternative splicing isoforms have a molecular weight of 68 455 Da and 69 179 Da, respectively. This is consistent with the less abundant band observed on the SDS-PAGE result (Figure 1). Our MS/MS results confidently identified novel peptides specific to these two isoforms, providing unambiguous identifications of these two splicing isoforms (Figure 4B and C).

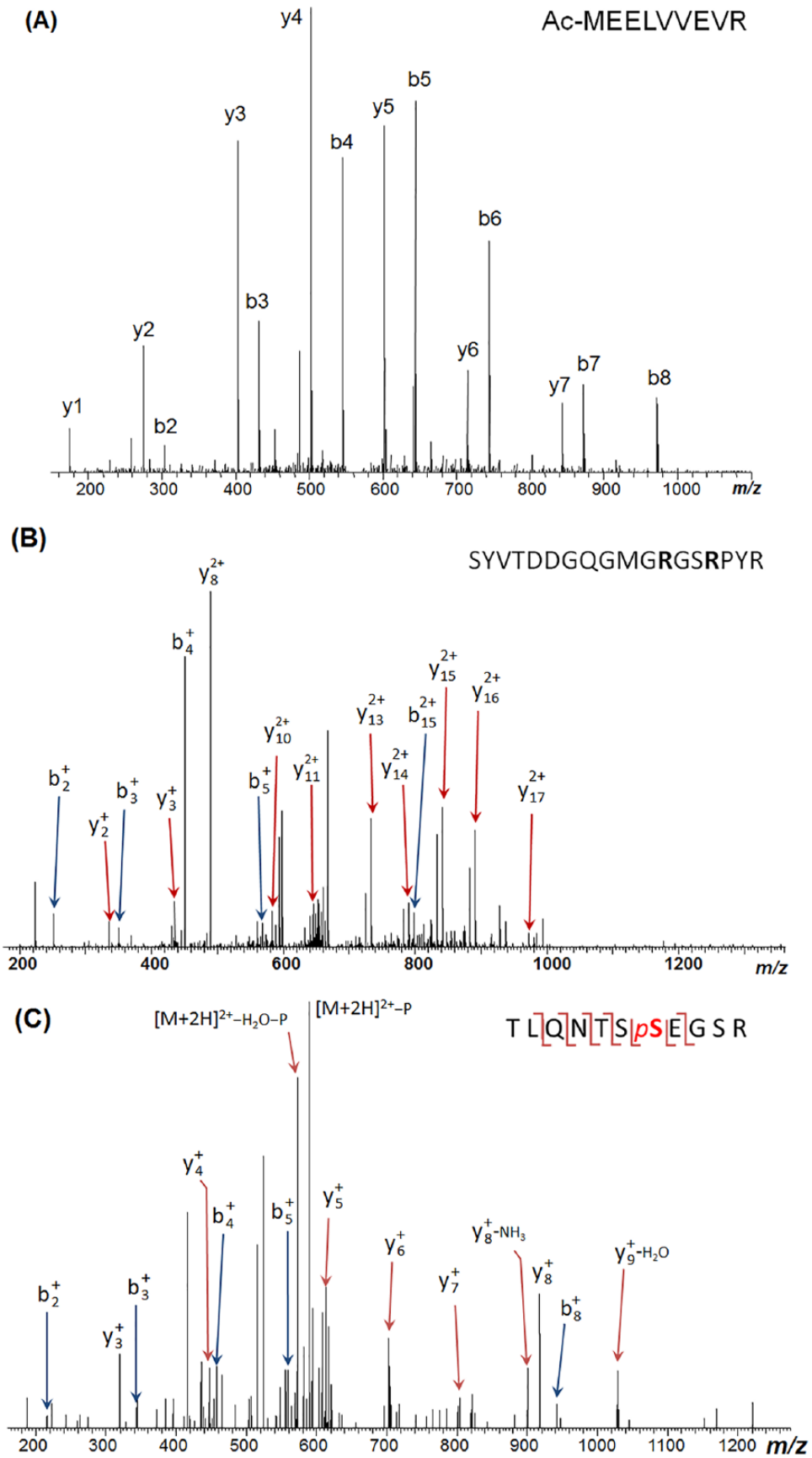

Figure 3. MS/MS data of identified FMRP sites with post-translational modifications: (A) the N-terminal acetylated peptide Ac-MEELVVEVR; (B) methylation at Arg471 and 474, SYVTDDGQGMG**R**GS**R**PYR. Bold font Arg residues indicate methylation; and (C) phosphorylation at Ser600, peptide TLQNTSpSEGSR.

