Estimation of Estradiol in Mouse Serum Samples: Evaluation of Commercial Estradiol Immunoassays

Daniel J. Haisenleder, Aleisha H. Schoenfelder, Elizabeth S. Marcinko, Lisa M. Geddis, and John C. Marshall

Ligand Assay and Analysis Core Facility, University of Virginia Center for Research in Reproduction, Charlottesville, Virginia 22908

The University of Virginia Center for Research in Reproduction Ligand Core performed an evaluation of nine commercial estradiol (E2) immunoassays for use with mouse serum. The evaluation had two components. 1) Recovery Studies: a mouse pool was spiked with E2 concentrations across the assay range, and percent recovery and parallelism to the assay standard curve were determined. 2) Correlation Studies: serum pools were collected from intact females, ovariectomized (OVX) and OVX-E2 treated mice and E2 assayed, then measured by gas chromatography/tandem mass spectrometry (GC/MSMS) for comparison to a gold standard method. Recovery results showed that E2 recovery from spiked mouse pools varied greatly (from <18% to >640%) among kits tested. However, three kits (DiaSorin Radioimmunoassay, Siemens Double Antibody RIA, and CalBiotech Enzyme Immunoassay) showed reasonable recoveries and parallelism. Data collected from the Correlation Study showed that values from intact, OVX and OVX-E2-treated mouse pools varied by several fold *vs.* GC/MSMS for most of the kits tested. The DiaSorin RIA and CalBiotech Enzyme Immunoassay Kits showed the best correlation to GC/MSMS. Unfortunately, while this evaluation was ongoing, the DiaSorin Kit was discontinued. In summary, the CalBiotech Kit was the only available assay tested that demonstrated good E2 parallelism to the assay standard curve and accuracy *vs.* a gold standard method (*i.e.* GC/MSMS). Also of note, the CalBiotech assay is sensitive and requires minimal sample volume. Therefore, based on these findings the CalBiotech E2 assay has been implemented for use in mouse serum samples within the Ligand Core. **(*Endocrinology* 152: 4443–4447, 2011)**

The University of Virginia Center for Research in Reproduction, Ligand Assay and Analysis Core (web site: http://www.medicine.virginia.edu/research/institutes-and-programs/crr/labfacilities-page) performs reproductive hormone assays for National Insstitute of Child Health and Human Development Reproductive Science Branch-supported investigators across the country. Various tests offered by the Core are used for serum or plasma samples from human, rat, mouse, and other mammalian species. Serum/plasma matrix, defined as the profile of proteins (including binding proteins), lipids, and other endogenous compounds that circulate within any mammalian species, can adversely influence the validity of an assay for samples from other species (1). Therefore, it is the policy of the Ligand Core to validate assays for each species. For most cases, multispecies assays are limited to steroids but also include some specific gonadal peptides (*e.g.* activins, inhibins, anti-Mullerian hormone) (2, 3).

One of the higher volume tests that is performed by the Ligand Core is mouse estradiol (E2). Since 2006, the method used to measure E2 in mouse serum has been the Beckman Coulter Coated-Tube (CT) RIA. In 2009, some investigators expressed concerns about the accuracy of the assay. In response to these concerns, we reevaluated the kit and found reduced performance for mouse serum.

This manuscript describes the evaluation of commercial kits as potential replacement methods for mouse E2. There are numerous commercial E2 kits, most of which are

ISSN Print 0013-7227 ISSN Online 1945-7170
Printed in U.S.A.

doi: 10.1210/en.2011-1501 Received July 13, 2011. Accepted August 29, 2011.
First Published Online September 20, 2011

Abbreviations: CT, Coated tube; CV, coefficient of variation; E2, Estradiol; EIA, enzyme immunoassay; GC/MSMS, gas chromatography/tandem mass spectrometry; OVX, ovariectomized; QC, quality control.

designed for human serum samples. The kits selected were based on a literature search that showed that greater than 80% of all published reports between 2007 and 2009 that presented serum E2 data in mice used one of the nine kits examined for this study.

