

Proteomics and Its Role in Nutrition Research^{1,2}

Junjun Wang,^{*†} Defa Li,^{*} Lawrence J. Dangott,[‡] and Guoyao Wu^{†3}

^{*}National Key Laboratory of Animal Nutrition, China Agricultural University, Beijing, China 100094 and [‡]Protein Chemistry Laboratory and [†]Department of Animal Science, Texas A&M University, College Station, TX 77843

ABSTRACT There are ~100,000 proteins in humans with various physiological functions. The complement of proteins in the organism as well as their interactions is defined as the proteome. Its analysis (proteomics) by highly specific, sensitive, and accurate MS has been made possible with matrix-assisted laser desorption ionization or electrospray ionization of proteins and large peptides. Currently, the most commonly used proteomics technologies involve either specific digestion of proteins (the bottom-up approach using 2-dimensional polyacrylamide gel electrophoresis and multidimensional protein identification technology) or direct analysis of intact proteins after their chromatographic separation (the top-down approach and surface-enhanced laser desorption ionization). Proteomics holds great promise for discoveries in nutrition research, including profiles and characteristics of dietary and body proteins; digestion, absorption, and metabolism of nutrients; functions of nutrients and other dietary factors in growth, reproduction, and health; biomarkers of the nutritional status and disease; and individualized requirements of nutrients. The proteome analysis is expected to play an important role in solving major nutrition-associated problems in humans and animals, such as obesity, diabetes, cardiovascular disease, cancer, aging, and intrauterine fetal retardation. *J. Nutr.* 136: 1759–1762, 2006.

KEY WORDS: • proteome • nutrients • metabolism • growth • health • disease

Since the human genome project was launched in 1989, there have been revolutionary developments in life science technologies characterized by high throughput, high efficiency, and rapid computation. Thus, advanced tools are now available for the analysis of DNA, RNA, protein, low-molecular-weight metabolites, and large data sets in nutrition research (1–4). Protein expression is the functional outcome of gene transcription and translation; thus, it has long been a focus of extensive

biology research. Such studies identified the crucial roles of proteins in cell structure and diverse biological processes, including signal transduction and nutrient utilization. Using proteomics, which is defined as the analysis of the proteome (the complement of proteins in cells, tissues, organs, and physiological fluids as well as their interactions), researchers can simultaneously display and determine thousands of proteins in a study sample and identify their changes in response to physiological, pathological, and nutritional alterations (1,2). Although still in its infancy, proteome analysis holds great promise for discoveries in nutrition research.

Proteomics technologies

Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE)⁴ and various chromatography techniques have long been applied as useful tools for protein separation (2). However, only with the invention of soft ionization techniques, matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI) in the late 1980s to vaporize peptides and proteins, has their analysis become possible using MS (1). The separation of ionized peptides or proteins in the high vacuum of the MS is based on the differences in their mass-to-charge ratios, with time-of-flight (TOF) and electric ion trap as major mass-selective analyzers in MALDI and ESI instruments, respectively. Over the past 20 y, MS has played a major role in the development of proteomics technologies because of its high speed, sensitivity, specificity, mass resolution capability, and mass accuracy for protein identification and characterization (1–3). The superior power of MS in the proteome analysis can be further enhanced when MS (as a detector) is combined with LC-GC or when tandem MS (MS-MS) is employed. There are currently 4 major proteomics technologies that involve either specific digestion of proteins [2D-PAGE MS and multidimensional protein identification technology (MudPIT)] or direct analysis of proteins after their chromatographic separation [the top-down approach and surface-enhanced laser desorption ionization (SELDI)] (Fig. 1). The choice of technology depends on the goal of study and facility availability. However, all methods require sample preparation, protein separation, MS analysis, and protein identification. The last mentioned is based on peptide mass fingerprints or spectra of fragmented ions of intact proteins and a search of the published databases (2).

2D-PAGE MS. This is the classical and most widely used proteomics technology. Proteins are separated in the first dimension involving isoelectric focusing, and in the second dimension involving SDS-PAGE, according to their isoelectric points (net charges) and sizes (molecular weights), respectively. The invention of immobilized pH gradients has allowed narrow pH intervals (e.g., 4–7 vs. the standard 3–10) and improved the reliability and reproducibility of protein separation by isoelectric

¹ Manuscript received 17 April 2006.