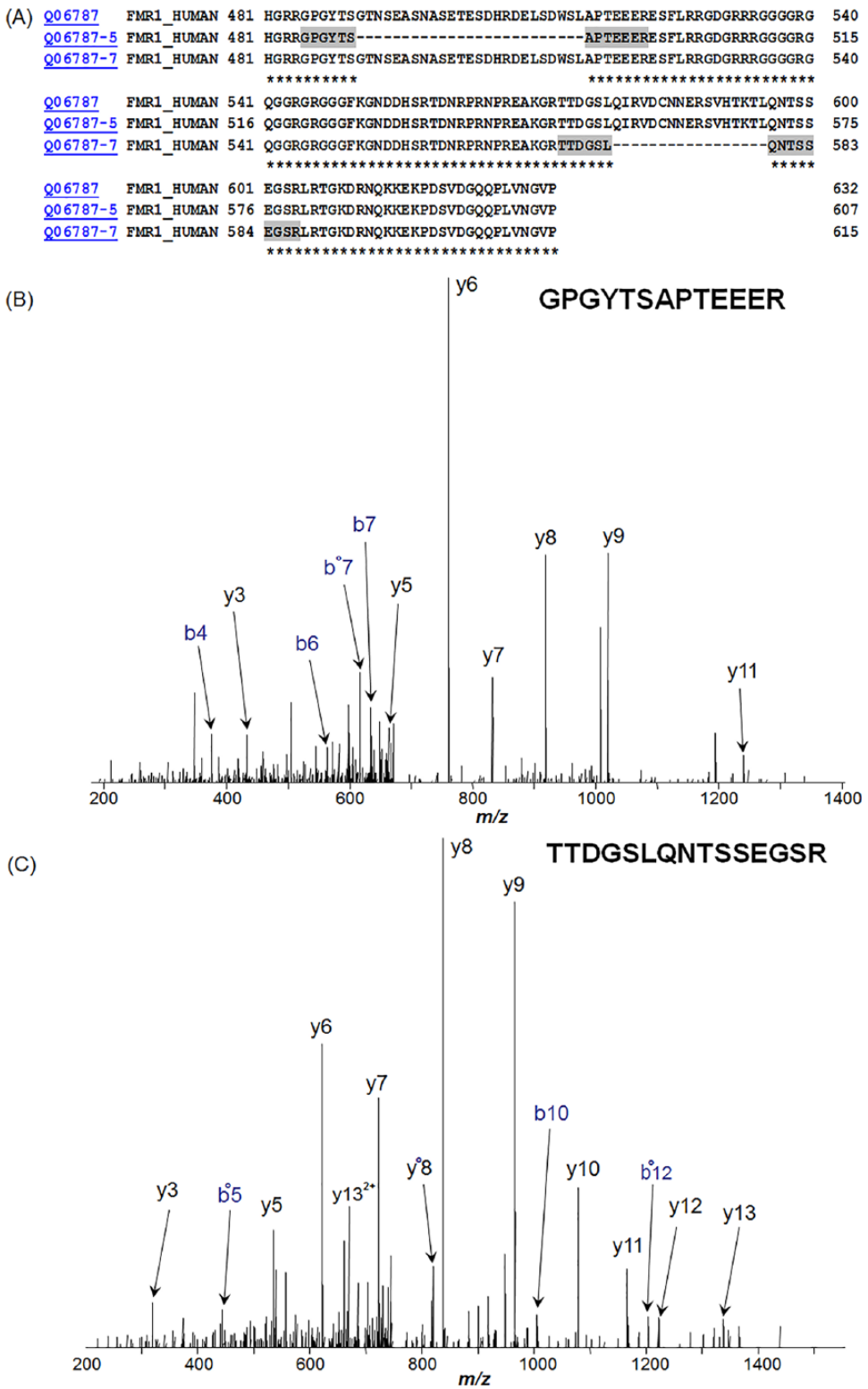

Figure 4. Identification of the FMRP isoforms 4 and 7. (A) Sequence alignment of identified FMRP isoforms by ClustalW2 algorithm. The peptides unique to the FMRP isoforms 4 and 7 are shade highlighted; (B) MS/MS data of the peptide unique to the FMRP isoform 4 (Q06787-5, *FMR1*_HUMAN); and (C) MS/MS data of the peptide unique to the FMRP isoform 7 (Q06787-7, *FMR1*_HUMAN). The square superscript denotes the ion with a loss of water.

The human *FMR1* gene has 17 exons that can undergo alternative splicing, resulting in potentially 20 different FMRP isoforms, 11 of which have been reported. The significance of the various FMRP isoforms has not been thoroughly determined but could be related to the different mRNA-binding affinities of the alternatively spliced isoforms. Therefore, the presence of these isoforms may capture underlying regulatory mechanisms related to FMRP activities. All FMRP isoforms have some

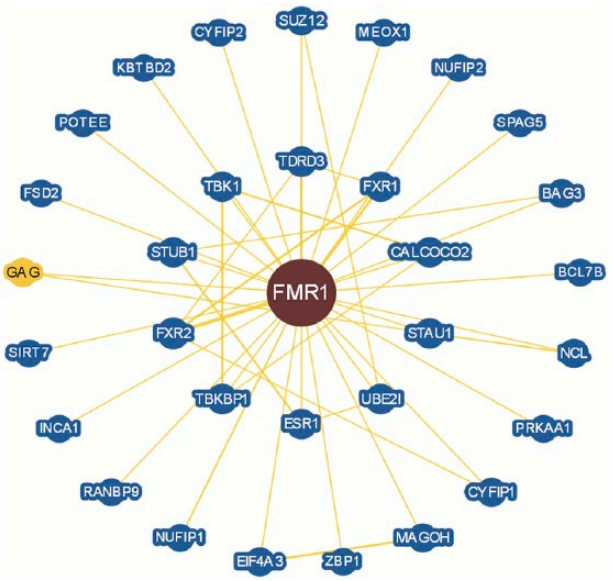

Figure 5. Protein interaction network of human FMRP, as determined by BioGRID. A number of bona fide FMRP-interacting proteins, including FXR1, FXR2, TDRD3, and NCL, were identified.

conserved structural motifs, including the KH and RGG box domains, which are essential for the FMRP function. However, truncation sites are in close proximity to the C-terminal region of the RGG box and therefore may result in different RNA-binding activities. Evans et al[7] reported that three FMRP isoforms had differential in vitro binding affinity for G-quadruplex RNA, suggesting the modulation of RNA-binding ability by naturally occurring sequence