Data Article

iTRAQ-Based proteomic dataset for bovine pre-ovulatory plasma and follicular fluid containing high and low Estradiol

P.A. Afedi a, E.L. Larimore b, R.A. Cushman c, D. Raynie a, G.A. Perry b,*

a Department of Chemistry and Biochemistry, South Dakota State University, Brookings, South Dakota, USA
b Department of Animal Science, South Dakota State University, Brookings, South Dakota, USA
c USDA-ARS U.S. Meat Animal Research Center, Clay Center, Nebraska 68933, USA

ABSTRACT

This is isobaric tags for a relative and absolute quantification (iTRAQ)-Based Proteomic Data on bovine plasma (PL) and follicular fluid (FF) containing high and low pre-ovulatory circulating concentration of estradiol (E2). The PL and FF were collected from nine beef cows that were identified to initiate a new follicular wave on day -4 during synchronization. Follicular dynamics and ovulatory response were monitored using transrectal ultrasonography. Blood samples were collected at slaughter and FF was aspirated from dominant follicles (DF; > 10 mm). Estradiol concentrations in PL and FF were measured by radioimmunoassays. Plasma and FF were labeled as containing high E2 (PL HE2 and FF HE2) or low E2 (PL LE2 and FF LE2). Abundant proteins (albumin, IgG, IgA, and alpha-1-antitrypsin) were depleted from the four PL and FF samples. Peptides were labeled with iTRAQ reagents and analyzed using 2-dimensional liquid chromatography ESI-based mass spectrometry. Proteins were identified and quantified using SEQUEST™ search engine embedded in Proteome Discoverer. The proteins matched with at least one unique peptide at minimum 95% confidence were considered positive identifications. Protein expression levels were determined by

https://doi.org/10.1016/j.dib.2021.106998
2352-3409/© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Specifications Table

| Subject                  | Animal Science |
|-------------------------|----------------|
| Specific subject area   | Folliculogenesis and ovulation, reproductive associated proteome in follicular fluid (FF) and plasma (PL) |
| Type of data            | Tables |
| How data were acquired  | Isobaric tags for a relative and absolute quantification [iTRAQ]-labeled strong cation exchanged (SCX) fractionated peptides were analyzed by Thermo-Fisher Finnigan™ LTQ™ mass spectrometer equipped with a nano electrospray source (New Objective, Woburn, MA, USA) and coupled with a nano-LC separation system (Eksigent nanoLC 1D-plus). The LC system is equipped with an autosampler (Spark Holland 920 Endurance Autosampler). The LTQ mass spectrometer was operated in a data-dependent mode. The acquired MS/MS LTQ data were correlated to theoretical fragmentation patterns of tryptic peptide sequences in bovine protein fasta database using SEQUEST™ search engine embedded in Proteome Discoverer (version 2.1; Thermo Fisher Scientific). Identified proteins were further analyzed using Protein Analysis Through Evolutionary Relationships (PANTHER, Version 14.1) and Database for Annotation, Visualization and Integrated Discovery (DAVID, Version 6.8). |
| Data format             | EXCEL |
| Parameters for data collection | iTRAQ-labeled SCX fractionated peptides were analyzed by nano-LC-MS/MS using Thermo-Fisher Finnigan™ LTQ™. A multistep 4 h gradient using solvent A (water premixed with 0.1% formic acid) and solvent B (acetonitrile premixed with 0.1% formic acid) at a flow rate of 200 nL/min was used. The LTQ mass spectrometer was operated in a data-dependent mode. The full MS spectra were acquired in positive mode within a range of 400–1800 m/z. The Q activation and time was set respectively at 0.7 eV and 0.1 ms. The maximum ion injection times used were 50 ms for the MS scan and 120 ms for the MS/MS scans. The automatic gain control target settings were 3.0 × 10⁶ for the MS scan mode and 1.0 × 10⁴ for the MS/MS scan mode. The normalized collision energy was 29.5% and the isolation window employed was 2 m/z. The dynamic exclusion settings were repeat count 2, exclusion duration 25 s, exclusion list size 500, exclusion mass width low 0.5, and exclusion mass width high 1.5. |

