

Potential of Proteomics as a Bioanalytic Technique for Quantifying Sleepiness

Nirinjini Naidoo, Ph.D.

Center for Sleep and Circadian Neurobiology, Division of Sleep Medicine, University of Pennsylvania, Philadelphia, PA

Currently, there are no simple ways to assess the degree of sleep loss in individual subjects. Proteins constitute greater than 98% of all molecules in the cell. They participate in physiological interactions, explain all of posttranslational modifications that occur in cellular micro-environments and thus are potential candidates for identification of a biomarker of sleepi-

ness. A variety of proteomic techniques are available which are being used in a current sleep deprivation study of monozygotic and dizygotic twin pairs we are performing.

Citation: Naidoo N. Potential of proteomics as a bioanalytic technique for quantifying sleepiness. *J Clin Sleep Med* 2011;7(5):Supplement S28-S30.

Sleep loss is common in the American population.¹ Sleep deprivation can result from a period of acute sleep loss or from insufficient sleep day after day. Sleep loss has a number of consequences. It leads to what has been termed wake-state instability. This results in lapses in performance and also compromises other aspects of cognitive function including executive attention and working memory. Sleep loss also has important metabolic and cardiovascular consequences. Epidemiological studies indicate an association between sleep loss and increased rates of obesity, type-2 diabetes and an increased risk of cardiovascular disease.¹

Currently, however, we do not have simple ways to assess the degree of sleep loss in individual subjects, thus we embarked on a study to identify a molecular biomarker for sleep drive. Biomarkers are usually associated with, or should yield information about, the cellular and molecular mechanisms of disease or a state. Biomarkers could be a direct cause of a disease or a state, a secondary player in disease initiation or progression, or merely be signal of a pathological condition. There are several ways to monitor molecular changes in the search for a biomarker. The best established methods are the genomic methods, which basically include cDNA microarrays. However, the use of proteomics (the large scale study of proteins) is now increasing, and metabolomics is a technique that holds promise for future investigation. In contrast to proteomics, metabolomics is the study of low molecular weight molecules. These include metabolites, such as peptides, amino acids, lipids and carbohydrates. The metabolome is believed to represent only about 2500 small molecules. The complexity of the various methodologies increases as one progresses from genomics and functional genomics to proteomics and metabolomics. In general, each methodology has advantages and disadvantages. Unfortunately, good stable platforms are uncommon, and these investigative techniques are really all works in progress.

Why should we study the proteome (set of proteins encoded by the genome) instead of the genome? Although the human genome project has generated substantial data, it has been ob-

served that only about 10% of the genome actually contributes directly to disease pathogenesis and an even smaller fraction of these genes are actual potential targets of therapeutics.² It should be emphasized that one of the goals of biomarker studies, besides identifying an altered state and diagnosis, is to generate therapeutics.

Why look at proteins and proteomics? Proteins constitute greater than 98% of all molecules in the cell; they govern metabolic processes and directly dictate cellular fate. Proteins participate in physiological interactions and explain all of posttranslational modifications that occur in cellular micro-environments. While the genome is very stable and relatively unchanging, the proteome is very dynamic. It changes from minute to minute and responds to tens of thousands of intra- and extra-cellular environmental signals. While gene sequence largely determines a protein's chemistry and behavior, the number and identities of other proteins presently in the same cell also influence it. This could be both an advantage and a detriment because more variations and dynamics increase the complexity of proteomic analyses.

