TOXICOLOGICAL SCIENCES **95(1)**, 74–81 (2007)
doi:10.1093/toxsci/kfl079
Advance Access publication August 17, 2006

A Proteomic (SELDI-TOF-MS) Approach to Estrogen Agonist Screening

Calvin C. Walker,*,[1] Kimberly A. Salinas,* Peggy S. Harris,* Sherry S. Wilkinson,* James D. Watts,† and Michael J. Hemmer*

US Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Gulf Ecology Division, Gulf Breeze, FL 32561; and †Senior Environmental Employee, Gulf Ecology Division, Gulf Breeze, FL 32561

Received May 25, 2006; accepted August 8, 2006

A small fish model and surface-enhanced laser desorption/ionization time-of-flight mass spectrometry were used to investigate plasma protein expression as a means to screen chemicals for estrogenic activity. Adult male sheepshead minnows (*Cyprinodon variegatus*) were placed into aquaria for seawater control, solvent control, and treatments of 17β-estradiol (E2), methoxychlor (MXC), bisphenol-A (BPA), 4-*tert*-pentylphenol (TPP), endosulfan (ES), and chlorpyriphos (CP). Fish plasma was applied to weak cation exchange (CM10) ProteinChip arrays, processed, and analyzed. The array produced approximately 42 peaks for E2 plasma and 30 peaks for solvent control plasma. Estrogen-responsive mass spectral biomarker peaks were identified by comparison of E2-treated and control plasma spectra. Thirteen potential protein biomarkers with a range from 1 to 13 kDa were up- or downregulated in E2-treated fish and their performance as estrogenic effects markers was evaluated by comparing spectra from control, estrogen agonist, and nonagonist stressor–treated males and normal female fish plasma. One of the biomarkers, mass-to-charge ratio 3025.5, was identified by high-resolution tandem mass spectrometry as *C. variegatus* zona radiata protein, fragment 2. The weak environmental estrogens MXC, BPA, and TPP elicited protein expression profiles consistent with the estrogen expression model. Estrogen-responsive peaks were not detected in plasma from fish in the seawater, vehicle, ES, or CP treatments. No difference was found between plasma protein expression of seawater control and solvent control fish. We show that water exposure of fish to estrogen agonists produces distinct plasma protein biomarkers that can be reproducibly detected at low levels using protein chips and mass spectrometry.

Key Words: **EDC; protein profiling; SELDI; fish; sheepshead minnow; estrogen.**

Endocrine hormones regulate a number of physiological processes in vertebrates, including homeostasis and the regulation of reproductive, developmental, and growth processes (Kavlock *et al.*, 1996). Exogenous compounds that disrupt physiological processes controlled by the endocrine systems of humans, domestic animals, and wildlife (Colborn *et al.*, 1993; Etienne and Jemmali, 1982; Rotchell and Ostrander, 2003) are termed endocrine disrupting compounds. Such chemicals may be naturally occurring (e.g., hormones in waste water; phytoestrogens such as the soy isoflavones genistein and daidzein; and estrogenic mycotoxins such as zearlenone) or synthetic (e.g., pharmaceuticals such as 17-α-ethinylestradiol; pesticides such as methoxychlor [MXC]; and industrial chemicals such as bisphenol-A [BPA] and alkylphenols).

The U.S. Environmental Protection Agency (EPA) has a legislative mandate to test an inventory of thousands of chemicals for endocrine disrupting effects (U.S. EPA, 1998). The challenge for the EPA resides in how to prioritize, screen, and identify endocrine activity of these many candidate compounds without employing animal and time-intensive methodologies. Current endocrine testing methods are animal intensive and lack the throughput necessary to screen large number of environmental chemicals for adverse effects. Current assays also analyze for a single endocrine mode of action (MOA) at a time. An assay, which can evaluate multiple endocrine MOAs from a single exposure is highly desirable. Further, there is a need for diagnostic biomarkers of estrogenic environmental contamination in many wildlife species.

Small fish species have often been used to investigate the toxicity of various compounds. Small size results in much lower cost and space requirements when compared with mammalian and avian studies, and fish may be treated via water exposure and do not require individual dosing. Like all vertebrates, fish possess receptors for endocrine hormones and have been used as surrogate vertebrate models to identify compounds with endocrine disrupting activity (Ankley *et al.*, 2003; Hemmer *et al.*, 2001) and as indicator organisms for endocrine effects in the aquatic environment (Parks *et al.*, 2001). Wild fish are exposed to estrogenic compounds from a variety of sources including domestic sewage effluents, agricultural and feed lot runoff, and industrial wastewaters. Much of the endocrine disruption research in fish has focused on the effects of estrogens and such effects are well documented both

[1] To whom correspondence should be addressed at Gulf Ecology Division, US Environmental Protection Agency, 1 Sabine Island Drive, Gulf Breeze, FL 32561. Fax: (850) 934-2402. E-mail: walker.calvin@epa.gov.

in laboratory and field studies (Rotchell and Ostrander, 2003). Endpoints used to assess fish for estrogen-related disruption include partial and full reproduction tests, behavioral, morphologic, and histologic changes, serum hormone levels, and molecular markers such as gene induction, and the expression of proteins such as vitellogenin (VTG) and zona radiata proteins (ZR) in male fish (Oppen-Berntsen *et al.*, 1999; Rotchell and Ostrander, 2003).