Materials and Methods

Commercial E2 kits examined

The following E2 kits were tested for this study: Beckman Coulter CT RIA and Beckman Coulter Ultra-Sensitive RIA (Beckman Coulter, Inc., Webster, TX); ALPCO Enzyme Immunoassay (EIA) (ALPCO Diagnostics, Salem, NH); BioVendor EIA (BioVendor LLC, Candler, NC); CalBiotech EIA (Spring Valley, CA); DiaSorin RIA (DiaSorin, Inc., Minneapolis, MN); Siemens Double Antibody RIA and Siemens Coated-Tube RIA (Siemens USA, Los Angeles, CA); IBL-CisBio RIA (CIS-Bio International, Toronto, Ontario, Canada).

Recovery studies

A commercial mouse serum pool (catalog no. M5905; Sigma Chemical Co., St. Louis, MO) was spiked with various E2 concentrations, and percent recovery and parallelism to the assay standard curve were determined. For each assay, the serum pool was spiked either with E2 across the assay range or vehicle (to determine endogenous E2 levels in the pool). Samples were run in duplicate, and each assay was repeated to confirm results. E2 recovery from each spiked sample was determined by subtracting E2 values in vehicle-spiked controls from E2-spiked samples. Data presented in Tables 1 and 2 are mean values from two separate assays.

Correlation studies

Serum pools were collected (kindly provided by Dr. Sue Moenter, University of Michigan, Ann Arbor, MI) from ovary-intact females, ovariectomized (OVX), and ovariectomized + E2-treated mice (10–20 mouse serum samples per pool) and were assayed in each kit, and then measured by gas chromatograph/tandem mass spectrometry (GC/MSMS; Taylor Tech, Princeton, NJ) for comparison to a gold standard method (4).

Assay performance

Performance data for each commercial E2 kit were provided by the manufacturer, as listed below.

Beckman Coulter CT RIA. Sensitivity: 9 pg/ml; precision: 2.4% (intraassay), 7.3% (interassay); specificity (cross-reactivity): estrone, 3.4%; estriol, 0.84% (other endogenous steroids tested were undetectable).

Beckman Coulter Ultra-Sensitive RIA. Sensitivity: 2.5 pg/ml; precision: 8.9% (intraassay), 12.2% (interassay); specificity: estrone, 2.4%; estriol, 0.64% (other endogenous steroids tested were undetectable).

ALPCO EIA. Sensitivity: 3 pg/ml; precision: 6.4% (intraassay), 7.6% (interassay); specificity: estrone, 0.2%; estriol, 0.05% (other endogenous steroids tested were undetectable).

BioVendor EIA. Sensitivity: 10 pg/ml; precision: 7.7% (intraassay), 8.7% (interassay); specificity: estrone, 1.3%, estriol, 1.6%, progesterone, 0.1%, cortisol, 0.1%.

CalBiotech EIA. Functional sensitivity (see below): 3 pg/ml; precision: 3.1% (intraassay), 9.9% (interassay); specificity: progesterone, 0.0002%, androstenedione, 0.0001%, testosterone, 0.0002%, cortisol, 0.0001% (other endogenous steroids tested were <0.0001% or undetectable). Sensitivity and precision data were collected by the Ligand Core as described below.

Functional sensitivity, defined as the lowest analyte concentration that can be reliably measured [*i.e.* the lowest concentration with a coefficient of variation (%CV) <20%]. Quality controls (QC) were serially diluted, and four replicates of each dilution were run in a single assay. This procedure was repeated on four separate days. Mean ± SD and %CV calculated from all data points for each diluted pool and functional sensitivity were determined.

Intraassay Precision. Low, medium, and high QC were run within a single assay (10 replicates per QC), and mean ± SD and %CV were calculated.

Interassay precision. QC (low, medium, and high) run in duplicate over time (five to six individual assays) and mean ± SD and %CV were calculated.

DiaSorin RIA. Sensitivity: 4 pg/ml; precision: 3.8% (intraassay), 5.1% (interassay); specificity: estrone, 0.63%; estriol, 0.65%; testosterone, 0.001% (other endogenous steroids tested were <0.01%).

