

Aus dem
Veterinärwissenschaftlichen Department
der Tierärztlichen Fakultät der Ludwig-Maximilians-Universität München

Angefertigt unter der Leitung von Univ.-Prof. Dr. Eckhard Wolf

Angefertigt an der externen Einrichtung LAFUGA
(Dr. Georg J. Arnold)

Qualitative and Quantitative Proteome Analyses of Bovine Oocytes and Early Embryos

Inaugural-Dissertation
zur Erlangung der tiermedizinischen Doktorwürde
der Tierärztlichen Fakultät der Ludwig-Maximilians-Universität
München

von
Myriam Demant
aus Bonn

München 2012

Gedruckt mit Genehmigung der Tierärztlichen Fakultät
Der Ludwig-Maximilians-Universität München

Dekan: Univ.-Prof. Dr. Braun

Berichterstatter: Univ.-Prof. Dr. Wolf

Korreferent/en: Priv.-Doz. Dr. André
Univ.-Prof. Dr. Braun
Prof. Dr. Kaltner
Univ.-Prof. Dr. Dr. Dr. habil Sinowatz

Tag der Promotion: 11. Februar 2012

*Für meine Eltern
Dipl. Math. Bernd und
Dipl. Math'n Brigitte Demant*

TABLE OF CONTENTS

1 INTRODUCTION	1
1.1 THE IMPACT AND DRAWBACKS OF ASSISTED REPRODUCTION TECHNIQUES	1
1.2 THE COW AS A MODEL FOR FEMALE HUMAN REPRODUCTION	2
1.3 OOCYTE AND EARLY EMBRYONIC DEVELOPMENT	2
1.3.1 Oogenesis and folliculogenesis.....	2
1.3.2 Preimplantation embryonic development	7
1.4 THE IMPACT OF PROTEOME ANALYSES ON REPRODUCTIVE RESEARCH	10
1.5 THE TWO STRATEGIES FOR PROTEOME ANALYSES	11
1.5.1 2D gel based techniques	11
1.5.2 Nano LC-MS/MS based techniques.....	12
1.6 HOLISTIC PROTEOME ANALYSES OF MAMMALIAN OOCYTES AND PREIMPLANTATION EMBRYOS	15
1.6.1 Proteome analyses of oocytes	15
1.6.2 Proteome analyses of preimplantation embryos	19
1.7 AIM OF THE THESIS	21
2 MATERIALS AND METHODS	23
2.1 SAMPLE GENERATION	23
2.1.1 Generation of GV oocytes	23
2.1.2 Generation of early embryos.....	23
2.1.3 Generation and processing of <i>in vivo</i> aspirated oocytes by ovum pick-up (OPU)	24
2.2 NANO-LC-MS/MS ANALYSES	25
2.2.1 Qualitative proteome profile of 900 GV oocytes.....	25
2.2.1.1 Lysis.....	25
2.2.1.2 1D polyacrylamide SDS gel electrophoresis.....	25
2.2.1.3 Gel slicing and tryptic in-gel digestion	25
2.2.1.4 1D-LC MS/MS analysis.....	26
2.2.1.5 Database search and data analysis.....	26
2.2.1.6 Gene ontology analysis	27
2.2.2 Qualitative and quantitative analyses from small sample amounts of oocytes and embryos.....	27
2.2.2.1 Sample preparation and iTRAQ labelling	27
2.2.2.2 1D and 2D liquid chromatography (LC).....	28
2.2.2.3 Mass spectrometry	28
2.2.2.4 Database search.....	28
2.2.2.5 Generation of exclusion lists for quantitative iTRAQ analysis.....	28
2.2.2.6 Data analysis for identification and quantification.....	29
2.2.2.7 Targeted quantification by Selected Reaction Monitoring (SRM).....	29
2.3 2D GEL BASED SATURATION DIGE ANALYSIS	30
2.3.1 Sample lysis	30
2.3.2 Labelling of proteins with fluorescent Cy5 Dyes for analytical gels	31
2.3.3 Labelling of proteins with fluorescent Cy5 Dye Cy3 for the preparative gel	31

2.3.4 2D polyacrylamide SDS gel electrophoresis	31
2.3.5 Scanning and evaluation of 2D saturation DIGE gels	32
2.3.6 Identification of spots from differentially abundant proteins.....	32
2.3.7 Comparative David gene ontology analysis of the combined data from differentially abundant proteins of morulae and blastocysts	33
2.4 BUFFERS, MEDIA AND SOLUTIONS	34
2.4.1 Buffers and media for sample generation	34
2.4.2 Buffers and solutions for analyses	34
2.5 CHEMICALS, KITS AND PREFABRICATED BUFFERS	36
2.6 SOFTWARE	38
2.6.1 Analyses software	38
2.6.2 Instrument specific software	39
2.7 INSTRUMENTS	39
2.8 CONSUMABLES.....	41
3 RESULTS.....	43
3.1 PREPARATION OF GV OOCYTES.....	43
3.2 QUALITATIVE PROTEOME PROFILE OF BOVINE GV OOCYTES	43
3.3 PROTEIN IDENTIFICATION FROM LIMITED SAMPLE AMOUNTS.....	47
3.4 PROTEIN QUANTIFICATION OF LIMITED SAMPLE AMOUNTS BY ITRAQ.....	49
3.5 COMPARISON OF <i>IN VIVO</i> AND <i>IN VITRO</i> MATURED OOCYTES FROM COWS OF DIFFERENT AGE GROUPS BY NANO LC-MS/MS ITRAQ ANALYSIS.	52
3.6 QUALITATIVE AND QUANTITATIVE PROTEOME ANALYSIS OF OOCYTES, 2-CELL STAGE EMBRYOS AND MORULAE	56
3.6.1 Validation of abundance alterations between oocytes, 2-cell stage embryos and morulae.....	59
3.7 DIFFERENTIAL PROTEOME ANALYSES OF MORULAE AND BLASTOCYSTS	61
3.7.1 2D saturation DIGE analysis	61
3.7.2 Nano LC-MS/MS iTRAQ analysis.....	67
3.7.3 Comparison of results from 2D DIGE analysis and from nano LC-MS/MS analysis	71
3.7.4 David gene ontology analysis of differentially abundant proteins.....	73
4 DISCUSSION	74
4.1 QUALITATIVE PROTEOME ANALYSIS OF 900 BOVINE GV OOCYTES	74
4.2 PROTEIN IDENTIFICATION AND QUANTIFICATION OF LIMITED SAMPLE AMOUNTS BY NANO LC-MS/MS	76
4.3 PROTEIN QUANTIFICATION OF LIMITED SAMPLE AMOUNTS BY ITRAQ.....	78
4.4 COMPARISON OF <i>IN VIVO</i> AND <i>IN VITRO</i> MATURED OOCYTES FROM COWS OF DIFFERENT AGE GROUPS BY NANO LC-MS/MS ITRAQ ANALYSIS	80
4.4.1 Glyoxylase 1, Ubiquitin and Bisphosphoglycerate mutase are candidate proteins for affecting the developmental potential of <i>in vitro</i> matured oocytes.....	81
4.5 QUALITATIVE AND QUANTITATIVE PROTEOME ANALYSIS OF OOCYTES, 2-CELL STAGE EMBRYOS AND MORULAE	84

4.5.1 The decrease of YBX2 in morulae could rely on its maternal RNA storage function	86
4.5.2 The decrease of IMP3 in morulae could be involved in first cell determinations.....	86
4.6 DIFFERENTIAL PROTEOME ANALYSES OF MORULAE AND BLASTOCYSTS	87
4.6.1 Saturation DIGE analysis of morulae and blastocysts revealed several protein isoforms of different abundance alterations.....	88
4.6.2 The iTRAQ analysis led to 141 quantified and identified proteins.....	89
4.6.3 The complementary characteristics of the saturation DIGE and the iTRAQ analyses are reflected by the comparison of morulae and blastocysts	89
4.6.4 Gene ontology clustering of proteins indicates a switch from a catabolic to an anabolic stage as well as cell proliferation during the morula to blastocyst transition.....	91
4.6.5 Proteome analysis of morulae and blastocysts reflects increasing translation during the morula to blastocyst transition	92
4.6.6 Creatin kinase B (CKB) and Annexin A6 are promising candidates in the molecular studies of the morula to blastocyst transition.....	93
4.6.7 The increase of the LGALS3 protein, RACK1 and Ion transport associated proteins could reflect the embryo remodelling during blastocyst development	94
4.6.8 The protein NPM1 could be involved in first lineage decisions	95
4.6.9 The reduction of redox enzymes may alter HIF dependent gene regulation.....	95
5 PERSPECTIVES.....	98
6 SUMMARY	99
7 ZUSAMMENFASSUNG	101
8 INDEX OF FIGURES.....	104
9 INDEX OF TABLES.....	106
10 INDEX OF ABBREVIATIONS.....	107
11 REFERENCE LIST	109
14 ACKNOWLEDGEMENTS.....	118

1 Introduction

1.1 The impact and drawbacks of assisted reproduction techniques

Assisted reproduction techniques (ARTs) become more and more important both in veterinary and in human medicine. While in veterinary medicine the focus of ARTs clearly lays in the improvement of characteristic features of productive livestock, in human medicine ARTs mainly address infertility treatment. ARTs generally comprise the *in vivo* or *in vitro* maturation of oocytes, followed by their *in vitro* fertilization and cultivation of the developed embryos until they are transferred into the uterus. This procedure is generally referred to as “**in vitro production of embryos**” (IVP) [1, 2].

Although the first human test-tube baby was already born in 1978 [2], ARTs in human medicine are still associated with several major drawbacks: Due to questions of efficiency it is still common to mature oocytes *in vivo*, which needs the women’s exposure to supraphysiological levels of gonadotrophins. This can lead to the so called “**ovarian hyperstimulation syndrome**” (OHSS) [3]. Furthermore multiple gestations induced by the common implantation of more than one embryo often represent a risk for the mother’s health and lead to pre-term delivery associated with low birthweight foetuses, suffering from neonatal and perinatal problems. Especially specific syndromes such as omphalocele, Beckwith–Wiedeman, Prader–Willi, Angelman and retinoblastoma have been observed in ART-derived babies with increased frequency and seem to be associated with epigenetic alterations induced by the IVF procedure [2].

In bovine the first calf from IVP was born in 1981 [4]. In 2007 already more than 200,000 embryos have been *in vitro* produced and transferred to recipients [5]. The primary target of IVP in bovine livestock is the improvement of selection intensity by increasing the number of offspring from a female cow of a desirable genotype. Additionally, ARTs are employed to increase pregnancy rates in herds with low fertility. However, also in cattle IVP is associated with developmental abnormalities in embryos, foetuses and calves such as increased rates of early embryonic death and abortion, production of large size foetuses and calves as well as abnormalities of placental development, summarized as the **large offspring syndrome** (LOS) [1]. These drawbacks are in large parts not understood and possibilities to avoid the occurrence of these problems are limited. Most common methods successfully applied to assess the embryo’s developmental competence are based on morphological criteria and on the determination of cleavage and blastocyst rates [6]. However, to improve the understanding of biochemical processes underlying oocyte maturation and early embryonic development, comprehensive analyses of the events taking place on the molecular level are indispensable.

1.2 The cow as a model for female human reproduction

The bovine model system reflects crucial parameters of the female human reproductive biology. Both species are monovulatory and non-seasonal polycyclic [7]. In both species the duration of pregnancy is about nine months. Further similarities can be found in size and anatomical structure of the ovaries. The follicle containing cortical region lies outwards and ovulation can occur at any point over the ovarian surface [7]. Moreover, the mechanisms of follicular wave emergence, selection of a dominant follicle and ovulation are fundamentally similar in cattle and women [8]. Even pathologic conditions, like follicular cysts, luteinized anovulatory follicles and lactation- or stress-related suppression of follicle growth and ovulation, occur in human and cows [7]. Using the cow as a model for female human reproduction has especially led to the discovery of follicular wave development in women [9, 10]. Also, examples like the use of the bovine in studies of reproductive aging in women, demonstrate that it is increasingly accepted as a model for human female reproduction [11, 12].

1.3 Oocyte and early embryonic development

Mammalian oocyte and early embryonic development comprises several crucial steps. A short description of these steps placing an emphasis on bovine development will be given in the following chapters.

1.3.1 Oogenesis and folliculogenesis

The development of female gametes is usually referred to as oogenesis: **primordial germ cells** (PCGs), specified to become the founder population of the germline, are initially identifiable in the posterior primitive streak before they begin to become a polarized morphology and extend cytoplasmic protrusions as they initiate their migration through the primitive streak into the adjacent posterior embryonic endoderm (yolk sac wall) where they can already be found at the 28th day of pregnancy in the cow [13]. Subsequently they migrate in the hindgut during its anterior extension and reach the mesodermal tissues, followed by their bilateral migration into the genital ridges, which are the precursors of the gonads [14]. A scheme of this oocyte migration in the mouse embryo is shown in Fig. 1. PCGs can be distinguished from somatic cells by their large round shape and the high expression levels of tissue-nonspecific alkaline phosphatase as well as by high levels of *Pouf5f1 mRNA* [14, 15]. It is known from the mouse embryo that the forming of PCGs is partially induced by “Bone morphogenetic protein” (BMP) signalling. This induction leads to the transcriptional regulation of epiblast cells, mediated by

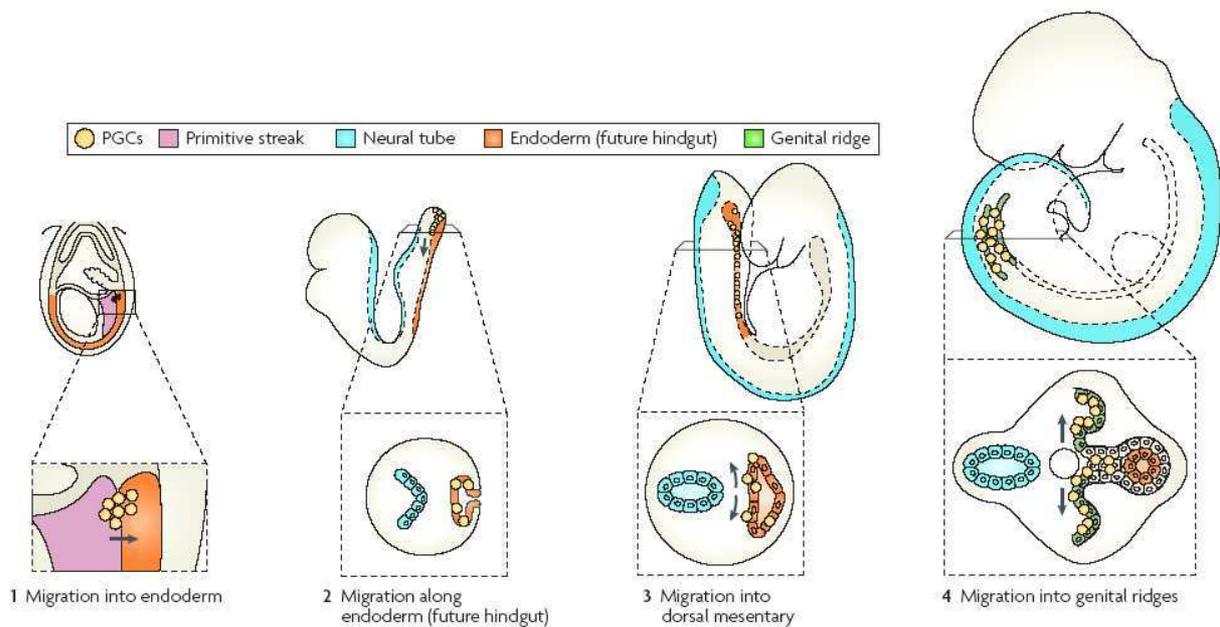


Fig. 1: Stages of primordial germ cell migration

PGCs in mice, specified in the proximal epiblast, migrate from the primitive streak to the endoderm (future hindgut) at embryonic day 7.5 (E7.5; step 1). A close-up is shown. At E8, PGCs migrate along the endoderm (step 2). At E9.5, PGCs migrate bilaterally towards the dorsal body wall (step 3). At E10.5, PGCs reach the genital ridges to form the embryonic gonad (step 4). Source of picture and legend: Richardson, B.E. and R. Lehmann, Mechanisms guiding primordial germ cell migration: strategies from different organisms. *Nat Rev Mol Cell Biol.* 11(1): p. 37-49.

the transcriptional repressor “B lymphocyte induced maturation protein1” (BLIMP1) which promotes the expression of PGC specific genes such as *stella* (also known as *Dppa3*) and represses the expression of somatic cell genes, in particular members of the *Hox* gene family [14]. The knowledge about how PCG migration is initiated in mammals is rather limited up to today, yet, a regulation by “Interferon induced transmembrane protein 1” (IFITM1) is controversially discussed and also the Receptor tyrosine kinase (KIT) and its ligand Steel also called “KIT ligand” (KITLG) are required for general PGC motility [14]. Removal of these factors leads to PGCs that migrate in the proper direction but at a greatly reduced rate. During the migration process, in particular the gut seems to be important for proper migration since removal of the Y box 17 (SOX17) transcription factor prevents expansion of the hindgut endoderm and PGCs fail to migrate to the genital ridges [14]. Moreover, the proteins “Stromal cell-derived factor 1” (SDF1), expressed in the genital ridges, and the C-X-C chemokine receptor type 4 (CXCR4), expressed in the PCGs, are required for later stages of migration to the genital ridge. Finally, also disruption of E-cadherin function results in PCGs which stop their migration just before reaching the outside gonad [14].

During the migration phase a period of mitosis and proliferation occurs [13] and simultaneously DNA methylation is significantly reduced and imprinted genes become biallelically expressed [16, 17]. By this decrease of epigenetic marks, germ cells are distinguished from somatic cells.

The oocyte development is continued by multiple changes in chromatin structure. DNA remethylation occurs and genomic imprints are acquired in a sex specific manner, meaning that DNA methylation is targeted specifically to paternally and maternally DNA-methylated “imprinting control regions” (ICRs). ICRs are cis active DNA sequences which regulate the imprinting of specific gene clusters. The majority of these genes is known to be involved in foetal and placental growth as well as somatic differentiation [18, 19]. Interestingly, the imprints are re-established during distinct time windows in the male and female germlines. While in male germs cells the remethylation takes place during prenatal sperm development, in female germ cells this process happens during postnatal oocyte maturation [18].

At the 57th day of the bovine pregnancy the germ cells form the so-called “oogonia”, initially located in the cortical region of the ovary. They are surrounded by somatic cells, which are the progenitors of granulosa cells. Subsequently, the mitotically active oogonia, containing KIT, interact with adjacent mesonephric cells, containing the “KIT ligand”, to form the ovigerous cords isolating the oogonia from the interstitium. In the following, tight junctions between the oogonia and the mesonephric cells that subsequently become pre-granulosa cells are formed. Then the female germ cells pass through the leptotene, zygotene and pachytene stage until they become arrested in the diplotene stage of the first meiotic division as so called “primary oocytes”. During the first meiotic division at least 80 % of the germ cells, but not their pre-granulosa cells, undergo apoptosis. In the following, also the somatic cells proliferate, surround the oocytes and form primordial follicles which are enclosed in basal membranes. Connective tissue grows between them and separates them from each other. First primordial follicles can be found at the 90th day of pregnancy in the cow [13, 20, 21].

The primordial follicles are located in the ovarian cortex where they are stored during the complete fertile lifespan. Although many of the processes, controlling activation of the primary follicles, remain to be elucidated it is assumed that inhibitory factors from the ovary as well as the anti-Müllerian hormone play a role in inhibiting too early follicle activation [21]. With the occurrence of oestrous cycles several resting primordial follicles are activated independently of gonadotrophins. The mechanisms regulating follicle activation are still only poorly understood [13, 20].

Activated follicles start to grow and become to the so-called primary follicles [13]. The oocytes increase three to ten folds in volume of smooth endoplasmic reticulum, mitochondria, ribosomes and lipid droplets, and the zona pellucida, absent in primordial follicles, is formed [21]. During this growing period the oocytes already start to acquire their developmental competence [22]. The developmental competence includes characteristics like (i) the ability to

resume meiosis, (ii) to cleave upon fertilization, (iii) to develop into a blastocyst, (iv) to induce pregnancy and (v) to generate a healthy offspring [23].

By intense mitotic proliferation of granulosa cells which change their shape to cubical, the follicles become multilayered, marking the secondary also called preantral follicle. The primary and small secondary follicles are sometimes also named “committed follicles”. Within the secondary follicle, the oocyte reaches its final size. In the end of the secondary follicle stage, surrounding cells start to develop into the “theca folliculi”. By the secretion of liquor folliculi into intercellular spaces, a lumen (antrum) is formed and the secondary becomes to the tertiary follicle (antral follicle), while it develops from the gonadotrophins-independent stage over a stage of gonadotropin-responsiveness to a stage of gonadotropin-dependence [21]. The transition from the stage of gonadotropin-responsiveness to the stage of gonadotropin-dependence is known to be at least partly induced by the “Insulin-like growth factor 1” (IGF1). Interestingly the supply of IGF1 is primarily controlled by the “**IGF binding proteins**” (IGFBPs) rather than by its total concentration [21]. After the switch to the stage of gonadotropin-dependence the further follicular development is controlled by the neuroendocrine system [13, 24]. During the development from the secondary to the tertiary follicle also the granulosa cells differentiate into two anatomically and functionally distinct lineages; the mural granulosa cells that line the wall of the follicle which have a steroidogenic role and the cumulus cells, forming an intimate life-support association with the oocyte. Simultaneously the oocyte is shifted to the follicle edge, forming a small protrusion called the cumulus oophorus, and the theca folliculi develops into an inner layer, the “theca interna”, and an outer layer the “theca externa” [13]. The granulosa cell differentiation is known to be driven by the oocyte secreting so called “**Oocyte secreted factors**” (OSFs). These factors, closely related members of the TGF- β superfamily, are the “Growth differentiation factor 9” (GDF9) and the “Bone morphogenetic protein 15” (BMP15) [21, 25]. They regulate several important functions of granulosa and cumulus cells; including regulation of cellular growth, enhancement of cell survival, modulation of steroidogenesis, regulation of the expansion of cumulus cells and the metabolism of cumulus cells [25, 26]. In particular the OSFs direct the lineage of its neighbouring granulosa cells towards the cumulus cell phenotype and prevent their

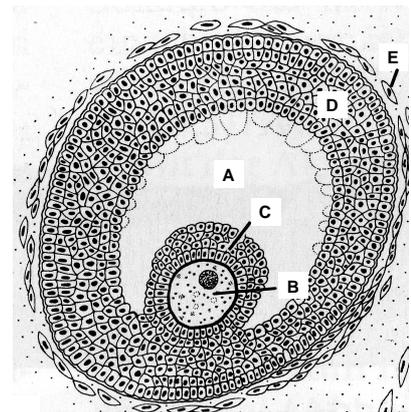


Fig. 2: Small antral/tertiary follicle
 (A) antrum, (B) the oocyte is located on the follicle edge within the cumulus oophorus surrounded by (C) cumulus granulosa cells, (D) mural granulosa cells, (E) theca cells; Source of picture: Rüsse I, S.F., Lehrbuch der Embryologie der Haustiere. 1998

luteinisation (transition to granulosa lutein cells after ovulation). The cumulus cells in turn are believed to regulate the special oocyte microenvironment to keep it distinct from the rest of the follicle [21]. A typical small tertiary follicle is shown in Fig. 2.

From the occurrence of the tertiary follicle onwards, a **dominant follicle (DF)** is selected which can in the following either undergo atresia or become ovulated. This process happens in so-called “waves of folliculogenesis”, which emerge in cows at regular intervals. In the bovine cycle mostly two or three of these waves are observed [24, 27, 28]. Only the last follicular wave leads to an ovulatory follicle. A follicle cohort typically consists of 5–20 tertiary follicles being larger than 5 mm in diameter. The wave emergence is correlated with a transient increase in FSH concentrations, inducing cellular growth and proliferation of granulosa cells [24, 29, 30]. There are hints that this FSH action is mediated by an increase of the “Pregnancy associated plasma protein-A” (PAPP-A) which induces increased rates of proteolytic degradation of the IGFBP protein thus leading to higher IGF1 activity [21]. Moreover, like in the earlier stage of follicular development OSFs also play a critical role by inhibiting luteinisation and promotion of mitotic activity of granulosa cells. As the DF follicle emerges from the cohort, it increases in size, which leads to an increase in oestradiol and inhibin concentrations due to the fact that these hormones are produced in granulosa cells. The intensive production of oestradiol is a defining characteristic of the DF which occurs already prior to visible differences in follicle size [24, 31]. In the following, the increase of oestradiol concentrations in concert with inhibin suppresses the FSH release from the pituitary gland by a direct inhibitory action on the expression of mRNA for FSH β , while the selected DF becomes increasingly responsive to LH through the augmented presence of **LH receptors (LH-R)** [24, 30, 32, 33]. This puts the DF on a beneficial growing position due to the facts that (i) the FSH concentration is pushed below the threshold needed to sustain other tertiary follicles and that (ii) the synthesis of oestradiol is dependent on the LH stimulated production of androgens in the theca cells and their subsequent aromatisation to oestradiol in granulosa cells. Oestradiol in turn induces further proliferation of granulosa cells what again leads to an enhancement of the follicle’s dominance. Since oestradiol also provokes an enhanced LH release from the pituitary gland this process leads to an LH peak which induces (i) the resumption of meiosis in the nucleus involving the dissolution of the nuclear (**germinal vesicle (GV)**) membrane and condensation of chromatin as well as separation of the homologous chromosomes and emission of the first polar body, (ii) the final growth spurt and (iii) ovulation of the oocyte when the metaphase of the second meiotic division is reached. The second meiotic division is completed after fertilization [13]. A scheme of the model for folliculogenesis, originally generated and based on the ewe by Scaramuzzi et al. [34], similar to that from the cow is shown in Fig. 3.

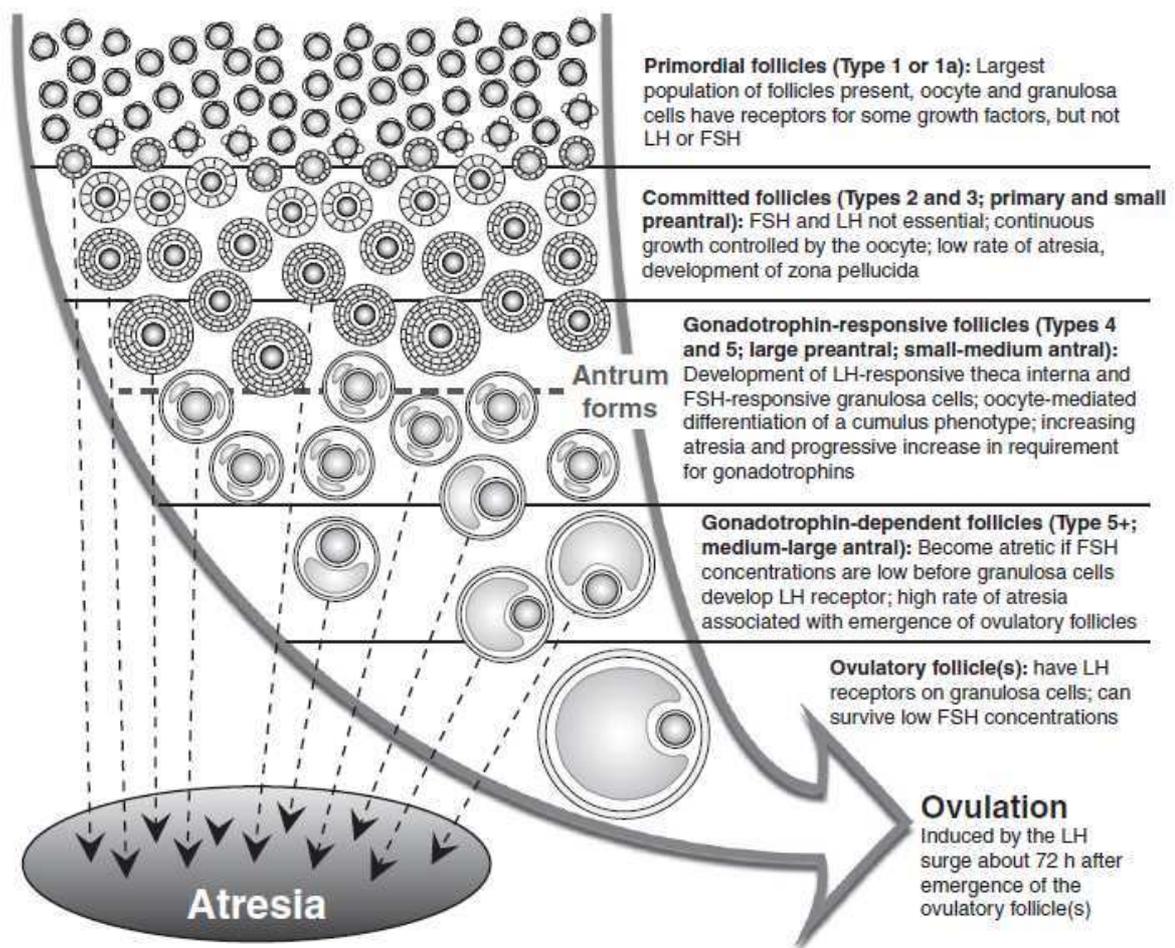


Fig. 3: A model for folliculogenesis initially generated and based on the ewe by Scaramuzzi et al.