The low abundance and low molecular mass range of the plasma proteome hold great potential as a source of diagnostic biomarkers (Guerrier *et al.*, 2005; Mehta *et al.*, 2003–2004; Tirumalai *et al.*, 2003). A relatively new technique, surface-enhanced laser desorption and ionization time-of-flight mass spectrometry (SELDI), provides a means to exploit low abundance proteins and peptides as MOA-specific biomarkers. SELDI can be thought of as an expansion of the matrix-assisted laser desorption and ionization time-of-flight mass spectrometry (MALDI) method (Hutchens and Yip, 1993). Both techniques use laser desorption and time-of-flight (TOF) mass spectrometry to determine protein and peptide molecular mass, as indicated by their mass-to-charge ratio (*m*/*z*). For SELDI, sample preparation, fractionation, and TOF analysis can be performed directly on the selective chip surface, whereas MALDI analysis typically requires separate preparative and fractionation steps prior to sample application onto a MALDI plate.

A comparison of protein expression in two groups can be used to discover differences in up- or downregulated proteins between groups. The molecular biomarkers or fingerprints derived from such analyses can be used to classify samples into two or more categories, i.e., a ''disease'' state versus a ''normal'' state, and can be used as a screening or diagnostic test of MOA-specific activity. Further, protein biomarkers discovered through SELDI analysis can be isolated and purified using the molecular weight and physico-chemical property information derived from SELDI and identified via high-resolution tandem mass spectrometry (MS/MS) or tryptic digests.

In this study we examine the use of a small fish and plasma protein expression for detecting estrogen agonist activity in a vertebrate model. The model's performance in correctly classifying chemicals by estrogenic activity was tested using self-validation and plasma from male fish exposed to weak environmental estrogen agonists, nonestrogenic chemical stressors, and unexposed adult female fish. We show for the first time the ability to classify chemicals by MOA using a small fish model and SELDI protein profiling.

MATERIALS AND METHODS

Study design. Experiments were conducted to determine if plasma protein expression differed between unexposed male fish and male fish exposed to estrogen agonists. Plasma protein expression in male fish exposed to the native estrogen ligand, 17β-estradiol (E2), was compared to nonexposed control fish to discover estrogen-responsive plasma protein biomarkers. The performance of these biomarkers for detecting estrogenic activity was evaluated by self-validation, and by testing plasma samples from male fish exposed to weak, nonsteroidal estrogen agonists, nonestrogenic chemical stressors, unexposed adult male and adult female minnows. In addition, an E2 concentration-response study was performed and protein identification of one biomarker was conducted using high-resolution MS.

Fish collection, care and exposure chemicals. Adult male sheepshead minnows (SHM) (*Cyprinodon variegatus*) were collected by trap and small bag seine from Santa Rosa Sound, FL. The fish were maintained with a 16 hour light to 8 hour dark photoperiod in 100-l tanks receiving a continuous flow of aerated filtered seawater at 20 ± 2 ppt and 25 ± 3°C for a minimum 17 days prior to exposure. Fish were fed Tetramin flake food to satiation at least twice daily except weekends when fed once daily. The utilization of fish in this study was compliant with ''U.S. Government Principles for the Utilization and Care of Vertebrate Animals Used in Testing, Research and Training'' (U.S. Office of Science and Technology Policy, 1985) and approved by the Gulf Ecology Division's Animal Care and Use Committee. Test chemicals were obtained from the following sources: E2, BPA, and 4-*tert*-pentylphenol (TPP) from Sigma (St Louis, MO) and MXC, endosulfan (ES), and chlorpyrifos (CP) from Chem Service (West Chester, PA). All compounds had a manufacture's stated purity of 97%. Chemical stock solutions were prepared using laboratory grade triethylene glycol (TEG) (Fisher Scientific, Pittsburgh, PA) as the solvent.

Biomarker discovery and validation chemical fish exposure. Adult male SHM were exposed to a single concentration of each of the following validation treatments: E2 (0.2 μg/l), BPA (1000 μg/l), TPP (100 μg/l), MXC (6 μg/l), ES (0.6 μg/l), CP (80 μg/l), TEG (12 μl/l), and seawater only (SWM). A time 0 sample ($N = 10$) was taken prior to loading, and fish were randomly divided between the chemical and control treatments with each exposure tank receiving 15 fish. A group of 15 adult reproductively active female SHM (SWF) was kept in an aquaria receiving clean seawater from the dosing system to serve as a positive control. The rationale for inclusion of the validation treatments and list of abbreviations are included in Table 1. Concurrent with the validation chemical exposures, 32 male fish were exposed to E2 (0.2 μg/l) and 30 males were exposed to TEG (12 μl/l) for biomarker discovery. Exposure chambers were 135-l glass aquaria containing 100 l of test solution. An intermittent seawater flow of 20 l/h was maintained to each treatment using a 1-l dosing apparatus. Exposure water test concentrations were maintained by injection of test stock solutions at 12-μl/l seawater/cycle using multiple two-channel Hamilton Microlab 500C dispensers fitted with 25-μl syringes. For a given test concentration, 1 l of seawater containing test solution flowed directly into the exposure tank with each cycle of the dosing apparatus. Exposure aquaria were maintained in a temperature-controlled water bath and under a fixed photoperiod of 16 hour light to 8 hour dark. The duration of exposure was 7 days for E2,

TABLE 1
Treatments for Method Validation Studies

Treatment	Abbreviation	Concentration (μg/l)	Inclusion rationale
17β-estradiol	E2	0.2	Native estrogen
Bisphenol-A	BPA	1000	Xenoestrogen
4-*tert*-pentylphenol	TPP	100	Xenoestrogen
Methoxychlor	MXC	6	Xenoestrogen
Endosulfan	ES	0.6	Stressor control
Chlorpyrifos	CP	80	Stressor control
Triethylene glycol	TEG	12[a]	Solvent control
Seawater (adult males)	SWM	NA	Male negative control
Seawater (adult females)	SWF	NA	Female positive control

[a]Unit: μl/l.