² Supported by the National Institute of Environmental Health Sciences grant P30-ES09106, the National Basic Research Program of P.R. China (No. 2004CB117503), and National Natural Science Foundation of P.R. China (No. 30121004).

³ To whom correspondence should be addressed. E-mail: g-wu@tamu.edu.

⁴ Abbreviations used: 2D-PAGE, 2-dimensional polyacrylamide gel electrophoresis; DIGE, difference gel electrophoresis; ESI, electrospray ionization; MALDI, matrix-assisted laser desorption ionization; MudPIT, multidimensional protein identification technology; SELDI, surface-enhanced laser desorption ionization; TOF, time-of-flight.

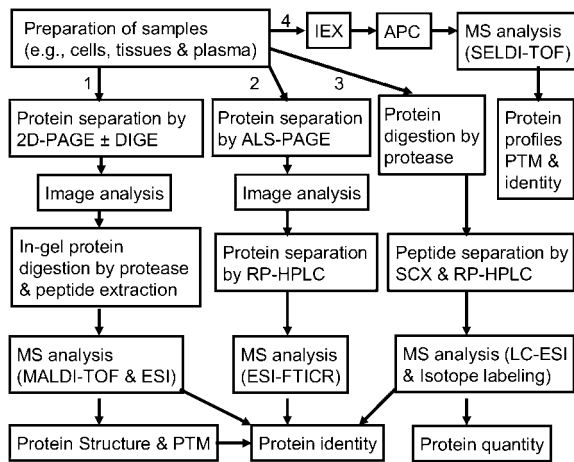


FIGURE 1 Workflows of commonly used proteomics technologies: 1. 2D-PAGE MS; 2. top-down approach; 3. MudPIT; 4. SELDI. Abbreviations used: ALD, acid-labile detergent; APC, affinity protein capture; FTICR, Fourier transform ion cyclotron resonance; IEX, ion-exchange chromatography; PTM, post-translational modifications; RP-HPLC, reversed-phase HPLC; SCX, strong cation exchange.

focusing (3). After the 2D-PAGE separation, proteins are stained with colloidal Coomassie blue, silver, or fluorescent dyes. Most recently, fluorescence-based difference gel electrophoresis (DIGE) was developed to eliminate gel-to-gel variation, which involves the addition of fluors 1 and 2 to the control and treatment samples, respectively (2). Computerized quantitative image analysis is performed using special software programs (2). In-gel digestion is then performed enzymatically (usually with trypsin) or chemically (e.g., cyanogen bromide), followed by MALDI-TOF MS analysis of the generated peptides. The advantages of the 2D-PAGE method are the powerful technology for protein separation, relative simultaneous quantification of proteins on gel images, and identification of protein isoforms and post-translational modifications (e.g., phosphorylation, hydroxylation, methylation, glycosylation, acetylation, and oxidation). The disadvantages include limited success for the separation of proteins (e.g., membrane proteins and some plasma proteins) with extremes in isoelectric points, molecular weights, abundance, and hydrophobicity, and an inability to provide automation or absolutely quantitative information.

MudPIT. This method involves initial digestion of proteins by a specific protease (usually trypsin). The generated peptides are separated using strong cation exchange and reversed-phase HPLC, followed by MS analysis (5). Because the ESI ionization source is compatible with a liquid sample, ESI-MS/MS is often conveniently coupled with HPLC. Along with MudPIT, isotope labeling (including $^{15}\text{N}/^{14}\text{N}$, $^{18}\text{O}/^{16}\text{O}$, or isotope-coded affinity tags) can be used to differentially label proteomes in control and treated samples, which yields quantitative proteomics information (2). Another quantitative method (Absolute Quantitative Analysis) involves the addition of a ^{13}C -labeled peptide to a protein digestion mixture for determining peptide recovery during sample processing (2). MudPIT not only overcomes the shortcomings of 2D-PAGE MS but also provides the following advantages: elimination of the time-consuming step for protein separation; high sensitivity and requirement of small sample size; and versatile mechanisms for peptide separation. A major disadvantage of MudPIT is an inability to readily provide information on protein isoforms or post-translational modifications.