(continued on next page)
The search parameters for protein identification were fixed cysteine modifications of +57 Da for carbamidomethyl-cysteines, +144 Da for lysine-iTRAQ labeling and N-terminal peptides; dynamic modifications allowing +16 with methionines for methionine sulfone and +144 Da for Y-iTRAQ labeling; restricted to trypsin digested peptides and allowed for two missed cleavages; precursor mass range was 350–5000 Da; peptide mass tolerance of 2.5 Da and fragment mass tolerance of ±0.8 Da; target false discovery rate (FDR) strict was 0.01 and FDR relaxed was 0.05; most confident centroid was selected for peak integration method and a 0.25 Da integration window tolerance was allowed.

**Description of data collection**

Cycling non-lactating beef cows were synchronized by an injection of GnRH (100 mcg as 2 mL of Factrel, intramuscularly) on day -7 and prostaglandin F2α (PGF2α) (25 mg as 5 mL of Lutalyse intramuscularly) on day 0. Blood samples and follicular fluid were collected from nine cows determined to initiate a new follicular wave at day -4. Estradiol (E2) concentrations in PL and FF were measured and animals were classified based on circulating E2 concentrations in PL at 36 h after PGF2α as either high E2 or low E2. Four PL and FF samples were classified namely PL HE2, PL LE2, FF HE2, and FF LE2. The four samples were depleted using the High Capacity Multiple Affinity Spin Cartridge (MARS Hu-6HC) # 5188-5341 (Agilent Technologies, CA, USA). Fifty μg of each of the four depleted bovine samples were pooled and vacuum dried. The dried peptides were reduced, alkylated, and digested. 4-plex iTRAQ reagents were added individually to the four digested PL and FF protein samples. The labeled peptides were mixed and the combined lysate mixture was divided into 13 SCX fractions. The SCX peptide fractions were analyzed by nano-LC-MS/MS using Thermo-Fisher FinniganTM LTQ™. Proteins were identified and quantified using Proteome Discoverer and SEQUEST™ search engine.

**Data source location**

South Dakota State University
Brookings, South Dakota
United States

**Data accessibility**

With the article

**Related research article**

P. A. Afedi, E. L. Larimore, R. A. Cushman, D. Raynie, G. A. Perry,
iTRAQ-Based Proteomic Analysis of Bovine Pre-ovulatory Plasma and Follicular Fluid, Domestic Animal Endocrinology.

https://doi.org/10.1016/j.domaniend.2021.106606 [1].

**Value of the Data**

- iTRAQ-labeling and nano-LC-MS/MS allowed for identification and relative quantification of proteins in bovine plasma and follicular fluid containing high and low pre-ovulatory circulating concentrations of E2 with high accuracy.
- The functional classifications of identified proteins give insights on their roles in follicle development. This information can be the basis for further hypothesis driven studies of functions of reproductive associated proteins in bovine.
- The iTRAQ ratio values from the paired comparisons are an important resource that can be used as a guide in the selection of proteins for further hypothesis driven studies on reproductive success in bovine.

1. **Data Description**

Proteomic approach was used to profile the proteome in PL and FF containing high or low E2. Abundant proteins, albumin, IgA, IgG, and alpha-1-antitrypsin in four PL and FF samples were depleted. Fifty μg of each sample was reduced, alkylated, and trypsin digested. The four samples were labelled individually with 4-plex iTRAQ reagents and fractionated by SCX. Protein identifications and relative quantifications were then performed via nano-LC-MS/MS. Fig. 1 is the overview of the the iTRAQ-based SCX nano-LC/MS/MS method used. 231 proteins were identified matched to 1142 peptides including 793 unique peptides. Each protein was identified by at least
Fig. 1. Overview of workflow for identification and relative quantification of proteins in bovine plasma and follicular fluid containing high and low pre-ovulatory levels of E2.