BPA, TPP, ES, and CP treatments and 10 days for the MXC treatment and SWM controls. Temperature (mean 24.8, SD 0.28°C) and salinity (mean 18.9, SD 0.2 ppt) were monitored daily, with dissolved oxygen (mean 6.5, SD 0.9 ppm) and pH (range 8.1–8.6) measured at the beginning, midpoint, and end of the exposures. Fish were fed a maintenance diet of Tetramin flake food daily.

E2 concentration-response exposure. Adult male SHM were exposed to nominal concentrations of 0.05, 0.1, 0.2, 0.5, and 1.0 μg E2/l using the intermittent dosing system described above. Prior to loading the aquaria, 10 male fish were impartially selected from the test population to serve as a pretest background sample (time 0). The remaining fish were randomly assigned to exposure and TEG control treatments with each aquarium receiving 28 fish. The exposure duration was 7 days. Environmental conditions were maintained as described for the biomarker discovery and validation study above. Temperature (mean 24.8, SD 0.4°C) and salinity (mean 20.0, SD 0.4 ppt) were monitored daily; dissolved oxygen (mean 5.8, SD 1.4 ppm) and pH (range 7.6–8.1) were measured at the beginning of the exposure, 24 h after introduction of the fish, and when the fish were sampled.

Plasma collection. Fish were anesthetized in a solution of tricaine methane sulfonate (Sigma-Aldrich) until incapacitated. For each fish examined, total length, wet weight, and liver weight were recorded and the sex verified by direct observation of the gonad. The caudal peduncle was severed with a razor blade and blood collected from the caudal vein and artery using heparinized capillary tubes. The plasma was separated by centrifugation at 13,700 × g for 3 min followed by storage in individual labeled cryovials at −70°C until analyzed.

Chemical analysis. Water samples for the concentration-response study were collected from exposure aquaria 96 h after the start of chemical flow and on days 3 and 7 of the experiment. Exposure concentrations of E2 were measured by competitive enzyme-linked immunosorbent assay (Cayman Chemical, Ann Arbor, MI) following the manufacturer's instructions with the following modifications. Standard curves were prepared using estradiol in seawater (20 ppt salinity) at concentrations of 0.016, 0.031, 0.063, 0.125, 0.25, 0.5, and 1.0 μg/l. Samples and standards were analyzed in triplicate wells in a single run for each sampling period. The 0.05, 0.1, and 0.2 μg E2/l samples were assayed directly without dilution. Samples from the 0.5 and 1.0 μg E2/l aquaria were diluted 2× and 4× in seawater, respectively, to fall within the midpoint of the standard curve. Water concentrations for the biomarker discovery and validation chemical exposures conducted with E2, BPA, TPP, MXC, ES, and CP were not measured and are based on nominal concentrations.

Plasma fractionation. Plasma samples from the E2 concentration-response study fish ($n = 5$ per treatment) and samples used for biomarker discovery ($n = 15$) were individually analyzed. However, plasma obtained from the validation chemical exposure treatments was pooled prior to freezing by combining variable amounts of plasma from 12 to 15 individuals per treatment. Plasma samples were robotically prepared for SELDI-TOF-MS analysis using a Biomek 2000 with an integrated Micromix 5 shaker (Beckman Coulter, Fullerton, CA) in 96-well format. Weak cation exchange (CM10) ProteinChips were placed into a ProteinChip array holder (Bioprocessor, Ciphergen, Fremont, CA) and activated with low stringency CM10 binding buffer (0.1M sodium acetate, pH 4.0). A 10-μl aliquot of plasma was mixed with 190 μl of CM10-binding buffer (plasma dilution of 1:19), duplicate 100-μl aliquots of diluted plasma were randomized across ProteinChips within an experiment and mixed for 30 min at room temperature on a Micromix 5 platform shaker. After mixing, the diluted plasma was removed, the spots washed with CM10-binding buffer, rinsed quickly with MilliQ water and allowed to air dry. A laser energy absorbing molecule solution (2 μl of a 25% sinapinic acid in 50% acetonitrile [ACN] and 0.5% trifluoroacetic acid [TFA] [v/v]) was applied to each spot twice and allowed to air dry. Because biomarker discovery, validation, and concentration-response samples were analyzed at different times, a temporal normalization standard consisting of archived frozen aliquots of salmon plasma (EastCoast Bio, Inc., North Berwick, ME) was included with each analytical run. The salmon plasma was placed on one spot of each ProteinChip to compare and normalize sample relative intensities to the salmon plasma total ion current (TIC) across experiments. The salmon plasma standards and SHM plasma samples were processed identically.

Data acquisition and processing. Prepared ProteinChip arrays were analyzed by a SELDI PBS-IIc ProteinChip System with autoloader (Ciphergen Biosystems) using Ciphergen ProteinChip software, version 3.2.1. The TOF mass spectra were generated by averaging 65 laser shots in positive mode with a laser intensity of 190 and detector sensitivity of 7. Data acquisition parameters were optimized to detect peaks in the range of 1–10 kDa with high mass set at 15 kDa. The instrument was mass calibrated for each protein analysis experiment using the All-in-One peptide molecular weight standards placed on NP20 ProteinChip arrays (Ciphergen Biosystems).