Folliculogenesis is illustrated as a cascade of development during which follicles emerge from a pool of primordial follicles to enter a process of growth and development that is continuous and ends in either atresia or ovulation. In most mammals this process is approximately linear to the gonadotrophin-responsive stage and, especially in ruminants, becomes wave like in the gonadotrophin-dependent stage. Source of picture and legend: Scaramuzzi, R.J., et al., Regulation of folliculogenesis and the determination of ovulation rate in ruminants. *Reprod Fertil Dev* . 23(3): p. 444-67.

1.3.2 Preimplantation embryonic development

During mammalian early embryonic development the zygote, respectively the early embryo, cleaves completely, leading to evenly divided blastomeres. The first three cell divisions occur within the oviduct. In the 8-cell stage, the bovine embryo is transported into the uterus where the further embryonic development takes place [13]. The phase between the 8- and the 16-cell stage is crucial for bovine embryonic development, since the **maternal-embryo transition (MET)** takes place [35, 36]. During the MET, the embryo's own genome is activated leading to increased transcription activity. A failure in embryonic genome activation leads to irreversible arrest of the embryo, because it can no longer support basic metabolic processes. The further embryonic development relies on its own genome [36, 37]. During the time of MET often a phenomenon known as "developmental block" occurs. It becomes very obvious during *in vitro*

embryo production. Many *in vitro* produced embryos fail to develop beyond the 8-cell stage, what is considered to be the reason for blastocyst rates of only 30 % to 40 % of *in vitro* fertilized oocytes. The embryos ability to perform the MET is likely correlated with oocyte quality. Although the mechanisms behind the developmental block are not fully understood, it is believed that they are based on inability (i) to overcome the chromatin repression and/or (ii) to react to injuries caused by the environment [23, 36].

During the next cell divisions, the embryo forms into a morula, which usually occurs on day five **post insemination** (p.i.) in the cow [13]. Within the morula, blastomeres gain a tighter contact to each other by an increase of intercellular adhesions and adopt a more flattened morphology, a process known as compaction [38]. It is unclear how compaction is initiated [39]. However, it is already known from the mouse model that E-cadherin, a major component of adherens junctions, switches to regions of cell-cell contact at the 8-cell stage [40]. Furthermore, compaction can be inhibited by removal of Ca^{2+} ions or addition of E-cadherin specific antibodies. Interestingly, compaction is also processed when mRNA or protein synthesis is blocked from the mouse 4-cell stage onwards, indicating that necessary factors have been produced earlier in development. This suggests that posttranslational mechanisms play an important role in the induction of compaction, possibly by maintaining the E-cadherin complex in an inactive state until it is needed for compaction [39].

At the morula stage, the first cell fate decisions are made between inside and outside located cells. During the following blastocyst development, the outer cells develop into **trophectoderm** (TE) cells, which are progenitor cells of the placenta while the inner cells form the **inner cell mass** (ICM) from which the embryo develops. Furthermore, a fluid-filled cavity, the blastocoel, is formed by merging of intercellular spaces and water movement into the embryo. The blastocyst usually occurs around day seven p.i. in the cow [13, 38]. The knowledge about how these cell lineages develop during the preimplantation period is considered to have a major impact for increasing the success of ARTs. The understanding of these processes is rather limited until today although early information on cell fate specification has already been provided by cell-lineage studies performed in ascidians at the turn of the last century by Conklin. Until today, much research concerning mammalian lineage decisions has been performed on the mouse model [38, 39]. Some of these insights will be described in the following:

It is generally assumed that the forming of cell polarities is of fundamental impact for cell fate decisions [38]. With the increase in cell adhesions blastomeres rapidly polarize along the axis perpendicular to cell contact so that outward facing (apical) regions become distinct from

inward facing (basolateral) regions [39]. The polarization involves the redistribution of membrane, cytoskeletal and organelle components [41]. Cell nuclei move to a basolateral position [42] while the endosomes [41] as well as cytoplasmic actin [43] and the membrane protein ezrin [44] become localized apically. Microvilli accumulate at the apical pole and are almost completely eliminated basolaterally [45, 46].

The establishment of cell polarity in organisms like *C. elegans* and *Drosophila* has been widely studied. For example the so-called “par” (**partitioning defective**) genes are known to be involved in the establishment of polarity in the zygote of *C. elegans* [47, 48]. The proteins PAR3, PAR6 and aPKC are known to form a complex involved in the establishment of polarity in the oocyte, in epithelial differentiation and in asymmetric divisions of delaminating neuroblasts in *Drosophila* [49]. Therefore, these proteins were also characterised in mouse early embryonic development and it was revealed that the polarity proteins PAR3 [50], PAR 6 [51], and “atypical protein kinase” [50] become localized to the apical domain of blastomeres, while the polarity protein PAR1 [51] accumulates basolaterally.

However, it is still unclear how *de novo* polarisation is initiated. On the one hand, various studies suggest that cell-cell contact is somehow important for the establishment of the apical and basolateral domains [52, 53]. On the other hand, polarisation is also observed in blastomeres isolated from cell contacts or prevented from compacting, coming along with a microtubule-mediated interaction between the nucleus and the cell cortex. Therefore, it is assumed that at least these two possibilities for polarization initiation exist [54].

The polarity of blastomeres seems to be important for early cell fate decisions since they are divided by mitosis either perpendicular or parallel to its axis of polarity, resulting in two separate groups of apolar inside and polar outside cells [55]. In the following the outer cells are becoming fully committed to the TE lineage and the blastocoel begins to form.

It is assumed that these first lineage decisions are made due to some cell fate determinants, being segregated specifically into, or out of, polarized cells during mitosis. For example the pluripotency markers Octamer 3/4 (OCT4, now POU5F1) [56] and NANOG [57] are restricted in their expression to the inside, future ICM cells. In contrast, the mRNA of the TE specific transcription factor “Caudal type homeobox 2” (CDX2) is localized to the apical domain of blastomeres at the 8-cell stage and becomes restricted to the outside, progenitors of the TE, cells in later stages [58, 59]. CDX2 is thought to repress POU5F1 and NANOG [60]. Embryos missing CDX2 cannot maintain epithelial integrity after blastocysts formation and fail to implant [60]. Knockout experiments suggested, that the transcription factor “TEA domain family member 4” (TEAD4) acts upstream of CDX2 in TE specification. TEAD4^{-/-} embryos do

not express CDX2, but do express POU5F1 and NANOG [61, 62]. Since they do not produce trophoblast stem cells, trophectoderm or blastocoel cavities, they are not able to implant into the uterine endometrium. However, they can produce embryonic stem cells. Disrupting the *tead4* allele after implantation leads to normal completion of embryo development [62]. Interestingly, in ES cells TEAD4 is not required for expression of CDX2 and other trophoblast genes as long as POU5F1 levels are reduced. Moreover, it is known that TEAD4 cannot act alone, instead it requires the “yes-associated protein 1” (YAP1) to act as a transcriptional activator. Although TEAD4 is similarly distributed in all cells of the embryo YAP1 is only localized to the nuclei of outer cells, indicating that these proteins can only cooperatively activate CDX2 expression in outside cells. The translocation into the nucleus was shown to be prevented through its phosphorylation by the Hippo pathway members LATS1 and LATS2. Given that cell contact can lead to Hippo pathway activation in cell cultures, the authors of these results supposed that increased intercellular cell contacts, like established in the embryo during compaction, leads to Hippo activation, followed by nuclear exclusion of YAP1 in inner cells. However, authors took into consideration that an involvement of cell polarity in YAP1 localisation is not ruled out by these results. Cell polarities could influence YAP1 localisation restricting the activity of Hippo signalling components, too. In support of this, dissociated blastomeres, without cell contact information and polarity, did not exhibit nuclear YAP1 [63].

1.4 The impact of proteome analyses on reproductive research

Several OMICS tools are available for embryonic analyses. While the use of genomic tools mainly focuses on the inherited genome itself and on chromosome aberrations, results from transcriptome, proteome and metabolome analyses can provide information about gene activity and reflect the environment composition due to epigenetic modifications and substrate availability [6]. Compared to a large number of genome and transcriptome approaches, only few holistic proteome studies concerning oocyte and embryonic development have been performed so far. Yet, results obtained by the analysis of mRNA cannot provide information about the instantaneous proteome status of a cell like the rate of translation or occurring post-translational modifications as well as protein secretion and proteolytic processes. Moreover, oocytes and early embryos contain not only a maternal messengerRNA (mRNA) but also a protein storage, which functions in fertilization and regulates preimplantation early embryo development [22]. The fact that only proteome analysis can provide insights into usage and assembly of this protein storage makes it even more promising. For example, it is still not known whether the above mentioned *Cdx2* expression on the apical embryonic cells of mice 8-cell embryos is also

reflected on the protein level [39]. Therefore, proteome techniques which can be applied for these efforts as well as the results obtained so far, will be described in the following chapters.

1.5 The two strategies for proteome analyses

For proteome analyses mainly two different strategies can be applied. On the one hand, the 2D gel based approach, which has the advantage of facilitating a very sensitive quantification and the effective detection of protein isoforms. On the other hand, the nano LC-MS/MS approach, which enables accurate and sensitive protein identification and has furthermore become more and more employed for protein quantification during the last years.

1.5.1 2D gel based techniques

The 2D gel based proteome analyses was independently introduced by Klose and O'Farrell in 1975 [64, 65]. On 2D polyacrylamide (PAA) gels proteins are separated in two dimensions according to their physicochemical parameters **isoelectric point (IP)** and **molecular weight (MW)**. Theoretically the separation strength of a typical 12 % PAA gel facilitates the separation of 100 proteins in each dimension, leading to 10,000 proteins in total. Practically proteins are not evenly distributed in the dimensions so that the number of proteins visible on 2D gels is around 2000 [66]. An important advantage of 2D gel electrophoresis is that different protein isoforms can usually clearly be separated.

For a long time protein visualization on 2D gels was limited to (i) the autoradiography of metabolically labelled proteins and (ii) to staining procedures, e.g. with silver [67], coomassie blue [68, 69] or sypro ruby [70] which could be either seen directly by the eye (silver, coomassie blue) or as fluorescent images after scanner-based absorbance spectroscopy (coomassie, sypro ruby). The silverstaining method is highly sensitive, yet due to its low dynamic range it is not appropriate for quantification [71]. Although the dynamic range of quantification was enhanced by the fluorescent sypro ruby staining, the quantification from 2D gels stayed associated with reproducibility problems which impaired inter-gel comparisons. In 1997, quantitative 2D gel electrophoresis was further improved by the introduction of the “**difference gel electrophoresis**” (DIGE) technique. It significantly enhanced the dynamic range of detection as well as the reproducibility of quantification on 2D gels [72]. In the 2D DIGE approach, fluorescent labels are attached to the proteins. The first labelling procedure introduced within the 2D DIGE approach is referred to as “minimal labelling”, since the fluorescent dyes are attached to only a small fraction of all ϵ -amino groups of lysines and free N-terminal residues, so that approximately 3 % of proteins in a sample are labelled. Currently

three fluorophores Cy2, Cy3 and Cy5 are available for minimal labelling. One of these dyes (usually Cy 2) is attached to a so-called “**internal pooled standard**” (IPS) consisting of an aliquot mixture from all samples included in the analysis. The remaining dyes are attached to different samples from the analysis. Prior to gel electrophoresis, samples and IPS are combined. Since the IPS is co-separated with each sample on one gel, it enables normalization and inter-gel comparisons leading to an enormous increase of reproducibility [66].

In 2003, an extremely sensitive modification of the DIGE concept, the “saturation labelling” which reduced the protein amounts needed for 2D gel based quantification by two orders of magnitude, was developed [73]. In contrast to the minimal labelling procedure, all proteins of a sample are labelled in the saturation DIGE approach. Instead to amino groups, the label is covalently linked to the sulfhydryl residues of cysteines. A crucial step for saturation DIGE analysis is the determination of an appropriate dye/protein ratio prior to the labelling procedure which must be strictly adhered to. In case of a too high dye/protein ratio (overlabelling) the ϵ amino group of lysines and free N-terminal residues are unspecifically labelled. This induces changes in the IP, leading to horizontal stripes on the gels. In case of a too low dye/protein ratio (underlabelling), proteins are incompletely labelled, what affects their molecular weight and induces vertical stripes on the gel. Currently only two fluorophores, Cy3 and Cy5, are available for saturation labelling, what doubles the number of gels to be prepared in contrast to a minimal DIGE experiment. The saturation labelling approach facilitates quantification of protein spots from gels containing only 500 ng protein, yet it has to be considered that for identification of differentially abundant spots preparative gels containing a protein amount of approximately 300-400 μ g protein are necessary [66, 74].

1.5.2 Nano LC-MS/MS based techniques

The history of mass spectrometry (MS) started at the end of the 19th century with the development of the first MS instrument by Thomson who was honoured by the Nobel Prize in 1906. In the following century, protein chemistry was mainly performed by N-terminal Edman degradation due to the lack of non-destructive ionisation techniques for the polar and zwitterionic peptides [75-77]. The first possibility enabling a “soft ionization” process was developed as “**fast atom bombardment**” (FAB) in the early eighties [78-80]. A decade later, protein identification by mass spectrometry was enabled through the invention of two non-destructive ionisation techniques for peptides, referred to as **electrospray ionisation** (ESI) [81] and **matrix-assisted laser desorption ionisation** (MALDI) [82]. Both approaches were published in the end of the eighties and made their inventors to Nobel Prize laureates in 2002. Today these

techniques are routinely employed in many labs for protein identification, representing a fast and sensitive method in many instances. A combination of **high performance liquid chromatography** (HPLC) for prefractionation and the so-called tandem mass spectrometry is generally referred to as LC-MS/MS [77].

A common strategy for LC-MS based proteome analysis is the so called “bottom-up” approach: In a first step, proteins are cleaved by sequence specific endopeptidases and the MS analysis is performed exclusively on the peptide level. Since this leads to an enormous increase of complexity it is often necessary to prefractionate the samples on the protein level, which is commonly performed by 1D SDS PAA gel electrophoresis [77]. In contrast, the so-called “top-down” approach enables a direct analysis of intact protein molecules [83].

One method for protein identification, applicable in cases where only a single or a few proteins are expected to be contained in a sample, is “**peptide mass fingerprinting**” (PMF). This technique is based on exact measurement of peptide masses, followed by a comparison of these values with theoretical peptide masses in protein databases [84, 85]. Since a PMF based protein identification cannot be applied for protein mixtures, additional information of the peptides is needed. Therefore, single peptides, referred to as “precursor ions”, providing a distinct MS spectrum are isolated and fragmented by collision with gas atoms or molecules which is called “**collision-induced dissociation**” (CID). The fragments of the precursor ions generate the so-called MS/MS spectra. These MS/MS spectra are then correlated with theoretical MS/MS spectra calculated from sequence databases. Due to the subsequent generation of MS and MS/MS spectra this method was named “Tandem MS” or “MS/MS” analysis [66].

Concerning protein quantification by nano LC-MS/MS, it is distinguished between “label free” approaches and such approaches wherein samples are labelled.

Label free approaches mainly comprise the method of “precursor ion intensity monitoring”, in which distinct precursor ions, for which reliable peptide matches have been established, are taken as a measure for the peptides’ abundances and the method of spectral counting. The “spectral counting” approach is based on the connection between a protein’s abundance and the number of MS/MS spectra assigned to an identified protein [66, 86, 87]

A common strategy in labelling approaches is (i) to label the samples with chemically identical reagents differing in their isotopic composition but not in their molecular weight, (ii) to pool samples after labelling and (iii) to perform LC-MS experiments with these pools. Since this procedure ensures that the retention time of differentially labelled peptides remains unaffected, peptides derived from different samples co-elute and reach the mass spectrometer at the same time. In the mass spectrometer, the signals from labels with different isotope composition are

detected individually and can then be quantitatively compared by their intensities [77]. There are several possibilities of sample labelling, differing in the level of sample procession at which the label is attached. If the label is introduced as isotope coded amino acid at the time of protein synthesis, the process is called metabolic labelling, also referred to as “stable isotope labelling by amino acids in cell culture” (SILAC) [88]. However, SILAC can be only applied to cell culture and small animal experiments due to cost reasons. In the so called “isotope-coded affinity tag” (ICAT) [89] and in the “isotope-coded protein labelling” (ICPL) [90] approach, the label is attached to the protein after sample lysis. The advantage of these two approaches is that labelled samples can be pooled prior to further digestion and preparation steps so that no more artificial differences are induced. Yet, a drawback in the ICAT system is, that the label is attached to cysteine residues which is problematic because only about 1.5 % of all amino acids within all known proteins are cysteines, so that often only very few peptides can be acquired for quantification. In the ICPL approach, the label is instead attached to lysine residues which indeed leads to more quantified peptides but comes with the problem that the major enzyme used for protein cleavage prior to mass spectrometry, namely trypsin, cannot cleave ICPL-modified lysine sites [77]. Hence, trypsin digestion leads to rather long peptides being difficult to analyse in the mass spectrometer. Using the so-called “isobaric tag for relative and absolute quantification” (iTRAQ) approach, the label is introduced on the peptide level on each lysine chain and on each N-terminal group so that the trypsin cleavage process is not impaired. It has been introduced 2004 by Ross et al. [91]. Advantages of the iTRAQ approach are (i) that in contrast to the ICPL approach the cleavage process is not biased since it is performed prior to the labelling procedure and (ii) that in contrast to the ICAT approach nearly all peptides are labelled which enhances the comprehensiveness of the quantitative analysis significantly [77].

For targeted, instead of holistic protein quantification, the “selected reaction monitoring” (SRM) technology can be employed. Although it has already been applied for small molecule quantification for several decades, it just started to be increasingly used for protein mass spectrometry in the last years [92]. Quantification by SRM can be performed in a mass spectrometer consisting of three quadrupole units usually referred to as “Triple Quad”. A peptide ion, the “precursor ion” is selected in the first quadrupole (Q1). It then enters the second quadrupole (Q2), where it undergoes collision-induced dissociation. In the third quadrupole (Q3), one or more fragment ions are selected for quantification by the detector [66]. This method can also be employed for absolute protein quantification, using isotopically labelled peptides which have been spiked into a sample in a known concentration prior to analysis [93, 94].

1.6 Holistic proteome analyses of mammalian oocytes and preimplantation embryos

There are numerous publications concerning oocyte and embryo development available, yet only few holistic proteome approaches have addressed mammalian oocytes and preimplantation embryos. Most likely, the reasons for the small number of proteomic publications can be found in the necessity of much higher amounts of sample material needed for a successful analysis as for example compared to transcriptomic experiments. A brief overview of these publications will be given in the following chapters.

1.6.1 Proteome analyses of oocytes

Since oocytes are cells with unique features (for example, they are able to undergo a transition from a fully differentiated stage to a stage of total pluripotency) and contain a protein storage which functions in fertilization and early embryonic development, the analysis of their proteome seems promising (see chapter 1.4). Therefore, several proteomic studies address both their qualitative proteome profile as well as quantitative alterations during the maturation process.

A very early protein analysis of mouse oocytes was published in 1977, when mass spectrometry was far away from being used for proteome analysis, thus limiting the authors in protein identification. Yet some important observations for oocyte biology were made by resolving [³⁵S]-radiolabelled proteins of different developmental oocyte stages on 1D PAA gels and subsequent imaging of the labelled proteins using fluorography. Oocytes were collected from juvenile (oocytes are growing and partly not able to undergo meiotic maturation) and from adult mice. The patterns of protein synthesis during different stages of growth and maturation were compared. The results demonstrated, that (i) the patterns of protein synthesis are very similar in individual oocytes at the same stage of growth or of meiotic maturation, indicating a high degree of biochemical homogeneity in a given population of isolated oocytes, (ii) that the linear increase in protein content of growing mouse oocytes is accompanied by significant qualitative changes of proteins synthesized, and that (iii) meiotic maturation is characterized by several discrete qualitative changes in the pattern of protein synthesis. Furthermore, it could be demonstrated by culturing oocytes with inhibitors of meiotic maturation that (i) protein synthesis is not required for germinal vesicle breakdown (GVBD) and that (ii) after GVBD has occurred, a programme of changes in protein synthesis is initiated, which takes place independently from the completion of nuclear progression to metaphase I [95].

In 2003, a 2D gel based study was performed to identify oolemmal mouse oocyte proteins. The authors separated the proteins from 2850 MII mouse oocytes on a 2D gel and identified some of the most intensive protein spots by mass spectrometry. In addition, 800 zona pellucida free

oocytes were surface labelled by biotinylation and also separated on a 2D gel. Spots were matched to the initial coomassie-stained reference gel and identified as molecular chaperones. The presence of HSP90, GRP94, GRP78 and Calreticulin on the oocyte surface was validated by immunofluorescence [96]. In 2010, a review was published by the same group in which it was stated that the potential functions of the molecular chaperones on the egg surface are still unclear. Yet they pointed out a report, which indicated that oolemmal Calreticulin likely plays a role in sperm–egg binding and signal transduction events during fertilization [97, 98].

In a quantitative analysis between GV and MII mouse oocytes by two silver stained 2D gels, one containing the proteins from 500 GV and one containing the proteins from 500 MII oocytes, 12 proteins appeared to be differentially expressed. Spot identification showed that TACC3, HSP 105, STI1, ADSS, Lipocalin and Lysozyme 1 were lower in abundance in MII oocytes, while one protein, TCTP was of higher abundance in MII oocytes. Since five proteins, PDCD6IP, Importin-a2, Nudix, Nucleoplasmin2 and Spindlin, were identified in both samples, but from spots of different mass and isoelectric points in the two gels, authors concluded that these proteins were posttranslationally modified during maturation [99]. Later it was revealed in different publications that TACC3 is required for microtubule anchoring at the centrosome [100] and that it is essential for spindle assembly and cell survival [101].

Another 2D gel based but qualitative mouse oocyte proteome study was published in 2008. From a gel of zona pellucida free MII mouse oocytes, 869 selected protein spots, corresponding to 380 unique proteins, were identified. A total of 90 protein spots, representing 53 unique proteins, were stained with Pro-Q Diamond, indicating that the contained proteins are in phosphorylated forms [102].

The next study also targeting qualitative mouse oocyte protein expression was performed in 2009. In an LC-MS/MS approach, 625 proteins were identified from 2700 zona free MII mouse oocytes proteins. By comparing the identified proteins to mRNAs known to be expressed at high levels both in oocytes and fertilized eggs, a subset of 76 proteins was pointed out. It contained nine proteins (MATER, STELLA, DNMT1, ZAR1, NPM2, PADI6, TLE6, TCL1, FILIA) of which the corresponding genes were described as “maternal effect genes” due to their characteristic as to be absolutely necessary for oogenesis, fertilization or early embryonic development. Consequently, these proteins were named “maternal effect proteins”. Additionally, identified proteins were compared to a proteome analysis of mouse embryonic stem cells (ESCs) and an overlap of 371 proteins was found. Besides the identification of some pluripotency markers, this group of proteins included many uncharacterized proteins, which

were supposed to be good candidates for studying the mechanism of reprogramming by the authors [103].

For studying the mouse oocyte proteome at different developmental stages, the proteins from 7000 GV oocytes as well as 7000 MII oocytes and 7000 zygotes were qualitatively analysed in 2010. 2781 proteins were identified from GV oocytes, 2973 proteins from MII oocytes, and 2082 proteins in zygotes by LC-MS/MS analysis. An abundance comparison of the proteins identified from the three developmental stages to each other as well as to proteins identified from ESCs was performed by taking the identified peptides per protein as indication for its abundance. The comparison revealed that GV and MII oocytes are more similar than zygotes and ESCs. Compared with oocytes, ESCs expressed some specific or highly abundant proteins, mostly involved in metabolism. Oocytes and zygotes possessed specific protein families, which are involved in self-renewal and cell cycle regulation, more than ESCs. In a pathway analysis of the developmental stages, it was revealed that GV oocytes contained a greater number of metabolism-related proteins responsible for supporting oocyte maturation, while MII oocyte proteins were more involved in the regulation of cell cycle events and epigenetic modifications. Taken together it was pointed out that different protein compositions are correlated with oocyte characteristics at different developmental stages [104].

Very recently the catalogue of identified mouse MII oocytes was extended to 3699 proteins, identified from 1884 zona free mouse MII oocytes. This large number of identifications was facilitated by the use of a mass spectrometer of the latest generation (Orbitrap Velos XL), demonstrating how the technical development in mass spectrometry is correlated with insights in biology. This protein catalogue was comparable by size to an ESC protein catalogue. Hence, a valid comparison between the MII oocyte proteins to the proteins from ES cells could be performed, leading to 2556 proteins contained in both cell types. Since the oocyte has been proposed to hold enough reprogramming factors for up to 100 nuclei, this subset of proteins was screened for such factors by looking for proteins which are (i) localized in the nucleus, (ii) having chromatin as substrate, and (iii) acting catalytically. The list of 28 proteins which matched those criteria was named “Reprogrammome” by the authors, since they considered them to belong to the set of molecules that enable reprogramming [105].

The bovine oocyte proteome has been addressed by fewer publications. In 2010, a qualitative proteome analysis of bovine GV oocytes, containing 811 identified proteins, was published. Additionally, authors analysed cumulus cells and identified 1247 proteins. A qualitative comparison of both datasets revealed that 352 proteins were common to both cell types. A quantitative comparison based on a spectral count approach led to 371 proteins differentially

expressed between oocytes and cumulus cells. Gene Ontology and a pathway analysis showed that cumulus cells have higher numbers of proteins involved in cell communication, generation of precursor metabolites and energy, as well as into transport than GV oocytes [106].

The pattern of newly synthesized proteins during *in vitro* maturation of bovine oocytes was addressed by a study published in 2004. Oocytes were *in vitro* matured with [³⁵S]-methionin for four hour periods from time zero to 28 hours. Pools of ten oocytes were then prepared for 2D gel electrophoresis. For each time interval, three gels were obtained, digitalized, and analysed. Three major patterns of protein synthesis were observed during bovine oocyte maturation *in vitro*: one at the beginning of maturation (0–4 hr), another one in the middle (4–16 hr) and the last one after the completion of MI stage (16–28 hr) [107]. Although this study revealed differences in the qualitative pattern of newly synthesized proteins during maturation, protein identifications were not provided.

The improvement in the development of 2D gel based techniques as well as their impact on developmental biology has been impressively demonstrated by an analysis of the bovine oocyte *in vitro* maturation process. The application of the saturation DIGE technique enabled quantification from 2D gels containing 500 ng proteins corresponding to only five oocytes. Six biological replicates from *in vivo* and *in vitro* matured oocytes were analysed and enabled a valid statistical evaluation of differentially abundant spots. Ten differentially abundant proteins could be unambiguously identified from a preparative gel containing 2200 bovine GV oocytes, including the “Translationally controlled tumor protein”, enzymes of the Krebs and pentose phosphate cycles, Clusterin, 14-3-3 ϵ , Elongation factor-1 γ , and redox enzymes. Especially interesting was the new detection of three Glutathione S-transferase Mu 5 (GSTM5) isoforms whose abundances were moreover altered in opposing directions between GV and MII oocytes [74].

