Protein biomarkers of alcohol abuse

Mariana P Torrente, Willard M Freeman, and Kent E Vrana*
Department of Pharmacology, Pennsylvania State University College of Medicine, 500 University Drive, Hershey, PA 17033, USA

Abstract
Alcohol abuse can lead to a number of health and social issues. Our current inability to accurately assess long-term drinking behaviors is an important obstacle to its diagnosis and treatment. Biomarkers for chronic alcohol consumption have made a number of important advances but have yet to become highly accurate and as accepted as objective tests for other diseases. Thus, there is a crucial need for the development of more sensitive and specific markers of alcohol abuse. Recent advancements in proteomic technologies have greatly increased the potential for alcohol abuse biomarker discovery. Here, the authors review established and novel protein biomarkers for long-term alcohol consumption and the proteomic technologies that have been used in their study.

Keywords
alcohol abuse and dependence; biomarkers; DIGE; ethanol; mass spectrometry; proteomics

Alcohol abuse and dependence are very prevalent disorders that can lead to a multitude of health and social issues. Heavy drinking is not only linked to liver disease, but is also causally linked to cancer and is known to aggravate illnesses such as diabetes and cardiovascular disease [1]. Moreover, alcohol abuse strongly correlates with a higher incidence of family problems, assaults, vehicular accidents and financial problems [2]. In total, the costs of alcohol abuse approach US$185 billion/year in the USA alone [3]. Furthermore, an estimated 3.8% of all global deaths and 4.6% of global disability-adjusted life-years are attributable to alcohol [4].

The inability to accurately assess alcohol drinking behaviors presents a significant barrier to the diagnosis and treatment of alcohol abuse. Brief surveys – including AUDIT-C [5] and CAGE [6] – aimed at quantifying alcohol intake through self-report of drinking behavior are still the “gold standard” for assessing alcohol use patterns [7]; however, these approaches have limited diagnostic utility, especially in those cases where individuals are motivated to deny or minimize the magnitude of drinking behavior in order to mitigate personal, professional or legal consequences of alcohol abuse [8], or in the context of altered mental states or mental illnesses [9,10].

Although acute alcohol consumption can be easily detected through the measurement of the blood and breath levels of ethanol itself, this measurement provides no indication about patterns of chronic alcohol use that are more directly related to the diagnosis of alcohol
abuse and dependence \[10,11\]. Moreover, quantifiable biomarkers to unequivocally assess alcohol consumption retrospectively across days or weeks have remained more elusive \[11,12\]. Currently, biomarkers for chronic alcohol abuse used in the clinic include nonprotein markers (e.g., mean corpuscular volume \[MCV\], ethyl glucuronide \[EtG\] and 5-hydroxytryptophol \[5-HTOL\]), as well as protein markers (e.g., carbohydrate-deficient transferrin \[CDT\] and \(\gamma\)-glutamyl transferase). However, these have yet to become highly accurate and as widely accepted as objective tests for other diseases \[10\]. In light of these shortcomings, there is a crucial need for the development of more sensitive and specific markers of alcohol abuse. Recent advancements in ‘omic’ technologies (i.e., genomics, proteomics and metabolomics) have greatly increased the capability for biomarker discovery. In this review, the authors will first provide a brief overview of currently available alcohol abuse biomarkers. Second, they will discuss the proteomic technologies that have been applied to the discovery and characterization of alcohol abuse biomarkers. Finally, the authors will focus their attention on noninvasively obtainable protein biomarkers, both established and novel, with a particular emphasis on work utilizing proteomic technologies.

**What is a biomarker? What alcohol abuse biomarkers are available?**

The term ‘biomarker’ is frequently used to describe any statistically significant biochemical or molecular change between two populations – in this case, populations with different drinking behavior \[11\]. A biomarker is defined by the NIH as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacologic responses to a therapeutic intervention” \[13\]. The difference between these two definitions lies in the former being the difference between means of two populations, while the NIH definition establishes that a biomarker must be informative for individual subjects, such that they would allow highly confident classification of individuals \[11\]. Thus, in the context of alcohol abuse, a biomarker would be an accurate indicator of an individual’s alcohol consumption over a particular period of time. It is important to note that a biomarker is defined by an association with a particular condition; it does not need to have causal or mechanistic relevance \[14,15\].

Alcohol biomarkers have important applications in medicine and public safety \[16\]. In the clinic, they not only provide an objective parameter of alcohol consumption to help diagnose alcohol abuse, but can also be used to track the progress of diseases related to alcohol abuse. Increasingly, alcohol biomarkers are used as objective measures of the performance of alcohol abuse treatment; accurate biomarkers could be used either to verify or even replace patient self-reports \[17\]. In public safety applications, alcohol biomarkers would be valuable to monitor abstinence in certain high-risk individuals such as pregnant women \[18\], and those previously convicted of a crime or persons in occupations that affect the well-being of the general public, such as patient care or air transportation \[17\]. Indeed, the single most preventable form of mental retardation is fetal alcohol syndrome. A critical advance would be the ability to accurately monitor pregnant women for abusive drinking behaviors.