Spectral data were uploaded to CiphergenExpress Software, version 3.0.6 for processing and analysis. Raw data processing by CiphergenExpress included mass calibration, baseline subtraction, TIC normalization, spectra alignment, and peak (cluster) detection. All spectra were aligned to a reference spectrum within a 0.3% mass acceptance window prior to clustering. Peak clusters, peaks of similar m/z across similar groups of spectra, were autodetected using CiphergenExpress Software with signal-to-noise ratio > 10 on first pass, > 2 on second pass, adding estimated peaks, and 0.3% mass window. Peak clusters were visually inspected as a quality control check. The relative intensity of each baseline-subtracted spectrum was normalized to the TIC (1–15 kDa) of the standard salmon plasma spectra from the E2 biomarker discovery samples. Normalization of peak intensities was used to account for spectrum-to-spectrum variations of a set of spectra due to sample processing, instrument, and day-to-day variation. Spectra with a normalization factor < 0.5 or > 2.0 were not included in final analysis. Peak clusters obtained from the E2 biomarker discovery spectra were saved and applied to spectra from the validation chemicals and concentration-response exposures to ensure proper peak assignment and to obtain peak intensity test data.

Biomarker discovery, data analysis and statistics. CiphergenExpress software was used to compare plasma protein profiles from 0.2 μg E2/l-treated fish and TEG control fish to discover peaks, which are differentially expressed between the two groups. Reference values for plasma expression levels of each protein of interest in "normal" male SHM were obtained using data derived from TEG control plasma (Solberg, 1987). These reference values were used with EXCEL (Microsoft Corp. 2003) to determine an interpercentile reference interval for each protein of interest by taking the mean peak intensity ± 3 SD. These "normal" response intervals are bounded by an upper and lower 99.7 percentile. Observed peak intensity values of samples were interpreted as estrogenic or nonestrogenic responses by comparison with the "normal" reference interval limits. Estrogen-responsive protein peaks were considered a biomarker for estrogen activity only if they were capable of discriminating between the two groups with 100% sensitivity and 100% specificity. Following biomarker discovery, the performance of each diagnostic protein peak to correctly classify samples as estrogenic or nonestrogenic was assessed with spectral data from the E2 biomarker discovery (self-validation), E2 concentration-response samples, and validation samples. Performance of the technique to correctly classify samples was evaluated by sensitivity and specificity determinations. Sensitivity, or the ability of an assay to correctly classify estrogen agonist–treated fish, was determined by dividing the number of true positives by the total number of positive results, expressed as a percentage. Specificity, or the ability of an assay to correctly classify nonestrogen-treated fish, was determined by dividing the number of true negatives by the total number of negative results, expressed as a percentage.

Identification of protein biomarkers. Representative plasma samples were fractionated prior to protein identification using S Ceramic HyperD F Anion Exchange Columns (Ciphergen Biosystems) with five pH elution buffers (50mM sodium citrate, pH 3, 100mM sodium acetate, pH 4, 100mM sodium acetate, pH 5, 50mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, pH 7, 50mM Tris-HCl, pH 9) and an organic solvent elution solution (33% isopropanol/17% ACN/0.1% TFA solution). Each fraction was spotted onto a normal phase (NP20) ProteinChip by adding 1 μl of sample to each spot, then

adding 1 μl of a 50% solution of saturated sinapinic acid twice, allowing spots to dry between applications. The ProteinChips were analyzed using manual spectra collection settings in the low-resolution PBSIIc TOF mass spectrometer to determine which fraction(s) contained proteins of interest for further analysis by MS/MS. The organic fraction provided a high abundance of the peak of interest (3027.3 *m/z*) and was spotted on a NP20 ProteinChip by adding 1 μl of sample to each spot, then immediately adding 1 μl of a saturated solution of α-cyano-4-hydroxycinnamic acid in 50% ACN, 0.5% TFA and mixing on the chip. The prepared ProteinChip was analyzed using a QSTAR-XL high-resolution tandem mass spectrometer equipped with Analyst QS software version 1.1 (Applied Biosystems, Inc., Foster City, CA) and a Ciphergen protein chip interface (Ciphergen PCI 1000). The instrument was externally mass calibrated using adrenocorticotropic hormone$_{(18–39)}$ parent and four fragment ions as calibrants, and all spectra were acquired in positive ion mode. The peptide of interest (3027.3 *m/z*) was isolated in TOF mode and fragmented and analyzed directly using MS/MS with argon for collision-induced dissociation. The raw MS/MS spectra were processed using the Anaylst QS software, submitted to the MASCOT database search engine (Matrix Science, Ltd, Boston, MA, MSDB 20050701, Timestamp 28 September 2005), and the database was queried for peptide fragment matches. The following search parameters were used: MS/MS ion search with monoisotopic mass values, no enzyme, unrestricted protein mass, ± 10 ppm peptide mass tolerance, ± 0.2-Da fragment mass tolerance, MALDI-QUAD-TOF instrument type, and taxonomic restriction to Chordata.

RESULTS

Exposures

Average measured concentrations of E2 for the concentration-response exposure are presented in Table 2. The coefficients of variation for these assays were $\leq$ 15%, and the correlation coefficient for each standard curve was $\geq$ 0.99. Mortalities occurring during the concentration-response and validation chemicals exposures were $\leq$ 13% with the exception of the ES stressor treatment where mortality reached 20%.

Protein Biomarkers Discovery

The SELDI process produced mass spectra with 42 peaks in biomarker discovery E2 plasma and 30 peaks for TEG control plasma. Thirteen of these peaks (2952.7, 2965.8, 3004.7, 3027.3, 3226.5, 5226.5, 5250.6, 5351.7, 5376.7, 8106.0, 9199.3, 9352.3, 12,965.0 *m/z*) correctly classified (self-validated) the biomarker discovery E2 and the TEG control samples with 100% sensitivity and 100% specificity and were selected for further evaluation. Twelve proteins were significantly upregulated, and one protein of 12,965 *m/z* was significantly down regulated in the 0.2-μg E2/l mass spectra when compared to spectra of TEG control fish.