one unique peptide at high confidence; false discovery rate (FDR) of 5% or lower. Supplemental Table 1 shows details of the 231 identified proteins which include protein sequence coverages, protein scores, molecular weights, and isoelectric points. Of the 231 identified proteins, 212 originate from both plasma and FF, 12 from PL, and 7 from FF. VENNY 2.1 was used to graphically depict the overlap of proteins from PL and FF. Paired comparisons were performed using Proteome Discover and 103 proteins were found to be up- or down-regulated as defined by fold change of >2.0 or <0.5. Protein Analysis Through Evolutionary Relationships (PANTHER) was used to classify the molecular functions, biological process, and cellular component of the identified proteins. Protein Analysis Through Evolutionary Relationships (PANTHER) and Database for Annotation, Visualization and Integrated Discovery (DAVID) were used to assign specific molecular function, biological process, and cellular component to individual proteins. The outcome of this is included in Supplemental Table 1.

2. Experimental Design, Materials and Methods

2.1. Animals’ experimental design

All procedures were approved by the South Dakota State University Institutional Animal Care and Use Committee. Samples in the present study were collected from a previous study to characterize changes in steroidogenic enzymes and FF steroid concentrations [2]. Briefly, 32 cycling non-lactating beef cows (age = 4 to 10 yrs; mean BCS of 5.5) were synchronized by an injection of GnRH (100 mcg as 2 mL of Factrel, intramuscularly; Zoetis on day -7 and prostaglandin F2α (PGF2α) 25 mg as 5 mL of Lutalyse intramuscularly (Zoetis, Florham Park, NJ, USA) on day 0.
Estrus was monitored every 3 h from PG on day 0 until hour 33 and at slaughter (hour 36 to 42) with the aid of Estrofect (Western Point, Inc, Apple Valley, MN, USA) estrus detection aids. Ovaries of all cows were examined on day -7, -4, and 0 by transrectal ultrasonography using an Aloka 500 V ultrasound with a 7.5 MHz linear probe (Aloka, Wallingford, CT, USA) to assess follicular dynamics and ovulatory response. Ten cows that were determined to initiate a new follicular wave by day -4 (as determined by a cohort of growing follicles all less than 9mm) were slaughtered by captive bolt and exsanguination on day 2 (hour 36 to 42) for ovary collection. All cows had a CL present and a new follicular wave that had been initiated between days -7 and -4. This would mean all follicles were growing in the presence of progesterone and would be at a similar stage of growth. Ovaries were only collected from 9 out of 10 cows, however, as one was determined to have a follicular cyst at time of slaughter and was excluded. Thus, a total of nine cows were used for further study and data analysis. Estradiol (E2) concentrations in PL and FF were measured by radioimmunoassays (RIA) according to procedures as previously described [3]. Animals were classified based on circulating E2 concentrations in PL at 36 h after PGF2α as either high E2 (peak PL estradiol 9.07 ± 0.89 pg/mL; PL: n = 4, FF concentrations were 1,565 ± 196.0 ng/mL) or low E2 (peak PL estradiol 3.22 ± 0.79 pg/mL; PL: n = 5, FF concentrations were 398 ± 175.0 ng/mL). Animals were classified into high and low E2 classification according to Jinks et al. [4] (Peak estradiol > 8.4 pg mL at 48 h after PGF2α). Concentrations of E2 and changes in the steroidogenic pathway have previously been reported by Larimore and coworkers [2].

2.2. Follicular fluid (FF) and blood samples collection

All animals were harvested in the South Dakota State University Meat Laboratory. Within minutes of slaughter the reproductive tract was collected. Follicular fluid (FF) was aspirated from dominant follicles (DF; >10 mm in diameter) and the GCs were separated from the FF by centrifugation (1,000 × g for 1 min). The FF was place in RNase Free Tubes (USA Scientific), snap frozen in liquid nitrogen and stored at −80 °C until ready for analysis. Blood samples were collected at exsanguination to provide better comparison to the FF collected at slaughter. To obtain plasma, blood collected at slaughter was placed in EDTA vacutainer tubes (Beckman Dickerson) and centrifuged at 1,200 × g for 30 min at 4 °C. The plasma supernatant was snap frozen in liquid nitrogen and stored at −80 °C until ready for proteomic analysis.