As the utility of a biomarker lies in its diagnostic power, the ability of a biomarker to correctly classify subjects is of utmost importance. To quantify a biomarker’s ‘usefulness’, sensitivity and specificity are the most commonly used concepts. Sensitivity measures the fraction of actual positives that are correctly identified (i.e., the percentage of alcohol abusers who are correctly identified as such). On the other hand, specificity measures the fraction of negatives that are correctly identified (i.e., the percentage of nonabusers who are correctly identified as such). In an ideal situation, a test would have both 100% sensitivity and specificity, but this is generally unattainable for any test due to individual differences in genetics, environment, comorbidities and drinking phenotype. Although the goal for a
biomarker is to achieve the highest possible sensitivity and specificity, one could picture certain scenarios in which either one or the other is more important. For instance, if a biomarker was solely intended to initiate an investigation into a patient’s drinking behavior or referral to a treatment specialist, sensitivity at the expense of a degree of specificity would be tolerable (i.e., detecting all cases of alcohol abuse, while incorrectly including some nonabusers). Conversely, if a test were to have important consequences for the individual, specificity would be most important to avoid false-positives (i.e., erroneously classifying a parolee as an alcohol abuser, resulting in revocation of parole, even if this would fail to detect alcohol abuse in some parolees) [11]. It is important to acknowledge, however, that while biomarkers can contribute to a psychiatric diagnosis of an alcohol use disorder, these biological indices need to be supplemented with behavioral assessments.

Extensive research has been aimed at developing objective measurements of alcohol intake. Remarkable success has been achieved in measuring acute alcohol consumption, which can be easily detected through measurement of the blood and breath levels of ethanol itself. These tests, now ubiquitously used by police and medical staff all over the world, are unfortunately not capable of conveying information about patterns of chronic alcohol use over several days or weeks that are directly related to the diagnosis of alcohol abuse and dependence [10,11]. Thus, a reliable diagnostic tool able to retrospectively examine alcohol use over a long period of time is still lacking. Small molecules, proteins or protein adducts could theoretically be valuable. In the remainder of this section, the authors focus on notable biomarkers of chronic alcohol use that can be easily and noninvasively obtained from blood, plasma, urine or hair (Table 1).

Nonprotein alcohol abuse biomarkers

MCV

MCV, or simply red blood cell size, increases in subjects with alcohol consumption and normalizes slowly, after 2–4 months of abstinence [19]. The exact mechanism behind this effect is unknown, but ethanol appears to have direct hematotoxic effects [20]. The sensitivities and specificities of MCV are around 30–75 and 60–90%, respectively [9]. Furthermore, MCV is also elevated in patients with liver diseases, vitamin B12 or folic acid deficiency, hematological diseases, reticulocytosis or hypothyroidism [19,20], and thus, its utility as an alcohol biomarker is fairly limited.

EtG

EtG, a minor metabolite of ethanol formed by the conjugation of ethanol with activated glucuronic acid by glucuronosyl transferase [9,19], is a biomarker that is useful in determining alcohol intake in blood for up to 8 h, and in urine for up to 80 h after heavy ethanol consumption [19,21,22]. Incidental exposure to ethanol-containing products and even yeast and sugar may result in false-positives [23], while false-negatives can occur with certain urinary tract infections [21]. Recently, EtG found in hair has been proposed as a long-term marker of ethanol use [24,25]. EtG measurements for alcohol abuse in hair have a relatively high sensitivity and specificity, of 70–90 and 80–95%, respectively [9,24,25]. EtG found in nails – determined by liquid chromatography (LC)–mass spectrometry (MS) – has also been proposed as a potential alcohol abuse biomarker [26]. However, hair and nail EtG are still new tests and concerns about other sources of alcohol resulting in false-positives remain an issue. In addition to EtG, ethanol may also be metabolized into ethyl sulfate. Both the compounds are ethanol-specific metabolic products and can be used together to detect recent alcohol use with improved sensitivity [27,28].
**Total serum sialic acid**

The levels of total serum sialic acid (TSA) have been proposed as a marker for heavy alcohol consumption [29,30]. As chronic alcohol use prevents the glycosylation of many proteins (e.g., fibrinogen, complement proteins and transferrin), TSA significantly increases during alcohol abuse [31,32] and normalizes during alcohol withdrawal [30]. TSA is also elevated in saliva and urine in heavy alcohol drinkers [33]. TSA can be measured by two methods. The first involves hydrolyzing conjugated sialic acid residues from serum proteins with a strong acid or by the enzymatic action of neuraminidase, and then these are measuring by visible light absorbance. Alternatively, protein-conjugated sialic acid residues in serum samples can be hydrolyzed, and then followed by sialic acid purification and quantification using HPLC [29]. Although these methods are relatively simple, their use is not widespread in clinical laboratories [29]. Furthermore, some cancers [34], cardiovascular disease [35] and other pathologies [36,37] can also result in increased serum TSA. Free sialic acid has also been investigated as a biomarker for alcohol abuse, but it was not found to be a better marker than TSA [31]. The sensitivities and specificities of TSA have been found to be between 48–82 and 18–96%, respectively [19,30,38].

**5-HTOL & 5-hydroxyindole-3-acetic acid**

The serotonin metabolite 5-HTOL is a normal constituent of urine; its concentration increases dramatically after alcohol intake. Elevated 5-HTOL can be detected for 5–15 h (depending on dose) after ethanol exposure [39]. The ratio of the 5-HTOL to another serotonin metabolite, 5-hydroxyindole-3-acetic acid (5-HIAA), can also be used to verify the presence of ethanol in the body. The 5-HTOL:5-HIAA ratio also remains elevated for several hours after ethanol intake [19,40]. The 5-HTOL:5-HIAA ratio was found to have 100% sensitivity 4 h after a moderate dose of ethanol [41], but the reliability of this marker has been found to decrease fairly rapidly after 7 h [42]. This short time frame limits the diagnostic utility of these measures for assessing the history of ethanol abuse. Moreover, the HPLC-based methods currently used to determine the 5-HTOL:5-HIAA ratio are difficult to translate into routine clinical practice [19].