TABLE 2
Nominal and Mean Measured Concentrations (± SD) of E2 from Three Water Sampling Periods during Flow-Through Concentration-Response Exposures of Male SHM (*Cyprinodon variegatus*)

Nominal concentration (μg/l)	Mean measured water concentration (μg/l)	Percent of nominal
0.05	0.055 (0.018)	110
0.1	0.12 (0.02)	120
0.2	0.22 (0.02)	111
0.5	0.40 (0.09)	80
1.0	0.91 (0.29)	91

Biomarker Performance

The ability of the 13 discriminator plasma protein peaks to correctly classify fish as estrogen or nonestrogen exposed was further evaluated by testing the model with spectral data from fish exposed to weak estrogen agonists (TPP, BPA, MXC), unexposed adult SWF, unexposed male SHM (TEG, SWM), nonestrogenic chemical stressors (ES, CP), and the E2 concentration-response exposure. All fish plasma samples from all weak estrogen agonist treatments were correctly classified as showing estrogenic effects with 100% sensitivity for all 13 discriminator peaks. For normal SWF plasma, 11 diagnostic peaks correctly classified the plasma samples as estrogen responsive, but two discriminator peaks (5351.7, 12,965 *m/z*) did not correctly classify the samples. All male fish from the nonestrogenic chemical stressors treatments and all negative control male fish (SWM and TEG) were correctly classified as nonestrogenic with 100% specificity. Differences in plasma protein expression levels between estrogen agonist and control treatments are visually apparent in the spectra (Fig. 1).

Within the E2 concentration-response samples, the ability of the 13 discriminator proteins to correctly classify plasma samples from each concentration level as estrogenic or nonestrogenic varied by protein and E2 concentration. Eight protein biomarkers had 100% sensitivity for all estrogen concentrations tested and seven protein biomarkers performed with 100% specificity for the negative controls. Overall, three proteins (2965.8, 3004.7, 9352.3 *m/z*) performed with both 100% sensitivity and 100% specificity (Table 3). The estrogen predictive model was able to detect low concentration effects with > 80% sensitivity for 10 of 13 discriminator peaks (Table 3). A concentration-related increase in plasma protein expression was observed for 12 of the 13 diagnostic peaks. Interestingly, a single peak at 12,965 *m/z* exhibited a strong negative concentration response across the E2 concentrations tested (Fig. 2). The E2 concentration-response results (Table 3) show decreasing sensitivity with decreasing E2 water concentration. Figure 2 illustrates the concentration-response relationships for the 13 protein biomarkers.

Protein Identification

The 3027.3 *m/z* biomarker was identified as SHM ZR protein, C-terminal fragment 2 (accession no. Q6GYP4_CYPVA) using MS/MS fragmentation and MASCOT peptide fragment database searches. A probability-based Mowse score of 30 was obtained for the match, where an individual ion's score of > 15 indicates peptides with significant homology and

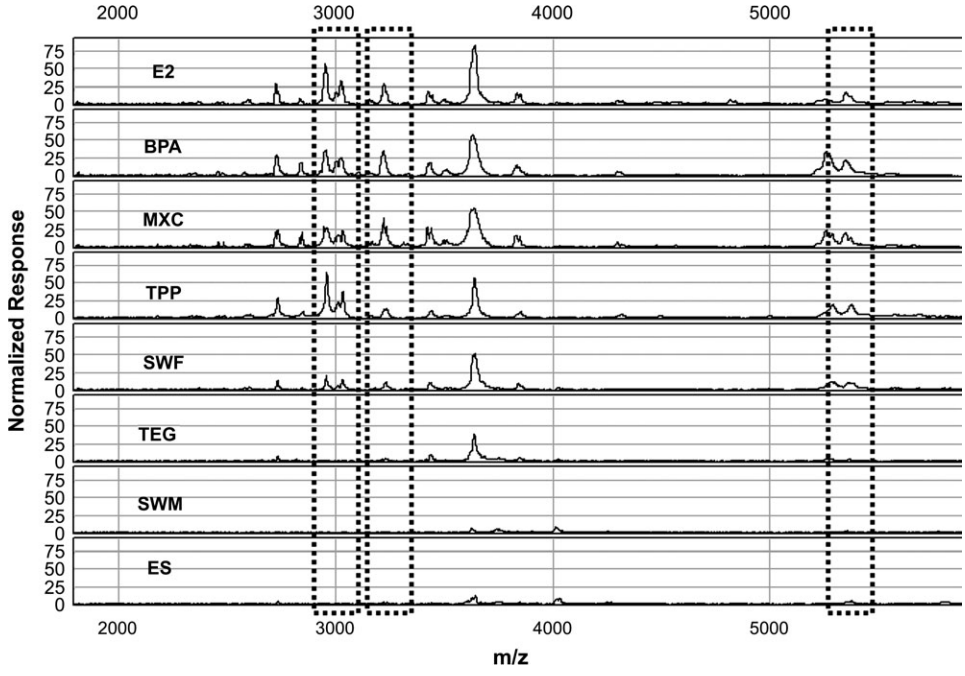

FIG. 1. Representative plasma mass spectra across treatments: male 17β-estradiol 0.2 μg/l E2, male BPA 1000 μg/l, male MXC 6 μg/l, male TPP 100 μg/l, female SWF, male TEG control, male SWM, and male ES 0.6 μg/l. Biomarkers are noted by boxed areas; *m/z*, from left to right, are 2952.7–3027.3, 3226.5, and 5226.5–5376.7, respectively.

> 53 indicates identity or extensive homology with $p < 0.05$. Figure 3 (supplementary data) shows the high-resolution mass spectra of the 3027.3 *m/z* peak acquired using MS/MS. A more accurate *m/z* of 3025.5 was determined for this biomarker using the higher resolution QSTAR MS/MS instrument.