Top-down approach. This method involves the separation of proteins first using an acid-labile detergent in gel electrophoresis and then using reversed-phase HPLC, followed by the analysis of intact proteins by MS (often ESI-MS/MS) (6). Most recently, Fourier transform ion cyclotron resonance has been used as the mass-selective analyzer to eliminate the problem that protein ions with different masses but the same mass-to-charge ratios exhibit the same cyclotron frequency (1). The advantages of the top-down approach are its applicability to membrane proteins that are soluble in acid-labile detergent; dynamic ranges of measured proteins; and its ability to provide information on protein isoforms and post-translational modifications. This is particularly important when the genes of interest have not been sequenced. Although the top-down approach is rapidly gaining recognition, it is currently not supported by software packages for the processing of complex data sets and may not be able to sequence all amino acid residues in proteins containing the cyclic heme (7).

SELDI. This technology involves the separation of proteins by ion-exchange or LC and antibody- or substrate-based affinity capture of one or more proteins of interest on a Protein Chip Array directly from the original source material (8). The chip surfaces function to fractionate and enrich subpopulations of proteins from complex protein mixtures (3). The captured proteins are analyzed by laser desorption/ionization-based MS-TOF (3). The major advantages of SELDI are simple sample preparation, reduction of sample complexity, suitability for low abundance proteins (e.g., transcription factors and a majority of cellular proteins), and rapid protein profiling. However, this technology is currently applicable only to proteins with maximum molecular weight ≤ 20 kDa and provides relatively lower mass accuracy than the 2D-PAGE MS method.

Application of proteomics to nutrition research

Proteomics has emerged as a revolutionary discovery tool in nutrition research. Its multiple areas have been greatly advanced by the use of this powerful technology. These include profiles and characteristics of dietary and body-fluid proteins; digestion, absorption and metabolism of nutrients as well as their functions in growth, reproduction, and health; and individualized requirements of nutrients (Table 1).

TABLE 1

Application of proteomics to nutrition research

Composition and characteristics of dietary proteins
Digestion and absorption of nutrients in the gastrointestinal tract
Nutrient metabolism (synthesis and catabolism) and its regulation
Interorgan transport of nutrients
Organell-, cell-, and tissue-specific metabolism of nutrients
Discovery of novel metabolic pathways and the mechanisms of their regulation
Functions of nutrients and phytochemicals in growth, reproduction, and health
Signal transduction and cellular defense against oxidative stress
Cell proliferation, differentiation, and apoptosis
Gene expression in response to nutrients and other dietary factors
Fetal and postnatal growth, development, and health
Dietary prevention and intervention of disease
Protein profiles and characteristics in cells, tissues, and physiological fluids
Biomarkers and individualized requirements of nutrients

Composition and characteristics of dietary proteins. The composition and characteristics of dietary proteins are major determinants of their nutritional values and potential pathogenic effects. Traditionally, dietary proteins were determined primarily using the Kjeldahl procedure and acid hydrolysis, which do not yield information about true protein, amino acid sequence, or some amino acids (e.g., glutamine, asparagine, and tryptophan). Such invaluable data can be readily provided from proteomics analysis. Notably, a recent study involving 2D-PAGE MALDI-MS revealed that the different effects of dietary soy isolates on humans (e.g., plasma lipids) in U.S. and European clinical studies are related to the differences in composition of the soy proteins used (e.g., 7S globulin products and intact 11S globulin subunits) (9). In addition, polymorphisms of dietary proteins (e.g., β -lactoglobulins A and B in cow's milk) may explain the formation of subtly but functionally distinct peptides that have remarkably different allergenicity in humans (4).

Digestion and absorption of nutrients in the gastrointestinal tract. The nutritional values of dietary nutrients and other factors depend on their digestion and absorption in the gastrointestinal tract. However, knowledge about digestive enzymes and epithelial-cell nutrient transporters remains incomplete, and it is critical for designing new ways to enhance the entry of low-molecular-weight nutrients into the portal vein. A recent proteomic analysis of the rat small-intestinal proteome indicates the presence of previously unrecognized proteins involved in intestinal molecular chaperones, cytoskeleton plasticity, and vitamin transporters, such as gastrotropin, filamin- α , and vitamin D-binding protein precursor (10). In addition, a MALDI-TOF MS study revealed the presence of 80 proteins in the myenteric plexus-longitudinal muscle layers of each of the rat intestinal segments (jejunum, ileum, and colon); these proteins may play a novel role in intestinal function (including digestion and absorption) (11). Moreover, the ESI-MS-MS technology was used to identify upregulation of 25 proteins and downregulation of 18 proteins in intestinal epithelial cells in response to endotoxin or pathogenic bacteria (12), thereby providing an explanation for impaired digestion and absorption of dietary nutrients under inflammatory conditions.