**Phosphatidylethanol**

Phosphatidylethanol (PEth) represents a unique phospholipid formed only in the presence of ethanol by the action of phospholipase D [19]. Since the formation of PEth is specifically dependent on ethanol, the diagnostic specificity of PEth as an alcohol biomarker is theoretically 100%. While PEth is formed in all cells by phospholipase D, for the purpose of serving as a biomarker of alcohol consumption, it is sampled in the cells of the blood where it can be readily accessed and measured. The half-life of PEth in blood is approximately 4 days [43]. Remarkably, PEth has been found to have a sensitivity of between 94.5 and 100%, and specificity of 100% [44-47]. Despite such high performance, the existing methods for detecting PEth may still be too challenging for routine clinical usage [19]. Moreover, its utility for the diagnosis of alcohol abuse is limited somewhat by the short-term nature (days) of this marker.

**Protein alcohol abuse biomarkers**

**CDT**

Transferrin is a liver protein that participates in iron transport. This protein exists in forms containing up to nine sialic acid residues, with four (tetrasialotransferrin) being the most common [48]. CDT refers to the minor varieties of transferrin with lower degrees of glycosylation, including asialo-, monosialo- and disialotransferrin, which contain zero, one and two sialic residues, respectively. Ethanol consumption increases serum concentrations of CDT, especially asialo- and disialo-transferrin [19]. Ethanol (or its metabolite
Acetaldehyde) appears to affect glycosylation/sialylation in the Golgi apparatus of hepatocytes, although the particular enzyme affected has not yet been discovered [49]. Trisialotransferrin is not included in CDT measurements as alcohol consumption has been found not to affect its levels [9].

CDT levels can be determined by electrophoretic, chromatographic and immunological methods, as well as more recently by MS [19,50-52]. Levels of CDT remain increased for 1.5–2 weeks [9]. The sensitivities and specificities of CDT are approximately 60–70 and 80–95%, respectively [9]. Serum CDT levels may be influenced by other conditions unrelated to alcohol use, such as anorexia nervosa [53] and pregnancy [54]. Additionally, CDT is altered by iron deficiency, chronic illnesses and menopausal status. False-negatives are associated with female gender, episodes of lower level alcohol use and acute trauma with blood loss [55]. Furthermore, CDT levels have been found to remain high in some individuals, even 6 weeks after the cessation of drinking [56]. Despite these limitations, CDT is currently considered to be the most useful single marker of alcohol misuse, and it is the only one approved by the US FDA for the detection of heavy alcohol consumption [7].

**γ-glutamyl transferase**

**γ-glutamyl transferase (GGT)** is a membrane-associated enzyme that transfers a glutamyl group onto certain amino acids. Although it is produced in many tissues, including the spleen, kidneys, pancreas, biliary tree, heart, brain and seminal vesicles, only liver GGT is detectable in blood [9,19]. GGT levels become elevated after alcohol intake; its sensitivity and specificity are approximately 40–60 and 80–90%, respectively [9]. The half-life of GGT is 14–26 days with a return to normal levels after 4–5 weeks of abstinence [57]. Its usefulness as an alcohol marker is limited due to the fact that many other conditions can cause an increase in GGT levels. In fact, GGT has been proposed for use as a marker for cardiovascular disease and Type 2 diabetes [58]. However, since the test for GGT remains very inexpensive and is conveniently included in routine laboratory testing, it remains the most commonly used marker for indicating acute alcohol-induced liver damage [9].

**Alanine aminotransferase/aspartate aminotransferase**

Alanine aminotransferase (ALT) and aspartate aminotransferase (AST) are enzymes that transform α-ketoacids into amino acids [19]. ALT is mainly present in hepatic tissue; AST (also known as serum glutamic oxaloacetic transaminase) is found predominantly in the liver but is also found in a large number of other tissues. Similar to GGT, elevated levels of these enzymes indicate generalized liver damage [9,59]. Predictably, the sensitivities for both of these enzymes in the context of alcohol abuse are low and vary greatly [19]. Nevertheless, akin to GGT, ALT and AST are commonly used as their determination is easy and inexpensive [59].

**β-hexosaminidase**

β-hexosaminidase (β-HEX) is a lysosomal hydrolase that is involved in the metabolism of carbohydrates and gangliosides in the liver. After heavy alcohol consumption, lysosomes are damaged and release the enzyme into the blood stream [29]. β-HEX, in both serum and urine, has long been known to be a very sensitive biomarker for chronic alcohol use [48,60]. The half-life of β-HEX in serum is approximately 6.5 days [19]. The sensitivity of serum and urinary β-HEX activity has been reported to be 69–94 and 81–85%, respectively, while the specificity of serum and urinary β-HEX activity is 91–98 and 84–96% [19]. However, elevated serum β-HEX occurs in patients with hypertension, diabetes, cirrhosis, myocardial infarction, in pregnancy and after oral contraceptive use [48,61]. Furthermore, the β-HEX assay is difficult to obtain in the USA, so clinicians have little experience with it [62].
Acetaldehyde adducts

Acetaldehyde is a product of oxidative ethanol metabolism. Its concentration after alcohol intake is highly variable with a life-time of approximately 3 h [63]. Circulating acetaldehyde has long been known to react with various proteins, resulting in aldehyde-protein adduct formation [64]. Some of these adducts have been detected up to 3 weeks after alcohol consumption. MS-based approaches aimed at detecting these for their use as biological markers of alcohol abuse have recently been developed [63]. Alternatively, circulating antibodies against acetaldehyde adducts have been directly measured as biomarkers of alcohol intake. The sensitivity and specificity of these are 65–73 and 88–94%, respectively [65].