DISCUSSION

The SELDI process was used as a biomarker discovery tool for routine analysis of samples for estrogenic MOA-specific activity, and for protein identification when coupled with

TABLE 3
Sensitivity (% True Positive) and Specificity (% True Negative) of 13 Plasma Estrogen Biomarkers in E2 Concentration-Response Test Set

	Sensitivity (%)					Specificity (%)	
	E2 exposure concentrations (μg/l)					Time zero	Solvent control
m/z	0.05	0.1	0.2	0.5	1.0	0	TEG
2952.7	100	100	100	100	100	80	100
2965.8	100	100	100	100	100	100	100
3004.7	100	100	100	100	100	100	100
3027.3	60	60	100	100	100	100	80
3226.5	100	100	100	100	100	80	100
5226.5	100	100	100	100	100	80	100
5250.6	100	100	100	100	100	80	100
5351.7	80	100	100	100	100	100	100
5376.7	40	100	100	100	100	100	100
8106.0	60	60	100	100	100	100	100
9199.3	100	100	100	100	100	80	100
9352.3	100	100	100	100	100	100	100
12,965.0	80	100	80	100	100	100	100

Note. ($N = 5$ per concentration.)

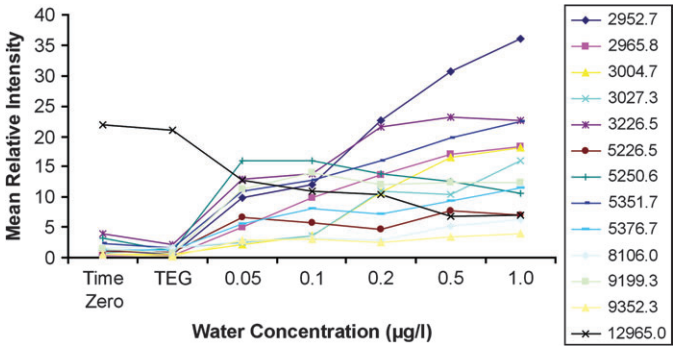

FIG. 2. E2 concentration-response curves for 13 plasma protein biomarkers listed at right (*m*/*z*).

tandem MS. We discovered several estrogen-responsive plasma protein biomarkers in the estuarine test species *C. variegatus* and determined that SELDI analysis of fish plasma proteins to classify chemicals according to their estrogenic properties performs with a high degree of sensitivity, specificity, and reproducibility. The development of statistically valid reference limits for protein biomarkers from untreated male fish plasma allowed for an uncomplicated comparison of plasma protein expression between untreated and estrogen agonist–treated fish. Reference limits are routinely used in clinical chemistry to identify the level at which a biomarker value falls outside of a ''normal'' range (Solberg, 1987) and is of particular value when working with relatively small ($N < 100$) data sets. SELDI-TOF-MS data can also be analyzed using other statistical techniques and pattern recognition algorithms (Wu *et al.*, 2003; Yasui, *et al.*, 2003). However, the reference limits method employed in this study highlights trends and unusual values in the data that could hold biological significance. The data are processed in spreadsheet columns of biomarker response values, and unexpected results can be easily traced back to individual fish and correlated with other forms of experimental data (abnormal histology, abnormal size, weight, liver weight, sickness, etc.).

A classification model has minimal utility unless it can be demonstrated to perform well in predicting unknown samples. External validation with samples representative of the type to be analyzed in real world application is commonly used to verify the performance of a predictive model to acceptably classify samples. Plasma from fish treated with weak estrogen agonists, nonagonist stressors, and normal male and female fish were used for external validation of the predictive model's performance. Three of the biomarker proteins (2965.8, 3004.7, and 9352.3 *m*/*z*) performed with 100% sensitivity and 100% specificity at classifying all samples in every data set. The general decrease in sensitivity with decreasing E2 concentration observed in the concentration-response samples results from a decreasing protein biomarker response (or increasing for 12,965 *m*/*z*) with decreasing concentration such that response levels approach and become indistinguishable from the normal reference intervals. A no observed effect level (NOEL) for E2 induced protein expression (i.e., highest water concentration which fails to produce diagnostic peak intensity greater than ''normal'' reference limits) is less than the lowest concentration examined (0.05 µg E2/l) for all 13 protein biomarkers and could not be determined from the water concentrations tested (Fig. 2). Future efforts will focus on determining the E2 NOEL using this technique.

Validation with estrogen agonist and nonagonist stressor treatments produced unexpected specificity results. In the E2 concentration-response controls (time 0, TEG) for six of the discriminator peaks, one out of five samples in these controls was misclassified as estrogenic (i.e., false positive). Examination of the raw data revealed one individual sheepshead minnow in the time 0 and one in the TEG were responsible for the aberration. The time 0 false positive results for 5 of 13 biomarkers may have resulted from misidentification of a female fish as male. The peak response for the misclassified TEG fish was just above the reference limit for the biomarker and the response was relatively low compared to the mean intensity of true positives. This response level likely represents chromatographic variability associated with responses approaching instrumental detection limits. Protein expression levels in control male plasma were below the detection limits of the SELDI assay for most of the upregulated diagnostic proteins. Where a true peak (i.e., response of > 3 times baseline noise) was present in the E2 treatment samples spectra but not in the control spectra the CiphergenExpress algorithm assigns a ''placeholder'' peak in control spectra. Therefore, the reference values for the ''normal'' response levels for these peaks in control males are simply an estimation of the chromatographic noise in the mass area for each peak and are not biologically based. A larger number of reference values (samples) of control males collected over time would provide a more realistic reference interval and reference limits estimate, and likely would improve specificity and decrease the false positive rate. Although the TEG false positive result occurred with only one biomarker out of 13 and in only one out of five fish, it illustrates a limitation to using the reference limit form of analysis.

