The Potential for Proteomics in Understanding Neurodevelopmental Disorders

Kelly L Wormwood1, Izabela Sokolowska1, Jeanne P Ryan2, Stephanie Russell3, Costel C Darie2 and Alisa G Woods1,2*

1Biochemistry & Proteomics Group, Department of Chemistry & Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY, 13699-5810, USA
2SUNY Plattsburgh Neuropsychology Clinic and Psychoeducational Services 101 Broad Street, Plattsburgh, NY, 12901, USA

Abstract

The field of proteomics and use of mass spectrometry as well as other proteomic techniques is continually growing, based on the interest of researchers for searching for potential biomarkers for numerous diseases and disorders. Currently protein biomarkers are not standardly used in clinical practice for diagnosing and monitoring neurodevelopmental disorders. The ability to predict the cause of a medical problem and aid in its treatment would be greatly increased with a biomarker or set of biomarkers. Most studies performed on neurodevelopmental disorders have been in the genomic field rather than in the proteomic realm. However, the few studies that have been performed in proteomics show promise for candidate neurodevelopmental protein biomarkers, both as research tools and in clinical use. Here, we describe potential protein biomarkers for neurodevelopmental disorders and the importance of proteomics in understanding neurodevelopmental disorders.

Keywords: Proteomics; Neurodevelopmental; Autism; SLOS; ADHD; Cholesterol

Introduction

Measuring biomarkers to predict or identify illness is a strategy that is increasingly used in many fields, including in the field of psychiatry and neurodevelopmental disorders [1-3]. The United States Food and Drug Administration (FDA) define a biomarker as an objective measurement of a normal or pathologic biological process, or an objective measurement that indicates response to a treatment [4,5]. Most biomarker studies of neurodevelopmental disorders concentrate on gene changes. Indeed, the recent August 2013 PubMed search “neurodevelopmental genomic” yields 985 articles but the search “neurodevelopmental proteomic” yields 26 articles (Figure 1). When searching for specific neurodevelopmental disorders (SLOS, autism, fragile X), there is also a very noticeable and significant increase in results for genomic studies versus proteomic studies (Figure 1).

Proteins are the active effectors of biological processes [6-11]. Protein biomarkers are understudied in neurodevelopmental disorders, and may have several advantages. The identification of protein biomarkers or biomarker signatures could provide targets for future therapeutics, and protein targets are more accessible than gene therapies. Protein biomarkers may give clues to the causes of disorders, may validate current theories and may provide a new method for treatment-monitoring. Protein biomarkers could be used in preclinical studies and clinical trials to monitor the effects of treatments. We may ultimately be able to understand the common pathways generally found in specific disorders through study of protein biomarkers [4,12]. Such analyses can be accomplished using proteomics and specifically mass spectrometry (MS) [4,5,13,14].

MS is a method that can reliably and comprehensively identify proteins and putative biomarkers [4,5,12-22]. MS is already in use for newborn screening for several other disorders, including phenylketonuria, congenital hypothyroidism and cystic fibrosis [23]. All mass spectrometers contain three essential sections including an ion source, mass analyzer and detector. The sample is ionized, using a specific method, by the ion source (e.g. electrospray ionization), producing ions. These ions then pass through the mass analyzer which separates the ions based on their mass to charge ratio (m/z). The ions are then detected (e.g. by an electron multiplier or a multichannel plate detector). A mass spectrum is then produced, which is a plot of ion abundance versus m/z.

In order for the sample to be analyzed by the mass analyzer, it first has to be ionized. This process takes place at the ion source, and the specific method is dependent upon the pH at which ionization occurs. There are two possible types of ionization including positive ionization (in which the compounds are analyzed under acidic conditions and are protonated) and negative ionization (in which the compounds are analyzed under basic conditions and are de-protonated). Positive ionization is used often when analyzing proteins and peptides [22]. Depending how the samples are ionized (whether the sample is liquid or solid), there are two ionization methods: Electrospray ionization (ESI), in which the sample is ionized from liquid droplets and Matrix Assisted Laser Desorption Ionization (MALDI) source, in which the sample is co-crystalized with a small molecule that can be easily ionized with laser beams [22]. There is no better or worse ionization method (nor a perfect mass spectrometer). Each method is best suited for a set of protein and peptide samples. Ultimately, these methods complement with each other [22].