Sialylation of Apo J

Apo J, or clusterin, is a highly sialylated protein thought to be involved in the exchange of lipids between different lipoproteins [29]. Chronic ethanol exposure decreases the sialylation of plasma Apo J [66]. The sialic acid index of Apo J (SIJ) refers to the ratio of moles of sialic acid per mole of Apo J protein. The number of sialic acids on Apo J is determined by immunoaffinity purification of Apo J, followed by hydrolysis of the sialic acid moieties and the spectrophotometric measurement of the amount of sialic acid. The amount of Apo J is determined by standard biochemical assays, and then SIJ is calculated [19,29,66]. SIJ levels are reduced with chronic alcohol intake and they then return to normal levels over a period of 8 weeks, with an approximate half-life of 4–5 weeks [66]. The large number of sialic acid residues on Apo J (28 mol sialic acid/1 mol Apo J) may allow for greater sensitivity among levels of drinking compared with CDT [29]. Moreover, SIJ has displayed a very high sensitivity and specificity in pilot studies (90–92% and almost 100%, respectively) [66], but additional work is needed to fully evaluate the usefulness of SIJ as an alcohol abuse marker. This is especially true given the complex procedure needed for its analysis.

Fatty acid ethyl esters

Fatty acid ethyl esters represent ester conjugates between fatty acyl chains (such as oleic acid, steric acid and palmitic acid) and ethanol (reviewed in [12,17]). These alcohol metabolites have been reported to be present in the blood for up to nearly 100 h in heavy drinkers, but currently have unknown specificity and sensitivity. Fatty acid ethyl ester may prove to have unique value through their measurement in meconium (first feces of the newborn; thereby providing evidence of drinking during pregnancy), as well as in the hair of the alcohol abuser (providing long-term insight into drinking behavior).

Cytokines

Cytokines are proteins implicated in cellular communication and activation; these proteins regulate processes such as inflammation, cell death, cell proliferation, cell migration and healing mechanisms. Circulating cytokines such as TNF-α, IL-1 and IL-6 are found to be elevated in both chronic and acute alcohol-induced liver disease (reviewed in [67]). Serum levels of TNF-α are higher in alcoholics than in the general population, regardless of alcohol consumption level [68]. A significantly increased production of IL-1β, IL-6, IL-12 and TNF-α has been observed among chronic alcoholics without liver disease and active alcohol intake. Interestingly, abnormally low levels of inflammatory cytokines were found in patients with alcohol liver cirrhosis who were actively drinking, while no significant changes in cytokine levels were observed in patients with alcohol liver cirrhosis who were in alcohol abstinence [69]. As the measurement of serum TNF-α levels has recently become routine in clinical practice [67], it is possible for circulating cytokines to aid in the diagnosis of alcohol abuse; however, given their broad biological role, it is unlikely that cytokines will be used as standalone alcohol biomarkers.
**Biomarker combinations**

While it may be tempting to think of biomarkers as single molecules, a growing body of evidence indicates that panels of biomolecules in combination may function best in terms of sensitivity and specificity. A ‘biomarker profile’ has been defined as a combination of distinct biomarkers, which when analyzed by a defined formula provides a diagnostic classification in regards to a specific condition/disease state [14]; in this case, alcohol abuse. For instance, the combination of CDT and GGT is more accurate than CDT alone [70,71]. Overall sensitivity and specificity are increased to 60–90 and 80–95%, respectively [9]. The Antilla Index has been used to combine CDT and GGT values into a mathematical formula to factor each test and establish a cutoff for an abnormal result, increasing sensitivity without sacrificing specificity [9,72]. CDT has also been combined with MCV, showing sensitivities and specificities of 60–95 and 80–95% [9]. Similarly, a combination of CDT, GGT, MCV and the small molecules homocysteine and folate has also been shown to have a higher sensitivity than the single markers [73].

Despite the number of putative alcohol biomarkers, aside from CDT, none of these have become widely accepted. This points to two needs in alcohol abuse research: more and larger validation studies to demonstrate biomarker utility in standard clinical settings and the need for the development of better markers of alcohol abuse. Recent advancements in ‘omic’ technologies have greatly increased the capacity for biomarker discovery.

**Proteomic methods for alcohol abuse biomarker discovery**

The ‘postgenome era’ has brought with it a wide array of new technologies. At present, the study of genomics, transcriptomics, proteomics and even metabolomics is not only relatively straightforward but also amenable to high-throughput analysis. Proteomics is defined as the analysis of many or all of the proteins in a given sample. Such analysis could entail studying thousands of proteins in a single-cell population [74]. The central premise of proteomics is that the comprehensive characterization of the proteins in a cell, tissue or organ will provide insights into the physiological status of the system [75]. Proteomics can be divided into three major branches or divisions based on the methodological considerations: the first one, structural proteomics, is the study of the physical arrangement of amino acids into a protein; this typically involves technologies such as x-ray crystallography and NMR spectroscopy. The second branch, functional proteomics, addresses the actual physiologic activity of proteins (e.g., enzyme activity, protein–protein interactions or interactions with other biomolecules) – normally using classic biochemistry approaches. Finally, expression proteomics focuses on the patterns of protein expression and modification in health and disease. This division has exploded with the advent of new high-throughput technologies for protein separation, quantification and identification [75].