The only external validation sample that did not have 100% sensitivity and specificity was the SWF. Two of the diagnostic peaks did not correctly classify the female plasma sample as estrogenic. Females were included as a positive control to demonstrate expression of the estrogen-responsive diagnostic proteins in plasma of mature females. The SWFs used in this experiment were trapped in the wild at different ages, and the data obtained from a sample of 15 pooled females. After plasma collection, gross pathology indicated mature ovaries with eggs present in all females. Therefore, the two false negative results may be attributable to the natural variability and cyclical endogenous estrogen levels inherent in wild fish populations.

A large challenge when using protein biomarker techniques is the isolation and subsequent identification of proteins. Identification of protein biomarkers discovered using low resolution SELDI may be accomplished through the use of high-resolution tandem MS, which is specifically adapted to analyze SELDI chips. Using this approach we identified one discriminatory peak as a fragment of ZR protein, and identification of additional biomarker proteins is ongoing. Estrogen agonists induce the biosynthesis of ZR proteins in fish and other oviparous species (Arukwe *et al.*, 2002), and ZR proteins appear in plasma earlier than VTG (Celius and Walther, 1998). Male and juvenile fish usually have low or no detectable plasma ZR proteins, so elevated levels of these proteins in males or juveniles indicate exposure to chemical(s) with estrogenic activity (Arukwe *et al.*, 1997; Larsson *et al.*, 1994).

It is important to note that results presented in this study were obtained from a single type of chip surface and wash conditions. The SELDI fractionation approach physically separates proteins into classes based upon chip surface (such as cationic/anionic or hydrophilic/hydrophobic) and wash conditions and greatly facilitates the characterization of proteins in a biological sample. Therefore, the protein biomarkers isolated in the present study represent only a small fraction of the entire inventory of proteins in the plasma. The use of additional types of chip surfaces and wash conditions, which utilize the inherent binding characteristics of a sample's proteins, are useful to further fractionate the sample and likely generate many more protein fingerprints and biomarkers. This approach will be applied to additional MOAs, such as antiestrogens androgens, and other biological models, to develop diagnostic protein biomarkers, and is currently ongoing.

Advances in ''omics'' technologies, such as genomics, proteomics, and metabolomics, have enabled applications for the discovery of gene and gene product expression biomarkers and profiles using a simple but potent design—comparing expression of two biological classes to identify biomarkers that are differentially expressed. All of these methods seek to discover a molecular biomarker or pattern to classify biological samples based on gene or gene product expression. However, biomarker discovery for diagnostic, prognostic, and research uses can consume much time and resources. Hence, it is imperative that the most efficient approaches are employed for biomarker discovery. Each of the ''omic'' technologies has inherent strengths and the biomarker which performs best in terms of addressing a defined problem will ultimately determine the type of technology selected (MacGregor, 2003). Genomics, proteomics, and metabolomics can be applied to tissues but proteomics and metabolomics have a particular advantage over genomic approaches in that they can be based on accessible biofluids such as plasma that do not contain mRNA. Further, proteomic and metabolomic approaches better represent the true functional responses of an organism to a given environmental stressor. In the case of a MOA that results in relatively large changes in protein expression, such as estrogen agonist activity, protein expression may offer the most effective approach.

The general process presented here allows for the exciting possibility that libraries of MOA-specific protein biomarkers and patterns may be developed. Once such libraries are developed, animals can be exposed to a chemical, their plasma or other tissue analyzed by the rapid protein biomarker profile process, and the resultant protein expression profile(s) compared to a protein expression library for determination of multiple MOA effects from a single treatment. Detection of multiple MOAs through protein expression profiling presents a viable option for ultimately replacing many of the single MOA based *in vivo* mammalian tests currently used in safety assessment and would greatly reduce the animal numbers needed for such assessments. Although the diagnostic proteins expressed may be specific to the species tested, the general technique can be transferred to any species of interest. For example, ZR proteins such as the one identified in this study are conserved among teleost fish (Epifano *et al.*, 1995; Oppen-Berntsen *et al.*, 1999) and such conservation may permit the direct application of this technique in additional fish species by examining protein homologs of *C. variegates* protein biomarkers.

In conclusion, plasma protein profiling using a small fish model and SELDI-TOF-MS provided a sensitive and specific high-throughput method to screen chemicals for estrogenic activity. A novel protein chip process was used to discover, isolate, analyze, and identify plasma protein biomarkers from estrogen agonist–treated fish. The technique performed with a high degree of sensitivity, specificity, and reproducibility for classifying male fish as either normal or exposed to an estrogen agonist. In the future, tissue protein biomarkers and profiling will be an essential element of toxicoproteomics and will have a direct impact on toxic disease diagnosis, safety assessment, and environmental monitoring.

SUPPLEMENTARY DATA

Supplementary data are available online at http://toxsci.oxfordjournals.org/.

ACKNOWLEDGMENTS

This work was supported in part by a Cooperative Research and Development Agreement (CRADA CR1-26XQX55-20203F) with the American Chemistry Council.

REFERENCES

Ankley, G. T., Jensen, K. M., Makynen, E. A., Kahl, M. D., Korte, J. J., Hornung, M. W., Henry, T. R., Denny, J. S., Leino, R. L., Wilson, V. S., *et al.* (2003). Effects of the androgenic growth promoter 17-β-trenbolone on

fecundity and reproductive endocrinology of the fathead minnow. *Environ. Toxicol. Chem.* **22,** 1350–1360.