MS-based proteomics is an excellent technique for protein biomarker identification for numerous different reasons. Using different proteomic approaches, one is able to detect and determine multiple properties of a specific sample [8,11,13,15,18,24-26]. Virtually all information about a protein/peptide such as their mass, sequence information and the charge state can be determined using
MS-based proteomics methods [13,20,27-30]. The first step in a proteomics experiment is sample fractionation, which can be done by non-denaturing biochemical approaches such as chromatography or centrifugation, or by denaturing methods such as acetone precipitation or electrophoresis (Figure 2). Once fractionated, the samples can be either further fractionated (i.e., by electrophoresis, followed by digestion with an enzyme (i.e. trypsin), or directly digested in solution. In both cases, the final outcome is a peptide mixture that is further analyzed by MS. Intact proteins, either fractionated or not, are also analyzed by MS. In order to accomplish MS analysis, the proteins or peptides must first be ionized by MALDI-MS or ESI. One advantage of ESI is that, when it is coupled with a high performance liquid chromatography (HPLC), the peptides or proteins are first fractionated by HPLC and then analyzed by MS, thus giving the MS more time to analyze more peptides/proteins. This specific method of coupling HPLC with MS, also called liquid chromatography-tandem mass spectrometry or LC-MS/MS, is very powerful, because the peptides are analyzed individually according to their elution time, rather than being analyzed simultaneously. This is a great advantage because it allows for a large number of identifications to occur in one experiment and is also a basis for improved quantitative analysis. This method is also able to provide protein information for thousands of proteins in a relatively short amount of time, which is very useful in biomarker discovery [6,7,12,22].

In LC-MS/MS, in addition to being fractionated over the HPLC, peptides are also analyzed in MS mode where the peaks that correspond to peptides are scanned by MS, their charge state is identified, and then the peaks are selected for fragmentation, fragmented by collision with an inert gas (i.e. Argon) and then analyzed by MS/MS. This procedure usually leads to identification of the amino acid sequence of a peptide that corresponds to the peak selected for fragmentation and fragmented, which in turn leads to identification of the protein the sequence of which contains that peptide, in a process called data processing and analysis. The data acquired by a mass spectrometer are usually processed into a form that is compatible with a database search (e.g. Mascot search engine). The end result of the proteomics experiment is identification of a protein or a list of proteins that are potential hits.

In addition to qualitative information (e.g. identification of proteins), MS analysis may also provide quantitative information about a particular protein or set of proteins. For example, some methods use absolute quantification through use of the isotopes. Among such methods are isotope-coded affinity tag (ICAT) [31],

![Proteomic vs. Genomic Studies for Neurodevelopmental Disorders](image1)

![Proteomic vs. Genomic Studies for Biological Disorders](image2)

Based on search conducted 08/22/13

*Figure 1: PubMed search results for neurodevelopmental disorders in genomics and proteomics.*
Stable isotope labeling by amino acids in cell culture (SILAC) [32], absolute quantitation (AQUA) [33], where quantification is performed at the cell/tissue, protein, peptide or MS sample. Other methods use chemical modification at the protein, peptide or MS sample level, such as isobaric tags for relative and absolute quantitation (iTRAQ) or differential gel electrophoresis (DIGE) [34]. Other methods monitor one (single reaction monitoring/SRM) or more (multiple reaction monitoring/MRM) [35], peptides and then proteins using addition of either label-based AQUA peptides or label-free internal standards (analog peptides), or even or spectral counting [36]. Spectral counting and label-free quantitation also gain in popularity. Overall, all these quantitation methods allow detection, identification and quantification of proteins or peptides.

A summary of a proteomics experiment, including data analysis is shown in Figure 2 and the principle of LC-MS/MS is shown in Figure 3.

Altogether, proteomics can reveal expression levels, complex protein-interaction networks and provide the opportunity to determine modifications at the protein level (post-translational modifications). Proteomics, in comparison with molecular genetics, also provides the opportunity to determine modifications at the protein level, making it potentially more closely related to pathophysiological processes underlying the clinical phenomenology of neurodevelopmental and other similar disorders [37,38].