2.3. Depletion and coomassie bradford assay quantitation

Depletion of the different bovine fluids (PL HE2, PL LE2, FF HE2, and FF LE2) were performed using the High Capacity Multiple Affinity Spin Cartridge (MARS Hu-6HC) # 5188-5341 (Agilent Technologies, CA, USA) as previously reported [5]. Protein content in depleted samples were quantified using Coomassie (Bradford) assay kit (Pierce Biotechnology, IL, USA) according to manufacturer’s instruction. Bovine serum albumin (Thermo Scientific, IL, USA) was used as a standard for making a calibration curve. Standard (2 mg/mL) was diluted with 50 mM TEAB and the calibration curve covered a concentration range of 0.00–1500 μg/mL.

2.4. Reduction, alkylation, digestion, and iTRAQ labeling of proteins

Fifty μg of each of four depleted bovine samples (PL LE2, PL HE2, FF LE2, and FF HE2) containing high or low E2 were pooled and vacuum dried (Labconco, Kansas, MO, USA) at 4 °C. The dried proteins were prepared with iTRAQ reagents according to manufacturer protocol (Applied Biosystems, Foster City, CA, USA). Briefly, the dried proteins were re-suspended in 25 μL 1M TEAB solution. The proteins were denatured with 1 μL 2% SDS solution, reduced with 2 μL
50 mM tris-(2-carboxy) ethylphosphine hydrochloride (TCEP), and alkylated with 1 μL of freshly prepared 84 mM iodoacetamide solution and incubated in dark room temperature for 30 min. Each sample (50 μg) was digested overnight at 37 °C with 10 μL of freshly prepared 1 μg/μL TPCK treated trypsin solution, P/N 4370282 (Ab Scienx Pte Ltd, MA, USA). iTRAQ reagents 114, 115, 116, and 117 were each resuspended in 70 μL ethanol and added individually to the four digested PL and FF protein samples. The samples were incubated at room temperature for 1 h and the reaction quenched by adding 100 μL of HPLC grade water and incubating at room temperature for 30 min. The iTRAQ-labeled peptides were vacuum dried at 4 °C, cleaned with Zeba™ desalting spin columns according to vendor instructions (Pierce, Rockford, IL, USA) and then vacuum dried at 4 °C. Samples were stored at −80 °C until ready for analysis.

2.5. Off-line strong cation exchange (SCX) separation

The iTRAQ-labeled peptides were separated using Cation Exchange Cartridge, # 4326695, Opti-Lynx Quick Connect Hardware, # 4326688, and Cation Exchange Buffer Pack, # 4326747 (Applied Biosystems, Foster City, CA, USA USA). The iTRAQ-labeled dried peptides were each diluted 10-fold with Load Buffer (10 mM KH₂PO₄ in 25% ACN; pH 3) and the four samples combined in one new vial. Equilibration of the SCX cartridge was done with 1 mL of Clean Buffer (10 mM KH₂PO₄ in 25% ACN / 1 M KCl; pH 3) followed by 2 mL Load Buffer. The mixed iTRAQ labeled peptides were then loaded onto the SCX cartridge. Excess iTRAQ reagents and salts in the cartridge were removed by washing the cartridge with 1 mL of Load Buffer. The bound peptides on the cartridge were eluted by sequential injection of 500 μL of a series of salt solutions: 0, 30, 40, 50, 60, 70, 85, 100, 130, 160, 350, 500, and 1000 mM KCl in Load Buffer. The flow rate of elution was ~1 drop/second. Each eluted fraction and wash solution was cleaned with Zeba™ desalting spin columns according to vendor instructions (Pierce, Rockford, IL, USA). Purified peptides were vacuum dried at at 4 °C (Labconco, Kansas, MO, USA) and stored at −80 °C until further analysis.