There are several proteomic approaches that one can use to identify a biomarker. The most common is 2D gel electrophoresis (2-DE) coupled with mass spectrometry (MS). The general methodology is as follows: a sample for analysis is fractionated and separated by 2-DE; the protein spots are detected by image analysis and then selected for identification. Identified protein spots are then digested to generate peptide fragments that are then subjected to MS. Matrix assisted laser desorption time of flight (MALDI-TOF) or electron ion spray (ESI) are the 2 major types of MS used. Subsequently, the mass to charge ratios of the peptide fragments are detected – these M/Z ratios are then searched against databases of known M/Z ratios to generate a protein identification. A second more unambiguous method uses tandem MS or MS/MS. With this approach, the peptide fragment is further cleaved by tandem MS to generate the primary sequence of the peptide and protein. DIGE (differential in gel electrophoresis) is a more recent de-

velopment in 2-DE allowing the comparison of 2 samples or 2 conditions to perform differential expression analysis through cyanine dye labeling within the same gel much like microarrays. Other methods include the use of HPLC coupled with mass spectrometry (LC MS), protein arrays, chemical arrays, antibody arrays and the original proteomic technique of 2-hybrid assays. Recently, we have successfully used DIGE and mass spectrometry to identify several proteins that change with age and sleep deprivation in mice cerebral cortex thus demonstrating the feasibility of using these techniques to identify a sleepiness biomarker in humans.³

As previously noted, we have recently begun a study to identify a biomarker for sleepiness in humans. Thus, the data being presented herein are very preliminary. In a 36 h sleep deprivation and performance study carried out in 56 monozygotic and 42 dizygotic twin pairs, we determined that the behavioral response to sleep loss has high heritability, i.e., 0.80. However, there is a large difference between individuals in how affected they are by sleep loss; some individuals are relatively resistant while others are markedly affected.⁴ Therefore, we subsequently obtained plasma samples every 4 h during baseline (normal sleep/wake), during sleep deprivation and then recovery sleep from 10 individuals (only one member of any twin pair) who had the lowest behavioral response to sleep deprivation, i.e., few lapses on psychomotor vigilance testing (PVT), and 10 individuals (only one member of any twin pair) who had the highest behavioral response to sleep deprivation (high responder) for microarray and proteomic studies. Thus, 2 groups of subjects are represented in the study—a group of high responders to PVT and a group of subjects that are considered low responders to PVT during sleep deprivation.

Greater specifics about our study protocols are as follows. After giving informed consent, subjects had a medical history and physical exam. Afterward, they completed depression, alcohol questionnaires, morning and evening questionnaires, and gave blood for a complete blood count, chemzymes and for DNA extraction. They did a urine drug screen and had an overnight, unattended sleep study. Next, they completed a sleep diary and wore an actigraph for two weeks. They then came back to the sleep center where they were admitted to the Clinical and Translational Research Center (CTRC) in the early evening. The history and physical examination were repeated and the actigraphy data reviewed. On day one in the CTCRC, they initiated baseline sleep one hour before normal bedtime. On day two, they started 38 h of sleep deprivation beginning one hour after normal rise time (at 8:00 a.m.). They had blood sampling before starting sleep deprivation and were administered a PVT every two hours. On day three, they had blood sampling after ending the period of sleep deprivation, during recovery sleep at their normal bedtime and upon discharge on day four.

The PVT assessments were done in two-hour intervals. Each trial was ten minutes in duration with multiple reaction time tests at random intervals during each 10-min trial. We have 19 PVT tests where we extracted performance lapses for each trial with a reaction time of greater than 500 milliseconds. The demographics of the subjects were as follows; men and women between the ages of 18 and 50 yr. They were Caucasian and African American. The low responders had a rate of increase in PVT lapses in the range between 0.01 to 0.062 and a mean and

standard deviation of 0.043 ± -0.04 . For the high responders the range was 1.9 to 0.324 with the mean and standard deviation being 0.62 ± -0.51 .

To make this study design a little less complex, given that we have 19 time points and 2 important independent variables: Sleep deprivation status (i.e., a baseline, sleep deprivation and recovery days) and Response to sleep deprivation (i.e., high PVT responders and low PVT responders), we decided to pool samples at 3 timepoints. Pooling samples also reduces biological variability. Thus we have pools at 12 h, 24 and 36 h of deprivation in group one, which are the low responders during the baseline day and during the sleep deprivation day. Then in the high responders, we have a similar set up where we pool the blood samples from 12 h, 24 h and 36 h. We then carry out protein separation by carrying out immunodepletion, HPLC, mass spectrometry and protein identification and bioinformatics. In our plasma proteomic strategy, we will be assessing proteins that change expression over 36 h of sleep deprivation in both groups. The study will include a discovery strategy using pooled samples at 12 h, 24 h and 36 h of sleep deprivation and a validation protocol based on individual subjects. It is anticipated that proteins or biomarkers related to sleep drive should change expression progressively (either up or down) in relationship to increasing sleep drive with sleep loss. We also expect to see differences in the protein expression between the high and low responder groups.