The pig oocyte proteome was addressed by a 2D gel study published in 2004. Proteins from 600 GV oocytes were separated on a 2D gel from which 35 protein spots were identified. Additionally, from the GV, the MI and the MII stage, 200 oocytes each were separated on 2D gels followed by silver staining. The quantitative comparison of these spots revealed six differentially abundant proteins, however, only one of these spots could be unambiguously identified [108]. The same group deepened these insights into porcine oocyte *in vitro* maturation by another 2D gel based study published in 2007. Four replicates of GV oocytes were *in vitro* matured and labelled with [³⁵S]-methionin at time points 0 (GV stage), 24 h (MI stage) and 40 h (MII stage) during *in vitro* culture. Autoradiography revealed between 240 and 120 protein spots at the indicated developmental stages, demonstrating the complexity of de novo protein

biosynthesis during oocyte maturation. In total 16 spots were found to be significantly altered in intensity between the GV and the MII stage, among which four spots showed an increased rate of biosynthesis and the contained proteins were therefore supposed to play an essential role in meiosis. The role of one of these spots, identified as “Ubiquitin C-terminal hydrolase-L1” (UCH-L1), was further studied and it was revealed that UCH-L1 inhibitor treated oocytes were impaired in their capacity to reach the MII stage during meiosis. In addition, the inhibition of UCH-L1 resulted in elevated M-phase promoting factor (MPF) activity and a low level of Monoubiquitin. The metaphase I to anaphase transition needs a decrease in MPF activity by degradation of its regulatory subunit cyclin B 1 over the ubiquitin–proteasome pathway. This supported the hypothesis that UCH-L1 might play a role in metaphase I-anaphase transition by regulating ubiquitin-dependent proteasome mechanisms [109].

Just recently, a nano LC-MS/MS based quantitative proteome analysis of high and low quality porcine oocytes was published. High quality oocytes were produced by *in vitro* maturation with gonadotrophins and low quality oocytes by *in vitro* maturation without gonadotrophins. The oocytes were labelled with different isobaric tags and subjected to mass spectrometry. In total 503 proteins were identified, from which 16 proteins differed in their abundance between two replicates of low and high quality oocytes. From the fact that differences between the first and the second replicate were significantly more pronounced in the first replicate, authors concluded a general quality difference between the oocytes for each replicate based on conditions and timing of oocyte collection. To find out whether the differentially abundant oocyte proteins are secreted, the maturation media were also qualitatively analysed by nano LC-MS/MS, leading to 110 identified proteins which were not detected in a control medium. Eight of these proteins belonged to the group of differentially abundant proteins between low and high quality oocytes. More abundant proteins in the high-quality oocyte proteome included “Kelch-like ECH-associated protein 1” (an adaptor for ubiquitin-ligase CUL3), “Nuclear export factor CRM1” and “Ataxia-telangiectasia mutated protein kinase”. Dystrophin (DMD) was more abundant in low-quality oocytes and was also identified from the media [110].

1.6.2 Proteome analyses of preimplantation embryos

Up to today the number of published data of holistic proteome analyses from preimplantation mammalian embryos is rather limited. This might be due to the time consuming process of collecting sufficient amounts of sample material. A method to address the proteome from very limited sample amounts is to employ the “surface enhanced laser desorption ionisation time-of-flight” (SELDI-TOF) technique. It is a variation of the MALDI technique (see chapter 1.5.2) but

has the disadvantage of a low resolution and accuracy. Therefore, this technique is indeed capable to recognize patterns of proteins expression, but cannot provide valid protein identifications and is regarded as to have a very limited potential to enhance biological knowledge [66]. For this reason, publications based on the SELDI-TOF technique will not be described in detail. Examples of its application are the comparisons of *in vivo* and under different oxygen concentrations *in vitro* cultured mice embryos [111] as well as the comparison of early, expanded and degenerated human blastocysts [112].

A very early approach to analyse bovine embryo proteins was published in 1989. Although protein identifications could not be provided, the results had an impact on the understanding of bovine embryo biology. The authors cultured bovine oocytes and preimplantation embryos with [³⁵S]-methionin and separated their protein lysates of different developmental stages by 1D-PAA-SDS gel electrophoresis. The protein patterns detected by autoradiography were compared. While minor differences were detected between the lanes from oocytes, zygotes, 2-cell-stage and 4-cell stage embryos the pattern of those stages is obviously different from the embryos collected when they showed between 8-16 blastomeres. The patterns of the embryos collected between the 16- to 32-cell stages and blastocysts were similar to each other but both were distinct from those produced by embryos before the 8-cell stage. Additionally, embryos of different developmental stages were incubated with [³H]-uridine to measure rates of RNA synthesis. Incorporation of radiolabelled uridine into RNA was first detected at the 16-cell stage. These results suggested to the authors that protein synthesis is programmed by maternal mRNA up to the 8-cell stage but switches to mRNA derived from the zygote genome between the 8- and 16-cell stages. This suggestion was confirmed by several publications later on and allowed to time the MET to occur between the 8- and 16-cell stage [35].

Seventeen years later, when proteome techniques were far more developed, the patterns of newly synthesized proteins during embryo development were addressed again by radiolabelling of proteins: An above mentioned study [107] concerning patterns of newly synthesized proteins during bovine *in vitro* maturation, was continued by identifying proteins which are not only translated throughout *in vitro* maturation of bovine oocytes but also in the 2-cell, 4-cell and 8-cell stage of *in vitro* cultured bovine embryos. Three replicates of 2-cell, 4-cell and 8-cell stage embryos were cultured with α -amanitin in order to repress transcription and with [³⁵S]-methionin and [³⁵S]-cystein to reveal newly translated proteins. Embryos were lysed and separated by 2D-gel electrophoresis. Autoradiography of these gels revealed 291 (2-cell stage), 373 (4-cell stage) and 252 (8-cell stage) spots on each gel. Of these spots 70, 83 and 28 contained proteins which are exclusively translated at the 2, 4 and 8-cell stage respectively, while 123 spots were continuously present during all three stages. These spots were matched to

the 92 spots from the previously performed *in vitro* maturation experiment (see previous chapter), revealing that 46 spots are shared in position. Of these 46 proteins, 32 spots could be matched to their position on a preparative gel from 3000 bovine GV oocytes. Ten of these spots were successfully identified by MALDI-TOF (HSC71; HSP70; CYPA; UCH-L1; GSTM5; CCT5; E-FABP; 2,3-BPGM, Ubiquitin-conjugating enzyme E2D3; and Beta-actin/Gamma-actin). Due to their characteristic features of being newly synthesized in all the above mentioned stages, the authors named them “**maternal house keeping proteins**” (MHKPs) [113].

A dedicated proteomic comparison between porcine IVF and parthenogenetic zygotes was published in 2009. Protein lysates of 6000 IVF zygotes and 6000 parthenogenetic zygotes were separated on 1D-PAA-SDS gels, from which gel slices were analysed by nano LC-MS/MS. In total 735 proteins were identified. Only 51.3 % (377) proteins were identified in both the IVF and PA zygotes, indicating major alterations of their protein composition. The relative abundance of proteins in each sample was estimated by a spectral count approach. 90 % of the proteins turned out to be differentially abundant with an abundance ratio of > 2 , among which abundance ratios of > 10 were observed for 410 proteins. To verify the quantitative measurements, SRM assays confirmed the tendency of abundance alteration for nine proteins (JAK2, STAT1, STAT2, Calpastatin, Complement cytolysis inhibitor, Plakoglobin, Serotransferrin, Epsilon-globin, Acetyl-CoA C-acetyltransferase) from the spectral count approach. These results provided putative markers of embryo quality and strengthened the hypothesis that, besides genomic imprinting, aberrant protein expression in PA embryos could be a reason for their developmental failure [114].

1.7 Aim of the thesis

The aim of this doctoral thesis was to analyse both the qualitative and quantitative proteome profile of oocytes and of preimplantation embryos. Due to the fact that the cow represents an excellent model for human reproduction, all analyses were performed on bovine samples.

To reflect the situation of ARTs embryos were produced *in vitro*. In a first step a qualitative protein profile of GV oocytes was produced to obtain an insight into the protein storage of oocytes. In the following, proteomic approaches were scaled down to face the small sample amounts available for comparative analyses of a sufficient number of biological replicates from oocytes and early embryos. This enabled to address several crucial steps of embryonic development on the protein level. Since the *in vitro* maturation process of oocytes is associated with major drawbacks [115], the differences between *in vivo* and *in vitro* matured oocytes were analysed. To encompass the maternal embryo transition (see chapter 1.3.2), 2-cell stage

embryos were compared with morulae. During the morula to blastocyst transition, the first cell differentiations become morphologically obvious, while major changes in the embryonic structure occur (see chapter 1.3.2). Therefore, this developmental period was addressed in a comprehensive way, by analysing morulae and blastocysts using two complementary approaches. An illustration of this subsequent proteome analysis of bovine oocytes and early embryonic stages as performed in this doctoral thesis is shown in Fig. 4.

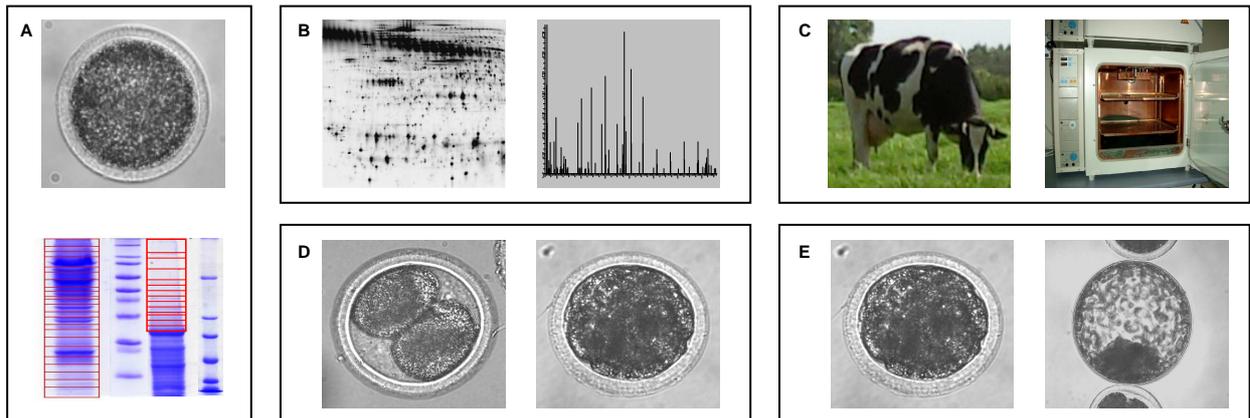


Fig. 4: Illustration of the subsequent proteome analyses of bovine oocytes and early embryonic stages as performed in this doctoral thesis

(A) Qualitative proteome profile of 900 GV oocytes; (B) Method adaptation to the analysis of small sample amounts; (C) Comparison of *in vitro* and *in vivo* matured oocytes (D) Comparison of 2-cell stage embryos and morulae; (E) Comparison of morulae and blastocysts; Embryo pictures were taken by Myriam Reichenbach.

2 Materials and Methods

2.1 Sample Generation

2.1.1 Generation of GV oocytes

Ovaries were collected from a local slaughterhouse and stored for approximately three hours in phosphate buffered saline (PBS) at 30°C. To remove blood and debris ovaries were washed three times in PBS at 30°C. Cumulus oocyte complexes (COCs) were obtained by aspirating 2–8 mm follicles with a 20-gauge needle and a vacuum pressure of approximately 100 mm Hg. COCs which had at least one layer of compact cumulus cells were selected and denuded mechanically by vortexing for four minutes. Denuded oocytes were washed three times in PBS + 0.1 % Polyvinylpyrrolidone (PVP) before they were transferred into a sample cup. The remaining buffer was extracted by suction with a capillary. Buffer free oocytes were frozen on dry ice and stored at -80°C until analysis.

2.1.2 Generation of early embryos

Embryos were prepared according to a modified procedure as primarily described by Berg and Brem [116]. COCs were obtained from slaughterhouse ovaries as described above. COCs which had at least one layer of compact cumulus cells were selected and washed three times in oocyte maturation medium. The COCs were matured for 22 hours in four-well dishes in 400 µl droplets of oocyte maturation medium in an incubator with 5 % CO₂ and 39.0°C. Matured COCs were washed three times in fertilization medium and transferred to 400 µl droplets of fertilization medium in four-well dishes. Frozen-thawed spermatozoa were prepared in a procedure called “swim up” according to a method developed by Parish et al [117]. For all experiments, sperm from the same bull (Mindel) was used. Sperm straws containing 20*10⁶ sperms were thawed for ten seconds at 37°C. Approximately 5*10⁶ thawed sperms were layered under 1 ml sperm talp medium in centrifugal tubes and incubated at 5 % CO₂ and 39.0°C for capacitation. After one hour incubation time, the supernatant containing the healthy and mobile sperm was collected and centrifuged for 10 minutes with 1800 g at 26 °C. The supernatant was removed and about 5*10⁶ sperms were co incubated with 20 to 40 oocytes in 400 µl fertilization medium for 18 hours at 5 % CO₂ and 39 °C. Presumptive zygotes were mechanically denuded by vortexing for four minutes, washed three times in synthetic oviductal fluid (SOF) and transferred to 400 µl droplets of SOF under mineral oil in four-well dishes. They were incubated with 5 % CO₂ and 5 % O₂ at 39°C. 2-cell-stage embryos were collected 27 hours post fertilization, morulae were collected five days post fertilization (day of fertilization = day 0) and blastocysts were collected

seven days post fertilization. Morulae were collected when they had at least 32 blastomeres. Blastocysts were collected both in the early and the expanded stage. All embryos were microscopically evaluated to their blastomeric structure prior to collection. Embryos were washed three times in PBS + 0.1 % PVP before they were transferred into a sample cup. The remaining buffer was extracted by suction with a capillary. Buffer free embryos were frozen on dry ice and stored at -80°C until analysis.

2.1.3 Generation and processing of *in vivo* aspirated oocytes by ovum pick-up (OPU)

The generation of *in vivo* aspirated oocytes was performed by the group of Prof. Wolf at the Moorversuchsgut Badersfeld. For the aspiration of GV oocytes, OPU was carried out in three and four day intervals. For the aspiration of MII oocytes, OPU of superstimulated cows was carried out in four to five weeks intervals. For the exact procedure of OPU please refer to the process described in the publication of Machado et al. [118].

For superstimulation animals were treated according to the stimulation scheme in Table 1. Animals received a “controlled intravaginal progesterone-releasing device” (CIDR) between days 7-10 of the oestrous cycle (day 0 = oestrus) for nine days. On day six after the application of the CIDR the twice daily injection of “follicle stimulating hormone” (FSH) at 12 hours intervals in decreasing doses was started and carried on for four days. The total dose was 500 I.U. for heifers and 750 I.U. for cows. PGF_{2α} was injected twice at day eight and day nine of the FSH injections (61 and 49 h prior to OPU). The CIDR was removed at the second PGF_{2α} administration. Additionally 18 to 20 h before the particular OPU, 5 ml of a GnRH analogue was injected.

Table 1: Scheme of superstimulation for the generation of *in vivo* matured oocytes

Day of stimulation	1 to 5	6	7	8	9	10	Day of OPU
Administered drug	CIDR	CIDR	CIDR	CIDR	CIDR		
		FSH	FSH	FSH	FSH		
				PGF _{2α}	PGF _{2α}	GnRH	

The *in vitro* maturation (IVM) of GV oocytes was performed as described in chapter 2.1.2. Matured oocytes were microscopically evaluated for the presence of a polar body and mechanically denuded by pipetting. Denuded MII oocytes were washed three times in PBS + 0.1 % PVP before they were transferred into a 0.5 ml sample cup. The remaining buffer was extracted by suction with a capillary. Buffer free oocytes were frozen on dry ice and stored at -80 °C until analysis.

2.2 Nano-LC-MS/MS analyses

2.2.1 Qualitative proteome profile of 900 GV oocytes

2.2.1.1 Lysis

70 μ l Laemmli buffer was added to 900 GV oocytes. Oocytes were pre-lysed by pipetting up and down before they were heated at 95°C for five minutes. In the following, they were frozen on dry ice for five minutes and sonicated in a water bath for five minutes. The procedure of freezing and sonication was repeated three times in total. The lysate was centrifuged for 15 minutes with 14,000 g. The supernatant was collected and stored at -80°C.

2.2.1.2 1D polyacrylamide SDS gel electrophoresis

SDS gel electrophoresis was performed using a 0.5 cm 4 % stacking gel and a 12 %, respectively a 6 % separation gel. The overall gel size was 7 cm (L) \times 8.5 cm (W) \times 0.75 mm. On each gel, half of the protein lysate of 900 GV oocytes was separated. Separation was performed for 20 minutes at a voltage of 100 V and for additional 60 min at 200 V in SDS running buffer. The gels were stained overnight with coomassie staining solution and destained for at least eight hours with coomassie destaining solution.

2.2.1.3 Gel slicing and tryptic in-gel digestion

The gel lanes were cut into 23 (12 % gel), respectively 14 (6 % gel) gel slices using a scalpel. The positions and size of the gel slices are indicated in Fig. 5. Slices were transferred in 1.5 ml sample cups filled with 500 μ l bidest H₂O. To reduce and block the cystein residues gel slices were incubated with 45 mM dithioerythritol (DTT) and 50 mM ammonium hydrogen carbonate (NH₄HCO₃) for 30 min at 55 °C followed by two 15 minutes incubation steps with 100 mM iodacetamide in 50 mM NH₄HCO₃. Prior to digestion, gel slices were washed twice for 15 min in 50 mM NH₄HCO₃ and minced with a pipette tip. Tryptic hydrolysis was performed overnight at 37 °C in 30 μ l 50 mM NH₄HCO₃ with 100 ng modified porcine trypsin per gel slice. The supernatant was collected and preserved. The peptides were further extracted by a ten minute treatment with 100 μ l 70 % acetonitril (ACN). The NH₄HCO₃ and ACN fractions were pooled and dried using a vacuum centrifuge. Samples were stored at -80 °C. Prior to LC-MS/MS analysis, the peptides were solved in 0.1 % formic acid (FA).

2.2.1.4 1D-LC MS/MS analysis

The 1D-liquid chromatography (1D-LC) was performed on a multi-dimensional nano-liquid chromatography system. Peptides of each gel slice were loaded separately on a trap column at a flow-rate of 10 μ l per minute and subsequently separated with an analytical column (reversed phase column I) at a flow rate of 260 nl/min and the following consecutive gradients: I) 0–30 % B in 80 min, II) 30–60 % B in 30 min, III) 100% B for 10 min (A: 0.1 % FA, B: 84 % ACN and 0.1 % FA). Mass spectrometry was performed on an LTQ mass spectrometer (Mass spectrometer I) online coupled to a nano-LC system (nano HPLC I). For electrospray ionization a “distal coated silica tip” and a needle voltage of 1.4 kV was used. The MS method consisted of cycles of one full MS scan (Mass range: 300–2000 m/z) and three data dependant MS/MS scans. MS/MS spectra were generated in the “collision induced dissociation” (CID) cell of the mass spectrometer with 35 % collision energy. Dynamic exclusion was set to 180s.

2.2.1.5 Database search and data analysis

The MS/MS data were searched with the mascot search software against the ipi_BOVIN_v362 database using the following parameters: i) Enzyme: Trypsin, ii) Fixed Modification: Carbamidomethyl (C), iii) Variable modifications: Oxidation (M); iv) Peptide tolerance 2 Da, v) MS/MS tolerance 0.8 Da, vi) Peptide charge 1+, 2+ and 3+, vii) Instrument ESI-TRAP and viii) Allow up to 1 missed cleavages. For statistical validation of MS/MS based peptide and protein identifications, the scaffold software was used. Scaffold was set up to accept peptide identifications greater than 95 % probability as specified by the “peptide prophet” algorithm [119]. Protein identifications were accepted if they met the scaffold protein probability of 99 % and were identified by at least two peptides. Protein probabilities were assigned by the “protein prophet” algorithm [120]. To further confirm the probability calculated by scaffold, a so-called decoy version of the ipi_BOVIN_v362 database consisting of randomized sequences was generated. The database was also searched by the mascot search software using the same parameters as described above. Any protein hit derived from the decoy database was regarded as false-positive identification.

2.2.1.6 Gene ontology analysis

Prior to gene ontology (GO) analyses all protein IPI accession numbers were transformed into Uniprot accession numbers. The GO analysis was performed by a GO tool which clusters proteins according to terms provided by the GO project [121]. The mapping and analysis tools are provided online by the Uniprot community (UniProt Gene Ontology / UniProt KB, <http://www.uniprot.org/>).

2.2.2 Qualitative and quantitative analyses from small sample amounts of oocytes and embryos

2.2.2.1 Sample preparation and iTRAQ labelling

Lysis was carried out by adding 8 M urea buffer to all samples in a concentration of 0.128 μ l per oocyte/embryo. Samples were pre-lysed by pipetting up and down followed by 15 minutes of sonication. To reduce and block the cystein residues samples were incubated with 45 mM DTT in 50 mM NH_4HCO_3 for 30 min at 55 °C, followed by 15 minutes incubation with 100 mM iodacetamide in 50 mM NH_4HCO_3 . Samples were diluted to a ratio of 1:8 with the NH_4HCO_3 to obtain a pH value of around eight suitable for activating trypsin. Modified porcine trypsin was added in a concentration of five ng per oocyte. Samples were incubated overnight at 37 °C before they were dried in a vacuum centrifuge. Samples were stored at -80 °C.

For quantitative iTRAQ analysis samples containing 25 embryos/oocytes each were additionally cleaned up using Zip Tips according to the manufacturer's protocol. Zip Tips are pipette tips containing C18 material which binds peptides so that superfluous salts and urea which would disturb the labelling reaction can be washed off.

The labelling procedure was performed according to the manufacturer's protocol downscaled to small sample amounts. Samples were dissolved in dissolution buffer to a concentration of 168 nl per oocyte/embryo. The iTRAQ reagent contained in a single vial (amount is not provided by the manufacturer) was reconstituted in 70 μ l high purity ethanol and vortexed for one minute. 1 % of the reagent amount contained in a single iTRAQ vial was added to the samples per oocyte/embryo. After one hour incubation time samples were dried in a vacuum centrifuge and stored at -20 °C until analysis.

iTRAQ labelled samples used for 1D LC-MS/MS analyses were cleaned up by "vivapure S mini M" devices according to the manufacturer's protocol. In a 2D LC-MS/MS approach this was not necessary since the clean up procedure is performed by the "strong cation exchange" column.

2.2.2.2 1D and 2D liquid chromatography (LC)

Prior to LC-MS/MS analysis the peptides were solved in 0.1 % FA. For 1D-LC MS/MS analysis, peptides were loaded on a trap column at a flow-rate of 10 μ l per minute and subsequently separated with an analytical column (Reversed Phase column II) at a flow rate of 260 nl/min and the following consecutive gradients: I) 0–30 % B in 80 min, II) 30–60 % B in 30 min, III) 100 % B for 10 min (A: 0.1 % FA, B: 84 % ACN and 0.1 % FA).

For 2D-LC MS/MS analysis peptides were additionally prefractionated using a “strong cation exchange” (SCX) column. Peptides were eluted from the SCX column with five different ammonium chloride (NH₄Cl) solutions of (i) 10, (ii) 25, (iii) 50, (iv) 100 and (v) 500 mM.

2.2.2.3 Mass spectrometry

Mass spectrometry was performed on an LTQ Orbitrap XL mass spectrometer (mass spectrometer II) online coupled to the nano-LC (nano HPLC I) system. For electrospray ionization, a “distal coated silica tip” and a needle voltage of 1.4 kV was used. The MS method consisted of cycles of one full MS scan (Mass range: 300–1800 m/z) and five data dependent MS/MS scans. MS/MS spectra were generated in the CID cell of the mass spectrometer with 35 % collision energy. Dynamic exclusion was set to 180s.

2.2.2.4 Database search

The MS/MS data were searched with Mascot against the ipi_BOVIN_v362 database using the following parameters: i) Enzyme: Trypsin, ii) Fixed Modification: Carbamidomethyl (C), iii) Variable modifications: Oxidation (M); iv) Peptide tolerance 50 ppm, v) MS/MS tolerance 0.8 Da, vi) Peptide charge 1+, 2+ and 3+, vii) Instrument ESI-TRAP and viii) Allow up to 1 missed cleavages. For the database search of the iTRAQ labelled samples the parameter ii) (Fixed Modification) was modified and set to “iTRAQ (K)” and “iTRAQ (N-term)”.

2.2.2.5 Generation of exclusion lists for quantitative iTRAQ analysis

Exclusion lists (ELs) were generated by the proteome discoverer software. Time intervals were set to 30 minutes during the optimisation experiments (see chapter 3.4) and to ten minutes within the comparison of *in vivo* versus *in vitro* matured oocytes (see chapter 3.5) as well as within the comparison of morulae and blastocysts (see chapter 3.7.2). Each exclusion list was loaded into the Xcalibur software which is used for MS acquisition.

2.2.2.6 Data analysis for identification and quantification

For protein identification, the scaffold software was used as described in chapter 2.2.1.5 with the additional modification that peptides had to be detected within a 10 ppm accuracy window, reflecting the high mass accuracy of the LTQ Orbitrap XL mass spectrometer.

The quantification by “spectral count” applied to the analysis of 2-cell stage embryos and morulae (see chapter 3.6) was performed by the scaffold software. Only proteins for which (i) MS/MS spectra had been detected in each stage, (ii) the corresponding peptides were detected with at least ten MS/MS spectra in at least one of the two stages and (iii) which had a “spectral count abundance ratio” of at least 2 were considered as differentially abundant.

For iTRAQ quantification, the scaffold’s quantification “Q+” module was used. For the comparison of *in vivo* versus *in vitro* derived oocytes, the peptide probability for quantification was set up to accept only peptides identifications greater than 80 % probability within 10 ppm mass tolerance and the minimum dynamic range was set to 1 %. For the comparison of morulae and blastocysts, the peptide probability for quantification was set up to accept only peptides identifications greater than 95 % probability within 10 ppm mass tolerance and the minimum dynamic range was set to 5 %.

For all analyses the “normalization to the entire dataset average” feature was applied. All quantification data were exported to Microsoft office excel. To find out which proteins were identified and quantified in all runs the Microsoft office access software was used. For the comparison of *in vivo* versus *in vitro* derived oocytes the average fold change was calculated in Microsoft office excel. Proteins with an average log 2 fold change of $> |1|$ were considered as differentially abundant proteins. For the comparison of morulae and blastocysts the calculations of student’s t-test p-values and the generation of a Volcano plot were performed by the R-Software. Proteins which had a p-value of < 0.05 and an average log 2 fold change of $> |0.6|$ were considered as differentially abundant proteins.

2.2.2.7 Targeted quantification by Selected Reaction Monitoring (SRM)

Targeted quantification of IMP3 and YBX2 was performed on a QTRAQ 5500 (mass spectrometer III) online coupled to a nano-liquid chromatography system (nano-HPLC II). Proteotypic peptides selected for quantification of IMP3 were the peptides IPVSGPFLVK (IMP3_1) and FTEEIPLK (IMP3_2). Proteotypic peptides selected for quantification of YBX2 were the peptides GAEANVTGPGGVPVK (YBX2_1) and TPGNPATAASGTPAPLAR (YBX2_2). Samples were loaded on a RP trap column at a flow-rate of 5 μ l per minute and subsequently separated with an analytical column at a flow rate of 280 nl / min and the

following consecutive gradients: I) 0–30 % B in 80 min, II) 85 % B for 10 min, (A: 0.1 % FA, B: 100 % ACN and 0.1 % FA). The exact peptide transitions can be looked up in Table 2. The method for targeted quantification of these three proteins has been previously established by Thomas Fröhlich and Daniela Deutsch in the proteomics facility of LAFUGA.

Table 2: Transitions for YBX2 and IMP3 Quantification

Peptide name	Sequence	Charge state	Q1 [m/z]	Q3 [m/z]	Ion	Dwell time [ms]	Collision Energy [V]
YBX2_2	TPGNPATAASGTPAPLAR	2	825.43	940.52	y10	100	43
YBX2_2	TPGNPATAASGTPAPLAR	2	825.43	869.48	y9	100	46
YBX2_2	TPGNPATAASGTPAPLAR	2	825.43	1011.56	y11	100	46
YBX2_1	GAEAAANVTGPGGVPVK	2	712.38	811.47	y9	100	35
YBX2_1	GAEAAANVTGPGGVPVK	2	712.38	910.54	y10	100	35
YBX2_1	GAEAAANVTGPGGVPVK	2	712.38	343.23	y3	100	33
IMP3_2	FTEEIPLK	2	488.77	728.42	y6	100	26
IMP3_2	FTEEIPLK	2	488.77	830.43	b7	100	26
IMP3_2	FTEEIPLK	2	488.77	599.38	y5	100	26
IMP3_1	IPVSGPFLVK	2	528.83	747.44	y7	100	27
IMP3_1	IPVSGPFLVK	2	528.83	846.51	y8	100	27
IMP3_1	IPVSGPFLVK	2	528.83	660.41	y6	100	30
YBX2_1_aqua	GAEAAANVTGPGGVPVK	2	716.38	819.47	y9	100	35
YBX2_1_aqua	GAEAAANVTGPGGVPVK	2	716.38	918.54	y10	100	35
YBX2_1_aqua	GAEAAANVTGPGGVPVK	2	716.38	351.23	y3	100	33
IMP3_1_aqua	IPVSGPFLVK	2	532.83	755.44	y7	100	27
IMP3_1_aqua	IPVSGPFLVK	2	532.83	854.51	y8	100	27
IMP3_1_aqua	IPVSGPFLVK	2	532.83	668.41	y6	100	30

2.3 2D gel based Saturation DIGE analysis

2.3.1 Sample lysis

Lysis of morulae and blastocysts for analytical gels was carried out by adding 4 μ l DIGE lysis buffer to each pool of 25 embryos (containing 2.25 μ g protein), so that the final protein concentration was 0.55 μ g/ μ l, which is the optimal protein concentration for the labelling reaction with Cy5Dye saturation Dyes. Lysis of GV oocytes for preparative gels was carried out by adding DIGE lysis buffer in a concentration of 0.03 μ l per oocyte to pools of 186 to 495 GV oocytes.