Nutrient metabolism and its regulation. There is growing interest in the role of proteomics in advancing our knowledge about nutrient metabolism and its regulation. Using MALDI-TOF MS, Yan et al. (13) reported major differences in cardiac glycolytic or mitochondrial pathways between young and aging monkeys or between males and females, which helps explain an aging-associated gender difference in the risk of cardiovascular disease. Also, proteomics identified up- and downregulated proteins (including vimentin and glucose-regulated protein 78) in insulin-treated adipocytes (14) and transcription factors in mammalian cells (3). Further, the levels of hepatic enzymes involved in glycolysis, gluconeogenesis, fatty acid oxidation, and amino acid metabolism vary greatly between lean and obese diabetic mice, which can be normalized with peroxisome proliferator-activated receptor activators (15). The findings from these studies greatly expand our knowledge about the regulatory networks for nutrient metabolism.

Functions of nutrients in growth and health. Diets for animals and humans contain a complex mixture of both nutrients and nonnutrient factors. Although adequate intakes of all nutrients are essential for maintenance, growth, reproduction, and health, an excess or deficiency in one or more of the essential nutrients

can result in metabolic disorders. Thus, there is increasing interest in the roles of nutrients and other dietary factors in cell physiology and health. Results of a recent MALDI-TOF-MS analysis indicated that glutamine alters the proteome of human intestinal cells, including the proteins that regulate amino acid, lipid, and vitamin A metabolism (16). Proteomics analysis also showed that the levels of hepatic lipid-metabolic enzymes and prooxidative proteins are increased in mice fed a high-fat diet (17), and hepatic protein profiles are altered in rainbow trout in response to dietary intake of proteins (18). Using a comprehensive proteomic approach, Li et al. (19) reported that energy restriction promotes proper protein folding and function, thereby maintaining a sufficient rate of glucose metabolism and retarding age-related retinal degeneration. Interestingly, dietary deficiency of copper (9), iron (10), folate (20), or zinc (21) markedly influences expression of intestinal and hepatic proteins related to cellular redox regulation, lipid metabolism, protein phosphorylation, DNA synthesis, and nutrient transporters. Notably, proteomic studies determined that dietary supplementation with genistein (the major isoflavone of soy) increases the expression of GTP cyclohydrolase-I [a key protein related to nitric oxide synthesis (22)] in the rat mammary gland in association with a reduction in cell proliferation and susceptibility to cancer (23). These studies help establish molecular mechanisms for the roles of nutrients and other dietary factors in growth, reproduction, and health.

Protein profiles and characteristics in physiological fluids. The human genome consists of 24,000–30,000 genes, which may generate ~100,000 proteins due to mRNA splice variants, protein processing, and post-translational modifications (2). Protein profiles and characteristics in physiological fluids are excellent indicators of nutritional status and protein post-translational modifications. Because the blood pool is readily accessible for noninvasive sampling, protein profiles in plasma/serum can be used as biomarkers to evaluate the adequacy of specific nutrients, diagnose disease, and monitor therapeutic response. There is evidence showing that nutrition alters plasma and body-fluid proteomes in humans and animals. For example, dietary supplementation with α -tocopherol increases plasma apolipoprotein A1 isoforms in normal healthy subjects (24), whereas a marked decrease in plasma levels of 3 proteins occurs in retinol-deficient rats (8). In addition, inadequate provision of dietary vitamin B-12 induces profound changes in the proteome of the rat cerebrospinal fluid, thus linking vitamin B-12 with neurological functions (25).

Individualized requirements of nutrients. The interactions between the genome and environment (e.g., nutrition) determine protein expression in organisms. Most genes exhibit small sequence differences or polymorphisms among individuals. Single nucleotide polymorphisms are the most common form of human DNA sequence variation and occur every 100–300 bases along the 3-billion-base human genome (2); they are responsible for the differences in both metabolism and sensitivity to dietary treatments among individuals. For example, the responses of C57BL/6J and C3H/HeJ mice to a high-fat diet differ markedly in the hepatic proteome (17). In addition, compared with the copper-tolerant Cambridge sheep, the copper-sensitive North Ronaldsay sheep respond to a high level of dietary copper with little expression of the cytosolic NADP-linked isocitrate dehydrogenase but upregulation of lysosomal cathepsin D (26). These novel findings explain a striking difference in the susceptibility to copper-induced oxidative stress and mitochondrial damage between the 2 breeds of sheep. Similarly, there are differences in the susceptibility of individual

subjects to various forms of cancer, cardiovascular disease, diabetes, and obesity, as well as in sensitivity to dietary treatments (27). Proteomics data will help guide the development of individualized requirements of nutrients to optimize health and reduce the risk of disease.