Siemens Double Antibody RIA. Sensitivity: 3 pg/ml; precision: 4.0% (intraassay), 4.6% (interassay); specificity: estrone, 12.5%; estriol, 0.24% (other endogenous steroids tested were undetectable).

Siemens Coated-Tube RIA. Sensitivity: 8 pg/ml; precision: 5.3% (intraassay), 6.4% (interassay); specificity: estrone, 10%; estriol, 0.32%; testosterone, 0.001%; dihydrotestosterone, 0.004% (other endogenous steroids tested were <0.001% or undetectable).

IBL-CisBio RIA. Sensitivity: 9 pg/ml; precision: 5.4% (intraassay); 8.2% (interassay); specificity: estrone, 1.8%; estriol,

TABLE 1. E2 recovery from mouse pools spiked with various concentrations of E2 (Beckman Coulter CT Kit)

Original validation (2006)			Repeat validation (2009)		
E2 added (pg/ml)	E2 recovered (pg/ml)	% Recovery	E2 added (pg/ml)	E2 recovered (pg/ml)	% Recovery
500	481	96	500	460	92
250	228	91	250	210	84
125	135	108	125	110	88
63	71	112	63	48	76
32	37	115	32	19	59

TABLE 2. E2 recovery from mouse pools spiked with various concentrations of E2

Siemens double antibody			Siemens CT RIA			Beckman Coulter CT RIA		
E2 added (pg/ml)	E2 recovered (pg/ml)	% Recovery	E2 added (pg/ml)	E2 recovered (pg/ml)	% Recovery	E2 added (pg/ml)	E2 recovered (pg/ml)	% Recovery
200	193	97	700	1016	145	500	460	92
100	73	73	350	529	151	250	210	84
50	36	72	175	255	146	125	110	88
2	20	80	88	173	198	63	48	76
12.5	10	80	44	81	184	32	19	59
Beckman Coulter Ultra-Sensitive			**ALPCO EIA**			**CalBiotech EIA**		
E2 added (pg/ml)	E2 recovered (pg/ml)	% Recovery	E2 added (pg/ml)	E2 recovered (pg/ml)	% Recovery	E2 added (pg/ml)	E2 recovered (pg/ml)	% Recovery
100	33	33	100	57	57	150	262	175
50	12	24	50	20	40	75	128	170
25	10	40	25	10.4	42	38	63	167
12.5	6	48	12.5	3.7	30	19	32	171
6.3	6	95	6.3	<3	<48	9.5	16.3	172
IBL-CisBio RIA			**BioVendor**			**DiaSorin RIA**		
E2 added (pg/ml)	E2 recovered (pg/ml)	% Recovery	E2 added (pg/ml)	E2 recovered (pg/ml)	% Recovery	E2 added (pg/ml)	E2 recovered (pg/ml)	% Recovery
200	106	53	500	>3200	>640	200	287	143
100	28	28	250	1131	452	100	141	141
50	<9	<18	125	213	170	50	70	140
25	<9	*	63	125	198	25	34	136
12.5	<9	*	32	15	47	12.5	17	136
6.3	<9	*				6.3	8.2	130

*, Indeterminate.

1.2%; testosterone, 0.0012%; androstenedione, 0.0011% (other endogenous steroids tested were <0.001% or undetectable).

GC/MSMS. Sensitivity: 3.1 pg/ml; precision: 8.1% (intraassay; all sample pools for this study were run in a single assay).

Results

In 2006, the Ligand Core implemented the Beckman Coulter CT E2 RIA for use in mouse serum samples after validation studies that included a recovery and parallelism evaluation (as described in *Materials and Methods*). However, in 2009, a few investigators informed us that they were concerned about recent assay results showing values in ovary-intact proestrous mice that were below the limit of detection. In response, QC data were reviewed over the previous 3 yr, with no systematic drifts or shifts in assay performance observed. As a follow up, the Core repeated the recovery/parallelism study. The results of the original and repeat validation studies are presented in Table 1. For the original study, E2 recovery was maintained within 15% of expected values across the assay range. In contrast, the repeat validation study showed a drop in recovery at the lower end of the curve, resulting in a loss of parallelism to the standard curve for mouse serum samples. In response to the decrease in E2 recovery, the Core contacted the Beckman Coulter Technical Service Department to obtain possible explanations for the change in assay performance. Unfortunately, the kit manufacturer's technical representatives were unable to provide an explanation or resolution to the problem. Therefore, based on the updated validation data, the assay was put on hold, and plans were initiated to find a replacement method.