MS is largely the method of choice in proteomic biomarker discovery however knowledge-based protein identification can also be employed. Such methods can include Enzyme-linked Immunoabsorbant Assay.
Tandem Mass Spectrometer

Figure 3: LC-MS/MS. The protein sample is fractionated by HPLC, then it is ionized and enters the mass analyzer, where a series of peaks are analyzed in MS mode and one of them (i.e. 785.05) is verified for the charge state (doubly charged) and then selected for fragmentation and fragmented in the collision cell. The MS fragments are analyzed in a second analyzed (MS/MS) and then all MS and MS/MS peaks are interrogated by database search and data analysis.

Figure 4: (A) a 90 minutes HPLC TIC. (B) At elution time point ~20 minutes, a MS spectrum was recorded. One of the peaks (circled) with m/z of 574.43, double charged (insert), was selected for fragmentation and produced a MS/MS spectrum (C). Analysis of the fragment b and y ions led to identification of the peptide LVGGDNLCSGR, which is part of human CD5 antigen-like protein.
MS has been more extensively used to identify human protein biomarkers in ASD than in other neurodevelopmental disorders. Investigation of protein biomarkers in ASD largely commenced without a priori hypothesis, i.e., screening for any potential protein biomarker that emerged was possible with proteomic technology. However, proteomic studies in ASD have largely supported theories of ASDs that propose that oxidative stress, heightened immune responses and cholesterol disturbances may underlie some forms of ASD. Indeed, preliminary putative biomarkers have been identified in human blood sera taken from people with ASDs. For example, alterations in complement proteins [40-42] as well as in apolipoproteins, may be present [13,38,40]. Corbett et al. [40] found that five peptide components corresponding to four known proteins (Apolipoprotein (apo) B100, Complement Factor H Related Protein (FHRI), Complement C1q and Fibronectin 1 (FN1)) had a MASCOT score of 30 or greater for autism compared to controls. Complement proteins are components of an immune response cascade. It was also found that Apo B100 and Apo A4 were higher in children with high functioning autism compared to low functioning autism. These findings could lead to the conclusion that cholesterol metabolism may also be dysregulated in individuals with autism due the role apolipoproteins have in this process. This conclusion is supported by a broad body of research, discussed in greater detail by Woods et al. [5]. For example, dysregulation of cholesterol metabolism-associated genes was found in ASD during a study of ASD and non-ASD sibling pairs [26].

Only one study has undertaken proteomic analysis of saliva in idiopathic autism, reporting that several salivary proteins undergo hypo-phosphorylation in ASDs, including statherin, histatin 1 and acidic proline-rich proteins [43]. These results were largely interpreted as providing biomarkers for ASD identification, however, the functional consequences of hypophosphorylation in salivary proteins is unclear. The authors of this report explained that such deficiencies in hypophosphorylation could point toward similar deficiencies in other secretory proteins, such as those involved in central nervous system development.

A more recent proteomic study examining blood plasma used surface-enhanced laser desorption/ionization time-of-flight (SELDI TOF) mass spectrometry to examine differentially expressed peptides in the plasma of children with ASD compared with non-ASD control subjects. This study found increases in the peaks of three peptides in the plasma of children with ASDs. These peptides corresponded to fragments of the C3 complement protein, once again suggested heightened immune system response in individuals with ASDs [41]. Further proteomic studies utilizing comprehensive approaches and modern technology will help to confirm the existence of immune and cholesterol disturbances in ASDs.

We should note ASD studies to date are limited for several reasons and perhaps one of the most important reasons is technology. This is not only true of ASD studies in particular, but also studies of in the following disorders described in this manuscript. For example, previous studies used instruments that identified and quantified only a limited number of proteins in the full proteome. Since the time of these studies MS has rapidly advanced. The current technology could both confirm and complement these studies, driving ASD biomarker research forward. Furthermore, there is no comprehensive study that uses a multidimensional approach such as combining two or more types of instruments, more than one type of protein digestion (in-gel and in-solution), as well as more than one type of software for data analysis. Furthermore, studies tend to analyze only one type of human biomaterial. Analysis of two types of biomaterials for the identification of ASD protein biomarkers could validate protein biomarker changes and could point toward the method that is most informative.