variants. Although the endogenous splicing isoforms of FMRP have been reported in other mammalian species such as mice,[19] the relative abundance of various human FMRP isoforms have not been reported thus far. Our study suggests that the FMRP isoforms 4, 6, and 7 are potentially the more abundant FMRP splicing isoforms in the human HEK293 cells.

Although the detailed cellular role of FMRP remains to be fully elucidated, it has been recognized that FMRP regulates mRNA translation through an extended network of interactions with mRNAs and protein partners. Particularly, the N-terminus of FMRP family is a highly conserved region, which is involved in a number of protein-protein and protein-RNA interactions. It has been reported that the N-terminal first 134 amino acids contain an independently folded domain which is responsible for the dimerization of FMRP.[20] FMRP forms homo- and heterdimers in vivo, which is consistent with the MS detection of isoforms 4 and 7, along with the full-length isoform 6.[21] FMRP also forms heterodimers and heteromultimers with FXR1P and FXR2P,[22-25] two close homologs in the FMRP protein family. Fragile X-related protein 1 and FXR2P share very similar overall structure with FMRP with ~60% amino acid identity, and they both have two KH domains and one RGG box that can bind RNA.[25] FMRP, FXR1P, and FXR2P have also been reported to be co-expressed in some tissues, which suggests compensatory functions.[22,23] Our experiments also confidently identified FXR1P (P51114, FXR1_HUMAN, 69,720.9 Da) and FXR2P (P51116, FXR2_HUMAN, 74,224.0 Da) with sequence coverage of 36% and 33%, respectively.