One of the aims of proteomics is to identify biomarkers of disease [14]. Remarkably, a proteomic technique – 2D electrophoresis (2-DE) – was used over 25 years ago to show changes in the sera of humans suffering from alcoholism: proteins such as α₁-acid glycoprotein, IgA, α₁-antichymotrypsin, haptoglobin and Apo A-I lipoprotein were found to be elevated, while antithrombin III was found to be decreased [76]. A decade later, the same technique was used to search for biomarkers for fetal alcohol syndrome: eight proteins were found to be candidate biomarkers, again including α₁-antitrypsin and haptoglobin [77]. Protein identification by this technique remained laborious and has historically been limited to the most abundant proteins [78]. Fortunately, proteomic methods have greatly evolved since these early reports and now allow for comprehensive proteomic experiments to be carried out in mere hours. Such improvements in proteomics-based technologies have brought great expectations for the discovery of new protein biomarkers [79].

**Expert Rev Proteomics. Author manuscript; available in PMC 2013 June 01.**
authors will discuss some of the most commonly used modern proteomic approaches for the discovery of alcohol biomarkers.

**Protein separation**

The extent of proteome complexity demands highly efficient analytical platforms, using a combination of protein separation and identification. Two leading approaches for protein separation used in proteomics are electrophoresis and liquid chromatography; both of these allow for fractionation of complex mixtures according to the chemical and physical properties of the proteins involved [74].

**2-DE**—2-DE separates proteins in a sample based on their isoelectric point and their molecular weight [80]. In this technique, protein extracts are first applied onto a pH gradient gel and exposed to an electric current to induce the proteins to migrate through the gel; the distance and direction traveled by the proteins depends on their overall net electric charge. The proteins are then applied onto a second, size-exclusion gel and exposed to a second electric current flowing in a direction perpendicular to the first one. Under these conditions, the proteins migrate with the distance traveled at any time point depending on their molecular weight. With this strategy, potentially thousands of different proteins in an extract can be separated into individual spots with characteristic coordinates. Even proteins that differ only in their post-translational modification can often be distinguished. Protein spots can then be visualized by staining the gel with selective dyes. Subsequently, individual spots can be cut from the gel and the proteins extracted for further analysis. A variation of 2-DE, fluorescent 2D-DIGE is a form of gel electrophoresis where up to three different protein samples can be labeled with fluorescent dyes prior to 2-DE, such that two or three samples can be mixed and run in the same gel [81]. After electrophoresis, the gel is scanned with the excitation wavelength of each dye separately, so each sample is visualized separately. This technique can be used to track changes in protein abundance (e.g., control samples vs alcohol abuser samples). Since the proteins from the different sample types are resolved on the same gel, they can be directly compared, and thus 2D-DIGE overcomes limitations in 2-DE caused by intergel variation.

Although 2-DE is one of the workhorses of proteomics, it does have some caveats [82]. First, 2-DE can only separate approximately 1000–2000 proteins in a sample, and rare species will often fall below the level of detection. This concern is the subject of an excellent review on general biomarker development [79]. Besides low-abundance proteins, other groups of proteins that are difficult to analyze by 2-DE include very small and very large proteins, alkaline proteins and hydrophobic proteins. Second, ‘individual’ spots can sometimes contain two or more very similar proteins (with regards to their isoelectric point and mass), which can complicate subsequent analyses. Third, even with computer assistance, 2-DE is still a fairly laborious technique. Despite these limitations, 2-DE continues to be an important component of the proteomic toolbox.

**LC**—Another separation technique, LC separates mixture components in a column filled with a solid material (stationary phase). In LC, fractionation of a proteome sample may be accomplished at the protein or peptide levels [74]. Depending on specific protein or peptide characteristics, such as its size and electrical charge, each protein is retained by the stationary phase for a specific amount of time. Accordingly, proteins that interact more strongly with the solid material will remain in the column longer than proteins that interact weakly. Proteins are eluted from the column by a solution (mobile phase), under isocratic or more often gradient conditions [83]. Eluate fractions are collected as they leave the column, with each fraction containing one or more proteins or peptides. In HPLC, a pump provides high pressure to move the mobile phase and the sample through a densely packed column
with very small particles of the stationary phase. This allows for a better separation on columns of shorter length when compared with traditional, low-pressure LC. HPLC is frequently combined with MS; in this configuration, separated proteins or peptides are further characterized by MS (see below). As peptides have an overall hydrophobic character, the most utilized type of chromatography for proteomic experiments is reverse-phase LC, which employs a hydrophobic stationary phase. Peptide mixtures can also be separated chromatographically in 2D. For example, a biphasic tandem column configuration packed with strong cation exchange followed by reverse-phase materials can be used to separate the peptides first according to their charge, and then on the basis of their hydrophobicity. This technique is the central component of an approach known as multidimensional protein-identification technology [84].

LC and HPLC overcome many of the issues associated with 2-DE, such as difficulty of automation, low accessibility of membrane-bound proteins, and the detection of proteins with large molecular weights, high isoelectric points, strong hydrophobicity or low abundance. However, LC still suffers from a lack of reproducibility and reliable quantitation [74].

Protein identification

Two prominent approaches for protein identification used in proteomics are MS and affinity reagents (i.e., antibodies, aptamers and binding partners). The first method allows for unbiased, nondirected protein identification, while the second allows for directed, quantitative and high-throughput protein expression profiling.