All samples were pre-lysed by pipetting up and down. Samples were sonicated for five minutes in an ice containing water bath, followed by five minutes of freezing on dry ice. The procedure of sonication and freezing on dry ice was repeated three times in total. Samples for the analytical gels were immediately processed. All GV oocyte samples for the preparative gel were pooled and the final protein concentration was determined by a standard Bradford assay before they were stored at -80 °C.

2.3.2 Labelling of proteins with fluorescent Cy5 Dyes for analytical gels

15 µg of an “internal pooled standard” (IPS) was prepared by pooling 1.25 µg protein of each sample and was labelled with Cy5 Dye Cy3, the remaining 1 µg from each sample was labelled with Cy5 Dye Cy5.

The labelling was performed according to the manufacturer’s protocol. Cystein residues were reduced with tris (2-carboxyethyl) phosphine hydrochloride (TCEP) which was added in a concentration of 0.5 nmol / µg to each sample. Samples were incubated at 37 °C for one hour in the dark. Cy5 Dyes were reconstituted in N,N-dimethylformamide (DMF) to a 2 mM solution and added to the samples in a concentration of 1 nmol / µg. Samples were incubated at 37 °C for 30 minutes in the dark. The labelling reaction was stopped by addition of an equal volume “stopping buffer for analytical labelling”. Samples were immediately processed.

2.3.3 Labelling of proteins with fluorescent Cy5 Dye Cy3 for the preparative gel

The labelling was performed according to the manufacturer’s protocol. The sample of 400 µg GV oocyte proteins (corresponding to 4444 GV oocytes) was diluted with DIGE lysis buffer to a protein concentration of 2 mg/ml. Cystein residues were reduced with 10 µl of a 25 mM TCEP solution. The sample was incubated at 37 °C for one hour in the dark. A total of 400 nmol Cy5 Dye Cy3 was reconstituted in DMF to a 20 mM solution and added to the samples in a concentration of 1 nmol / µg. Samples were incubated at 37 °C for 30 minutes in the dark. The labelling reaction was stopped by diluting Samples with 165 µl “stopping buffer for preparative labelling”. Samples were immediately processed.

2.3.4 2D polyacrylamide SDS gel electrophoresis

For analytical gels IEF Immobiline DryStrips pH 4–7, 24 cm were rehydrated overnight in 450 µl rehydration buffer using a reswelling tray. Dry strips were overlaid with DryStrip Cover Fluid to prevent evaporation. Prior to isoelectric focusing (IEF) 0.25 µg of the Cy5 labelled samples were mixed with 0.25 µg Cy3 labelled IPS. IEF was performed using an Ettan IPGphor. Samples were applied to the first dimension gel using anodic cup loading. Focusing was done for a total of 38.25 kVh. Prior to SDS-PAGE, IEF strips were equilibrated for ten minutes in 15 µl equilibration buffer containing 200 µl saturated bromphenol blue (BPB) solution.

For the preparative gel IEF Immobiline DryStrips pH 4–7, 24 cm were rehydrated overnight with the labelled GV oocyte sample as produced in chapter 2.3.3. Focusing was done for a total

of 56.5 kVh. Prior to SDS-PAGE, IEF strips were equilibrated for ten minutes in 15 μ l equilibration buffer containing 200 μ l saturated BPB solution.

12 % SDS PAA gel solution was prepared and cooled to 8 °C prior to gel casting. Gels were casted in an Ettan Dalt twelve gel caster, overlaid with water and polymerized overnight. The overall gel size was 25.5 cm \times 19 cm \times 0.1 cm.

IPG strips were loaded onto the second dimension gels and overlaid with 0.5 % w/v agarose in SDS running buffer. Electrophoresis was performed using an Ettan Dalt six electrophoresis unit. Gels were run in groups of six with 10 mA per gel for one hour and with 40 mA per gel for six hours in 1 x SDS running buffer for the anodic chamber and 2 x SDS running buffer for the cathodic chamber.

2.3.5 Scanning and evaluation of 2D saturation DIGE gels

Gels were scanned using a Typhoon 9400 fluorescence scanner with a photomultiplier voltage of 670 V. Pixel size was set to 100 μ m. Images were visualized by scanning with a wavelength of 584 nm (Cy3) and 684 nm (Cy5) respectively.

Images were manually cropped to the same size by the image editor software prior to the evaluation by the De Cyder 6.5 software. Spots were detected with the software's "**differential in gel analyses**" (DIA) tool. The spot detection parameter "estimated number of spots" was set to 10,000 and spots with a volume lower than 30,000 were excluded. Spot matching and inter-gel intensity comparisons and the statistical analyses were performed with the software's "**biological variation analysis**" (BVA) tool. Spots which met the criteria of a p-value of lower than 0.05 in the student's t-test while the software's "FDR correction" feature is applied and which were moreover detected within all gel images were considered as spots from differentially abundant proteins.

2.3.6 Identification of spots from differentially abundant proteins

The preparative gel was scanned by the Typhoon 9400 fluorescence scanner and the fluorescence image was imported to the SPControl software used for set up of the spot picking robot. Spots from differentially abundant proteins were manually assigned and picked by the spot picker. The gel was rescanned by the Typhoon 9400 fluorescence scanner and controlled for the correct out-cuts of spots.

The spots were washed four times in 50mM NH_4HCO_3 . Tryptic hydrolysis was performed overnight at 37 °C in 20 μ l 50 mM NH_4HCO_3 with 100 ng modified porcine trypsin per spot.

The supernatant was collected and preserved. The peptides were further extracted by a ten minute treatment with 40 μ l 70 % ACN. The NH_4HCO_3 and ACN fractions were pooled and dried using a vacuum centrifuge. Samples were stored at -80°C . Prior to LC-MS/MS analysis, the peptides were solved in 0.1 % FA.

The Peptides were separated by a Reversed Phase column at a flow rate of 260 nl/min with the following consecutive gradients: I) 0–60 % B in 30 min, (II) 100 % B for 10 min (A: 0.1 % FA, B: 84 % ACN and 0.1 % FA).

Mass spectrometry was performed on an LTQ Orbitrap XL mass spectrometer (mass spectrometer II) online coupled to the nano-LC system (nano HPLC I) For electrospray ionization a “distal coated silica tip” and a needle voltage of 1.4 kV was used. The MS method consisted of cycles of one full MS scan (Mass range: 300–1800 m/z) and five data dependent MS/MS scans. MS/MS spectra were generated in the CID cell of the mass spectrometer with 35 % collision energy. Dynamic exclusion was set to 180s.

2.3.7 Comparative David gene ontology analysis of the combined data from differentially abundant proteins of morulae and blastocysts

The GO analysis was performed by the open access David software [122]. The GO clustering was performed according to the GOTERM_BP_FAT categorization. The threshold for the minimum numbers of proteins in a cluster was set to 2 and the maximum EASE score was set to 0.01.

2.4 Buffers, media and solutions

2.4.1 Buffers and media for sample generation

Buffer / Medium	Composition
Fertilization medium	50 µg Heparin sodium salt 110 µg Sodium Pyruvate 60 mg BSA in 10 ml TL Fertilisationsmedium
Oocyte maturation medium	5 % (v/v) Oestrus cow serum 50 µg FSH 50 µg LH in 10 ml TCM 199 Reifungsmedium
Phosphate buffered saline (PBS)	136.9 mM NaCl 2.7 mM KCl 6.5 mM Na ₂ HPO ₄ 1.5 mM KH ₂ PO ₄
Sperm Talp medium	1.1 mg Sodium Pyruvate 60 mg BSA in 10 ml TL Medium für Spermienkapazitation
Synthetic oviduct fluid (SOF)	1 % (v/v) MEM Non-essential amino acids solution 4 % (v/v) BME Amino acids solution 5 % (v/v) Oestrus cow serum 3.6 mg Sodium Pyruvate in 10 ml SOF Kulturmedium

2.4.2 Buffers and solutions for analyses

Buffer / Solution	Composition
8 M Urea buffer	8.0 M Urea 0.4 M NH ₄ HCO ₃
Agarose overlay solution	0.5 % (w/v) Agarose in SDS running buffer
Coomassie destaining solution	5% (v/v) Methanol 7% (v/v) Acetic Acid
Coomassie staining solution	50 % (v/v) Methanol 10 % (v/v) Acetic Acid 0.05 % (w/v) Coomassie R-250
DIGE lysis buffer	30 mM Tris 7 M Urea 2 M Thiourea 4% (w/v) CHAPS adjusted to pH 8.0 with HCl
Equilibration buffer	6 M Urea 30 % (v/v) Glycerin 2 % (w/v) SDS 0.1 M Tris-HCl pH 8.0
Gel solution for 1D gel electrophoresis (12% gel)	0.375 M Tris-HCl pH 8.8 12.0 % (w/v) Acrylamide/Bisacrylamide 0.1% (w/v) SDS 0.1% (w/v) APS 0.05% (v/v) TEMED

Buffer / Solution	Composition
Gel solution for 1D gel electrophoresis (6% gel)	0.375 M Tris-HCl pH 8.8 6.0 % (w/v) Acrylamide/Bisacrylamide 0.1 % (w/v) SDS 0.1 % (w/v) APS 0.05 % TEMED
Gel solution for 1D gel electrophoresis stacking gel (4% gel)	0.125 mM Tris-HCl pH 6.8 4.0 % (w/v) Acrylamide/Bisacrylamide 0.1 % (w/v) SDS 0.05 % (w/v) APS 0.10 % TEMED
Gel solution for 2D gel electrophoresis (12.5 % gel)	0.375 M Tris-HCl pH 8.8 12.5 % Acrylamide/Bisacrylamide 0.1% (w/v) SDS 0.1% (w/v) APS 0.014% TEMED
Laemmli buffer	62.5 mM Tris pH 6,8 10% (v/v) Glycerin 2% (w/v) SDS 5% (w/v) β -Mercaptoethanol
Rehydration buffer	13 mM DTT 7 M Urea 2 M Thiourea 4% (w/v) CHAPS 1% (v/v) Pharmalytes
SDS running buffer	25 mM Tris 192 mM Glycin 0,1 % w/v SDS
Stopping buffer for analytical labelling	7 M Urea 2 M Thiourea 4 % (w/v) CHAPS 2 % (v/v) Pharmalytes 130mM DTT
Stopping buffer for preparative labelling	361mM DTT 7 M Urea 2 M Thiourea 4 % (w/v) CHAPS 3 % (v/v) Pharmalytes

2.5 Chemicals, kits and prefabricated buffers

Substance	Manufacturer
3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS)	Sigma-Aldrich, Deisenhofen
Acetic Acid	Roth, Karlsruhe
Acetonitril (ACN)	Merck, Darmstadt
Acrylamide/bisacrylamide solution (37,5:1; 30% w/v)	Serva, Heidelberg
Agarose	Serva, Heidelberg
Ammonium chloride (NH ₄ Cl)	Merck, Darmstadt
Ammonium hydrogen carbonate (NH ₄ HCO ₃)	Riedel-de Haën, Seelze
Ammonium peroxide sulfate (APS)	Merck, Darmstadt
Bovine serum albumin (BSA)	Sigma-Aldrich, Deisenhofen
Bradford assay Kit	Thermo Scientific, Dreieich
Bromophenolblue (BPB)	Sigma-Aldrich, Deisenhofen
Calcium chloride (CaCl ₂)	Roth, Karlsruhe
Chromatography water	Merck, Darmstadt
Coomassie Brilliant Blue R250	Sigma-Aldrich, Deisenhofen
CyDye DIGE Fluor, Cy3 saturation dye	GE Healthcare, Freiburg
CyDye DIGE Fluor, Cy5 saturation dye	GE Healthcare, Freiburg
Disodium hydrogen phosphate (Na ₂ HPO ₄)	Merck, Darmstadt
Dissolution buffer	AB SCIEX, Darmstadt
Dithioerythritol (DTT)	Roth, Karlsruhe
Ethanol	Agentur für Rohstoffe, Hohenzell
Follicle stimulating hormone (FSH)	Sioux Biochemical inc., USA
Follicle stimulating hormone (FSH) for injection (Pluset®)	Pharmanovo, Spain
Formaldehyde (37%)	Sigma-Aldrich, Deisenhofen
Formic acid (FA)	Sigma-Aldrich, Deisenhofen
Glycerin	Roth, Karlsruhe

Substance	Manufacturer
Glycin	Roth, Karlsruhe
GnRH analogue (Receptal®)	Intervet, Unterschleißheim
Heparin sodium salt	Sigma-Aldrich, Deisenhofen
High purity ethanol	AB SCIEX, Darmstadt
Hydrogen chloride (HCl)	Roth, Karlsruhe
Iodacetamide	Sigma-Aldrich, Deisenhofen
iTRAQ™ Reagents	AB SCIEX, Darmstadt
Luteinizing hormone (LH)	Sioux Biochemical inc., USA
Magnesium chloride (MgCl ₂)	Merck, Darmstadt
Methanol	Agentur für Rohstoffe, Hohenzell
Modified porcine trypsin	Promega, Wisconsin, USA
N,N,N,N-Tetramethylethyldiamine (TEMED)	Roth, Karlsruhe
N,N-Dimethylformamide (DMF)	Sigma-Aldrich, Deisenhofen
Pharmalytes	GE Healthcare, Freiburg
Polyvinylpyrrolidon (PVP)	Sigma-Aldrich, Deisenhofen
Potassium chloride (KCl)	Merck, Darmstadt
Potassium dihydrogen phosphate (KH ₂ PO ₄)	Merck, Darmstadt
Sigma Wide Range marker	Sigma-Aldrich, Deisenhofen
Sodium chloride (NaCl)	Merck, Darmstadt
Sodium Dodecyl Sulfate (SDS)	Serva, Heidelberg
Sodium Pyruvate	Sigma-Aldrich, Deisenhofen
β-Mercaptoethanol	Sigma-Aldrich, Deisenhofen
Thiourea	Sigma-Aldrich, Deisenhofen

Substance	Manufacturer
Tetramethylethylenediamine (TEMED)	Roth, Karlsruhe
Tris	Roth, Karlsruhe
Tris-(2-carboxyethyl) phosphine hydrochloride (TCEP)	Sigma-Aldrich, Deisenhofen
Urea	Roth, Karlsruhe

2.6 Software

2.6.1 Analyses software

Name	Function	Manufacturer
DAVID	Gene Ontology analysis Pathway analysis	European Bioinformatics Institute (EBI), Swiss Institute of Bioinformatics (SIB) und Protein Information Resource (PIR)
DeCyder 6.5	Analysis of 2D DIGE gels	GE Healthcare, Freiburg
Image editor	Cropping of gel images	GE Healthcare, Freiburg
MASCOT Search	Database search for mass spectrometry	Matrix Science, Boston (MA), USA
Microsoft Office Access	Protein list comparison	Microsoft, Unterschleißheim
Microsoft Office Excel	Statistical analyses	Microsoft, Unterschleißheim
Proteome Discoverer 1.1	Database search for generating exclusion lists in mass spectrometry	Thermo Scientific, Dreieich
R	Statistical analyses	The R Foundation for Statistical Computing
Scaffold_3_00_06	Statistical evaluation of protein list; Calculation of protein fold changes	Proteome Software, Portland (OR), USA
UniProt Gene Ontology / UniProt KB	Gene Ontology analysis; Transformation of protein IDs	European Bioinformatics Institute (EBI), Swiss Institute of Bioinformatics (SIB) und Protein Information Resource (PIR)

2.6.2 Instrument specific software

Name	Operated Instrument	Manufacturer
Analyst 1.5.1	QTRAP 5500	AB SCIEX, Darmstadt
SPControl 3.1	PROTEINEER spli	Bruker Daltonics, Bremen
Typhoon Scanner Control, ImageQuant 5.2	Typhoon 9400	GE Healthcare, Freiburg
UNICORN 5.01	Ettan MDLC	Amersham Biosciences, Freiburg
Xcalibur (inc. Dynamic exclusion software)	Finnigan LTQ	Thermo Scientific, Dreieich
Xcalibur (inc. Dynamic exclusion software)	LTQ Orbitrap XL	Thermo Scientific, Dreieich

2.7 Instruments

Instrument	Name	Manufacturer
2D electrophoresis gel caster	Ettan DALTwelve gel caster	GE Healthcare, Freiburg
2D electrophoresis running chamber	Ettan DALTsix electrophoresis unit	GE Healthcare, Freiburg
2D gel glass spacer	no name provided	Glaserei Brendle, Neuaubing
2D gel glass plates	no name provided	Glaserei Brendle, Neuaubing
Blank Cassette inserts	Blank Cassette Inserts	GE Healthcare, Freiburg
Incubator for Digestion	Modell 400	Memmert, Schwabach
1D electrophoresis gel caster	Mini-PROTEAN 2	BioRad, München
1D electrophoresis running chamber	Mini-PROTEAN 2	BioRad, München
1D gel glass plates	glass plates	GE Healthcare, Freiburg
1D gel spacer	Plastic spacer	BioRad, München
Special accuracy weighing machine	Sartorius Basic Plus	Sartorius, Göttingen
Fluorescent-Scanner	Typhoon 9400	GE Healthcare, Freiburg
Freezer	HFU 686 Top (-80°C)	Heraeus, Hanau
Isoelectric focussing machine	IPGphor	Pharmacia Biotech, Freiburg
Refrigerator (+4°C)	Premium	Liebherr, Ochsenhausen

Instrument	Name	Manufacturer
magnetic stirrer	MR 2000	Heidolph, Schwabach
Mass spectrometer I	Finnigan LTQ	Thermo Scientific, Dreieich
Mass spectrometer II	LTQ Orbitrap XL	Thermo Scientific, Dreieich
Mass spectrometer III	QTRAP 5500	AB SCIEX, Darmstadt
Microwave	no name provided	Panasonic, Hamburg
Microcentrifuge	EBA 3S	Hettich, Tuttlingen
Nano-HPLC I	Ettan MDLC	Amersham Biosciences, Freiburg
Nano-HPLC II	Nano Acquity	Waters, Eschborn
pH meter	761 Calimatic	Knick, Berlin
Pick-Robot	PROTEINEER spII	Bruker Daltonics, Bremen
Power-Supply for 1D gels	Power Supply Model 200/2.0	BioRad, München
Power-Supply for 2D gels	EPS 3501 XL	Amersham Biosciences, Freiburg
Reswelling Tray	Immobiline Dry Strip Reswelling Tray	GE Healthcare, Freiburg
Shaker	Certomat U	B.Braun, Melsungen
Thermomixer	Thermomixer 5436	Eppendorf, Köln
Cooling unit for 2D gel electrophoresis	no name provided	Eppendorf, Köln
Sonication Cleaner	Sonorex RK 100	Bandelin, Berlin
Vacuum centrifuge	Vacuum Concentrator	Bachofer, Reutlingen
Vortexer	Vortex Genie 2	Bachhofer, Reutlingen
Scale	Sartorius 2357	Sartorius, Göttingen
Centrifuge I	GS-15R Centrifuge	Beckman, Krefeld
Centrifuge II	Heraeus Sepatech Megafuge	Heraeus, Hanau
Transferpettor caps	Transferpettor caps	Brand, Wertheim
Transferpettor	Transferpettor	Brand, Wertheim

2.8 Consumables

Product	Name	Manufacturer
4 well dishes	4 well Multidish	Nunc/Thermo Scientific, Dreieich
Aluminium foil	Alu-Folie	Roth, Karlsruhe
Aspiration pump	Electronic aspiration pump, modell 3014	Labotect, Göttingen
Canules 1.2 x 40mm	Sterican Einmal-Injektionskanülen	B.Braun, Melsungen
Canules, 0.45mm x 25mm	Sterican Einmal-Injektionskanülen	B.Braun, Melsungen
Collection plate	DigestPro MS 96 well collection plates	Intavis, Köln
Electrospray needle	Distal coated Silica Tips	New Objectiv, Woburn (MA), USA
Falcon Tube 15ml and 50ml	Falcon-Röhrchen	TPP, Trasadingen, Schweiz
Filter paper	Whatman Filterpapier	Whatman / Schleicher & Schuell, Dassel
IPG-strips for IEF	Immobiline DryStrips, pH 4-7, 24cm	GE Healthcare, Freiburg
LC Glass vials I	Glass vial, 1,5ml	Amersham Biosciences, Freiburg
LC Glass vials II	12x32 mm glass screw neck vial	Waters, Eschborn
Mineral oil for embryo culture	Mineral oil, embryo culture tested	Sigma
Oil for IEF	DryStrip Cover Fluid	GE Healthcare, Freiburg
Petri dishes 94mm	Petri dishes	Roth, Karlsruhe
Pipette tips (0,5 – 10 µl)	DIAMOND 10µl	Gilson, Villiers Le Bel, Frankreich
Pipette tips (2 – 200 µl)	PLASTIBRAND 200µl	Brand, Wertheim
Pipette tips (50 – 1000 µl)	PLASTIBRAND 1000µl	Brand, Wertheim
Plastic wrap	Saran	Dow, Schwalbach
Progesterone releasing device	Intravaginal progesterone-releasing device (CIDR®)	Pfizer, Berlin
Reversed Phase Column I	C18 PepMap100	LC Packings Dionex, Idstein
Reversed Phase Column II	Reposil-Pur-C18-AQ	Dr. Maisch GmbH, Ammerbuch-Entringen
Sample Cups (0.5 ml, 1.5 ml and 2 ml)	Safe-Lock Tubes	Eppendorf, Köln
Sample Cups for IEF	IPGphore Sample cups	GE Healthcare, Freiburg
SCX Column	Biobasic	Thermo Scientific, Dreieich

Product	Name	Manufacturer
Syringes, 10 ml	Omnifix 10ml	B.Braun, Melsungen
Trapping Column	C18 PepMap100	LC Packings Dionex, Idstein
Vivapure S mini M clean up devices	Vivapure S mini M clean up devices	Vivascience AG, Hannover
Wash dishes (40 mm x 12 mm)	40 mm tissue culture dish	Nunc/Thermo Scientific, Dreieich
Zip tips	Zip Tips μ -C18	Millipore, Schwalbach

3 Results

3.1 Preparation of GV oocytes

Several experiments required the availability of germinal vesicle (GV) oocytes, which were obtained by the aspiration of cumulus oocyte complexes (COCs) from slaughterhouse ovaries. One ovary delivered on average 10 COCs each. All oocytes were denuded and washed. In total, 5900 GV oocytes were prepared from 590 ovaries. Aliquots comprising 25 to 400 oocytes each were stored in 0.5 ml sample cups at $-80\text{ }^{\circ}\text{C}$. These samples are referred to as “denuded GV oocytes” throughout this work.

3.2 Qualitative proteome profile of bovine GV oocytes

For a qualitative proteome profile, 900 denuded GV oocytes were lysed. Half of the protein lysate (40.5 μg total protein), was prefractionated on a 12 % 1D SDS polyacrylamide (PAA) gel, while the other half was prefractionated on a 6 % 1D SDS PAA gel. Both gels were stained with coomassie. The 12 % 1D SDS PAA gel was sliced into 23 pieces and the 6 % SDS gel into 14 slices, as indicated in Fig. 5. Each slice was individually subjected to in-gel trypsinization and analysed separately by nano LC-MS/MS using an LTQ mass spectrometer.

In total 1,923,060 MS/MS spectra were acquired and subjected to Mascot data processing using the bovine IPI database. Mascot data were statistically analysed using the Scaffold software tool. Applying a protein identification probability of $> 99\%$ on the protein level and a minimum number of two different identified peptides per protein ($> 95\%$ peptide identification probability), 20,053 spectra could be assigned to 4,229 different peptides, corresponding to 791 different proteins from the 12 % gel. From the well separated high molecular mass region of the 6 % gel, 38 additional proteins could additionally be identified from 335 peptides (2510 additional spectra), leading to 829 identified proteins in total.

A decoy database search led to no hits in the decoy sequences, demonstrating a false discovery rate (FDR) far below 1 %. The 829 IPI accession numbers of identified proteins were transformed into accession numbers provided by the Uniprot protein database. In the Uniprot database, five levels of evidence for the existence of a protein are indicated: (i) “Evidence at protein level” applies to proteins that have been previously identified on the protein level, (ii) “Evidence at transcript level” applies to proteins of which the corresponding mRNA has been identified, (iii) “Inferred from homology” applies to proteins of which the existence is probable because clear orthologs exist in closely related species, and (iv) “Predicted” applies to proteins which have been predicted by DNA analyses, yet for which no evidence at protein, transcript, or

homology levels exists. The weakest term is (v) “Uncertain”. In total 739 of the 829 identified proteins could be transformed into Uniprot accession numbers. From these 739 proteins, 118 were classified as proteins with experimental evidence on the protein level, 514 as proteins with experimental evidence on the RNA level, 83 as entries inferred from homology from sequences of other species and 24 as predicted.

In addition, the entire set of proteins was subjected to Uniprot gene ontology analysis, the result of which is shown in Fig. 6. Proteins were assigned to the three main categories “cellular component”, “biological process” and “molecular function”. Within the “biological process” category, most proteins (47 %) are assigned to the cluster “metabolism”. In the “molecular function” category most proteins are assigned to the cluster “binding” (51 %) and to the cluster “enzymes” (41 %). Further analysis of the catalytic activity cluster revealed that most of the enzymes could be assigned to oxidoreductases (77) and hydrolases (108).

To find out in which pathways the identified proteins are involved, a so-called “David analysis” of KEGG Pathways was performed. This bioinformatics tool reveals pathways in which proteins from the input list show up more than they are expected to appear from a randomized protein list [123]. A total of 23 proteins were found in the KEGG Pathway term “Ribosome”, while 16 proteins belonged to the “oocyte meiosis pathway” (Fig. 7).

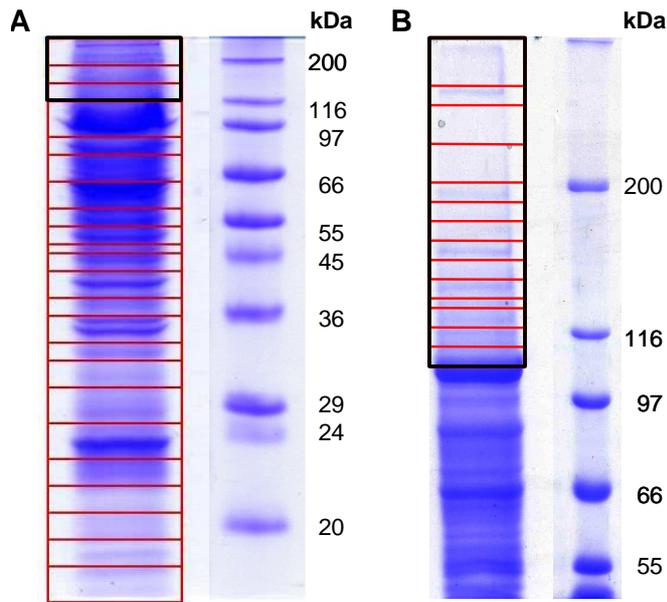


Fig. 5: 1 D SDS PAA gels for the prefractionation of a protein lysate from 450 denuded GV oocytes

(A) 12 % 1D SDS-PAA gel: left lane: protein lysate of 450 bovine GV oocytes, right lane: Sigma Wide Range Marker; (B) 6 % 1D SDS gel: left lane protein lysate of 450 bovine GV oocytes, right lane: Sigma Wide Range Marker; Red brackets: Indicate the fractions subjected to LC-MS/MS analysis; Bold rectangle: Indicates high molecular mass regions on the 6 % and 12 % gel, respectively.