**Fragile X syndrome**

Proteomics in Fragile X syndrome (FXS) has not been fully explored in human studies, and has the potential not only to elucidate information about this genetic disorder, but also to provide information about other related neurodevelopmental disorders. FXS results from mutations or loss of the fragile X Mental Retardation (FMR1) gene and protein, and consequently malfunction of numerous proteins associated with the central nervous system (CNS) [44,45]. FXS can result in cognitive deficits, epilepsy, autism, and attention deficit hyperactivity disorder (ADHD) [46]. Studies of ASD may be relevant to FXS, since FXS is currently the best understood genetic cause of autistic behaviors [47,48]. FXS accounts for 5% of autism cases [49]. Not all children who develop FXS develop ASDs, between 15-60% of those with FXS do [47-50]. The reasons for this are not clear. We might therefore expect some overlap in protein biomarkers between FXS and ASD. Although no current medications are approved for FXS, clinical trials in humans are currently underway for medications such as arbaclofen [51] and minocycline [52], which address FXS symptoms. Endpoints in these clinical trials are based on behavior rather than biological markers. Biological markers for FXS in clinical trials would supplement behavioral measurements by providing additional information as well as validation of behavioral measures. Biomarkers might be earlier indicators of drug efficacy and could provide quantitative measurements of efficacy. The identification of biomarkers for FXS is therefore greatly needed right now to advance clinical studies of FXS and treatments for FXS, and protein biomarkers hold particular promise.

Proteomic studies in FXS animal models have indeed provided evidence for alterations of several proteins associated with the central nervous system [53-57]. In a FXS mouse cerebellum cell the expression of the fragile X mental retardation protein (FMRP) was altered, creating a model for cellular deficits found in FXS. The investigators found 56 proteins differentially expressed in the transfected cells, including 16 proteins whose levels were decreased and 40 whose levels were increased [53]. Using SILAC, a study of primary neuronal culture taken from an FXS knock-out mouse demonstrated changes in more than 100 proteins, many of which were associated with nervous system function [58]. FXS drosophila models have demonstrated alterations in actin-binding protein profilin and microtubulin-associated protein futsch, utilizing SILAC [59]. A separate drosophila study used a proteomic approach involving DIGE analyses followed by MS protein identification. The investigators found changes in monoamine-associated enzymes phenylalanine hydroxylase (Henna) and GTP cyclohydrolase (Punch) [56]. Using an embryonic drosophila FXS model, DIGE and MS, Monzo et al. [55], found alterations in 28 proteins including CCT subunits [55]. Changes in Hsp60B-, Hsp68-, Hsp90-related protein TRAP1 were identified by a different group using DIGE and MS to identify protein differences in an FXS fly model [57]. In FXS knock-out mice hippocampal synapse proteins were analyzed using iTRAQ. Changes in 23 proteins were observed relative...
to wild-type animals, including increases in Basp1, Gap43 and Cend1 [54].

There are many different protein alterations observed in these studies, leading to some difficulty in identifying clear markers. The differences observed may have to do with the different proteomic techniques and different tissues used. Possibly some resolution could be provided by a focus on human materials and in vivo studies. In addition, markers identified using human biomaterials could have direct clinical utility.

Very little proteome data has been collected from human FXS biomaterials. One early study found acetylation defects of annexin-A1 in leukocytes taken from individuals with FXS [60]. MS analysis of human blood sera found that the FMR1 intron 1 is hyper-methylated in FXS syndrome, and the degree of hyper-methylation is predictive of cognitive impairment [61-63]. This latter finding illustrates another potential use of protein biomarkers: for predicting symptomology.

Overall, proteomic studies in FXS have focused predominantly on animal models. Although this has provided interesting information, further studies in human subjects are greatly needed. Preliminary data points toward possible involvement of annexins, but further studies focusing on the proteome analysis of human biomaterials in FXS may reveal even more protein biomarkers. It might be expected that proteins involved in central nervous system development and signaling will ultimately become implicated. Because FXS is sometimes (but not always) associated with autism, differences in biomarkers in FXS with and without autism may shed light on autism causes and may identify autism-specific markers.