MS—MS is an analytical technique used to identify and characterize proteins based on highly sensitive mass determinations. In proteomic applications, there are three main MS modalities: top down and middle down MS methods analyze intact proteins or large polypeptides, respectively [85]. More commonly used in the alcohol biomarker field, the bottom up methods enzymatically digests proteins into short peptides prior to MS analysis. Tryptic digestion is usually utilized to generate such peptides [85]. Subsequently, the masses of the resulting peptides are analyzed by one of the MS approaches described below. Following data collection, specialized software generates a list of the masses of all measured peptides that is then queried against databases of known and predicted proteins and the peptides they would produce if these proteins also were treated with trypsin. If a given number of the peptides from the unknown protein match to the peptides predicted for a known protein in a database, then the unknown protein has been identified. MS methods are very sensitive, highly accurate and can characterize proteins present in very small amounts [80].

Two MS approaches specifically employed in proteomic analyses to identify alcohol biomarkers include MALDI-TOF MS and electron-spray ionization combined with tandem MS (MS/MS) [80]. For MALDI-TOF MS, the digested proteins are mixed with an excess of ultraviolet-absorbing matrix. When irradiated with a laser beam of the appropriate wavelength, the excess matrix molecules sublime and carry the peptides into the gas phase. In this way, singly charged peptide ions are formed [86]. The ionized peptides then pass through accelerating grids and travel down a vacuum flight tube, with smaller ions travelling faster than larger ions. When the ions reach the end of the tube they strike a detector [87]. The ‘TOF’ required to reach the detector is used to calculate the masses of the peptides; these measured masses can be compared with databases of known proteins and their tryptic peptides for peptide identification. Indeed, the most advanced instruments estimate the peptide masses at extraordinarily high accuracy (in the ppm range; ±0.001 Da). A variation of the MALDI-TOF approach, SELDI-TOF MS has also been utilized in the past in the
search for alcohol biomarkers [81,88,89]. The SELDI-TOF-MS is a ‘hybrid’ between chromatography and MALDI-TOF-MS that employs a solid-phase chromatographic surface for binding proteins at a particular binding condition. There are several types of surfaces with different chromatographic properties, including hydrophobic, hydrophilic and ion exchange. These properties enable them to capture different subsets of proteins according to their physicochemical properties [90]. While attractive in principle, this approach lacks the requisite MS sensitivity to directly identify the peptide/protein and, thus, has enjoyed limited acceptance.

Alternatively, peptides can be ionized using ESI combined with MS/MS. The end of an LC column or a metal needle is held at a high electrical potential with respect to the entrance to the mass spectrometer, such that the peptides eluting from the chromatography column are electrostatically dispersed. This generates highly charged droplets that are normally positively charged in proteomics experiments. Once the droplets are airborne, the solvent evaporates and ions are now in the gaseous phase [86] and can be analyzed by MS/MS in the instrument of choice, such as a linear ion trap, an orbitrap or a hybrid linear ion trap/orbitrap MS. Although this approach is technically more complex than MALDI-TOF, ESI-MS/MS can be coupled online with HPLC, allowing for both preconcentration of the sample and analytical separation from matrix interferences, providing enhanced sensitivity and selectivity [91]. Furthermore, the generation of highly charged peptide ions allows for their fragmentation and the direct determination of the peptide sequence. Thus, more accurate peptide identification is achieved by looking at the peptide mass and the peptide sequence. However, compared with ESI-MS/MS, MALDI-TOF has the advantage of being able to accommodate more complex samples with a higher throughput [91].

**Affinity reagents**—In addition to the ‘open discovery’ approaches described above that rely on MS for protein identification, a number of methodological approaches have been developed that use affinity reagents to examine predefined sets of proteins. These affinity reagents range from traditional monoclonal antibodies to newer technologies such as aptamers. The advantage of this ‘directed discovery’ approach is that tens to hundreds of known proteins of interest can be simultaneously examined. This does, however, require *a priori* knowledge and protein-specific reagents. Sets of capture reagents can be arrayed on a surface in a manner analogous to nucleic acid microarrays, for example, capture arrays [92]. Alternatively, capture reagents can be attached to beads to create ‘suspension’ arrays. In this approach, each specific antibody or other capture reagent is attached to a microsphere with a unique fluorescent spectrum and many types of beads can be mixed together. After capture of the proteins of interest, the individual beads can be resolved by flow cytometry. Advantages of suspension arrays include fast, fluid-phase kinetics, high sensitivity, custom multiplexing and greater quantitative precision [93]. Microbead assay technology, often known by their trade name Luminex, has been approved by the FDA for use in diagnostics of other diseases.

These rapidly evolving technologies are already producing new leads into potential alcohol abuse biomarkers and promise to deliver further information toward the detection of alcohol dependence. In the following section, the authors will consider novel alcohol abuse biomarkers discovered through the various proteomic techniques discussed earlier.

**Current methods for alcohol abuse biomarker discovery: insight into novel markers**

In the past few years, modern proteomic techniques have provided insight into novel alcohol biomarkers. In 2004, Nomura *et al.* reported the first application of MS (SELDI-TOF) into the alcohol biomarker field, discovering fragments of fibrinogen αE (5.9-kDa peptide) and
Apo AII (7.8 kDa) to be downregulated during chronic alcohol use, and then significantly increased during alcohol abstinence [88]. By contrast, MS work by Freeman and coworkers detected Apo AII to be increased in the serum of ethanol self-administering monkeys [89]. More recently, Nomura’s group used the 5.9 kDa peak, combined with GGT and an additional 28 kDa protein, to screen for habitual drinkers with a sensitivity of 96.8% and a specificity of 60.9% [94]. The same research group has also used MALDI-TOF/TOF to detect upregulation of fibrinopeptide A (unmodified and phosphorylated) and downregulation of fibrinogen αC caused by chronic drinking [95]. Last year, Nomura’s laboratory used more traditional proteomic methods (i.e., SDS-PAGE and ELISA) to propose pigment epithelium-derived factor as a biomarker for excessive alcohol consumption; pigment epithelium-derived factor is increased in moderate-to-heavy drinkers compared with individuals without a drinking history [96]. Additional work involving LC–MS/MS has invoked other proteins, such as gelsolin, selenoprotein P, serotransferrin, tetranectin and hemopexin, to be potential biomarkers for alcohol abuse [97].