Summary and Perspectives

The field of nutrition research was transformed by the recent development of proteomics as a powerful discovery tool. Given the very large number of proteins as well as their exceedingly large dynamic ranges, diverse physicochemical properties, and complex interactions, a combination of various proteomics technologies will complement other techniques and provide comprehensive information about protein identification and characterization. The proteome analysis will be a useful technology with which to solve major nutrition-associated problems in humans and animals (including obesity, diabetes, cardiovascular disease, cancer, aging, and intrauterine fetal retardation) (27–32). Integrated with other advanced technologies (genomics, transcriptomics, metabolomics, and bioinformatics) and systems biology (4), proteomics will greatly facilitate the discovery of key proteins that function to regulate metabolic pathways and whose synthesis, degradation, and modifications are affected by specific nutrients or other dietary factors. This will aid in rapidly enhancing our knowledge of the complex mechanisms responsible for nutrient utilization (including food efficiency), identifying new biomarkers for nutritional status and disease progression, and designing a contemporary paradigm for dietary prevention and intervention of disease. Thus, the proteome analysis holds great promise in improving human health and enhancing the efficiency of animal agriculture.

LITERATURE CITED

- Fuchs D, Winkelmann I, Johnson IT, Mariman E, Wenzel U, Daniel H. Proteomics in nutrition research: principles, technologies and applications. *Br J Nutr.* 2005;94:302–14.
- Kim H, Page GP, Barnes S. Proteomics and mass spectrometry in nutrition research. *Nutrition.* 2004;20:155–65.
- Forde CE, McCutchen-Maloney SL. Characterization of transcription factors by mass spectrometry and the role of SELDI-MS. *Mass Spectrom Rev.* 2002;21:419–39.
- Corthesy-Theulaz I, den Dunnen JT, Ferre P, Geurts JM, Muller M, van Belzen N, van Ommen B. Nutrigenomics: the impact of biomics technology on nutrition research. *Ann Nutr Metab.* 2005;49:355–65.
- Washburn MP, Wolters D, Yates, III JR. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol.* 2001;19:242–7.
- Meng F, Cargile BJ, Patrie SM, Johnson JR, McLoughlin SM, Kelleher NL. Processing complex mixtures of intact proteins for direct analysis by mass spectrometry. *Anal Chem.* 2002;74:2923–9.
- Horn DM, Zubarev RA, McLafferty FW. Automated de novo sequencing of proteins by tandem high-resolution mass spectrometry. *Proc Natl Acad Sci USA.* 2000;97:10313–7.
- Linke T, Ross AC, Harrison EH. Profiling of rat plasma by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry, a novel tool for biomarker discovery in nutrition research. *J Chromatogr A.* 2004;1043:65–71.
- Gianazza E, Eberini I, Arnoldi A, Wait R, Sirtori CR. A proteomic investigation of isolated soy proteins with variable effects in experimental and clinical studies. *J Nutr.* 2003;133:9–14.
- Tosco A, Siciliano RA, Cacace G, Mazzeo MF, Capone R, Malorni A, Leone A, Marzullo L. Dietary effects of copper and iron deficiency on rat intestine: a differential display proteome analysis. *J Proteome Res.* 2005;4:1781–8.
- Marvin-Guy L, Lopes LV, Affolter M, Courtet-Compondu MC, Wagniere S, Bergonzelli GE, Fay LB, Kussmann M. Proteomics of the rat gut: analysis of the myenteric plexus-longitudinal muscle preparation. *Proteomics.* 2005;5:2561–9.
- Alpert C, Engst W, Guehler A, Oelschaeger T, Blaut M. Bacterial response to eukaryotic cells analysis of differentially expressed proteins using nano liquid chromatography-electrospray ionization tandem mass spectrometry. *J Chromatogr A.* 2005;1082:25–32.