Smith lemi利 optiz syndrome

Smith-Lemli-Opitz Syndrome (SLOS) is an inborn deacrease of cholesterol synthesis associated with ASD symptoms [64-67]. Although this may be conceived of as a disorder of metabolism, rather than a neurodevelopmental disorder, autism and mental retardation are highly prevalent in SLOS. For this reason, the proteomic study of SLOS may help shed light on our understanding of neurodevelopmental disorders. Prenatal SLOS is predicted based on malformations and intrauterine growth retardation in prenatal ultrasonography and reduced maternal free estriol in serum. Diagnosis is confirmed via sterol analysis in a chorionic villus biopsy or amniotic fluid [68]. SLOS symptoms improve rapidly (though incompletely) with cholesterol supplementation and children with ASD without SLOS have been identified with low total cholesterol [65]. Total or partial deficiency of Dhcr7 causes SLOS [69]. As mentioned, several developmental disorders are associated with SLOS, including incomplete myelination, and mental retardation [70] as well as autism [64-67]. Tissue cholesterol and total sterol levels are substantially depleted, and 7-dehydrocholesterol levels increase SLOS and the Dhcr7 gene. High levels of 7-dehydrocholesterol inhibits the gene Hmgcr which exacerbates the cellular cholesterol deficit further [71]. However, effects on the protein carriers of cholesterol or other protein biomarkers are relatively unstudied in SLOS.

MS analyses of SLOS and SLOS animal models have focused on analysis of sterols [72-76] but not on proteomic analyses. Proteomic identification of other components of cholesterol metabolism, such as apolipoproteins could be of value for understanding the consequences of this disorder and for identifying unique therapeutic targets. As stated earlier in this manuscript, preliminary putative biomarkers have been identified in blood and saliva taken from individuals from other neurodevelopmental disorders, for example, idiopathic ASD. To give one example of putative biomarker identification, in human blood sera taken from people with ASDs, alterations in complement proteins [40-42] as well as in apolipoproteins, may be present [5,40]. Proteomic analysis of SLOS biomarkers has the potential to reveal new information about the disorder. Saliva may be a convenient biomaterial for biomarker discovery that could be collected from individuals with SLOS and analyzed, particularly since saliva collection is less disturbing to study participants than a blood draw. In all, 2290 proteins have been found in saliva compared with 2698 proteins found in plasma. Nearly 40% of proteins believed to be candidate markers for diseases such as cancer, cardiovascular disease, and stroke can be found in whole saliva. Therefore, saliva presents a non-invasive, accessible bodily fluid for analysis [77].

Because cholesterol supplementation is only somewhat effective for treating SLOS symptoms, further analysis of SLOS biomarkers promises to identify possible treatment targets and to elucidate perturbations present in SLOS. Apolipoproteins may be particularly interesting biomarkers, since they have been found to be dysregulated in another neurodevelopmental disorder, ASD, and could be used to supplement and carry cholesterol in individuals with ASDs. Further investigation of apolipoproteins in SLOS may yield promising data.

Attention deficit hyperactivity disorder (ADHD)

Proteomic analysis of ADHD is currently quite limited. One proteomic investigation of 99 children with ADHD compared to 175 non-ADHD control subjects examined levels of plasma protein glycosylation. This study used a newly developed HPLC-based glycan analysis method. Specific changes in glycosylation were observed in the ADHD group, specifically, increased antennary fucosylation of biantennary glycans and decreased levels of specific complex glycans. In contrast, changes in glycosylation were not observed in a 86 individuals with ASD compared with 165 individuals without ASD [78]. Further proteomic investigation in ADHD specifically is greatly needed.

Conclusions

Several neurodevelopmental disorders have been studied using proteomic technology, including ASD and DS. FXS has been studied predominantly using animal models and proteomic analysis of human biomaterials holds particular promise. Although MS studies of the neurodevelopmental disorder SLOS have been conducted, such studies have focused on sterols and lipidomics, rather than proteomics. Proteomic analysis, particularly of apolipoproteins, may hold particular promise for greater understanding of this disorder. Few proteomic investigations in ADHD have been conducted. Overall, proteomic biomarker discovery in neurodevelopmental disorders may promise new methods for diagnosis, treatment monitoring and treatment target identification and in addition could advance our general understanding of the underlying biology of these disorders.

Acknowledgements

This work was in part supported by the David A. Walsh ’67 Fellowship, awarded to KLW & CCD. This work was also supported in part by Mary Joyce, Robert Matloff, Ken Sandler and by the SciFund challenge contributors.

References