Expanding on the notion of combination tests, recent studies by Freeman and Vrana have explored the possibility of using a panel of proteins, rather than a single protein, to assess an individual’s drinking behavior. A 17-plasma protein panel – discovered by Luminex analysis of 90 known plasma cytokines, growth factors and other proteins – correctly classified abusive drinking with 100% sensitivity and differentiated any level of drinking from alcohol abstinence with 88% accuracy in nonhuman primates [98]. In monkeys, 2D-DIGE was also used to quantify plasma proteins from within-subject samples collected before exposure to ethanol and after 3 months of excessive ethanol self-administration. Altered levels of serum amyloid A4, retinol-binding protein, inter-α inhibitor H4, Apo J (clusterin) and fibronectin were detected and confirmed by immunoblotting. Investigation of these target proteins in human subjects with excessive alcohol intake revealed increased levels of serum amyloid A4 and Apo J and decreased levels of fibronectin compared with controls [99].

One can envision using these blood alcohol biomarker proteins for sampling all body compartments and replacing traditional biomarkers testing the effects of alcohol in specific body systems (Figure 1). Furthermore, such a multianalyte biomarker panel could conceivably not only differentiate between abstinence and heavy drinking, but also between abusive and nonabusive drinking (Figure 1). In this regard, it is important to realize that the members of a multianalyte panel may be reflective of completely different processes. That is, some biomarkers will represent alcohol-induced toxicity, while others simply represent physiologically relevant alterations in response to light or heavy consumption.

**Expert commentary**

Proteomic techniques are powerful tools in the discovery, characterization and validation of new protein biomarkers for alcohol abuse. As the instrumentation and the computational power for this type of analysis has become more sophisticated, new protein changes in response to chronic alcohol consumption have been illuminated. Recent reports have also highlighted the need for biomarker panels, rather than single biomarkers, to be used to accurately assess an individual’s drinking behavior. The problem with unitary (single) markers of alcohol abuse is the fact that alcohol engages physiologies and engenders pathophysiological effects that are common to other toxicants and organ systems (e.g., liver disease). With growing usage of proteomic techniques in the context of alcohol abuse, evidence for biomarker panels is likely to increase. Furthermore, as proteomic experiments are now able to detect thousands of proteins in a single run, it is likely that the complexity of biomarker panels will increase. Additional complexity is likely to arise from the use of
quantitative proteomic strategies such as isobaric tags for relative and absolute quantitation and accurate mass tags.

Five-year view

The rise of proteomic techniques in the alcohol abuse field holds great promise, as well as great challenges. The size and complexity of proteomic datasets from multiple patients across multiple time points – which will be needed not only for biomarker discovery but also for their validation – will demand computational approaches able to sort massive amounts of data. Moreover, one of the most difficult challenges will encompass the development of the statistical approaches to interpret data involving biomarker panels. Further complexity will arise from the fact that future alcohol biomarkers must be able to differentiate between nondrinkers, light drinkers and excessive drinkers, as opposed to just differentiating between nondrinkers and excessive drinking. Many studies to date have focused on simpler nondrinking versus heavy drinking scenarios and thus do not reflect the range of alcohol intakes occurring in real clinical practice. Aside from these, perhaps the greatest challenge will be to translate complex proteomic experiments into cost-effective and straightforward diagnostics that can be used in clinical practice; in fact, novel biomarkers rarely make it from the bench to the bedside. A final hurdle to be cleared will be educating clinicians to make objective measures of alcohol consumption the gold standard when diagnosing alcohol use disorders.

Acknowledgments

MP Torrente gratefully acknowledges support from the NIH in the form of a F32 NRSA postdoctoral fellowship (5F32MH094071) as well as a Neuroscience Career Development Fellowship from the Society for Neuroscience. This work was also supported by grants from the NIH to KE Vrana and WM Freeman (AA016613-03). WM Freeman and KE Vrana report a patent pending on diagnostics of alcohol intake.

References

Papers of special note have been highlighted as:

- of interest
- of considerable interest


Key issues

- Alcohol abuse can lead to severe health and societal consequences.
- Our current inability to accurately assess long-term drinking behaviors is an important obstacle to the diagnosis and treatment of alcohol abuse.
- Existing biomarkers for alcohol abuse lack either accuracy and sensitivity or clinical practicality. Thus, there is a crucial need for the development of more sensitive and specific markers of alcohol abuse.
- Recent advancements in proteomic technologies have greatly increased the potential for biomarker discovery in alcohol abuse.
- Proteomic technologies used in alcohol biomarker research include 2DE, liquid chromatography, mass spectrometry and affinity reagents (protein microarrays).
- Recent proteomic experiments highlight protein panels, rather than single proteins, as the most accurate alcohol biomarker prospects.
- Future alcohol biomarkers need to be able to differentiate between a variety of drinking behaviors (abstinence vs light vs heavy drinking) and assess both average intake and drinking patterns (e.g., binge drinking).
- Future research should emphasize alcohol biomarker validation and translation to clinical practice.
Figure 1. Alcohol biomarkers must transition from single analyte markers to multianalyte panels able to discern between ranges of alcohol drinking

Traditional unitary biomarkers can be replaced by marker panels sampling all body compartments. Most traditional unitary biomarkers represent a single anatomical compartment. A multianalyte biomarker panel could discern between diverse drinking behaviors and patterns (e.g., nondrinkers, nonabusive drinkers and excessive drinkers), versus simply differentiating between nondrinkers and excessive drinkers. A unitary marker of alcohol use and abuse may not be able to discriminate between modest drinking and abusive drinking. On the other hand, the incorporation of additional analytes permits the separation of heavy and light drinking. Indeed, with multidimensional statistical methods, we can theoretically incorporate large numbers of analytes into a biomarker panel.