Arukwe, A., Knudsen, F. R., and Goksøy, A. (1997). Fish zona radiata (eggshell) protein: A sensitive biomarker for environmental estrogens. *Environ. Health Perspect.* **105,** 418–422.

Arukwe, A., Kullman, S. W., Berg, K., Goksøy, A., and Hinton, D. E. (2002). Molecular cloning of rainbow trout (*Oncorhynchus mykiss*) eggshell *zona radiata* protein complementary DNA: mRNA expression in 17β-estradiol- and nonylphenol-treated fish. *Comp. Biochem. Physiol. Part B* **132,** 315–326.

Celius, T., and Walther, B. T. (1998). Oogenesis in Atlantic salmon (*Salmo salar* L.) occurs by zonagenesis preceding vitellogenesis *in vivo* and *in vitro*. *J. Endocrinol.* **158,** 259–266.

Colborn, T., von Saal, F. S., and Soto, A. M. (1993). Developmental effects of endocrine-disrupting chemicals in wildlife and humans. *Environ. Health Perspect.* **101,** 378–384.

Epifano, O., Liang, L., and Dean, J. (1995). Mouse Zp1 encodes a zona pellucida protein homologous to egg envelope proteins in mammals and fish. *J. Biol. Chem.* **270,** 27254–27258.

Etienne, M., and Jemmali, M. (1982). Effects of zearalenone (F2) on estrous activity and reproduction in gilts. *J. Anim. Sci.* **55,** 1–10.

Guerrier, L., Lomas, L., and Boschetti, E. (2005). A simplified monobuffer multidimensional chromatography for high-throughput proteome fractionation. *J. Chromatogr. A* **1073,** 25–33.

Hemmer, M. J., Hemmer, B. L., Bowman, C. J., Kroll, K. J., Folmar, L. C., Marcovich, D., Hoglund, M. D., and Denslow, N. D. (2001). Effects of p-nonylphenol, methoxychlor and endosulfan on vitellogenin induction and expression in the sheepshead minnow, *Cyprinodon variegatus*. *Environ. Toxicol. Chem.* **20,** 336–343.

Hutchens, T. W., and Yip, T. T. (1993). New desorption strategies for the mass spectrometric analysis of macromolecules. *Rapid Commun. Mass Spectrom.* **7,** 576–580.

Kavlock, R. J., Daston, G. P., DeRosa, C., Fenner-Crisp, P., Gray, L. E., Kaattari, S., Lucier, G., Luster, M., Mac, M. J., Maczka, C., *et al.* (1996). Research needs for the risk assessment of health and environmental effects of endocrine disrupters: A report of the U.S. EPA-sponsored workshop. *Environ. Health Perspect.* **104**(Suppl. 4), 715–740.

Larsson, D. G. J., Hyllner, S. J., and Haux, C. (1994). Induction of vitelline envelope proteins by estradiol-17β in 10 teleost species. *Gen. Comp. Endocrinol.* **96,** 445–450.

MacGregor, J. T. (2003). The future of regulatory toxicology: Impact of the biotechnology revolution. *Toxicol. Sci.* **75,** 236–248.

Mehta, A. I., Ross, S., Lowenthal, M. S., Fusaro, V., Fishman, D. A., Petricoin, E. F. III, and Liotta, L. A. (2003–2004). Biomarker amplification by serum carrier protein binding. *Dis. Markers* **19,** 1–10.

Oppen-Berntsen, D. O., Arukwe, A., Yadetie, F., Lorens, J. B., and Male, R. (1999). Salmon eggshell protein expression: A marker for environmental estrogens. *Mar. Biotechnol.* **1,** 252–260.

Parks, L. G., Lambright, C. S., Orlando, E. F., Guillette, L. J., Jr, Ankley, G. T., and Gray, L. E. Jr. (2001). Masculinization of female mosquitofish in kraft mill effluent-contaminated Fenholloway river water is associated with androgen receptor agonist activity. *Toxicol. Sci.* **62,** 257–267.

Rotchell, J. M., and Ostrander, G. K. (2003). Molecular markers of endocrine disruption in aquatic organisms. *J. Toxicol. Environ. Health Part B* **6,** 453–495.

Solberg, H. E. (1987). Establishment and use of reference values. In *Fundamentals of Clinical Chemistry* (N. W. Tietz, Ed.), 3rd ed., pp. 197–212. W. B. Saunders Co., Philadelphia.

Tirumalai, R. S., Chan, K. C., Prieto, D. A., Issaq, H. J., Conrads, T. P., and Veenstra, T. D. (2003). Characterization of the low molecular weight human serum proteome. *Mol. Cell Proteomics* **2,** 96–103.

U.S. Environmental Protection Agency. (1998). Endocrine disruptor screening program; proposed statement of policy. *Fed. Regist.* **63,** 71542–71568.

U.S. Office of Science and Technology Policy. (1985). U.S. government principles for the utilization and care of vertebrate animals used in testing, research and training. *Fed. Regist.* **50,** 20864–20865.

Wu, B., Abbott, T., Fishman, D., McMurray, W., Mor, G., Stone, K., Ward, D., Williams, K., and Zhao, H. (2003). Comparison of statistical methods for classification of ovarian cancer using mass spectrometry data. *Bioinformatics* **19,** 1636–1643.

Yasui, Y., Pepe, M., Thompson, M. L., Adam, B.-L., Wright, G. L., Jr, Qu, Y., Potter, J. D., Winget, M., Thornquist, M., and Feng, Z. (2003). A data-analytic strategy for protein biomarker discovery: Profiling of high-dimensional proteomic data for cancer detection. *Biostatistics* **4,** 449–463.