2.6. Nano-LC-MS/MS analysis

Aliquots of the dried SCX peptide fractions were analyzed using the Thermo-Fisher Finnigan™ LTQ™ mass spectrometer equipped with a nano electrospray source (New Objective, Woburn, MA, USA) and coupled with a nano-LC separation system (Eksigent nanoLC 1D-plus). The LC system is equipped with an autosampler (Spark Holland 920 Endurance Autosampler). Each peptide fraction was re-suspended in 10 μL of water/ACN/formic acid (95%/5%/0.1%) and then 3 μL loaded onto IntegraFrit Sample trap, ProteoPep II™ C18, 300 Å, 5 μm, 100 μm x 25 mm (New Objective, Inc., Woburn, MA). The retained peptides were washed isocratically with water premixed with 0.1% formic acid pumped from channel 1A to remove any excess reagents. Peptide separation was performed on an IntegraFrit Analytical Column (ProteoPep II™ C18, 75 μm x 100 mm, New Objective, Inc., Woburn, MA) with a multistep 4 h gradient using solvent A (water premixed with 0.1% formic acid) and solvent B (acetonitrile pre-mixed with 0.1% formic acid) at a flow rate of 200 nL/min. The gradient started at 5% solvent B and was held for 5 min, then linearly increased to 50% solvent B at 205 min and to 95% solvent B at 213 min and finally held at 95% solvent B for 5 min before allowing to return to initial 5% solvent B at 223 min. Column re-equilibration with initial 5% solvent B was done for 17 min.

The LTQ mass spectrometer was operated in a data-dependent mode. The full MS spectra were acquired in positive mode within a range of 400–1800 m/z. The top four most intense ions in the acquired full mass scan were selected for followed pulsed Q dissociation (PQD) in LTQ. The Q activation and time was set respectively at 0.7 eV and 0.1 ms. The maximum ion injection times used were 50 ms for the MS scan and 120 ms for the MS/MS scans. One microscan of full MS was performed. The automatic gain control target settings were 3.0 × 10⁴ for the MS
scan mode and $1.0 \times 10^4$ for the MS/MS scan mode. The normalized collision energy was 29.5\% and the isolation window employed was 2 m/z. The dynamic exclusion settings utilized were repeat count 2, exclusion duration 25 s, exclusion list size 500, exclusion mass width low 0.5 and exclusion mass width high 1.5.

2.7. Protein identification and quantification

The LC-MS/MS raw data from nanoLC-LTQ were correlated to theoretical fragmentation patterns of tryptic peptide sequences in bovine protein fasta database using SEQUEST\textsuperscript{TM} search engine embedded in Proteome Discoverer (version 2.1; Thermo Fisher Scientific). The search parameters included were as follows: fixed cysteine modifications of +57 Da for carbamidomethyl-cysteines, +144 Da for lysine-iTRAQ labeling and N-terminal peptides; dynamic modifications allowing +16 with methionines for methionine sulfoxide and + 144 Da for Y-iTRAQ labeling; restricted to trypsin digested peptides and allowed for two missed cleavages; precursor mass range was 350–5000 Da; peptide mass tolerance of 2.5 Da and fragment mass tolerance of ±0.8 Da; target false discovery rate (FDR) strict was 0.01 and FDR relaxed was 0.05; most confident centroid was selected for peak integration method and a 0.25 Da integration window tolerance was allowed. The proteins matched with at least one unique peptide at minimum 95\% confidence were considered positive identifications. The FDR is a statistical measure of the certainty of the identifications and is calculated using F-values. The F-value calculation factors the primary score from the spectra library search (Dot Score) and the Dot Bias Score which is an indication of the number of spectra used in determining the Dot Score.