The Ligand Core ran recovery studies in eight kits, and findings are presented in Table 2 (the repeat Beckman Coulter CT RIA evaluation from Table 1 is included for comparison). Results show that E2 recovery from spiked mouse pools varied greatly (from <18% to >640%) among the kits tested. Only three kits (Siemens Double Antibody RIA, CalBiotech EIA, and DiaSorin RIA) showed good parallelism to the standard curve across the assay range.

Measured E2 concentrations in serum pools from ovary-intact female, OVX, and OVX+E2-treated mice are presented in Table 3. The Siemens CT RIA was not included in this Method Comparison Study, to conserve precious sample pool volumes and because the kit did not demonstrate good parallelism to the standard curve in the Recovery Study (Table 2). Results showed that values from intact, OVX, and OVX+E2-treated mouse pools varied

TABLE 3. Mouse E2 assay: method comparison (pg/ml)

Group	BV EIA	IBL RIA	BC-CT RIA	ALPCO EIA	S-DA RIA	BC US RIA	CB EIA	DS RIA	GC MSMS
Intact female	71.9	<9.0	<9.0	<3.0	<3.0	<2.5	13.7	7.8	4.0
	49.8	<9.0	<9.0	<3.0	<3.0	<2.5	8.7	<4.0	<3.1
	53.9	<9.0	<9.0	<3.0	<3.0	<2.5	8.3	5.9	<3.1
OVX	57.9	<9.0	<9.0	<3.0	<3.0	<2.5	6.7	<4.0	<3.1
	39.5	<9.0	<9.0	<3.0	<3.0	<2.5	8.4	<4.0	<3.1
	29.1	<9.0	<9.0	<3.0	<3.0	<2.5	4.3	<4.0	<3.1
OVX + E2	40.8	<9.0	<9.0	<3.0	<3.0	<2.5	9.9	7.0	8.9
	36.3	<9.0	<9.0	6.4	12.0	9.5	26.2	25.6	42.8
	56.1	14.1	34.3	19.8	62.5	40.3	73.2	75.5	99.6
	41.5	<9.0	<9.0	<3.0	<3.0	5.2	11.1	9.4	12.1
	72.4	<9.0	<9.0	<3.0	<3.0	<2.5	11.8	9.7	14.5

The assay results for each kit are presented in order of worst (*left*) to best (*right*) correlation to GC/MSMS.

BV EIA, BioVendor EIA; IBL RIA, IBL-CisBio RIA; BC CT RIA, Beckman Coulter Coated Tube RIA; ALPCO EIA, ALPCO Enzyme Immunoassay; S-DA RIA, Siemens Double Antibody RIA; BC US RIA, Beckman Coulter Ultra-Sensitive RIA; CB EIA, CalBiotech EIA; DS RIA, DiaSorin RIA.

by several fold among the kits tested. Only three kits were able to detect serum E2 in intact female mice (CalBiochem, DiaSorin, and BioVendor). However, values from the BioVendor EIA were considerably higher than data collected from other kits examined and did not correlate to samples measured by GC/MSMS [correlation coefficient (r^2) = 0.002]. DiaSorin RIA (r^2 = 0.983) and CalBiotech EIA (r^2 = 0.969) showed the best correlation to GC/MSMS.

Discussion

The present study demonstrates that commercial E2 kit performance can vary greatly for mouse serum, as far as recovery and accuracy. Our results showed that two kits (DiaSorin and CalBiotech) performed the best overall *vs.* the other kits examined. Both kits correlated very well to GC/MSMS, and E2-spiked mouse serum pools were parallel to each respective standard curve. Unfortunately, while this evaluation was ongoing, the Core was notified that the DiaSorin Kit was to be discontinued. Therefore, based on the present findings the CalBiotech EIA was implemented by the Ligand Core for measuring E2 in mouse serum. Also of note, the CalBiotech assay is sensitive (limit of detection = 3 pg/ml) and requires minimal sample volume (25 μl/well), which enhance the utility of the assay for the mouse, where limited samples volume is a critical issue.