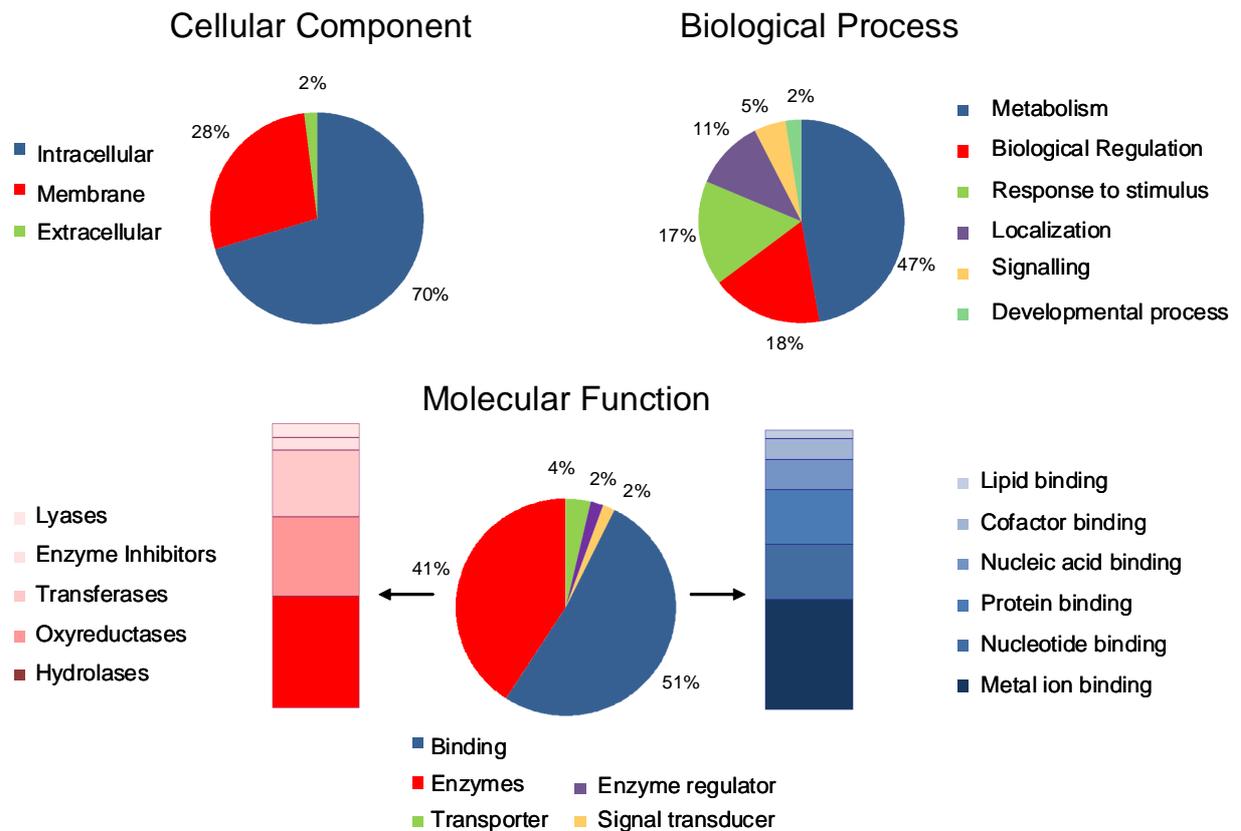


Fig. 6: Uniprot gene ontology analysis of 739 identified GV oocytes proteins

The “cellular component“, the “biological process” and the “molecular function” categories of the GO database from the Gene Ontology project, last updated on December 9, 2008, were used.

3.3 Protein identification from limited sample amounts

The identification of around 800 proteins usually requires protein amounts of about 40 to 100 µg. Since one oocyte or early embryo only contains approximately 90 ng protein, such an identification approach would require the generation of hundreds of embryos. Especially in cases where several biological replicates are needed, the preparation of so many embryos would be very cost intensive and time consuming. For this reason it was tested how many proteins can be identified from much smaller sample amounts corresponding to the protein amounts available from one oocyte up to 50 oocytes.

The protein amounts of one, five, ten, 25 and 50 denuded GV oocytes were analysed in a 1D LC-MS/MS approach. Proteins were accepted as correctly identified only when they had been identified with at least two different peptides, which were detected within a 10 ppm mass accuracy range, and had a 95 % peptide probability as well as a 99 % protein probability.

A total of 54 proteins could be identified from a single oocyte, 75 proteins from 5 oocytes, 99 proteins from ten oocytes and 106 proteins from 25 oocytes. In contrast, the protein amount from 50 oocytes dropped the number of identified proteins down to 62 (Fig. 8). To find out whether the number of identifications from protein amounts corresponding to 25 and 50 oocytes can be further enhanced, a two-dimensional LC-MS/MS analysis was performed by an additional prefractionation step of tryptic peptides on a “strong cation exchange” (SCX) column prior to the reversed phase (RP) separation. SCX chromatography uses the charge state of a molecule as a major separation criterion and represents an orthogonal separation step to RP chromatography. The “base peak ion chromatograms” (BICs) in Fig. 9 demonstrate that this prefractionation method led to an equal distribution of peptides over all salt fractions.

Using the 2D LC-MS/MS approach, 206 proteins could be identified from the protein amount of 25 oocytes. The protein amount of 50 oocytes did not lead to more identifications; instead, the number of identified proteins was dropped down to 198 (Fig. 8).

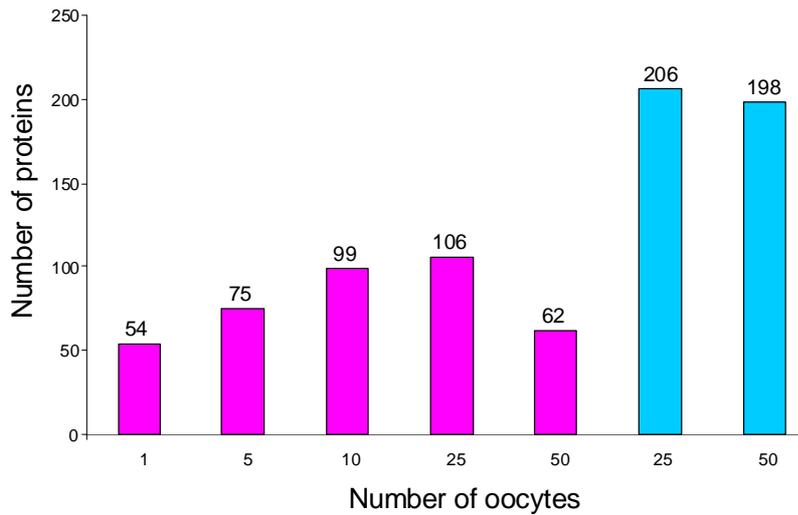


Fig. 8: Comparison of protein numbers identified from 1, 5, 10, 25 and 50 GV oocytes

X-axis: numbers of oocytes subjected to LC-MS/MS analysis, Y-axis: numbers of identified proteins (FDR < 1 %); Magenta: proteins identified by 1D LC-MS/MS analysis, Blue: proteins identified by 2D LC-MS/MS analysis.

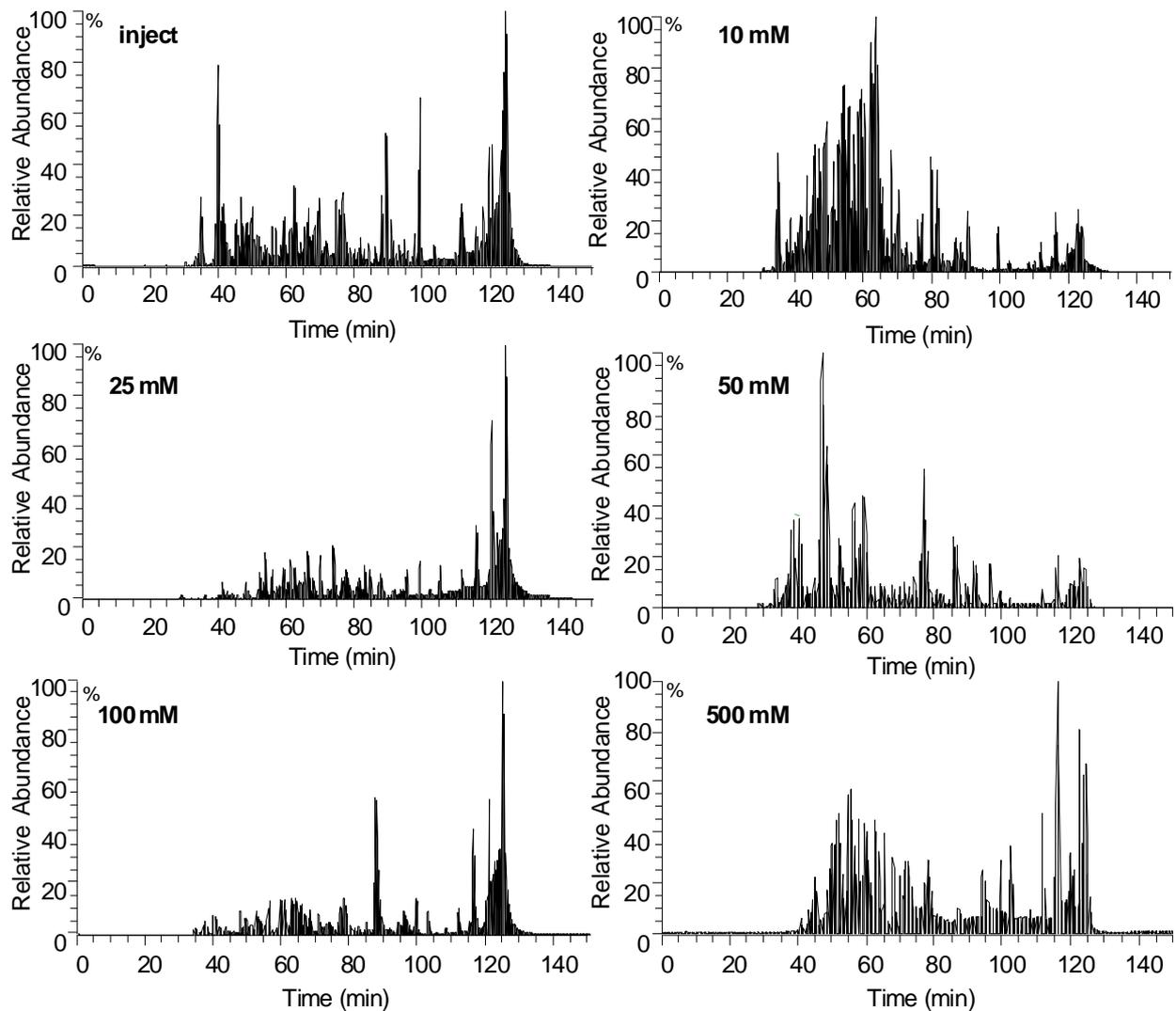


Fig. 9: Base Peak Ion Chromatograms (BICs) of individual salt steps from the analysis of 25 GV oocytes by 2D LC-MS/MS

X-axis: elution time; Y-axis: relative signal intensity.

3.4 Protein quantification of limited sample amounts by iTRAQ

A standard proteomic approach for quantification is to employ the “isobaric tag for relative and absolute quantitation” (iTRAQ) technique, which was applied for the quantitative analyses of small sample amounts. For quantification, it is necessary to obtain iTRAQ reporter ion signals generated by the “higher energy collision dissociation” (HCD) cell of the mass spectrometer. Due to longer scan times needed to acquire the HCD spectra, fewer spectra are acquired during an individual LC-MS/MS run. As a consequence, fewer proteins are expected to be quantified and identified during an individual LC run as compared to LC runs dedicated to identification only (see previous chapter). Therefore, an experimental series was performed to optimise the number of quantified proteins for a given amount of oocytes, based on sample splitting and the establishment of several consecutive “precursor mass exclusion lists”.

In total 350 oocytes from the “denuded GV oocyte pool” were lysed, tryptically digested and cleaned up by “vivapure S mini M” devices. From this pool containing 31.5 μg peptides, two aliquots were prepared and labelled with iTRAQ reagent containing either the 115 reporter tag or the 117 reporter tag. After labelling, the samples were pooled and served as a source for optimisation experiments. The pool is referred to as “iTRAQ labelled GV oocyte peptides” and contains 31.5 μg protein (corresponding to 350 oocytes). Protein quantification was based on reporter ion signal intensities in HCD spectra, and protein identification was based on both HCD and CID spectra derived from the Orbitrap XL instrument. As an example, Fig. 10 shows a typical HCD MS/MS spectrum of the peptide LSDGVALVK contained in “Heat shock 60 kDa protein”. The signal intensity ratio of reporter ions 115 and 117 is close to one, confirming the presence of equal amounts of 115 tag and 117 tag peptides in the “iTRAQ labelled GV oocyte peptides”.

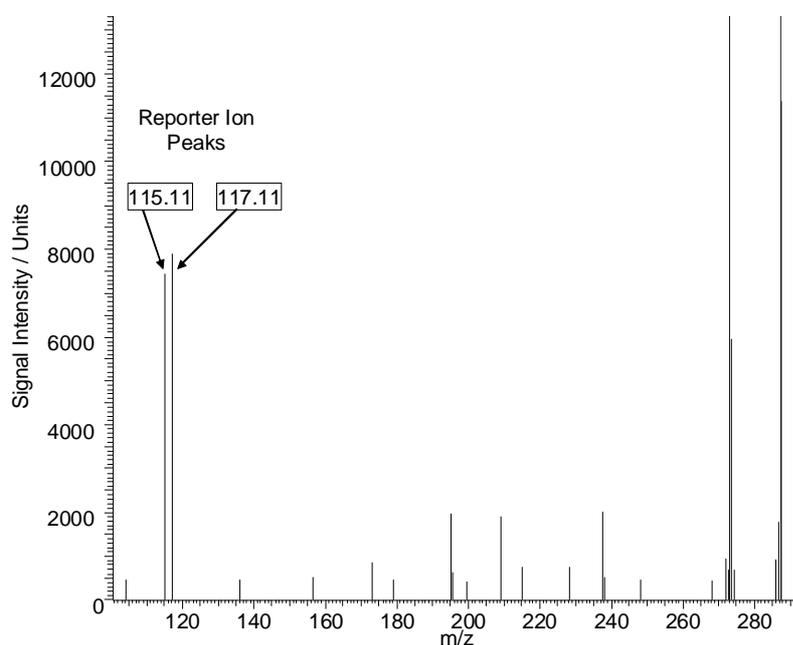


Fig. 10: HCD MS/MS spectrum of the Peptide LSDGVALVK derived from Heat shock 60 kDa Protein 1
X-axis: m/z values; Y-axis: signal intensity.

To find out how many proteins can be quantified and identified from small protein amounts, a 1D LC-MS/MS analysis was performed using the protein amount corresponding to 25 and 50 oocytes (2.25 μg respectively 4.5 μg) from the “iTRAQ labelled GV oocyte peptides”. From the protein equivalent of 25 and 50 oocytes a total of 67 respectively 86 proteins could be identified and quantified (green columns in Fig. 11). In order to enhance the number of identified and quantified proteins, the iTRAQ labelled proteins from 25 and from 50 oocytes were split up into two equal parts each. In both cases, after the first aliquot had been analysed by LC-MS/MS, an exclusion list (EL) was generated containing all peptide masses from which MS/MS spectra had been successfully acquired. This EL was applied to the analyses of the second aliquots, forcing the instrument to ignore these masses and instead acquire MS/MS spectra from so far unconsidered peptide precursors. The application of the EL enhanced the number of identified and quantified proteins from the protein amount of 25 oocytes to 107, and from the protein amount of 50 oocytes to 116 (blue columns in Fig. 11).

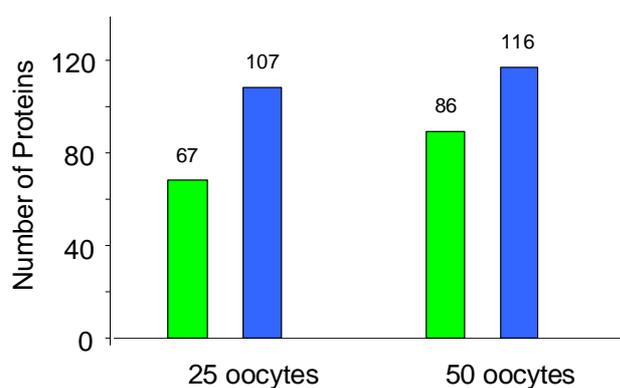


Fig. 11: Numbers of proteins identified and quantified by iTRAQ 1D nano LC-MS/MS analysis from 25 and 50 GV oocytes

X-axis: numbers of oocytes subjected to LC-MS/MS analysis, Y-axis: number of identified and quantified proteins (FDR<1 %); Green columns: no exclusion list applied, blue columns: exclusion list applied.

To find out whether the number of identified and quantified proteins can be further enhanced, the amount of “iTRAQ labelled GV oocyte peptides”, from 50 oocytes was divided into four equal parts. They were then analysed in four consecutive 1D LC-MS/MS runs and an exclusion list for each run was generated from all previous runs. Such a combination of four runs will be referred to as “run set” throughout this work. Fig. 12 shows the number of identified and quantified proteins from the merge of individual runs with all previous runs. As a maximum, 150 proteins were identified and quantified from the protein amount of 50 oocytes.

The same set up was also tested in a 2 D LC-MS/MS approach. Exclusion lists were generated for each run of the six salt fractions and applied to the analysis of the corresponding fraction in the following run. The protein equivalent of 50 oocytes from the “iTRAQ labelled GV oocyte peptides” was split into two equal parts which were subsequently analysed, while the six exclusion lists from the first run were applied to the second run.

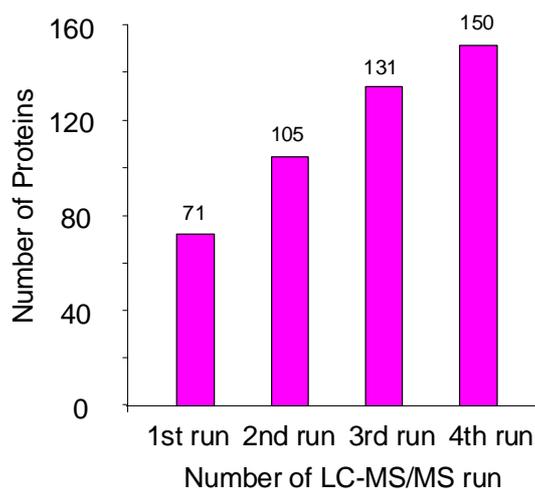


Fig. 12: Total numbers of identified and quantified proteins after each 1D LC-MS/MS run of a “run set”

X-axis: number of nano-LCMS/MS run, Y-axis: number of identified and quantified proteins (FDR < 1 %) from the merge of individual runs with all previous runs; in total 50 GV oocytes were analysed.

In total 300 proteins could be identified and quantified. To find out whether also this number can be further enhanced the amount of “iTRAQ labelled GV oocyte peptides” from 50 oocytes was divided into four equal parts and analysed in four consecutive 2D LC-MS/MS runs. For each run, six ELs (one per salt fraction) were generated from all previous runs and applied to the analyses of the corresponding fractions in the following run. This enhanced the number of identified and quantified proteins to 403. Since the analysis of one 2D LC-MS/MS run took 18 hours, only one 2D LC-MS/MS run could be performed per day. The analysis time for a complete 2D LC-MS/MS run set therefore comprised four days. Fig. 13 demonstrates how many proteins can be identified and quantified from 50 GV oocytes using the different methods of spectra acquisition and chromatographic separation.

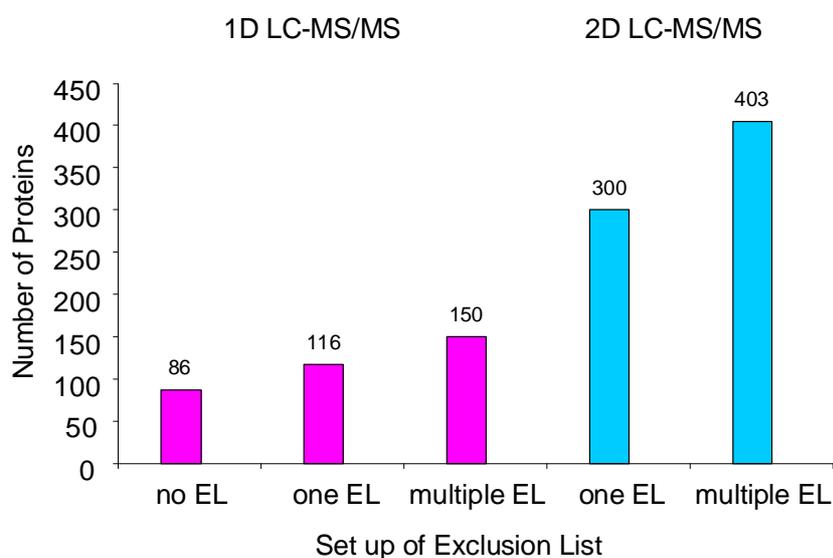


Fig. 13: Numbers of identified and quantified proteins from the protein amount of 50 oocytes using different nano LC-MS/MS approaches

X-axis: indicates whether no EL (no EL), one EL (one EL) or multiple ELs (multiple EL) were applied for nano LC-MS/MS analysis; Multiple EL: the application of three subsequent ELs in four consecutive nano LC-MS/MS runs (referred to as “run set”), Y-axis: numbers of identified and quantified proteins (FDR < 1 %); Magenta columns: 1D LC-MS/MS approach, Blue columns: 2D LC-MS/MS approach.

3.5 Comparison of *in vivo* and *in vitro* matured oocytes from cows of different age groups by nano LC-MS/MS iTRAQ analysis.

The *in vitro* maturation process of oocytes is known to produce oocytes of lower developmental competence compared to *in vivo* matured oocytes [115]. Therefore, proteome alterations between *in vivo* and *in vitro* matured oocytes were analysed by the nano LC-MS/MS iTRAQ technique.

COCs were obtained by aspiration from ovarian follicles from heifers (10-12 months) as well as from young cows (three years, first lactation) and old cows (> ten years) with and without hormonal superstimulation by FSH (referred to as “stimulated” and “unstimulated”). Each age group contained five cows. The aspiration procedure was performed by the group of Prof. Wolf at the Moorversuchgut Badersfeld and is referred to as “OPU (ovum pick-up) session” throughout this work.

From seven different OPU sessions, 95 GV COCs from unstimulated heifers, 120 GV COCs from unstimulated young cows and 179 GV COCs from unstimulated old cows were obtained. All GV COCs from unstimulated cows were matured *in vitro* for 22 hours to the MII oocyte stage. Oocytes were denuded and preserved at – 80 °C in aliquots of 7 to 39 oocytes.

From superstimulated cows, MII oocytes were collected in four different OPU sessions. In total 123 MII COCs from superstimulated heifers, 89 MII COCs from superstimulated young cows and 144 MII COCs from superstimulated old cows were obtained. All MII COCs from hormonally superstimulated cows were denuded immediately after aspiration and preserved at - 80 °C in aliquots of 14 to 47 oocytes.

A “biological replicate” within this chapter is defined as six pools containing 25 oocytes from each MII oocyte sample obtained from individual OPU sessions.

For nano LC-MS/MS analysis, each sample of 25 MII oocytes was lysed and divided into two parts, each containing the peptides from 12.5 oocytes. As a reference, a pool of 50 denuded GV oocytes was treated in the same way and divided into four parts, each containing the peptides from 12.5 oocytes.

The labelling was performed according to the scheme in Table 3, which allowed the multiplexed analysis of four different samples per run set. As established before (see chapter 3.4) each run set consisted of four consecutive 2D LC-MS/MS runs. Samples were combined according to the scheme in Table 3. This labelling and combination strategy facilitated, within an individual run set, the comparison of *in vivo* and *in vitro* matured oocytes obtained from an individual age group, as well as the comparison of oocytes obtained from different age groups. The denuded GV oocytes in each run set were used for standardisation efforts. Four run sets were needed for

the analysis of one biological replicate, consuming 16 days of LC-MS/MS instrument time. The complete analysis of all three biological replicates comprised 48 days.

A total of 74 proteins were identified and quantified from the three run sets no. one, 82 proteins from the three run sets no. two, 53 proteins from the three runs sets no. three and 116 proteins from the three run sets no. four in all biological replicates (Table 3).

Table 3: Labelling and run set scheme for one biological replicate of *in vivo* and *in vitro* matured oocytes analysed by nano LC-MS/MS iTRAQ analysis

Numbers 1-4 indicate the run sets; Numbers 114-117 indicate the iTRAQ reporter tag masses attached to a sample

Number of run set	Old cows / <i>in vitro</i> matured	Young cows / <i>in vitro</i> matured	Heifers / <i>in vitro</i> matured	Old cows / <i>in vivo</i> matured	Young cows / <i>in vivo</i> matured	Heifers / <i>in vivo</i> matured	GV Oocytes	Numbers of identified and quantified proteins
1		116	117			115	114	74
2	116	117			115		114	82
3	116		117	115			114	53
4				115	116	117	114	116

The mean values for all fold changes from the three biological replicates were calculated, and proteins which had an average fold change of < 0.5 or > 2 were considered as differentially abundant proteins.

In the comparison of *in vivo* versus *in vitro* matured oocytes, four differentially abundant proteins were found in the age group of young cows (Table 5) and in each age group of old cows and heifers, nine differentially abundant proteins were found (Table 4, Table 6). No differentially abundant proteins were found in the comparison of oocytes from different age groups from *in vivo* matured oocytes. In the comparison of oocytes from different age groups of *in vitro* matured oocytes, four proteins were found to be differentially abundant in oocytes from old cows versus heifers (Table 7). Also four proteins were differentially abundant between *in vitro* matured oocytes from young cows versus heifers (Table 8). In the comparison from *in vitro* matured oocytes from young cows and old cows, no proteins of differential abundance were detected. Fig. 14 shows the mean value of abundance alterations of three proteins altered between *in vivo* and *in vitro* matured oocytes.

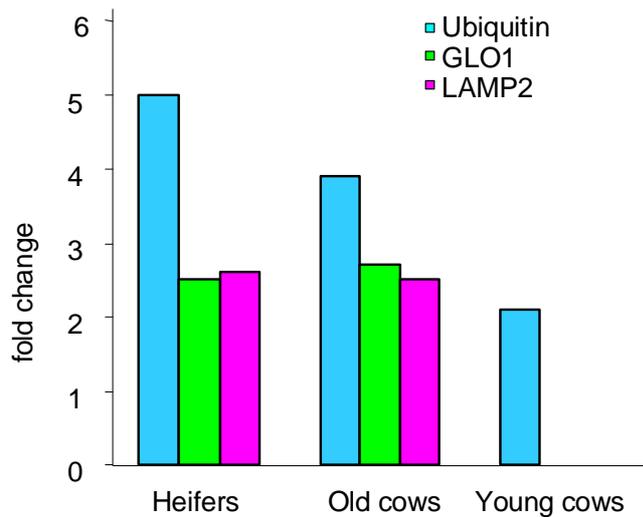


Fig. 14: Fold changes of the abundance of Ubiquitin, GLO1 and LAMP2 *in vivo* versus *in vitro* matured oocytes from different age groups

X-axis: age group of oocyte donor cows, Y-axis: fold change of protein abundance; Blue column: Ubiquitin, Yellow column: Glo1, Magenta column: LAMP2; GLO1 and LAMP2 were not altered in young cows.

Table 4: Differentially abundant proteins between *in vivo* and *in vitro* matured oocytes from heifers

Protein Name	Accession Number	Mean Value of Fold Change
Ubiquitin	IPI00786471	5.0
Pterin-4 alpha-carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor 1 alpha	IPI00689238	3.8
Lysosomal-associated membrane protein 2 isoform 2	IPI00842574	2.6
GLO1 protein	IPI00713484	2.5
Lysosome-associated membrane glycoprotein 1	IPI00705401	2.3
Glyceraldehyde-3-phosphate dehydrogenase	IPI00713814	0.5
22 kDa protein	IPI00702336	0.5
similar to transducin-like enhancer of split 6	IPI00716544	0.5
Bisphosphoglycerate mutase	IPI00706349	0.4

Table 5: Differentially abundant proteins and abundance ratios between *in vivo* and *in vitro* matured oocytes from young cows

Protein Name	Accession Number	Mean Value of Fold Change
Ubiquitin	IPI00786471	2.1
Adenosine deaminase	IPI00704185	2.0
Bisphosphoglycerate mutase	IPI00706349	0.5
Nucleoside diphosphate kinase B	IPI00686420	0.5

Table 6: Differentially abundant proteins and abundance ratios between *in vivo* and *in vitro* matured oocytes from old cows

Protein Name	Accession Number	Mean Value of Fold Change
Ubiquitin	IPI00786471	3.9
GLO1 protein	IPI00713484	2.7
lysosomal-associated membrane protein 2 isoform 2	IPI00842574	2.5
Phosphatidylethanolamine-binding protein 1	IPI00704735	2.5
Lysosome-associated membrane glycoprotein 1	IPI00705401	2.4
HAPLN3 protein	IPI00699203	2.3
Heat shock 70 kDa protein 1B	IPI00700035	2.1
10 kDa heat shock protein. mitochondrial	IPI00702858	2.0
22 kDa protein	IPI00702336	0.5

Table 7: Differentially abundant proteins and abundance ratios between *in vitro* matured oocytes of old cows and heifers

Protein Name	Accession Number	Mean Value of Fold Change
Lysosomal-associated membrane protein 2 isoform 2	IPI00842574	3.6
Ubiquitin	IPI00786471	2.6
Lysosome-associated membrane glycoprotein 1	IPI00705401	2.2
Phosphatidylethanolamine-binding protein 1	IPI00704735	2.0

Table 8: Differentially abundant proteins and abundance ratios between *in vitro* matured oocytes of young cows and heifers

Protein Name	Accession Number	Mean Value of Fold Change
Lysosomal-associated membrane protein 2 isoform 2	IPI00842574	2.7
Pterin-4 alpha-carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor 1 alpha	IPI00689238	2.3
HAPLN3 protein	IPI00699203	2.2
Isoform 1 of Proactivator polypeptide	IPI00718311	2.2

3.6 Qualitative and quantitative proteome analysis of oocytes, 2-cell stage embryos and morulae

To analyse both the qualitative proteome of 2-cell stage embryos and morulae as well as proteome alterations during the development from the 2-cell stage embryo to the morula stage, qualitative and quantitative analyses based on nano LC-MS/MS were performed.