The relative quantification of proteins in the samples was performed with the Proteome Discoverer (1.2). The quantification utilized the relative peak intensities of the iTRAQ reporter ions derived from MS/MS spectra of all unique peptides that represented each protein. iTRAQ ratios of the reporter ions were calculated using reporter ions representing any two samples. The final ratios obtained from the relative protein quantifications were normalized according to the median protein quantification ratio. The protein ratios were the median ratio of the corresponding peptide ratios. To determine protein expression levels, a fold change of $>2.0$ or $<0.5$ between any pair from the four sample types, i.e. PL LE2, PL HE2, FF LE2, and FF HE2 were respectively set for up- and down-regulated proteins. The paired comparisons made were PL HE2 and PL LE2, FF HE2 and FF LE2, PL HE2 and FF HE2, and PL LE2 and FF LE2. Functional analysis was performed for all identified proteins and for up- and down-regulated proteins.

2.8. Bioinformatic analysis of identified proteins

The PANTHER (Protein Analysis Through Evolutionary Relationships) classification system [6] and Database for Annotation, Visualization and Integrated Discovery (DAVID) [7] were used to categorize the identified proteins based on their molecular function (MF), biological process (BP), and cellular component (CC)/localization.

Ethics Statement

All procedures were approved by the South Dakota State University Institutional Animal Care and Use Committee (approval # 12-078A) and conformed to the Guide for the Care and Use of Agriculture Animals in Research and Teaching.

CRediT Author Statement

P. A. Afedi: Formal Analysis, Investigation, Writing – Original Draft, Writing – Review & Editing; E. L. Larimore: Conceptualization, Investigation, Writing – Review & Editing; R. A.
Cushman: Conceptualization, Investigation, Writing – Review & Editing, Supervision; D. Raynie: Formal Analysis, Writing – Reviewing & Editing, Supervision; G. A. Perry: Conceptualization, Methodology, Investigation, Resources, Writing – Review & Editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.

Acknowledgments

We thank the support of USDA/National Institute of Food and Agriculture, Hatch project 1004566. We also thank the Core Campus Mass Spectrometry Facility at South Dakota State University (SDSU) supported in part by the NSF/EPSCoR Grant No. 0091948 and NSF MRI Grant No. SS1000060 and the State of South Dakota.

Supplementary Materials

Supplementary material associated with this article can be found in the online version at doi: 10.1016/j.dib.2021.106998.

References

[1] P.A. Afedi, E.L. Larimore, R.A. Cushman, D. Raynie, G.A. Perry. iTRAQ-Based Proteomic Analysis of Bovine Pre-ovulatory Plasma and Follicular Fluid, Domest. Anim. Endocrinol. doi:10.1016/j.domaniend.2021.106606.
[2] E.L. Larimore, O.L. Amundson, G.A. Bridges, A.K. McNeel, R.A. Cushman, G.A. Perry, Changes in ovarian function associated with circulating concentrations of estradiol before a gnrh-induced ovulation in beef cows, Domest. Anim. Endocrinol. 57 (2016) 71–79.
[3] G.A. Perry, B.L. Perry, Effect of preovulatory concentrations of estradiol and initiation of standing estrus on uterine ph in beef cows, Domest. Anim. Endocrinol. 34 (2008) 333–338.
[4] E.M. Jinks, M.F. Smith, J.A. Atkins, K.G. Pohler, G.A. Perry, M.D. MacNeil, A.J. Roberts, R.C. Waterman, L.J. Alexander, T.W. Geary, Preovulatory estradiol and the establishment and maintenance of pregnancy in suckled beef cows1, J. Anim. Sci. 91 (2013) 1176–1185.
[5] P.A. Afedi, L. Jing, G.A. Perry, D.E. Raynie, Immunodepletion of multiple high-abundant proteins from bovine fluids, J Bioanal. Methods Tech. 1 (2020) 101–111.
[6] H. Mi, A. Muruganujan, J.T. Casagrande, P.D. Thomas, Large-scale gene function analysis with the panther classification system, Nat. Protoc. 8 (2013) 1551.
[7] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using david bioinformatics resources, Nat. Protocols 4 (2008) 44–57.