Although the DiaSorin and CalBiotech Kits performed better than the other assays tested, recovery from spiked samples was high (130–175%). It is likely that this observation reflects differences in serum matrix. Like most commercial steroid kits, the Diasorin RIA is designed for human samples and uses assay standards in human serum. The CalBiotech EIA is advertised as a Rat/Mouse kit, but also uses standards in human serum. Hormones (*e.g.* proteins, steroids) are measured in various biological fluids (*e.g.* serum, plasma, urine, cerebral spinal fluid, saliva, *etc.*) using immunoassay methods. The matrix of biological fluids plays an important role in assay performance (5), and it is recognized that an immunoassay designed for use in one biological fluid is not necessarily viable for others (6). Of importance to the present study, serum matrix can also play a significant role in assay validity for different species (5). A change in serum matrix can affect assay performance by altering antibody/antigen reaction kinetics or affecting antibody/antigen binding affinity, resulting in interference (falsely low signal) or cross-reactivity (falsely high signal) (1, 7, 8).

To varying degrees, the other kits tested for this study showed significant drawbacks relating to insufficient sensitivity to detect E2 in intact female mice, and/or poor parallelism to the assay standard curve. One kit in particular, BioVendor EIA, performed the poorest as far as E2 recovery and correlation to GC/MSMS, with significantly higher measured values.

In contrast to most of the kits tested, serum E2 was detectable in OVX samples in the CalBiotech EIA (Table 3). Two of the most likely explanations for this observation are 1) antibody cross-reactivity with another steroid, steroid metabolite, or serum compound; 2) serum matrix effects, as discussed above. We cannot say for certain which mechanism is responsible for the present observation. However, similar higher than expected values were noted in response to spiking a mouse serum pool with E2, which resulted in recoveries of 167–175% (Table 2) across the assay range. For the present recovery studies, endogenous E2 values in the vehicle-treated serum pool were subtracted from levels measured in all E2-spiked samples. If the assay was detecting a cross-reacting endogenous compound, that component would be constant across all E2-spiked samples and subtracted from the final calcu-

lated E2 recovery value. Therefore, we believe that serum matrix is the likely explanation for the detectable concentrations in OVX mice with the CalBiotech EIA.

The most significant limitation of the CalBiotech assay and all other methods tested (including GC/MSMS) is the inability to distinguish between values in ovary-intact *vs.* OVX mice. Due to the large volume of serum required for the method comparison study (Table 3), measuring E2 in serum pools from proestrous mice was not feasible for this evaluation. Each pool used for the method comparison included serum from 10–20 animals. Although it is likely that serum from proestrous animals were included in the ovary-intact group, their contribution to measured E2 values were diminished due to sample dilution. However, because serum E2 levels increase several fold on proestrus, it is probable that the assay will distinguish between high physiological and basal E2 levels in cyclic mice.

For most assays performed in the Ligand Core, we use Bio-Rad Tri-Level QC (Bio-Rad Laboratories, Inc., Hercules, CA) because the material is well characterized and can be purchased in large batches, which allow for the use of the same lot of reagent for up to 2 yr. However, because Bio-Rad QC are a human-derived product, a change in critical kit reagents that might adversely affect the viability of mouse serum in the assay could go unnoticed. Therefore, with the implementation of the new assay, we now include mouse serum-based QC, which allow for long-term monitoring of assay performance in a mouse serum matrix.