2-cell-stage embryos and morulae for the proteome analyses were generated from 35 slaughterhouse ovaries. On average, each ovary delivered ten COCs. COCs were matured *in vitro* and fertilized. The developed embryos were denuded and cultivated for nine hours (2-cell stage embryos) or for 5 days (morulae). Prior to collection, embryos were microscopically controlled for their blastomeric structure. Morulae were collected when they had at least 32 blastomeres. Both compacted and not compacted morulae were collected. On average, 90 % of all COCs developed into 2-cell stage embryos and 60 % of all COCs developed into morulae.

For a qualitative nano LC-MS/MS analysis, two pools of 25 2-cell-stage embryos each and two pools of 25 morulae each, representing two technical replicates, were generated. Embryos were lysed and tryptically digested prior to 2D LC-MS/MS analysis. Database search and statistical analysis were performed as described in chapter 2.2.2.4 and in chapter 2.2.2.6. The results from the two technical replicates were merged by the scaffold software. A total of 318 proteins were identified from 2-cell stage embryos and 348 proteins were identified from morulae.

254 of these proteins could be identified both from two cell stage embryos and morulae, while 64 proteins were exclusively identified from 2-cell stage embryos, and 94 proteins were exclusively identified from morulae. The datasets with the identified proteins from these embryonic stages were compared to the dataset of GV oocytes proteins described in chapter 3.2 (Fig. 15). Although the GV oocytes dataset contains far more proteins (829), 52 proteins identified from 2-cell stage embryos and 73 proteins identified from morulae were not contained in the oocyte dataset. All three datasets were subjected individually to a gene ontology analysis. This comparison is shown in Fig. 16. The percentage of proteins which are assigned to the different gene ontology categories revealed to be similar in all three stages.

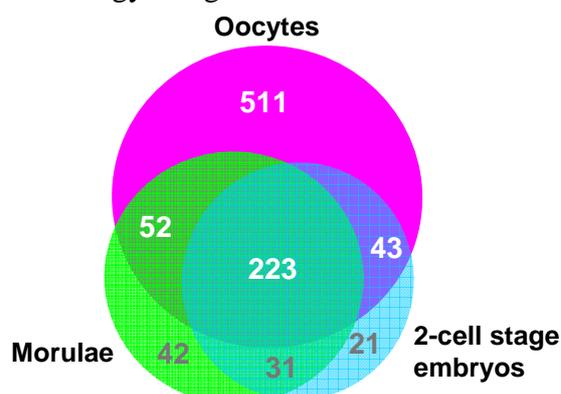


Fig. 15: Proportional Venn diagram of identified proteins from oocytes, 2-cell stage embryos and morulae

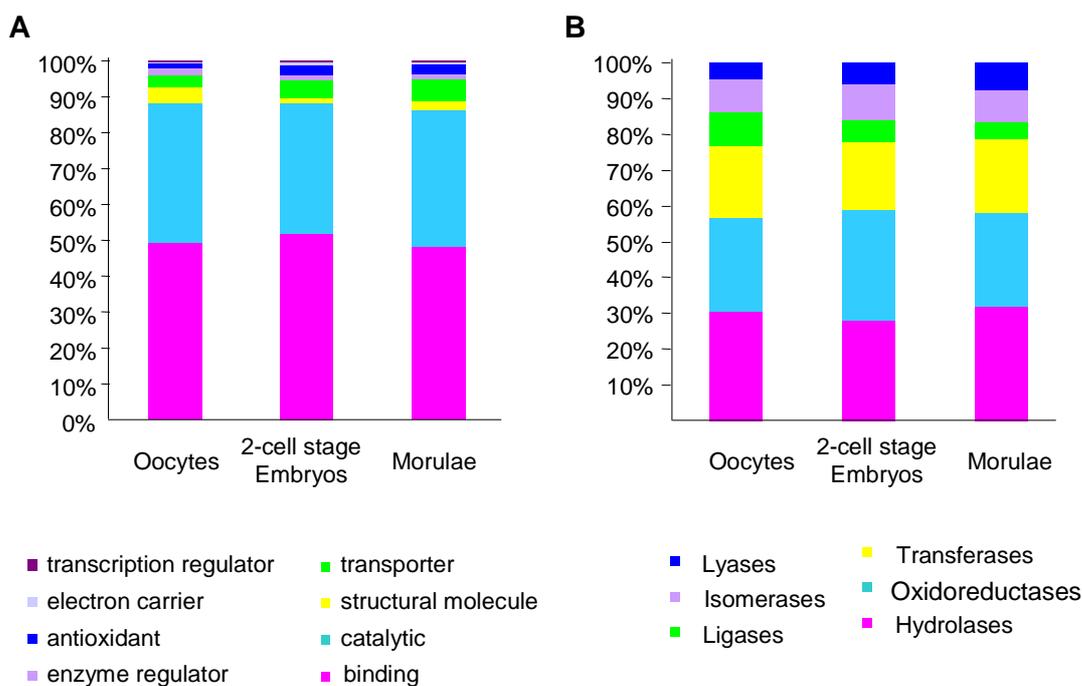


Fig. 16: Uniprot gene ontology analysis of the identified proteins from oocytes, two-cell-stage embryos and morulae

(A) “molecular function” category of the GO database from the Gene Ontology project, last updated on December 9, 2008; X-axis: developmental stage, Y-axis: percentage of proteins assigned to a GO cluster. (B) “catalytic activity” category of the GO database from the Gene Ontology project, last updated on December 9, 2008; X-axis: developmental stage, Y-axis: percentage of proteins assigned to a GO cluster.

The datasets from 2-cell stage embryos and morulae were subjected to a “spectral count” [87] analysis. The method of spectral counting enables quantitative comparisons between samples which have been analysed in two different LC-MS/MS runs (see chapter 1.5.2). The number of spectra which can be assigned to a certain proteotypic peptide – and thereby to a certain protein - is counted and normalized to the total number of spectra obtained from the entire sample. Normalized spectral count numbers between different samples are compared for relative quantification of the corresponding proteins. Only proteins for which (i) MS/MS spectra had been detected in each stage, (ii) the corresponding peptides were detected with at least ten MS/MS spectra in at least one of the two stages and (iii) which had a “spectral count abundance ratio” of at least two were considered as differentially abundant. These criteria were met by 28 proteins, which are listed in Table 9.

Table 9: Differentially abundant proteins between 2-cell stage embryos and morulae detected by nano LC-MS/MS spectral count analysis

Protein Name	Accession Number	Molecular Weight	Fold Change	Number of spectra in two-cell stage Embryos	Number of spectra in Morulae
Similar to Tubulin, alpha 1 isoform 1	IPI00839209	50 kDa	2.1	32	17
Similar to transducin-like enhancer of split 6	IPI00716544	47 kDa	0.5	31	72
Nuclear autoantigenic sperm protein	IPI00703854	84 kDa	3.8	27	8
YBX2 protein	IPI00705175	38 kDa	13	24	2
40 kDa protein	IPI00837636	40 kDa	3.3	21	7
Heat shock protein 105 kDa	IPI00705477	97 kDa	2.4	17	8
Similar to Polyadenylate-binding protein 1-like isoform 19	IPI00715420	68 kDa	3.6	16	5
S-formylglutathione hydrolase	IPI00689436	32 kDa	2.8	15	6
Isoform 1 of Proactivator polypeptide	IPI00718311	58 kDa	2.1	15	8
Insulin-like growth factor 2 mRNA binding protein 3 isoform 1	IPI00712553	64 kDa	2.1	13	7
Heat shock 70kDa protein 4	IPI00698179	95 kDa	2.2	12	7
Programmed cell death protein 5	IPI00714175	14 kDa	2.8	10	4
Chloride intracellular channel protein 4	IPI00717902	29 kDa	2.8	10	4
Similar to dynein, cytoplasmic, heavy polypeptide 1 isoform 3	IPI00726312	532 kDa	0.5	9	19
ADP/ATP translocase 3	IPI00705378	33 kDa	0.5	9	22
Similar to Hist1h4c protein	IPI00691248	11 kDa	0.3	9	29
similar to NALP13	IPI00709780	103 kDa	0.5	8	17
Canx protein	IPI00871078	68 kDa	0.5	7	15
Similar to karyopherin beta 1 isoform 2	IPI00710247	97 kDa	0.5	6	14
Myosin-10	IPI00709219	229 kDa	0.5	5	11
Superoxide dismutase [Mn], mitochondrial	IPI00692468	25 kDa	0.5	5	12
Isocitrate dehydrogenase [NADP] cytoplasmic	IPI00702781	47 kDa	0.4	5	13
Macrophage migration inhibitory factor	IPI00694142	12 kDa	0.4	4	10
Alpha-enolase	IPI00707095	47 kDa	0.3	4	14
Fructose-bisphosphate aldolase	IPI00852561	39 kDa	0.3	3	10
Elongation factor 2	IPI00707751	95 kDa	0.2	2	11
Similar to SCF complex protein cul-1 isoform 1	IPI00717794	89 kDa	0.1	1	11
53 kDa protein	IPI00714264	53 kDa	0.09	1	13

3.6.1 Validation of abundance alterations between oocytes, 2-cell stage embryos and morulae

Spectral count is a rather simple but fairly crude method of quantification, demanding a validation of the results by an independent approach. For this reason, two of the differentially abundant proteins have been validated by the SRM technique: i) Y-box protein 2 (YBX2) and the (ii) “Insulin-like growth factor 2 mRNA binding protein 3 isoform 1” (IF2B3/IMP3), being especially interesting with respect to early embryonic development. SRM stands for “selected reaction monitoring” and is a method for quantification of targeted proteins. Fig. 17 demonstrates the SRM principle: In a triple quadrupole mass spectrometer, the so-called “precursor ion” is selected to enter the second quadrupole (Q2) where it undergoes collision-induced dissociation (CID). In the third quadrupole (Q3), one or more fragment ions are selected for quantification by the detector. These validation experiments were mainly performed by Thomas Fröhlich and Daniela Deutsch.

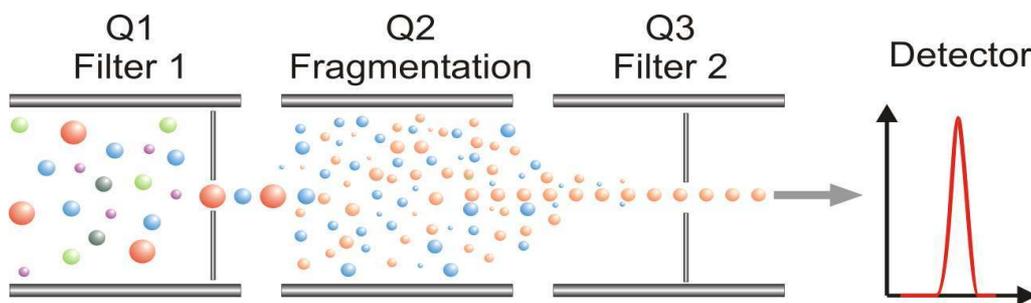


Fig. 17: Principle of SRM quantification

The targeted peptide ion (precursor ion) is selected in the first quadrupole (Q1) and enters the second quadrupole (Q2) where it undergoes collision-induced dissociation (CID). One or more fragment ions are selected in the third quadrupole (Q3) for quantification by the detector; Source: Nanoproteomics – Methods and Protocols, will be published in October 2011 by Humana Press (Springer, ISBN 978-1-61779-318-9).

Proteotypic peptides selected for quantification of IMP3 were the peptides IPVSGPFLVK (IMP3_1) and FTTEEIPLK (IMP3_2). Proteotypic peptides selected for quantification of YBX2 were the peptides GAEANVTGPGGVPVK (YBX2_1) and TPGNPATAASGTPAPLAR (YBX2_2). Three biological replicates of 25 2-cell stage embryos and 25 morulae each were generated, lysed and tryptically digested. From all replicates and stages, the peptides from ten embryos were analysed, and the value of the area under the peak was used for quantification. In Fig. 18, the mean values of the areas under the curve for the peptides IMP3_2 and YBX2_2 in each stage are shown. The fold changes for IMP3 and YBX2 in 2-cell stage embryos versus morulae are 0.5 and 0.1 respectively. In Fig. 19 an example from the quantification of fragment ions from the three proteins in ten 2-cell stage embryos is shown.

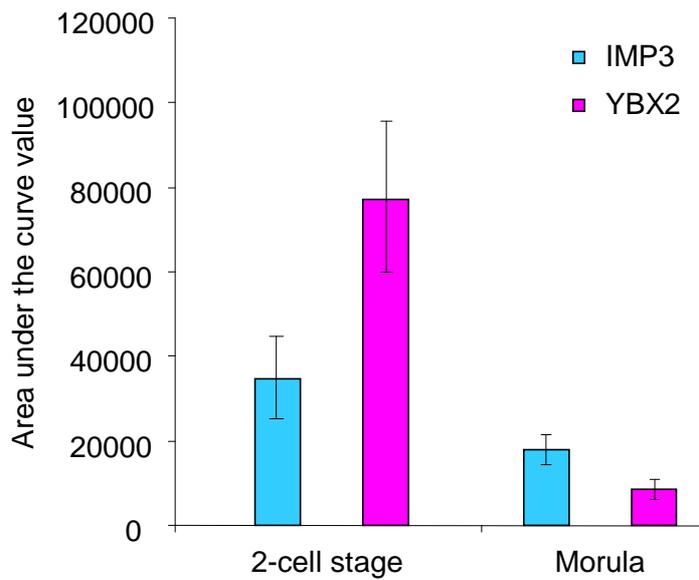


Fig. 18: Areas under the curve values from the SRM quantification of IMP3 and YBX2 in 2-cell stage embryos and morulae

Blue: Peptide FT EEIPLK (IMP3_2) from IMP3, Magenta: Peptide TPGNPATAASGTPAPLAR (YBX2_2) from YBX2; X-axis: developmental stage, Y-axis: Area under the curve value.

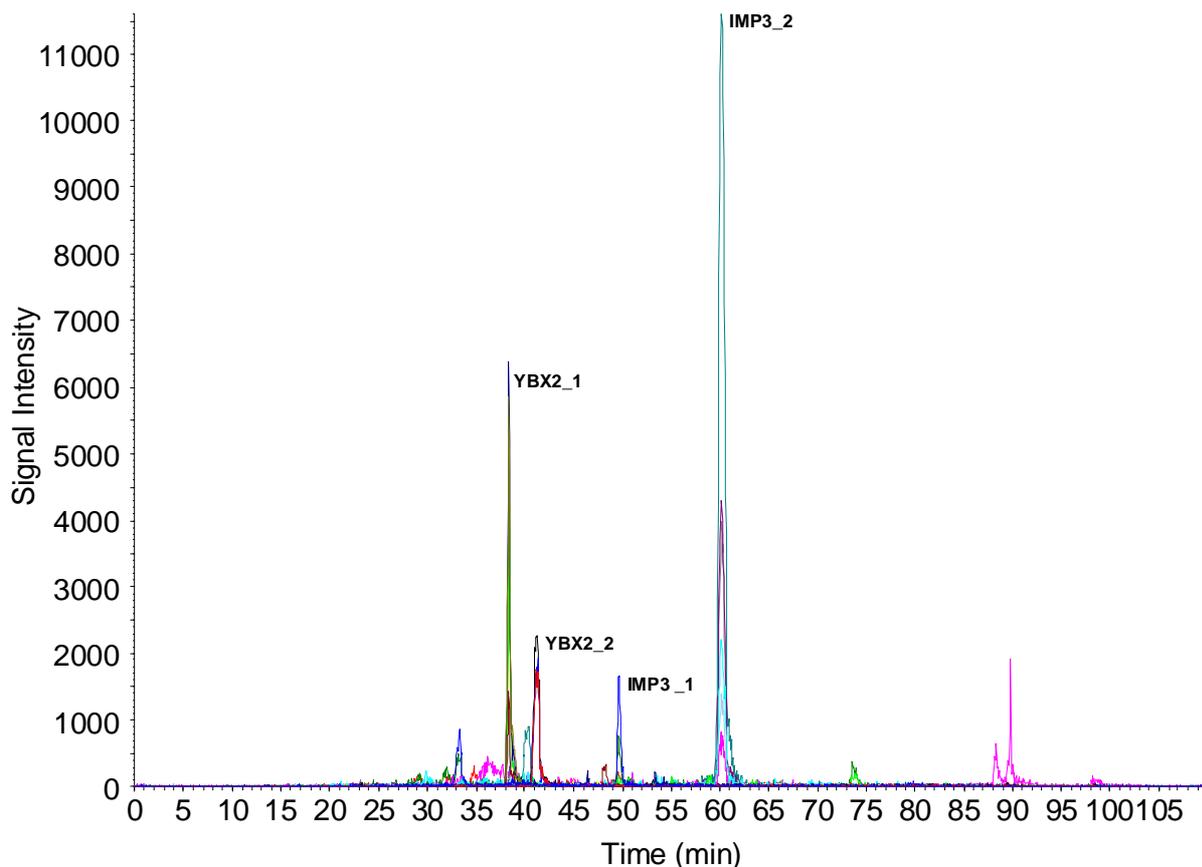


Fig. 19: Chromatogram of a multiplexed SRM analysis

Example from the quantification of fragment ions from the proteins IMP3 and YBX2 in ten 2-cell stage embryos; X-axis: elution time; Y-axis: signal intensity; peptides represented by the individual peaks are indicated right of the peaks: IMP3_1 (IPVSGPFLVK), IMP3_2 (FT EEIPLK), YBX2_1 (GAEEANVTGPGGVPVK), YBX2_2 (TPGNPATAASGTPAPLAR).

3.7 Differential proteome analyses of morulae and blastocysts

To analyse the proteome alterations during the morula to blastocyst transition, a qualitative comparison using both a 2D gel based and a nano LC-MS/MS based technique was performed. Morulae and blastocysts for differential quantitative proteome analysis were generated from 150 slaughterhouse ovaries. Each ovary delivered ten COCs on average. COCs were *in vitro* matured and fertilized. The zygotes were denuded and cultivated for five days (morulae) or for seven days (blastocysts). Prior to collection, embryos were microscopically controlled to their blastomeric structures. Morulae were collected when they had at least 32 blastomeres. Both compacted and not compacted morulae were collected. Blastocysts were collected when they showed a blastocoel as well as the partition in “inner cell mass” (ICM) and trophectoderm (TE) and were collected both in the expanded and not expanded stage. On average, 60 % of all COCs developed into morulae and 30 % of all COCs developed into blastocysts. In total, 12 biological replicates were prepared. One biological replicate contained 25 morulae and 25 blastocysts each, which originated from different animals and had been cultured in individual Petri dishes in groups of 40 embryos.

For the proteomic analysis, two complementary strategies were applied: i) the 2D gel based “saturation DIGE” technique and ii) the LC-MS/MS based iTRAQ technique. Six biological replicates of both stages were analysed by each strategy.

3.7.1 2D saturation DIGE analysis

For the 2D saturation DIGE analysis, six biological replicates corresponding to 25 morulae and 25 blastocysts (2.25 µg total protein) each were lysed. An internal pooled standard (IPS) was prepared consisting of 1.25 µg protein from each stage in a biological replicate (12 samples). 15 µg IPS were labelled with 7.5 nmol fluorescent Dye Cy3. The residual 1 µg protein of each sample was labelled separately with 0.5 nmol fluorescent Dye Cy5 each.

Two 2D gels (pH range four to seven) per biological replicate containing 250 ng IPS (protein amount corresponding to less than three embryos) and 250 ng of Cy5 labelled proteins each were prepared. In Fig. 20, representative Cy5 gel images of proteins from morulae and blastocysts are shown, demonstrating the high separation strength in both dimensions.

Image analysis, performed by the DeCyder 6.5 Software, led to the detection of at least 2948 signals per gel from which 2024 signals were matched and quantified in all gels. These signals can be considered as signals from real protein spots.

Spots detected on all gel images which had a student’s t-test p-values lower than 0.05 were considered as differentially abundant protein spots. Moreover, a FDR correction according to

Benjamini and Hochberg [124] was applied to further minimize the detection of false positive proteome alterations. A total of 61 protein spots fulfilled these criteria. Out of these, 30 spots were altered by a fold change of < 0.5 or > 2 . Fig. 21 shows a Cy 5 morula gel image in which the spots from differentially abundant proteins are marked.

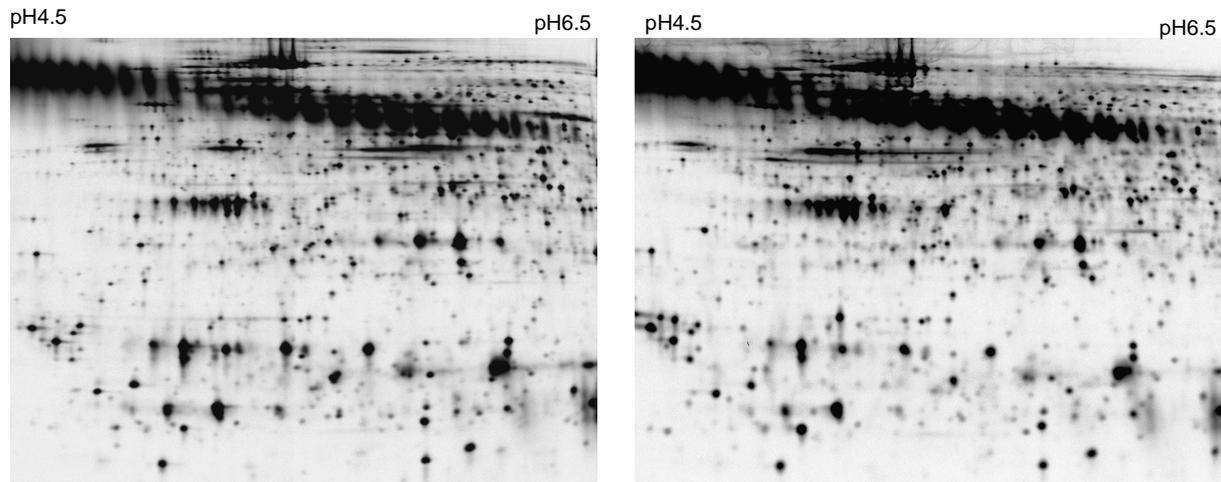


Fig. 20: 2D saturation DIGE analysis of morulae and blastocysts
Cy5 readouts of a morula (left panel) and a blastocyst (right panel) gel image

Fig. 22 contains representative 3D intensity shape plots and corresponding graph views of spots from differentially abundant proteins. Intensity and shape plots visualize spot intensities and spot boundaries, demonstrating the high reproducibility of quantification.

For protein identification from saturation DIGE gels, a preparative gel containing several hundreds of microgram labelled protein is mandatory, requiring 4000 to 5000 oocytes or embryos. Due to high costs of embryo generation, 4444 denuded GV oocytes were used as a source for protein identification. To ensure a comparable running behaviour between analytical and preparative gels, the GV oocyte protein samples were also labelled with saturation dyes.

The gel spots from the Cy3 readout of the preparative gel were matched to their positions on the analytical gels. The matching of protein spots from differentially abundant proteins was manually controlled, and 40 spots could be unambiguously assigned to the preparative gel. The spots were picked with a “spot picking” robot, tryptically digested and analysed in an Orbitrap XL mass spectrometer. Only Proteins which were identified with at least two peptides and a protein score of higher than 99 were accepted as correct identifications (corresponding to a mascot p-Value of < 0.05). In total 34 of the picked spots were abundant enough to lead to protein identification by an Orbitrap XL mass spectrometer.

From all identified spots 18 protein spots could be unambiguously assigned to a single protein (Table 10), 11 spots contained two proteins (Table 11) and five spots contained three or more

proteins (Table 12). Spots which could not be identified are listed in Table 13. In total, 11 proteins were identified from more than one spot, indicating that corresponding spots contain different isoforms of the proteins, for example different post translational modifications. Examples of such proteins detected by this experiment are the four different isoforms of “Aldo-keto reductase family 1 member B1” (AKR1B1), three Peroxiredoxin 2 (PRDX2) isoforms and two Prohibitin (PHB) and “Chloride intracellular channel protein 4” (CLIC4) isoforms.

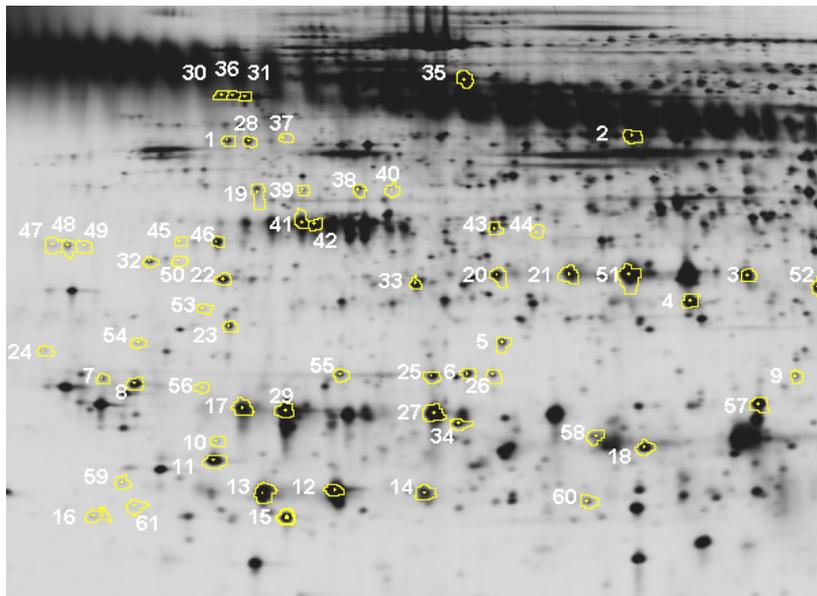


Fig. 21: De Cyder 6.5 Cy5 readout image from a morula gel

Yellow line: spot boundaries;
Numbers: Spots differentially abundant in morulae and blastocysts, corresponding to the numbers used in tables 10 to 13.

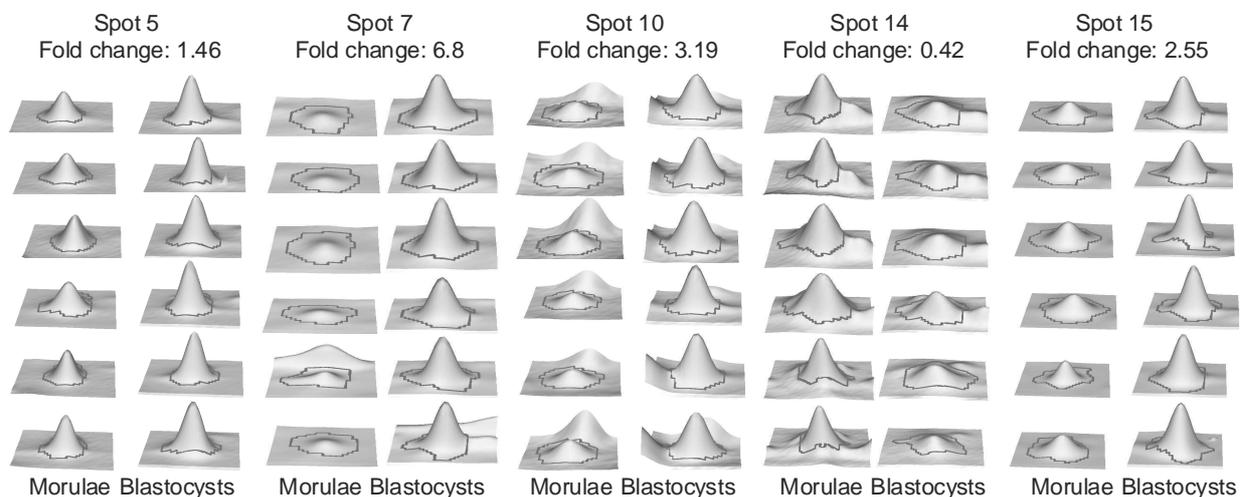


Fig. 22: Graphic representations of five protein spots differing in intensity between morulae and blastocysts
Intensity and shape plots of five differentially abundant proteins in all six biological replicates. Spot intensities are depicted as the height and spot borders as the shape of the peaks; Columns: individual spots as indicated, Rows: individual biological replicates.