We are using a 3D protein-profiling method for biological fluids in the discovery phase. This has been shown to detect more proteins than any other method used in the HUPO (human plasma proteomics) pilot project. It consists of immunodepletion, 1D- gel electrophoresis, nano-capillary ultra performance liquid chromatography, and tandem mass spectrometry. In the validation phase, we are planning to validate the most promising candidates by using MRM (multiple reaction monitoring) assays. Basically, we are looking to test about 20 candidates in all samples at all time points.

A real problem with doing proteomics and working with the plasma is that the plasma proteome is highly complex. It is the largest proteome. It has the largest number of proteins and it is the deepest, meaning it has the highest dynamic range (ten orders of magnitude).

There are very highly-abundant proteins in plasma; essentially, 10 proteins represent 90% of the plasma concentration. 50% of plasma is represented by albumin. Thus, the first step in our study is to carry out an immunodepletion protocol. Following immunodepletion, our samples are then prepared for mass spectrometry. They are first ethanol precipitated, then reduced, alkylated and run on 1-D gels. The gels are pixilated and trypsin digested and run on HPLC and subjected to mass spectrometry. We currently have several hundred proteins that are differentially expressed with sleep deprivation and after computational analyses and bioinformatics we expect to select a number of candidates for validation.

In the validation phase, we plan to use MRM analyses, using stable isotope-coded internal standard peptides to set up a calibration curve. Once that is completed, we will determine the magnitude of each of the proteins in the each of the 19 time points in the control and in the sleep-deprived state. It is anticipated that proteins related to sleep drive should change

expression progressively (either up or down) in relationship to increasing sleep drive with sleep loss.

To summarize, proteomics has the potential to yield multiple biomarkers. While it is certainly a very challenging process it has great promise to yield substantial information not only about state dependent protein expression levels but about the post-translational modifications and interactions of these proteins.

REFERENCES

1. Committee on Sleep Medicine Policy and Research Board on Health Sciences. Sleep disorders and Sleep Deprivation—An Unmet Public Health Problem. Washington, D.C.: National Academies Press, 2006.
2. Betz UA, Farquhar R, Ziegelbauer K. Genomics: success or failure to deliver drug targets? *Curr Opin Chem Biol* 2005;9:387–91.
3. Pawlyk AC, Ferber M, Shah A, Pack AI, Naidoo N. Proteomic analysis of the effects and interactions of sleep deprivation and aging in mouse cerebral cortex. *J Neurochem* 2007;103:2301–13.
4. Van Dongen HP, Baynard MD, Maislin G, Dinges DF. Systematic interindividual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. *Sleep* 2004;27:423–33.

ACKNOWLEDGMENTS

I would like to thank Dr. Allan Pack who initiated the biomarker for sleepiness study at the University of Pennsylvania, Erna Arnardóttir who is involved in the BSleep study, Dr. David Speicher for his collaboration with the mass spectrometry and NHLBI for funding. Editing of the conference proceedings supported by HL104874.

CORRESPONDENCE

Address correspondence to: Nirinjini Naidoo, Ph.D., Center for Sleep and Circadian Neurobiology, Division of Sleep Medicine, University of Pennsylvania, 2116 TRL, 125 S.31st Street, Philadelphia, PA 19104; Tel: (215) 746-4811; E-mail: naidoo@mail.med.upenn.edu

DISCLOSURE STATEMENT

Dr. Naidoo has indicated no financial conflicts of interest.