- Yan L, Ge H, Li H, Lieber SC, Natividad F, Resuello RRG, Kim SJ, Akeju S, Sun A, et al. Gender-specific proteomic alterations in glycolytic and mitochondrial pathways in aging monkey hearts. *J Mol Cell Cardiol.* 2004;37:921–9.
- Bluher M, Wilson-Fritch L, Leszyk J, Laustsen PG, Corvera S, Kahn CR. Role of insulin action and cell size on protein expression patterns in adipocytes. *J Biol Chem.* 2004;279:31902–9.
- Edvardsson U, von Lowenhillem HB, Panfilov O, Nystrom AC, Nilsson F, Dahllof B. Hepatic protein expression of lean and obese diabetic mice treated with peroxisome proliferator-activated receptor activators. *Proteomics.* 2003;3:468–78.
- Lenaerts K, Mariman E, Bouwman F, Renes J. Glutamine regulates the expression of proteins with a potential health-promoting effect in human intestinal Caco-2 cells. *Proteomics.* 2006;6:2454–64.
- Park JY, Seong JK, Paik YK. Proteomics analysis of diet-induced hypercholesterolemic mice. *Proteomics.* 2004;4:514–23.
- Martin SAM, Vilhelmsson O, Medale F, Watt P, Kaushik S, Houlihan DF. Proteomic sensitivity to dietary manipulations in rainbow trout. *Biochim Biophys Acta.* 2003;1651:17–29.
- Li D, Sun F, Wang KY. Protein profile of aging and its retardation by caloric restriction in neural retina. *Biochem Biophys Res Commun.* 2004;318:253–8.
- Chanson A, Sayd T, Rock E, Chambon C, Sante-Lhoutellier V, de Courcy GP, Brachet P. Proteomic analysis reveals changes in the liver protein pattern of rats exposed to dietary folate deficiency. *J Nutr.* 2005;135:2524–9.
- tom Dieck H, Doring F, Fuchs D, Roth HP, Daniel H. Transcriptome and proteome analysis identifies the pathways that increase hepatic lipid accumulation in zinc-deficient rats. *J Nutr.* 2005;135:199–205.
- Shi W, Meininger CJ, Haynes TE, Hatakeyama K, Wu G. Regulation of tetrahydropterin synthesis and bioavailability in endothelial cells. *Cell Biochem Biophys.* 2004;41:415–33.
- Rowell C, Carpenter DM, Lamartiniere CA. Chemoprevention of breast cancer, proteomics discovery of genistein action in the rat mammary gland. *J Nutr.* 2005;135:2953S–9.
- Aldred S, Sozzi T, Mudway I, Grant MM, Neubert H, Kelly FJ, Griffiths HR. α Tocopherol supplementation elevates plasma apolipoprotein A1 isoforms in normal healthy subjects. *Proteomics.* 2006;6:1695–703.
- Gianazza E, Veber D, Eberini I, Buccellato FR, Mutti E, Sironi L, Scalabrino G. Cobalamin (vitamin B12)-deficiency-induced changes in the proteome of rat cerebrospinal fluid. *Biochem J.* 2003;374:239–46.
- Haywood S, Simpson DM, Ross G, Beynon RJ. The greater susceptibility of North Ronaldsay sheep compared with Cambridge sheep to copper-induced oxidative stress, mitochondrial damage and hepatic stellate cell activation. *J Comp Pathol.* 2005;133:114–27.
- Davis CD, Milner J. Frontiers in nutrigenomics, proteomics, metabolomics and cancer prevention. *Mutat Res.* 2004;551:51–64.
- Korc M. Diabetes mellitus in the era of proteomics. *Mol Cell Proteomics.* 2003;2:399–404.
- Wu G, Meininger CJ. Regulation of nitric oxide synthesis by dietary factors. *Annu Rev Nutr.* 2002;22:61–86.
- Wu G, Jaeger LA, Bazer FW, Rhoads JM. Arginine deficiency in premature infants: biochemical mechanisms and nutritional implications. *J Nutr Biochem.* 2004;15:442–51.
- Wu G, Fang YZ, Yang S, Lupton JR, Turner ND. Glutathione metabolism and its implications for health. *J Nutr.* 2004;134:489–92.
- Wu G, Bazer FW, Cudd TA, Meininger CJ, Spencer TE. Maternal nutrition and fetal development. *J Nutr.* 2004;134:2169–72.