Table 10: Unambiguously identified spots from differentially abundant proteins between morulae and blastocysts from the 2D saturation DIGE analysis

No.	Protein Name	IPI Accession	fold change	p-value	Molecular Weight	Mascot Score	Number of peptides (Ions score>30)
1	Rab GDP dissociation inhibitor alpha	IPI00698594	0.48	2.30E-03	57256	122	3
2	PDIA3 protein	IPI00852512	0.61	3.30E-04	61600	2159	22
3	Aldo-keto reductase family 1, member B1	IPI00700920	0.50	1.00E-03	40734	292	4
4	L-lactate dehydrogenase B chain	IPI00760524	0.71	4.10E-04	40062	40062	12
5	Isoform Beta-3 of F-actin-capping protein subunit beta	IPI00688921	1.46	1.30E-04	39098	438	4
6	Prohibitin	IPI00688006	0.66	2.80E-05	30458	940	13
7	Isoform 2 of Tropomyosin alpha-3 chain	IPI00714405	6.80	3.60E-06	31704	641	6
8	Ubiquitin carboxyl-terminal hydrolase isozyme L3	IPI00705749	0.56	7.90E-05	28182	561	4
9	Glutathione S-transferase mu 3	IPI00702950	1.66	4.80E-03	30866	724	8
10	Calpain small subunit 1	IPI00696263	3.19	2.20E-06	29258	99	3
11	GLO1 protein	IPI00713484	0.57	1.00E-06	22769	305	4
12	Peroxiredoxin-2	IPI00713112	0.48	1.60E-05	25294	541	6
13	Peroxiredoxin-2	IPI00713112	0.53	1.50E-05	25294	331	4
14	Peroxiredoxin-2	IPI00713112	0.42	2.80E-05	25294	380	5
15	EIF5A Eukaryotic translation initiation factor 5A-1	IPI00704728	2.55	2.30E-06	19511	194	2
16	MYL12B Myosin regulatory light chain 12B	IPI00687409	3.50	8.10E-06	20352	334	3
17	Ubiquitin carboxyl-terminal hydrolase isozyme L1	IPI00697759	0.48	1.70E-04	33023	148	3
18	Peroxiredoxin-6	IPI00689857	0.53	3.20E-04	25723	414	4

Table 11: Spots from differentially abundant proteins between morulae and blastocysts of the 2D saturation DIGE analysis, from which two proteins were identified

No.	Protein Name	IPI Accession	Fold change	p-value	Molecular Weight	Mascot Score	Number of peptides (Ions score>30)
19	ATP synthase subunit beta, mitochondrial	IPI00717884	1.49	6.60E-03	56249	2136	15
	Hsp90 co-chaperone Cdc37	IPI00718101	1.49	6.60E-03	50603	486	5
20	Stomatin-like protein 2	IPI00705395	0.56	3.40E-05	40054	529	8
	Aldo-keto reductase family 1, member B1	IPI00700920	0.56	3.40E-05	40734	171	3
21	Aldo-keto reductase family 1, member B1	IPI00700920	0.58	1.00E-04	40734	399	6
	Stomatin-like protein 2	IPI00705395	0.58	1.00E-04	40054	112	3
22	48 kDa protein	IPI00717972	0.52	3.20E-04	53665	258	5
	FDPS protein	IPI00839514	0.52	3.20E-04	53681	258	5
23	Inositol monophosphatase 1	IPI00692819	0.65	3.30E-04	34070	404	4
	Eukaryotic translation elongation factor 1 delta	IPI00691738	0.65	3.30E-04	75032	148	2
24	Eukaryotic translation initiation factor 6	IPI00699022	2.40	4.50E-04	33219	185	2
	Clusterin	IPI00694304	2.40	4.50E-04	57804	124	2
25	Chloride intracellular channel protein 4	IPI00717902	0.63	2.30E-04	31398	538	6
	Prohibitin	IPI00688006	0.63	2.30E-04	30458	243	3
26	Chloride intracellular channel protein 4	IPI00717902	1.61	1.80E-03	31398	718	6
	Prohibitin	IPI00688006	1.61	1.80E-03	30458	137	2
27	22 kDa protein	IPI00702336	0.39	1.50E-06	24917	619	6
	Tumor protein D52-like 2	IPI00693729	0.39	1.50E-06	20594	125	2
28	Rab GDP dissociation inhibitor alpha	IPI00698594	0.58	3.20E-03	57256	886	9
	FK506 binding protein 8, 38kDa	IPI00687839	0.58	3.20E-03	49710	99	3
29	Ubiquitin carboxyl-terminal hydrolase isozyme L1	IPI00697759	0.46	1.60E-05	33023	231	4
	Rho GDP-dissociation inhibitor 1	IPI00716328	0.46	1.60E-05	24265	187	2

Table 12: Spots from differentially abundant proteins between morulae and blastocysts of the 2D saturation DIGE analysis, from which three or more proteins were identified

No.	Protein Name	IPI Accession	Fold Change	p-value	Molecular Weight	Mascot Score	Number of peptides (Ions score>30)
30	78 kDa glucose-regulated protein	IPI00717234	0.71	6.00E-03	73700	1709	18
	Zona pellucida sperm-binding protein 4	IPI00686205	0.71	6.00E-03	73281	435	4
	TDRKH protein	IPI00703713	0.71	6.00E-03	68662	353	4
	Zona pellucida sperm-binding protein 3	IPI00703432	0.71	6.00E-03	57272	242	3
	Protein kinase C and casein kinase substrate in neurons 2	IPI00708226	0.71	6.00E-03	59890	161	3
31	78 kDa glucose-regulated protein	IPI00717234	0.44	3.30E-05	73700	2887	22
	TDRKH protein	IPI00703713	0.44	3.30E-05	68662	449	6
	Zona pellucida sperm-binding protein 4	IPI00686205	0.44	3.30E-05	73281	298	3
	Protein kinase C and casein kinase substrate in neurons 2	IPI00708226	0.44	3.30E-05	59890	201	3
32	Similar to 40S ribosomal protein SA (Fragment)	IPI00699280	3.58	3.60E-06	34208	525	4
	Poly(ADP-ribose) glycohydrolase ARH3	IPI00692404	3.58	3.60E-06	44575	214	4
	SET translocation	IPI00728768	3.58	3.60E-06	32098	154	2
33	Suppressor of G2 allele of SKP1 homolog similar to Thioredoxin domain-containing protein 5 precursor	IPI00689292	0.39	2.20E-06	40739	1109	11
	UBFD1 protein	IPI00699038	0.39	2.20E-06	53848	272	7
		IPI00841388	0.39	2.20E-06	35140	171	2
34	Aldo-keto reductase family 1, member B1	IPI00700920	0.47	1.50E-06	40734	610	8
	Endoplasmic reticulum resident protein 29	IPI00702891	0.47	1.50E-06	29460	334	5
	Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial	IPI00699016	0.47	1.50E-06	45021	221	2
	22 kDa protein	IPI00702336	0.47	1.50E-06	24917	178	2

Table 13: Spots from differentially abundant proteins between morulae and blastocysts which were not identified

No.	p-value in student's t-Test	Fold Change	No.	p-value in student's t-Test	Fold Change
35	3.60E-05	0.48			
36	3.50E-05	0.51	49	3.70E-04	0.40
37	8.40E-04	2.37	50	8.50E-03	2.36
38	8.40E-07	3.45	51	2.50E-05	0.56
39	7.90E-05	1.77	52	7.10E-03	0.72
40	3.00E-03	2.34	53	1.00E-04	0.63
41	5.80E-03	0.58	54	8.40E-05	0.66
42	4.00E-04	1.73	55	2.20E-06	5.82
43	2.20E-06	12.75	56	5.50E-05	3.22
44	3.60E-06	7.3	57	1.30E-05	0.34
45	4.00E-04	1.7	58	2.30E-04	0.61
46	3.40E-06	0.52	59	1.60E-05	2.63
47	6.80E-05	0.33	60	1.30E-04	0.68
48	8.40E-07	0.17	61	9.50E-04	0.55

3.7.2 Nano LC-MS/MS iTRAQ analysis

For iTRAQ nano LC-MS/MS analysis, six biological replicates corresponding to 25 morulae and 25 blastocysts (2.25 µg total protein) each were lysed, tryptically digested and labelled with iTRAQ tags. The labelling was performed according to the scheme in Table 14, which allowed the multiplexed analysis of two replicates per run set. As established before (see chapter 3.4), each run set consisted of four consecutive 2D LC-MS/MS runs. In total three run sets corresponding to 12 days of LC-MS/MS measuring were analysed.

Table 14: Labelling and run set scheme of six biological replicates from morulae and blastocysts analysed by nano LC-MS/MS iTRAQ analysis

Numbers 114-117 indicate the iTRAQ reporter tag masses attached to a sample; Letters: M indicates a morula sample, B indicates a blastocyst sample; Numbers 1-6 indicate the numbers of biological replicates

	Label 114	Label 115	Label 116	Label 117
Run Set 1	M1	B1	M2	B2
Run Set 2	M3	B3	M4	B4
Run Set 3	M5	B5	M6	B6

In total, 141 proteins were identified and quantified from the three iTRAQ run sets comprising all six biological replicates. Their log₂ fold changes in morulae versus blastocysts were subjected to statistical analyses by the R software. The Volcano Plot in Fig. 23 shows the abundance ratios and p-values of all proteins from the analysis. Proteins which had a log₂ fold change of $> |0.6|$ - corresponding to an abundance ratio of 1.51 - and a p-value in the student's t-test of < 0.05 were considered as differentially abundant. 50 proteins fulfilled these criteria and are represented as red data points in Fig. 23. The corresponding protein IDs are listed in Table 15. A total of 28 proteins was strongly affected in abundance, as indicated by a fold change of < 0.5 or > 2 .

The reproducibility of quantification within all biological replicates is exemplified in Fig. 24 on behalf of the HCD MS/MS spectra of the peptide LDNLVAILDINR derived from transketolase. Intensity ratios of iTRAQ reporter signals are highly reproducible in all cases.

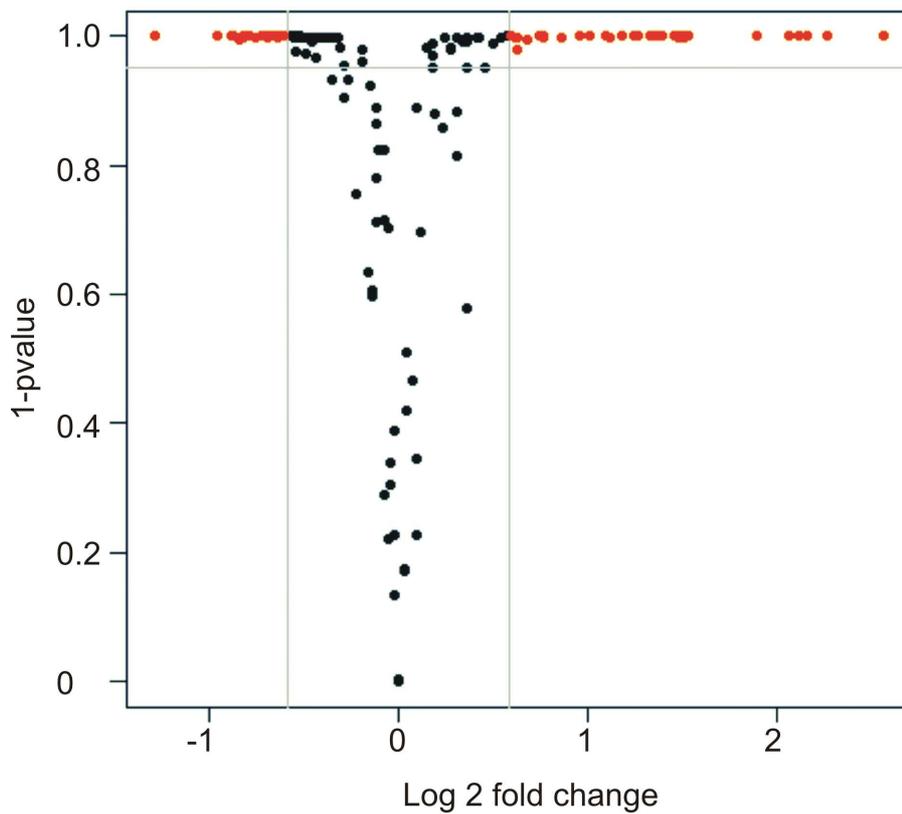


Fig. 23: Volcano Plot of 141 proteins identified and quantified from morulae and blastocysts

Red data points represent proteins which have a log₂ fold change of $> |0.6|$ and a p-value in the student's t-test of < 0.05 ; X-axis: log₂ fold change of proteins, Y-axis: 1-p-values of proteins.

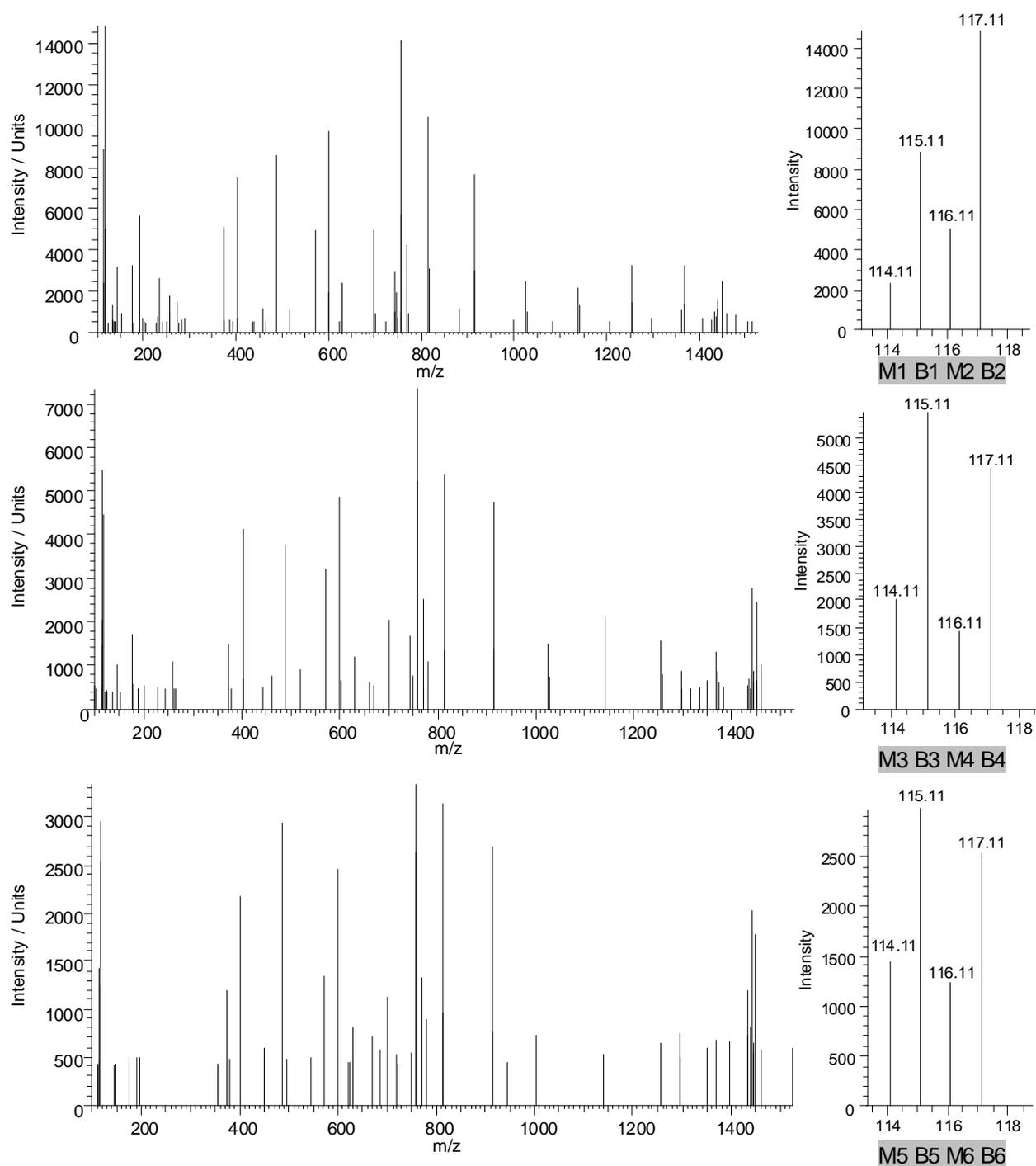


Fig. 24: Quality and reproducibility of HCD spectra

Left panels: HCD MS/MS spectra of the transketolase peptide LDNLVAILDINR from each of the three run sets of the iTRAQ nano LC-MS/MS analysis of morulae and blastocysts; Right panel: the iTRAQ reporter ion peaks are zoomed out; X-axis: m/z values of peaks from peptide fragments; Y-axis: peak intensity.

Table 15: Differentially abundant proteins between morulae and blastocysts detected by the LC-MS/MS iTRAQ analysis

Protein name	IPI Accession Number	Molecular Weight	average fold change	p-value
Creatine kinase B-type	IPI00716827	43 kDa	5.9	6.35E-06
Similar to histone cluster 1. H2ag	IPI00689821	14 kDa	4.8	2.94E-06
Intestinal alkaline phosphatase VI	IPI00703521	61 kDa	4.5	5.38E-05
Annexin A6	IPI00868644	76 kDa	4.3	5.62E-05
HIST1H1C protein	IPI00699808	21 kDa	4.2	1.65E-06
Similar to filamin B. beta (actin binding protein 278) isoform 7	IPI00688180	279 kDa	3.7	3.45E-05
Transketolase	IPI00904104	65 kDa	2.9	2.89E-06
40S ribosomal protein S25	IPI00691584	14 kDa	2.9	5.95E-04
60S ribosomal protein L27a	IPI00694444	17 kDa	2.9	1.72E-05
74 kDa protein	IPI00690432	74 kDa	2.8	1.38E-05
Guanine nucleotide-binding protein subunit beta-2-like 1 (RACK1)	IPI00700792	35 kDa	2.8	8.01E-06
75 kDa protein	IPI00826962	75 kDa	2.8	2.73E-03
LGALS3 protein	IPI00867041	28 kDa	2.8	1.44E-05
40S ribosomal protein S11	IPI00689811	18 kDa	2.6	1.24E-05
Nucleophosmin	IPI00689902	33 kDa	2.6	2.49E-05
Citrate synthase. mitochondrial	IPI00699260	52 kDa	2.5	8.20E-05
Sodium/potassium-transporting ATPase subunit alpha-1	IPI00705159	113 kDa	2.5	6.32E-05
40S ribosomal protein S3	IPI00685601	27 kDa	2.4	8.11E-05
Similar to Protein FAM151A	IPI00710809	64 kDa	2.4	2.56E-04
Ezrin	IPI00694641	69 kDa	2.3	1.35E-05
Polyadenylate-binding protein 1	IPI00705440	71 kDa	2.2	6.74E-04
Non-muscle myosin heavy chain	IPI00696012	228 kDa	2.1	1.46E-04
Heat shock protein HSP 90-beta	IPI00709435	83 kDa	2.0	1.32E-05
Elongation factor 2	IPI00707751	95 kDa	2.0	8.85E-05
Cytochrome c	IPI00717272	12 kDa	1.8	4.97E-04
Elongation factor 1-alpha 1	IPI00712775	50 kDa	1.7	2.03E-05
Voltage-dependent anion-selective channel protein 2	IPI00687246	32 kDa	1.7	4.84E-04
Isocitrate dehydrogenase [NADP] cytoplasmic	IPI00702781	47 kDa	1.7	2.38E-05
Actin. cytoplasmic 1	IPI00698900	42 kDa	1.6	3.62E-03
Cofilin-1	IPI00699700	19 kDa	1.6	1.90E-02
ATP synthase subunit delta. mitochondrial	IPI00707257	18 kDa	1.6	2.36E-03
ATP synthase subunit alpha. mitochondrial	IPI00694295	60 kDa	1.5	3.20E-04
ATP synthase subunit beta. mitochondrial	IPI00717884	56 kDa	1.5	1.42E-04
Protein disulfide isomerase family A. member 6 isoform 7	IPI00693090	65 kDa	0.7	1.42E-04

Protein name	IPI Accession Number	Molecular Weight	average fold change	p-value
Nucleoside diphosphate kinase B	IPI00686420	17 kDa	0.6	1.83E-03
Peroxiredoxin-1	IPI00686092	22 kDa	0.6	9.27E-05
Cell division protein kinase 5	IPI00689812	33 kDa	0.6	2.39E-04
Family with sequence similarity 62 (C2 domain containing). member A	IPI00713575	123 kDa	0.6	1.84E-03
L-lactate dehydrogenase A chain	IPI00716974	37 kDa	0.6	1.05E-04
PDIA3 protein	IPI00852512	57 kDa	0.6	2.37E-05
78 kDa glucose-regulated protein	IPI00717234	72 kDa	0.6	1.73E-03
Peroxiredoxin-2	IPI00713112	22 kDa	0.6	2.88E-05
Glutathione S-transferase mu 3	IPI00702950	27 kDa	0.6	4.74E-05
73 kDa protein	IPI00841750	73 kDa	0.6	2.96E-03
Inositol polyphosphate 1-phosphatase	IPI00686250	44 kDa	0.6	1.25E-05
Retinal dehydrogenase 1	IPI00692627	55 kDa	0.6	5.95E-03
similar to T04F3.1	IPI00686245	78 kDa	0.5	2.26E-04
22 kDa protein	IPI00702336	22 kDa	0.5	5.67E-05
Ubiquitin carboxyl-terminal hydrolase isozyme L1	IPI00697759	28 kDa	0.5	7.44E-06
Bisphosphoglycerate mutase	IPI00706349	30 kDa	0.4	1.66E-04

3.7.3 Comparison of results from 2D DIGE analysis and from nano LC-MS/MS analysis

Results of the LC-MS/MS based iTRAQ analysis and 2D gel based saturation DIGE analysis were compared. The comparison showed that seven proteins had been identified as differentially abundant in both approaches. They are listed together with their fold changes in Table 16. A total of six out of the seven proteins showed similar abundance alterations in both approaches. This is demonstrated on behalf of Peroxiredoxin 2 in Fig. 25 which shows a comparison of HCD MS/MS spectra of peroxiredoxin 2's peptide QVTINDLPVGR, and corresponding 3D intensity shape plots from the 2D DIGE analysis. In contrast, one protein spot identified as "Glutathione S-transferase mu3" is increased in the 2D DIGE analysis, while the majority of the corresponding peptides revealed to be decreased in abundance in the iTRAQ analysis.

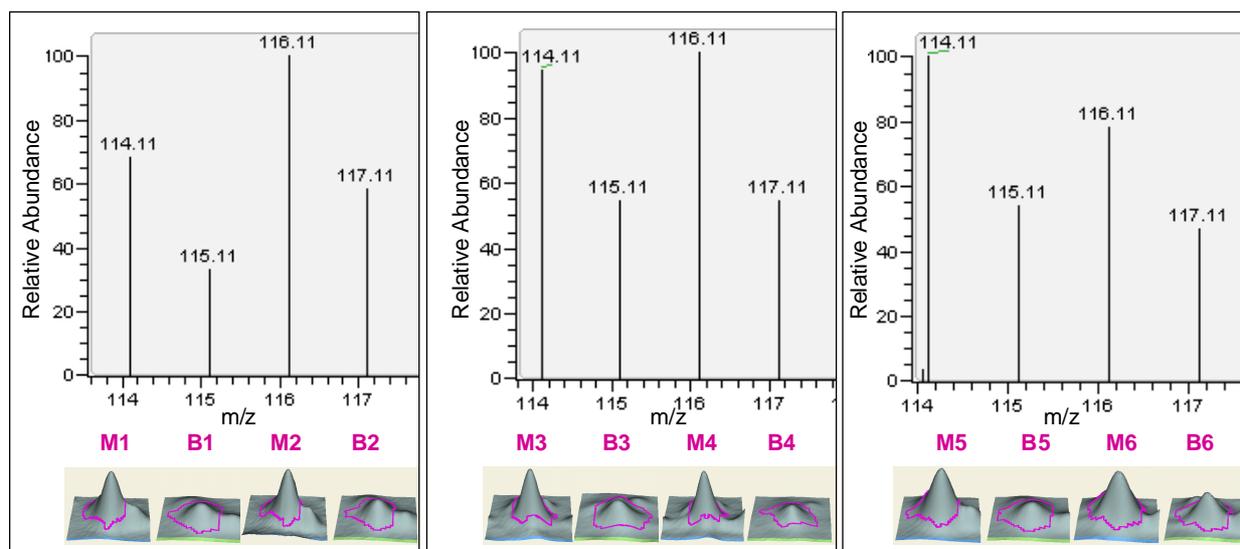


Fig. 25: Comparison of abundance alterations determined by DIGE spot intensity and iTRAQ reporter ion intensity

Upper graphs in panels: iTRAQ-reporter ion zoom out of HCD MS/MS spectra of Peroxiredoxin2's peptide QVTINDLPVGR from each of the three run sets of the iTRAQ nano LC-MS/MS analysis of morulae and blastocysts; X-axis: m/z values of peaks from peptide fragments, Y-axis: relative abundance of peak intensities; Lower graphs in panels: Intensity and shape plots of a differentially abundant spot identified as Peroxiredoxin 2; Left panel: replicates 1 and 2, middle panel: replicates 3 and 4, right panel: replicates 5 and 6.

Table 16: Differentially abundant proteins between morulae and blastocysts detected by both, the Saturation DIGE and the LC-MS/MS iTRAQ analysis

Spot No.	Protein name	IPI Accession number	Fold change in iTRAQ Analysis	Fold change in 2D DIGE Analysis
19	ATP synthase subunit beta, mitochondrial	IPI00717884	1.5	1.5
12,13,14	Peroxiredoxin-2	IPI00713112	0.6	0.5
30,31	78 kDa glucose-regulated protein	IPI00717234	0.6	0.6
2	PDIA3 protein	IPI00852512	0.6	0.6
9	Glutathione S-transferase mu 3	IPI00702950	0.6	1.7
34,27	22 kDa protein	IPI00702336	0.5	0.4
17,29	Ubiquitin carboxyl-terminal hydrolase isozyme L3	IPI00697759	0.5	0.5

3.7.4 David gene ontology analysis of differentially abundant proteins

Results from both the 2D DIGE and the iTRAQ analyses were merged, leading to a set of 42 different proteins with decreasing abundance alterations between morulae and blastocysts and to a set of 47 different proteins with increasing abundance alterations between morulae and blastocysts.

These two sets were individually subjected to David **gene ontology** (GO) analysis. The David GO clustering according to the GOTERM_BP_FAT categories leads to 40 different GO terms within the group of proteins showing decreased abundance alterations in blastocysts. A total of nine proteins are associated with catabolic processes while four proteins are associated with biosynthetic processes. Furthermore, 12 proteins are associated with redox processes.

Within the group of proteins showing increased abundance alterations, David GO clustering leads to 54 different GO Terms. A total of seven proteins are associated with biosynthetic processes and only two proteins with catabolic processes. A total of three proteins are assigned to cytoskeletal organisation and five proteins are associated with chromosome or nucleosome organization. A total of eight proteins are assigned to the GO term translation while seven of the proteins are assigned to Ion transport.