The accurate measurement of circulating E2 levels in the mouse is critical to the research of many reproductive investigators. Evidence from various studies suggests that GC/MSMS is a gold standard method (9–11). However, critical factors, including cost of instrumentation, technical expertise, sample volume required, and method sensitivity, make its utility as a research assay not feasible for mouse serum samples. To underline these points, the sample pools listed in Table 3 were sent to Taylor Tech for GC/MSMS analysis, based on a published report showing state of the art sensitivity and accuracy (4). However, to achieve a level of detectability that is similar to the Calbiotech EIA (*i.e.* 3 pg/ml) required 200 μl of serum per singlet measurement (the CalBiotech EIA uses 25 μl of serum per singlet and is therefore 8 times more sensitive than the present GC/MSMS system). Also, 200 μl of serum are difficult to obtain from mice (if serum is needed to measure other tests), and the cost of using GC/MSMS is prohibitive ($225 per singlet determination). Thus, the use of immunoassay technology to measure E2 in mice may not be state of the art, but if the method is properly characterized, we are confident that resulting data are scientifically valid.

In conclusion, based on these findings, the Ligand Core has implemented the CalBiochem EIA for the measurement of E2 in mouse serum. Also, the present study supports the concept that appropriate steroid assay validation studies are required before implementation for different species, and the use of species-specific QC are essential for monitoring E2 assay performance. Finally, although the CalBiotech assay is an acceptable method for the present, the development of enhanced sensitivity, highly specific, and economical assays for mouse E2 are needed in the future.

Acknowledgments

Address all correspondence and requests for reprints to: Daniel J. Haisenleder, Aurbach Medical Research Building, P.O. Box 801412, University of Virginia Health Sciences Center, Charlottesville, Virginia 22908. E-mail: djh2q@virginia.edu.

This work was supported by the Reproductive Science Branch of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development; SCCPIR U54-HD28934.

Disclosure Summary: No author of this manuscript has any conflicts of interest to disclose.

References

1. **Pratt JJ** 1978 Steroid immunoassays in clinical chemistry. Clin Chem 24:1869–1890
2. **Haisenleder DJ, Schoenfelder AH, Harrison M** Validation of the Diagnostic Systems Laboratories, Inc. (DSL) human inhibin (A and B) and free activin (A, B and AB) assays for the mouse and rat. Proc SCCPIR 3rd Annual Meeting, Portland, OR, 2007, p 27
3. **Haisenleder DJ, Schoenfelder AH, Marcinko E** 2009 Validation of the Beckman-Coulter DSL Human Anti-Mullerian Hormone Assay for Mouse and Rat. Proc SCCPIR 4th Annual Meeting, Chicago, IL, 2009, p 31
4. **Santen RJ, Demers L, Ohorodnik S, Settlage J, Langecker P, Blanchett D, Goss PE, Wang S** 2007 Superiority of gas chromatography/tandem mass spectrometry assay (GC/MSMS) for estradiol for monitoring of aromatase inhibitor therapy. Steroids 72:666–671
5. **Miller JJ, Valdes Jr R** 1991 Approaches in minimizing interference by cross-reacting molecules in immunoassays. Clin Chem 37:144–153
6. **Carlström K** 1996 Low endogenous estrogen levels - analytical problems and tissue sensitivity. Acta Obstet Gynecol Scand Suppl 163: 11–15
7. **Pratt JJ, Koope W, Woldring MG, Wiegman T** 1976 Effect of plasma proteins on the specificity of steroid immunoassays. Eur J Nucl Med 1:37–48
8. **Pratt JJ, Woldring MG, Boonman R, Kittikool J** 1979 Specificity of immunoassays. Heterogeneity of specificity of antibodies in antisera used for steroid immunoassay and the selective blocking of less specific antibodies, including a new method for the measurement of immunoassay specificity. Eur J Nucl Med 4:159–170
9. **Dehennin L** 1989 Estradiol determined in plasma by GC/MS with selected ion monitoring of mixed silyl ether-perfluoroacyl ester derivatives and use of various stable isotope-labeled internal standards. Clin Chem 35:532–536
10. **Nelson RE, Grebe SK, OKane DJ, Singh RJ** 2004 Liquid chromatography-tandem mass spectrometry assay for simultaneous measurement of estradiol and estrone in human plasma. Clin Chem 50:373–384
11. **Soldin SJ, Soldin OP** 2009 Steroid hormone analysis by tandem mass spectrometry. Clin Chem 55:1061–1066