5-HIAA: 5-hydroxyindole-3-acetic acid; 5-HTOL: 5-hydroxytryptophol; β-HEX: β-hexosaminidase; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; CDT: Carbohydrate-deficient transferrin; EtG: Ethyl glucuronide; FAEE: Fatty acid ethyl ester; GGT: γ-glutamyl transferase; MCV: Mean corpuscular volume; PEth: Phosphatidylethanol.
Traditional biomarkers of alcohol abuse.

<table>
<thead>
<tr>
<th>Alcohol biomarker</th>
<th>Source</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Time frame</th>
<th>Confounders</th>
<th>US FDA approval?</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCV</td>
<td>Blood</td>
<td>30–75</td>
<td>60–90</td>
<td>2–4 months</td>
<td>Liver diseases, vitamin B12 or folic acid deficiency, hematological diseases, reticulocytosis or hypothyroidism</td>
<td>No</td>
</tr>
<tr>
<td>EtG</td>
<td>Blood</td>
<td>70–90 (hair)</td>
<td>80–95 (hair)</td>
<td>8 h (blood)</td>
<td>None</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Urine</td>
<td></td>
<td></td>
<td>80 h (urine)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nails</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSA</td>
<td>Blood</td>
<td>48–82</td>
<td>18–96</td>
<td>NA</td>
<td>Cancer, cardiovascular disease</td>
<td>No</td>
</tr>
<tr>
<td>5-HTOL/5-HIAA</td>
<td>Urine</td>
<td>100</td>
<td>NA</td>
<td>5–15 h</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>PEth</td>
<td>Blood</td>
<td>94.5–100</td>
<td>100</td>
<td>4 days</td>
<td>None</td>
<td>No</td>
</tr>
<tr>
<td>CDT</td>
<td>Blood</td>
<td>60–70</td>
<td>80–95</td>
<td>1.5–2 weeks</td>
<td>Anorexia nervosa, pregnancy, iron deficiency, chronic illnesses and menopausal status</td>
<td>Yes</td>
</tr>
<tr>
<td>GGT</td>
<td>Blood</td>
<td>40–60</td>
<td>80–90</td>
<td>14–26 days</td>
<td>Liver damage, cardiovascular disease, diabetes</td>
<td>No</td>
</tr>
<tr>
<td>ALT/AST</td>
<td>Blood</td>
<td>18–58 (ALT)</td>
<td>50–95 (AST)</td>
<td>NA</td>
<td>Liver damage</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15–69 (AST)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β-HEX</td>
<td>Blood</td>
<td>69–94 (blood)</td>
<td>91–98 (blood)</td>
<td>6.5 days (blood)</td>
<td>Hypertension, diabetes, cirrhosis, myocardial infarction, in pregnancy and after oral contraceptive use</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Urine</td>
<td>81–85 (urine)</td>
<td>84–96 (urine)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acetaldehyde adducts</td>
<td>Blood</td>
<td>65–73</td>
<td>88–94</td>
<td>Up to 3 weeks</td>
<td>NA</td>
<td>No</td>
</tr>
<tr>
<td>Sialylation of Apo J</td>
<td>Blood</td>
<td>90–92</td>
<td>~100</td>
<td>Up to 8 weeks</td>
<td>NA</td>
<td>No</td>
</tr>
<tr>
<td>Fatty acid ethyl esters</td>
<td>Blood</td>
<td>NA</td>
<td>NA</td>
<td>100 h (blood)</td>
<td>NA</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Hair</td>
<td></td>
<td></td>
<td>2 months (hair)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cytokines</td>
<td>Blood</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Hair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDT + GGT</td>
<td>Blood</td>
<td>60–90</td>
<td>80–95</td>
<td>NA</td>
<td>See CDT and GGT</td>
<td>No</td>
</tr>
<tr>
<td>Alcohol biomarker</td>
<td>Source</td>
<td>Sensitivity (%)</td>
<td>Specificity (%)</td>
<td>Time frame</td>
<td>Confounders</td>
<td>US FDA approval?</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>------------</td>
<td>-------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>CDT + MCV</td>
<td>Blood</td>
<td>60–95</td>
<td>80–95</td>
<td>NA</td>
<td>See CDT and MCV</td>
<td>No</td>
</tr>
</tbody>
</table>

Biofluid sources, sensitivity and specificity values are listed for established alcohol biomarkers. US FDA approval status and confounding factors for each biomarker are also listed, as well as their diagnostic time frames (time period during which the marker is indicative of alcohol intake).

5-HIAA: 5-hydroxyindole-3-acetic acid; 5-HTOL: 5-hydroxytryptophol; β-HEX: β-hexosaminidase; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; CDT: Carbohydrate-deficient transferrin; EtG: Ethyl glucuronide; GGT: γ-glutamyl transferase; MCV: Mean corpuscular volume; NA: Not applicable; PEth: Phosphatidylethanol; TSA: Total serum sialic acid.