A number of other potential FMRP binding proteins were also identified in our experiments, including Bcl-2-associated transcription factor 1 (Q9NYF8, BCLF1_HUMAN, sequence coverage 33.26%), RNA-binding protein 10 (P98175, RBM10_HUMAN, sequence coverage 26.77%), tudor domain-containing protein 3 (Q9H7E2, TDRD3_HUMAN, sequence coverage 22.58%), double-stranded RNA-binding protein Staufen homolog 2 (E5RGT3, STAU2_HUMAN, sequence coverage 12.8%), nucleolin (H7BY16, NCL_HUMAN, sequence coverage 8.75%), and others. It is noteworthy that several of these proteins, including TDRD3, STAU2 and NCL, have been identified as bona fide interacting proteins for human FMRP. Figure 5 shows the protein interaction network of human FMRP as determined by BioGRID.[26,27] A catalog of co-identified proteins with sequence coverage above 15% is provided in the Supplemental Table 1. Detailed annotation of these proteins may yield more in-depth information regarding the interaction network of FMRP in the human cells.

Peptide quantification of FMRP and potential clinical implications

A special focus in our study was devoted to the quantification of tryptic peptides of FMRP, which may provide an important basis for the development of targeted FMRP quantification by multiple reaction monitoring (MRM) for measuring FMRP in human blood samples. Mass spectrometry-based targeted protein quantification has many advantages over conventional techniques such as enzyme-linked immunosorbent assay (ELISA) and Luminex assays, including higher analytical specificity for proteins with similar sequence variants, potential multiplexing, and higher analytical throughput. As a result, the development of a direct MS-based method for quantifying FMRP is of high interest for detecting FMRP in population-based research.

Table 1 shows several of the high abundance FMRP tryptic peptides that were identified in our MS experiments. Spectral counts are used in MS experiments to quantify peptides and proteins in the liquid chromatography—mass spectrometry (LC-MS/MS) experiments.[28,29] According to our results, peptides SFLEFAEDVIQVPR and MEELVVEVR provided the highest spectral counts and therefore represent the most sensitive peptides for targeted quantification analysis of FMRP in clinical samples. These results are consistent with those reported recently by Kusebauch and co-workers on large scale SRM-based targeted assay for human proteome.[30]

Although findings from this study are not likely to change current diagnostic practices in FXS (where FXS is confirmed

Table 1. Quantification of FMRP tryptic peptide abundance by the MS spectral count.

SEQUENCE	PROBILITY	MASCOT ION SCORE	SPECTRAL COUNT
SFLEFAEDVIQVPR	100%	69.34	75
MEELVVEVR	100%	70.36	52
VLVASSVVAGESQKPELK	100%	73.15	36
LIQEIVDK	99%	36.68	27
LQIDEQLR	99%	48.9	27
DINESDEVEVYSR	100%	84.36	19
VGPNAPEEK	100%	59.24	19
FPPPVGYNK	99%	32.39	18
EDLMGLAIGTHGANIQQAR	100%	68.42	17
DVHEDSITVAFENNWQPDR	100%	90.12	15
LDVPEDLR	99%	33.81	14
TTDGSLQIR	100%	70.09	13
SYVTDDGQGMGR	100%	58.75	12

Prob = probability, indicative of the peptide identification confidence; Mascot Ion score = Mascot algorithm metric for evaluating the identity of the peptide/protein; spectral count = based on total MS/MS data.

by CGG repeat length), more efficient measurement of FMRP and detection of peptides reflecting total FMRP (i.e. that are not isoform specific) could help to provide enhanced molecular characterization (and potentially with noninvasive assays, such as dried blood spots) that may inform the molecular genetic basis of clinically meaningful phenotypes in FXS. Specifically, given that FMRP deficiency leads to the neural changes resulting in the cognitive delays observed in FXS, accurate and efficient quantification of FMRP may hold potential for predicting phenotypic expression earlier in development, when behavioral interventions are most effective, or of understanding relationships between FMRP and clinical characteristics (e.g. severity of behavioral impairment, cognitive level) and/or relationships between FMRP expression and responses to targeted treatment directed at correcting the molecular deficit in FXS. In addition, the ability to reliably and accurately quantify FMRP will potentially allow an understanding of the role of relative FMRP deficits in other disorders such as ASD and schizophrenia, where more subtle deficits in FMRP may be evident.

Conclusions

We expressed the full-length human FMRP in the HEK293 cells and comprehensively characterized the purified proteins using bottom-up proteomics. Our MS results provided high sequence coverage of the expressed protein, indicating the successful expression of the 632 residue full-length FMRP isoform. Several functionally important PTM sites including Ser600 phosphorylation and Arg290/471/474 methylation were identified in the expressed FMRP. Two alternative splicing FMRP isoforms, as well as FMRP-interacting proteins, were co-purified and identified, providing information on the interaction network of FMRP. Peptides providing the highest spectral counts were also identified, which provides important reference for targeted analysis to quantify FMRP in clinical samples. Collectively, these results represent the first comprehensive MS characterization of FMRP purified from human cells, which could help advance the understanding of the role of FMRP in the pathogenesis of FXS.

Acknowledgements

The authors thank Drs Neil Kelleher and Paul Thomas at the Northwestern Proteomics for their assistance in the MS experiments.

Author Contributions

JZ was the lead author and performed the laboratory analyses and assisted with data interpretation, GW performed the protein expression and purification experiments, He performed the protein expression and purification experiments, ML assisted with data interpretation, EBK assisted with data interpretation, WEF assisted with study design and data analyses and interpretation. All authors reviewed the final manuscript.

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

1. Cook D, Nuro E, Murai KK. Increasing our understanding of human cognition through the study of fragile X syndrome. *Dev Neurobiol*. 2014;74:147–177.
2. Bardoni B, Schenck A, Mandel JL. The fragile X mental retardation protein. *Brain Res Bull*. 2001;56:375–382.

3. Kim M, Bellini M, Ceman S. Fragile X mental retardation protein FMRP binds mRNAs in the nucleus. *Mol Cell Biol*. 2009;29:214–228.
4. Jin P, Warren ST. Understanding the molecular basis of fragile X syndrome. *Hum Mol Genet*. 2000;9:901–908.
5. Jin P, Warren ST. New insights into fragile X syndrome: from molecules to neurobehaviors. *Trends Biochem Sci*. 2003;28:152–158.
6. Antar LN, Dictenberg JB, Plociniak M, Afroz R, Bassell GJ. Localization of FMRP-associated mRNA granules and requirement of microtubules for activity-dependent trafficking in hippocampal neurons. *Genes Brain Behav*. 2005;4:350–359.
7. Evans TL, Blice-Baum AC, Mihailescu MR. Analysis of the fragile X mental retardation protein isoforms 1, 2 and 3 interactions with the G-quadruplex forming semaphorin 3F mRNA. *Mol Biosyst*. 2012;8:642–649.
8. Polevoda B, Sherman F. N-alpha-terminal acetylation of eukaryotic proteins. *J Biol Chem*. 2000;275:36479–36482.
9. Van Damme P, Hole K, Pimenta-Marques A, et al. NatF contributes to an evolutionary shift in protein N-terminal acetylation and is important for normal chromosome segregation. *PLoS Genet*. 2011;7:e1002169.
10. Shechter Y, Burstein Y, Patchornik A. Selective oxidation of methionine residues in proteins. *Biochemistry*. 1975;14:4497–4503.
11. Levine RL, Moskovitz J, Stadtman ER. Oxidation of methionine in proteins: roles in antioxidant defense and cellular regulation. *IUBMB Life*. 2000;50:301–307.
12. Ceman S, O'Donnell WT, Reed M, Patton S, Pohl J, Warren ST. Phosphorylation influences the translation state of FMRP-associated polyribosomes. *Hum Mol Genet*. 2003;12:3295–3305.
13. Bartley CM, O'Keefe RA, Bordey A. FMRP S499 is phosphorylated independent of mTORC1-S6K1 activity. *PLoS ONE*. 2014;9:e96956.
14. Siomi MC, Higashijima K, Ishizuka A, Siomi H. Casein kinase II phosphorylates the fragile X mental retardation protein and modulates its biological properties. *Mol Cell Biol*. 2002;22:8438–8447.
15. Guo A, Gu H, Zhou J, et al. Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. *Mol Cell Proteomics*. 2014;13:372–387.
16. Blackwell E, Zhang X, Ceman S. Arginines of the RGG box regulate FMRP association with polyribosomes and mRNA. *Hum Mol Genet*. 2010;19:1314–1323.
17. Gary JD, Clarke S. RNA and protein interactions modulated by protein arginine methylation. *Prog Nucleic Acid Res Mol bBiol*. 1998;61:65–131.
18. Dolzhanskaya N, Merz G, Denman RB. Alternative splicing modulates protein arginine methyltransferase-dependent methylation of fragile X syndrome mental retardation protein. *Biochemistry*. 2006;45:10385–10393.
19. Brackett DM, Qing F, Amieux PS, Sellers DL, Horner PJ, Morris DR. ***Fmr1*** transcript isoforms: association with polyribosomes: regional and developmental expression in mouse brain. *PLoS ONE*. 2013;8:e58296.
20. Adinolfi S, Ramos A, Martin SR, et al. The N-terminus of the fragile X mental retardation protein contains a novel domain involved in dimerization and RNA binding. *Biochemistry*. 2003;42:10437–10444.
21. Tamanini F, van Unen L, Bakker C, et al. Oligomerization properties of fragile-X mental-retardation protein (FMRP) and the fragile-X-related proteins FXR1P and FXR2P. *Biochem J*. 1999;343:517–523.
22. Tamanini F, Willemsen R, van Unen L, et al. Differential expression of FMR1, FXR1 and FXR2 proteins in human brain and testis. *Hum Mol Genet*. 1997;6:1315–1322.
23. Agulhon C, Blanchet P, Kobetz A, et al. Expression of FMR1, FXR1, and FXR2 genes in human prenatal tissues. *J Neuropathol Exp Neurol*. 1999;58:867–880.
24. Schenck A, Bardoni B, Moro A, Bagni C, Mandel JL. A highly conserved protein family interacting with the fragile X mental retardation protein (FMRP) and displaying selective interactions with FMRP-related proteins FXR1P and FXR2P. *Proc Natl Acad Sci U S A*. 2001;98:8844–8849.
25. Zhang Y, O'Connor JP, Siomi MC, et al. The fragile X mental retardation syndrome protein interacts with novel homologs FXR1 and FXR2. *EMBO J*. 1995;14:5358–5366.
26. Stark C, Breitkreutz B-J, Reguly T, et al. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res*. 2006;34:D535–D539.
27. Breitkreutz B-J, Stark C, Tyers M. The GRID: the general repository for interaction datasets. *Genome Biol*. 2003;4:R23.
28. Liu HB, Sadygov RG, Yates JR. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem*. 2004;76:4193–4201.
29. Zhang J, Lanham KA, Heideman W, Peterson RE, Li L. Statistically enhanced spectral counting approach to TCDD cardiac toxicity in the adult zebrafish heart. *J Proteome Res*. 2013;12:3093–3103.
30. Kusebauch U, Campbell DS, Deutsch EW, et al. Human SRMAtlas: a resource of targeted assays to quantify the complete human proteome. *Cell*. 2016;166:766–778.