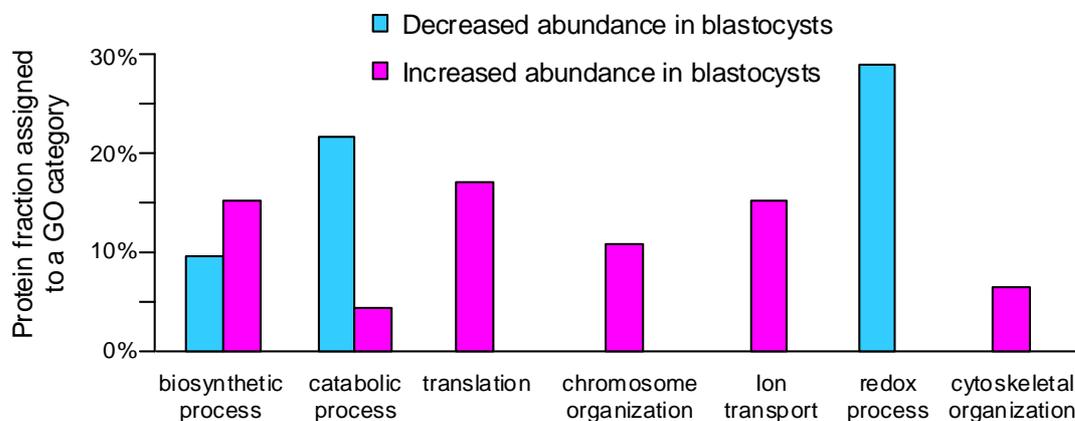


Fig. 26: Comparative “David” GO analysis

Results from the comparative David Gene Ontology analysis between increased (magenta) and decreased (blue) proteins in morulae versus blastocysts; X-axis: Gene Ontology category, Y-axis: percentage of increased/decreased proteins within the indicated category.

4 Discussion

4.1 Qualitative proteome analysis of 900 bovine GV oocytes

Developmental competence of oocytes impacts nearly all important steps in embryonic development [22, 23]. It is acquired during the growth of oocytes in folliculogenesis and during the oocyte maturation process [22]. Developmental competence is characterized by the oocyte's ability (i) to resume meiosis, (ii) to cleave following fertilization, (iii) to develop to the blastocyst stage, (iv) to induce a pregnancy and (v) to develop to term in good health [23]. During the growth phase of oocytes, large stores of mRNAs and proteins that function after fertilization are produced [22]. Deficiencies in mRNA and protein accumulation are thought to lead to failures in cytoplasmic maturation, which occurs besides the nuclear maturation process [22]. A part of the cytoplasmic maturation process is the molecular maturation process. It is assumed that specific mRNAs and proteins are produced and added to the oocyte's stockpile in the last few days before ovulation [23].

To get an overview of the kind of these stored proteins, a qualitative proteome profile was performed by analysing the proteins contained in GV oocytes. GV oocytes are easily available from slaughterhouse ovaries and have a large cytoplasmatic part containing approximately 90 ng protein. In contrast, MII oocytes are more difficult and time consuming to obtain and have the major disadvantage that their biology is altered by either *in vitro* maturation or by FSH stimulation of donor cows.

In the middle of the nineties it became possible to identify large numbers of proteins by the ongoing development of protein mass spectrometry. A powerful technique applied for this purpose is the nano LC-MS/MS technique, in which peptides are the basis of protein identification (see chapter 1.5.2). Usually, the protein subset of a cell, organ or organism, containing peptides which are easily ionisable, can be detected and identified by mass spectrometry. Protein identification by MS/MS works by the probability based comparison of experimentally acquired peptide MS/MS spectra with *in silico* generated theoretical spectra deduced from protein databases. At least two different peptides of a protein should be identified, since a protein identification based on just a single peptide may in many cases represent a false positive identification. The false discovery rate (FDR) of a dataset should be limited to 1 % [66]. This can be checked by a search, against a so-called decoy database, generated by scrambling the amino acids from each database entry in a random manner. If the number of protein IDs obtained from the decoy database search doesn't exceed 1 % of the number of IDs in the dataset containing the identified proteins, an FDR of lower than 1 % can be assumed [125]. The application of these criteria for protein identification became more and

more accepted during the last years. For example, an experiment leading to a protein dataset of bovine GV oocytes containing 1092 identified proteins published by an American group in 2007 was repeated in 2010. Although during the repeated experiment almost three times as many mass spectra as before were acquired, a decoy database search and the statistical evaluation dropped the number of identified proteins down to 811. From this list still only 329 GV oocyte proteins were identified by at least two peptides [106, 126].

To obtain a sufficient protein amount for qualitative nano-LC-MS/MS analysis, 900 GV oocytes (containing 81 μg protein) were lysed. To decrease the complexity, the protein lysate was split into two equal parts which were prefractionated on a 12 % and on a 6 % 1D SDS PAA gel (Fig. 5) prior to nano LC-MS/MS analysis. The criteria to be at least identified by two peptides while the dataset has an FDR of lower than 1 %, confirmed by a decoy database search, were fulfilled by 791 proteins from the oocyte lysate prefractionated on the 12 % SDS PAA gel. Additionally, 38 mostly high molecular weight (> 100 kDa) proteins were identified from the lysate prefractionated on the 6 % SDS PAA gel, due to the enhanced separation power for high molecular weight proteins on low percentage gels. In total 829 proteins were identified.

The existence of many proteins in genomic databases has not been validated by protein identification but is deduced from corresponding DNA/RNA sequences. In the Uniprot database, five levels of evidence for protein existence are indicated: (Uniprot Level I) “Evidence at protein level” applies to proteins that have been previously identified, (Uniprot Level II) “Evidence at transcript level” applies to proteins of which the corresponding RNA has been identified, (Uniprot Level III) “Inferred from homology” applies to proteins of which the existence is probable because clear orthologs exist in closely related species, and (Uniprot Level IV) “Predicted” applies to proteins which have been predicted by the DNA analyses yet for which no evidence at protein, transcript, or homology levels exists. The weakest term is (Uniprot Level V) “Uncertain”.

The level of protein existence of the 829 proteins from the GV oocyte dataset was checked in the Uniprot database. The evidence for the majority of the identified proteins (75 %) was not confirmed on the protein level in the bovine system prior to this experiment. For 515 proteins, there was at least experimental evidence on the RNA level (Uniprot Level II), but for 83 proteins, the existence was only assumed from homology with sequences of other species (Uniprot Level III) and 24 proteins were only predicted to exist (Uniprot Level IV).

To find out which biological and molecular functions might be executed by the identified proteins, they were clustered according to their gene ontology (GO) terms (Fig. 6). Proteins were assigned to the three main categories “cellular component”, “biological process” and

“molecular function”. In the biological process category most proteins (47 %) are assigned to the cluster “metabolism”. This is coherent with the fact that the oocyte’s ability to correctly control metabolism is of special importance for the developmental potential [22].

An 18 % fraction of the identified proteins is assigned to the cluster “biological regulation”: One can assume that within this group, proteins important for the correct initiation of oocyte meiosis, first cleavages and the correct **maternal embryo transition** (MET) can be found. This hypothesis is strengthened by the fact that it contains for example the proteins Peroxiredoxin 2 and 3. Peroxiredoxin 3 is already known to be altered during the bovine *in vitro* maturation process [74]. Peroxiredoxin 2 (PRDX2) was found to decrease during the morula to blastocyst transition (see chapter 3.7). Another protein found in this group is YBX2, which functions in storage maintenance of maternal mRNA and was found to be differentially abundant between 2-cell stage embryos and morulae (see chapter 3.6).

A 41 % fraction of the proteins contained in the “molecular function” category is assigned to the cluster catalytic activity, wherein most of the proteins could be assigned to oxidoreductases (77) and hydrolases (108). These proteins can also be considered to play a role for oocyte developmental competence since it is influenced by oxygen tension and by the concentration of ROS species [22]. Furthermore, redox enzyme variants were shown to be altered during the oocyte maturation process [74].

To find out in which pathways the identified proteins are involved, a David analysis of KEGG Pathways was performed. This bioinformatic tool reveals pathways in which proteins from the input list show up more than they are expected to appear from a randomized protein list [123]. A total of 23 proteins were found in the KEGG Pathway term “Ribosome”, indicating the ongoing translation of proteins eventually important for developmental competence. A total of 16 proteins were found to take part in the oocyte meiosis pathway (Fig. 7), demonstrating the accumulation of proteins important for this initial step of embryonic development.

In general, this experiment reflects the impact of generating qualitative protein lists. They can not only provide evidence for many proteins to exist but moreover give hints to pathways and processes which might be important for biological events, like for example the early embryonic development.

4.2 Protein identification and quantification of limited sample amounts by nano LC-MS/MS

As described in the previous chapter, the identification of ~ 800 proteins requires protein amounts of about 40 to 100 µg. Since one oocyte or early embryo contains only approximately

90 ng protein, such an identification approach would require the generation of hundreds of embryos. Especially in cases where several biological replicates are needed, the preparation of so many embryos would be very cost intensive and time consuming. For this reason, it was tested how many proteins can be identified from much smaller sample amounts corresponding to the protein amounts of one oocyte up to 50 oocytes. The same strict identification criteria as described in the previous chapter were applied, and a FDR < 1 % was confirmed by decoy database searches.

Even from a single oocyte, 54 proteins could be identified. Using 25 bovine oocytes led to the highest number of 106 identified proteins. In contrast, the protein amount from 50 oocytes dropped the number of identified proteins down to 62 (Fig. 8).

This decrease in the number of identified proteins, despite the higher protein amount subjected to nano LC-MS/MS analysis, can be explained by column overloading: Column overloading usually occurs when the peptide amount loaded onto a reversed phase (RP) column is beyond its capacity, leading to peak broadening and lower separation strength. As a consequence, the number of identified peptides, respectively proteins, is decreased.

To find out whether the number of identifications from protein amounts corresponding to 25 and 50 oocytes can be further enhanced, a two-dimensional (2D) LC-MS/MS analysis was performed by an additional prefractionation step of tryptic peptides on a “strong cation exchange” (SCX) column prior to RP separation. SCX chromatography uses the charge state of a molecule as a major separation criterion and represents an orthogonal separation step to RP chromatography. The base peak ion chromatograms (BICs) in Fig. 9 demonstrate that peptides were evenly distributed over all salt fractions. Due to the lower complexity of protein mixtures after SCX prefractionation and an extended time of spectra acquisition, the number of identified proteins from a sample amount corresponding to 25 oocytes could nearly be doubled to 206 compared to the 1D LC-MS/MS approach. Like in the 1D approach, it was not beneficial to analyse the larger protein amount of 4.5 µg corresponding to 50 oocytes. This protein amount dropped the number of identified proteins down to 198, most likely also due to column overloading occurring within the analysis of single salt fractions.

These experiments created the basis for (i) the adaptation of an iTRAQ quantification approach for small sample amounts (see chapter 4.3) and (ii) for the analysis of 2-cell stage embryos and morulae (see chapter 4.5).

4.3 Protein quantification of limited sample amounts by iTRAQ

Protein quantification by nano LC-MS/MS is relatively complex. The quantitative comparison of MS signals between different LC-MS runs is not reliable since even small irregularities in the composition of samples and replicates may influence retention times and ionization rates of single peptides. Therefore, a common strategy is (i) to label the samples with chemically identical reagents differing in their isotopic composition but not in their molecular mass, (ii) to pool samples after labelling and (iii) to perform LC-MS experiments with these pools. Since this procedure ensures that the retention time of differentially labelled peptides remains unaffected, peptides derived from different samples co-elute and reach the mass spectrometer simultaneously. Hence, signal intensities of corresponding peptides from different samples can be quantitatively compared and assigned to the sample by their isotopic label [77].

There are several possibilities of sample labelling differing in the level of sample procession (*in vivo*/ on the protein level/ on the peptide level) at which the label is attached. For a more detailed description of these different techniques please refer to chapter 1.5.2.

Using the so-called “isobaric tag for relative and absolute quantification” (iTRAQ) four plex approach, the label is introduced on the peptide level. It was the method of choice for these experiments due to the facts that the protein cleavage process is not biased and that nearly all peptides can be labelled in contrast to the ICPL and the ICAT approach (see chapter 1.5.2).

After lysis and tryptic digestion, the peptides were labelled with an isobaric tag on each lysine chain and on each N-terminal group. As shown in Fig. 27, this tag consists of a reporter group which differs in its mass between 114 and 117 Dalton and a balancer group, which has a mass between 31 and 28 Dalton so that the total tag mass is 145 Dalton in each case.

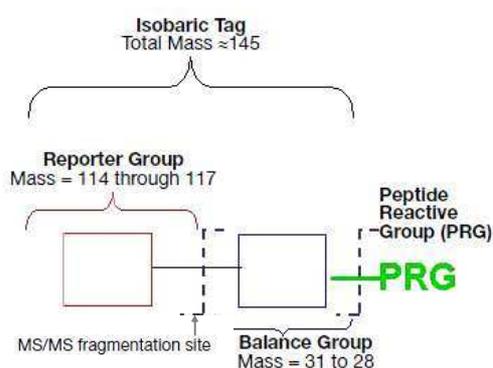


Fig. 27: Composition scheme of the iTRAQ tag

Source: Applied Biosystems iTRAQ™ Reagents - Chemistry Reference Guide, Applied Biosystems, 2004.

Up to four different samples can be labelled, each with a tag containing a different reporter group. Samples are then pooled and passed through an LC-MS/MS run. During fragmentation, reporter and balancer groups are split from the peptide in the so-called “higher energy collision dissociation” (HCD) cell of the mass spectrometer. The iTRAQ reagent reporter group ions that

are generated appear as peaks in the low-mass region between m/z 114 and m/z 118 of the MS/MS spectra. Because this region is free of other common fragment ions, signals found in this region are not influenced by random noise and their intensity relationships represent those of peptides from different samples.

Due to the fact that HCD spectra need an extended scan time in the mass spectrometer, fewer proteins are expected to be quantified and identified during an individual LC run compared to LC runs dedicated to identification only (see previous chapter). Therefore, an experimental series was performed to optimise the number of quantified proteins for a given amount of oocytes.

To find out how many proteins can be quantified and identified from small protein amounts, a 1D LC-MS/MS analyses was performed using 2.25 μg and 4.5 μg pooled iTRAQ sample corresponding to 25 and 50 oocytes, respectively. From the protein equivalent to 25 and 50 oocytes of the “iTRAQ labelled GV oocyte peptides”, 67 and 86 proteins could be identified and quantified, respectively. In order to enhance the number of identified and quantified proteins, the iTRAQ labelled proteins from 25 and from 50 oocytes were split into two equal parts each. In both cases, after the first aliquot had been analysed by LC-MS/MS, an exclusion list (EL) was generated, containing all peptide masses from which MS/MS spectra had been successfully acquired. This exclusion list was applied to the analyses of the second aliquots, forcing the instrument to ignore these masses and instead acquire MS/MS spectra from so far unconsidered peptide precursors. The application of the exclusion list enhanced the number of identified and quantified proteins from the protein amount of 25 oocytes to 107, and from the protein amount of 50 oocytes to 116 (Fig. 11). In order to further enhance these numbers, the pool of iTRAQ labelled GV oocyte peptides from 50 oocytes was divided into four equal parts, which were analysed in four consecutive 1D LC-MS/MS runs, and an exclusion list for each run was generated from all previous runs. Such a combination of four runs is referred to as “run set” throughout this work. As a maximum, 150 proteins were identified and quantified from the protein amount of 50 oocytes. Fig. 12 shows the numbers of identified and quantified proteins from the merge of individual runs with the previous runs. The fact that the number of identified and quantified proteins is enhanced about 45 % by the application of the first EL, but is only enhanced about 13 % by the application of the last EL, demonstrates a saturation effect after the 4th run. One can assume that splitting 50 oocytes in more than four parts followed by a subsequent analysis in a run set consisting of more than four runs would enhance the number of identifications to only a few more. Instead, this set up was tested in a 2D LC-MS/MS approach. Six exclusion lists for each run (one per salt fraction) were generated from all previous runs and applied to the analyses of the corresponding fractions in the following run. This enhanced the

number of identified and quantified proteins to 403. Since the analysis of one 2D LC-MS/MS run took 18 hours, only one 2D LC-MS/MS run could be performed per day. The analysis time for a complete 2D LC-MS/MS run set therefore comprised four days.

Fig. 13 shows the number of identified and quantified proteins from the protein amount corresponding to 50 GV oocytes using different approaches. It clearly demonstrates that the iTRAQ approach leads to the highest number of identified and quantified proteins when multiple exclusion lists in connection with a 2D LC-MS/MS set up are applied. This set up facilitated the analysis of *in vivo* versus *in vitro* matured oocytes (see next chapter) and was also applied to the analysis of morulae versus blastocysts like it is described in chapter 4.6.

4.4 Comparison of *in vivo* and *in vitro* matured oocytes from cows of different age groups by nano LC-MS/MS iTRAQ analysis

Assisted reproduction techniques using IVP become more and more important both in human and in veterinary medicine. However, *in vitro* produced embryos still show more developmental abnormalities than *in vivo* produced embryos. The abnormalities are reflected by increased rates of early embryonic death and abortion, production of large size foetuses, abnormal organ growth as well as failures in the placenta development summarized as the “large offspring syndrome” (LOS) [1]. During IVP of embryos the first step is oocyte maturation. This can be performed *in vivo* or *in vitro*. Especially in human medicine, it is desirable to mature oocytes *in vitro* in order to avoid, or to drastically reduce, the use of gonadotrophins for ovarian hyperstimulation [115]. However, IVM has the disadvantage of a significantly lower efficiency compared to the usage of *in vivo* maturation of oocytes in terms of “developed embryos per collected oocyte” and “offspring per collected oocyte” [115]. This is also the case for IVM of cattle embryos. While approximately 50-80 % of *in vivo* matured oocytes develop to the blastocyst stage, a typical blastocyst rate of 15-40 % is reached with IVM oocytes [1]. Although more and more publications suggest failures in epigenetic reprogramming as the main reasons for these phenomena, the molecular mechanisms underlying these drawbacks are still unresolved [1]. As already mentioned in chapter 4.1, the proteome analysis of oocytes is promising, since they contain a maternal protein storage which functions in fertilization and early embryonic development and might therefore play an important role for the developmental competence of oocytes [22].

To evaluate the alterations in the quantitative protein expression profile between *in vivo* and *in vitro* matured MII oocytes, a nano-LC-MS/MS iTRAQ analysis was performed. Given that the age of oocyte donors is a significant factor for the oocyte’s developmental competence [127],

oocytes were split into three groups according to their donor cow's age (heifers, young cows and old cows). Each age group contained five cows. All oocytes were aspirated *in vivo* in so called "OPU sessions". During one OPU session, *in vivo* aspiration from all five cows of an age group was performed. Oocytes were either aspirated from unstimulated cows in the GV oocyte stage or from FSH stimulated cows in the MII oocyte stage. GV oocytes from unstimulated cows were matured *in vitro* immediately after the OPU session and developed to the MII oocyte stage. Oocytes obtained during this experiment were not only used for proteome analyses but also for evaluation studies in which different cleavage and blastocyst rates were analysed [128]. For each of the age groups, three biological replicates were generated, each consisting of 25 *in vivo* and 25 *in vitro* matured MII oocytes. Usually, more than one OPU session was necessary to collect the 25 oocytes needed for a biological replicate. The experiment was performed by nano LC-MS/MS based iTRAQ quantification using a 2D LC-MS/MS approach in combination with the application of multiple exclusion lists, as established before (see chapter 4.3). For the labelling procedure, four different iTRAQ tags were available which enabled the multiplexed analysis of several samples, within one "run set" (for definition of "run set" please refer to chapter 4.3). Each analysis of one biological replicate comprised four run sets. As demonstrated in Table 3, this multiplexing facilitated for example the comparative analysis of stimulated oocytes from heifers, young cows and old cows in one run set as well as the comparison of *in vivo* and *in vitro* matured oocytes within individual age groups.

The numbers of proteins identified and quantified in each run set can be found in Table 3. Reasons for the relatively large differences (up to 40 %) in numbers of identified and quantified proteins per run set are most likely based on different methods of sample storage. Oocyte aspiration for this experiment was started at a date when the optimal way to store very small numbers of approximately ten oocytes per sample was not finally established. Therefore some samples were stored in small glass capillaries while others were stored in sample cups. Hence, different protein amounts may have been obtained during the lysis procedure causing the alterations in numbers of identified and quantified proteins.

4.4.1 Glyoxylase 1, Ubiquitin and Bisphosphoglycerate mutase are candidate proteins for affecting the developmental potential of *in vitro* matured oocytes

The comparison of *in vivo* versus *in vitro* matured oocytes revealed nine proteins to be differentially abundant in the heifers group (Table 4), and nine proteins in the old cows group (Table 6). Four proteins, LAMP1, LAMP2, GLO1 and Ubiquitin, are altered by a fold change of > 2 in *in vivo* versus *in vitro* matured oocytes in the heifers as well as in the old cows group.

Since *in vitro* matured oocytes are known to be impaired in their developmental competence, a fact which was moreover confirmed by the evaluation of blastocysts and cleavage rates of oocytes from this experiment [128], the higher levels of these proteins in IVM oocytes are considered to be involved in affecting their developmental potential.

Ubiquitin usually carries out two different activities, namely modifying proteins, which will in the following either be degraded by the proteasome and serves to attract other proteins for initiating signalling cascades [129]. Since no peptides besides the C terminal end of Ubiquitin were identified, it is not possible to decide whether Ubiquitin was identified as a free molecule or covalently bound to other proteins. Its abundance increase in IVM oocytes could indicate higher abundance levels of proteins that are either misfolded or need to be replenished. A similar phenomenon was regarded during an analysis of murine COPD lung tissues, in which animals with severe emphysema had a higher accumulation of ubiquitinated proteins as compared to lower affected animals [130].

Glyoxylase 1 (GLO1) is an enzyme responsible for the degradation of methylglyoxal [131]. Methylglyoxal represents the aldehyde form of pyruvic acid and is a side product of glycolysis. Considering that glucose metabolism plays a pivotal role in determining the developmental competence of oocytes [132], a higher activation of glycolysis in *in vitro* than in *in vivo* matured oocytes can be supposed. It was already stated by Rieger et al. in 1995 that high rates of glucose metabolism may be unfavourable for development, after it was found that the metabolism of glucose was significantly greater in embryos which reached the 16-cell stage a whole day later than those with a lower glucose metabolism [133]. Moreover, it is already known that poor maternal health, resulting in increased intra-follicular glucose levels, like it is the case in diabetic mothers, is associated with poor oocyte viability [132]. Therefore, the increase of GLO1 could be a hint to an altered metabolic status inducing a loss of developmental competence of *in vitro* compared to *in vivo* matured oocytes.

During the evaluation experiments of cleavage and blastocyst rates, *in vitro* matured oocytes from young cows had higher cleavage and blastocysts rates than those from the other age groups [128]. Therefore, further evaluation of differentially abundant proteins between *in vivo* and *in vitro* matured oocytes within this age group is of special interest with respect to the developmental competence. Four proteins were found to be differentially abundant between *in vivo* and *in vitro* matured oocytes from young cows (Table 5). One of these proteins is Bisphosphoglycerate mutase (BPGM), which was two times more abundant in *in vivo* than in *in vitro* matured oocytes. Moreover, it was also found to be of lower abundance in *in vivo* versus *in vitro* matured oocytes from heifers. Usually, BPGM is found in erythrocytes, where it

synthesises 2,3-bisphosphoglycerate, which regulates the oxygen affinity of haemoglobin [134-136]. In 2006, it was demonstrated that its expression is not only limited to erythrocytes, since it was also found to be synthesized in non-erythroid cells of the human placenta [137]. Interestingly, BPMG seems to be a so called “**maternal house keeping protein**” (MHKP). MHKPs were defined as proteins, which are newly synthesized from maternal mRNA during oocyte maturation and also during the development of preimplantation embryos up to the 8-cell stage [113]. This indicates that the translation maintenance of these proteins is of special importance for correct embryonic development. During meiotic maturation, the time of protein synthesis is limited to the period of metaphase I until it declines to a basal level in metaphase II [138]. Due to the lower levels of BPMG in *in vitro* matured oocytes, it can be assumed that translation of BPMG is impaired or that its degradation is enhanced during the *in vitro* compared to the *in vivo* maturation process. Since *in vitro* matured oocytes are of lower developmental potential than *in vivo* matured oocytes, BPGM could be a protein executing crucial functions during early embryo development. Therefore, the low abundance of BPGM in *in vitro* matured oocytes could be involved in affecting their developmental potential.

In the comparison of *in vitro* matured oocytes from different age groups, the protein “Lysosomal-associated membrane protein 2 isoform 2” (LAMB2) was higher abundant in heifers than in young cows (fold change 2.7) and in old cows (fold change 3.6). Since the developmental competence of oocytes depends not only on the maturation process (*in vivo* versus *in vitro*) but also on the age of oocyte donors [127], LAMP2 could be involved in altering the developmental potential. Due to the fact that during the evaluation of cleavage and blastocyst rates from *in vitro* matured oocytes a higher blastocyst rate in young cows than in heifers was observed, one could assume that high LAMP2 concentrations are detrimental to the developmental competence of oocytes [128]. This hypothesis would be supported by the fact that LAMP2 was also increased in *in vitro* matured oocytes from heifers and old cows compared to their *in vivo* matured counterparts. However, only a slight increase in blastocyst rates in oocytes from old cows compared to that from heifers was observed [128], although LAMP2 was increased nearly by the same fold change in heifers versus old cows as in heifers versus young cows. Therefore, it is also possible that higher LAMP2 concentrations in oocytes from heifers are not associated with the developmental competence of oocytes.

The evaluation experiment of cleavage and blastocyst rates revealed that age associated differences could not be observed in *in vivo* matured oocytes, in contrast to *in vitro* matured oocytes [128]. This is consistent with the finding from the proteome analysis, that no differentially abundant proteins were found in the comparison of oocytes from different age groups in the group of *in vivo* matured oocytes. This result suggests, that age associated

differences are either not as abundant on the protein level in *in vivo* matured oocytes, or that they are evened out by FSH superstimulation, which was used for their generation.

For the statistical evaluation of differentially abundant proteins, it would have been beneficial if more than three biological replicates were analysed. Since oocytes (i) were not only obtained for proteomic experiments but also for IVP studies and histological analyses, and (ii) since in case of *in vivo* matured oocytes, a collection interval of four to five weeks was mandatory to follow a correct FSH stimulation protocol, the period of oocyte aspiration endured for about 1.5 years. The generation of more than the three replicates would have required employing new animals, due to the fact that animals did not match the age group criteria anymore after 1.5 years of their use in the experiment. Due to these limitations in the statistical analysis protein candidates from this experiment will have to be further validated, for example by targeted quantification using the SRM technology.

4.5 Qualitative and quantitative proteome analysis of oocytes, 2-cell stage embryos and morulae

The development from a 2-cell stage embryo into a morula takes about three days. During this period, the embryonic genome is activated (**maternal embryo transition (MET)**) and the first cell fate decisions are made. Within a morula, the outer cells, which have contact to the Zona pellucida, are already determined to develop into the trophectoderm cells, which are the progenitor cells of the placenta, while the inner cells will form the **inner cell mass (ICM)**, from which the embryo develops (for detailed description see chapter 1.3.2). To detect protein alterations coming along or are even induced by these processes, 50 2-cell stage embryos and 50 morulae were compared by 2D nano LC-MS/MS analysis. A total of 318 proteins were identified from 2-cell-stage embryos and 348 proteins were identified from morulae. These protein lists represent the first qualitative proteomic dataset from bovine 2-cell stage embryos and from morulae in general. The impact of these datasets to biological knowledge has been already discussed in chapter 4.1.

The identified proteins from these embryonic stages were compared to those of oocyte proteins (Fig. 15). Although the oocyte dataset contains far more proteins (829), there are still 94 proteins which have been identified from morulae and/or from 2-cell stage embryos but not from oocytes. 223 proteins could be identified from all stages. The overlap of only 223 proteins in all three stages is a clear hint to intensive proteome alterations during the developmental period from the GV oocyte to the morula stage. A total of 52 proteins were detected in 2-cell stage embryos, but not in GV oocytes. These proteins could become either (i) accumulated

during the time of GVBD when translation is enhanced during the maturation process [138] or (ii) translated from stored mRNA after fertilization due to the fact that transcription is relatively silenced until the onset of MET [36]. Furthermore, 254 of the identified proteins could be found both in 2-cell stage embryos and morulae. A total of 64 proteins were exclusively identified in 2-cell stage embryos while 94 proteins were exclusively identified in morulae. These differences are coherent with the subsequent degradation of maternal proteins and with the MET, leading to transcription and translation of genes from the embryonic genome [36].

The datasets from GV oocytes, 2-cell stage embryos and morulae were subjected to a GO analysis, the result of which is shown in Fig. 16. The percentage of proteins which are assigned to the different gene ontology categories is similar in all three stages. As demonstrated in Fig. 16, proteins assigned to binding and catalytic activity represent the biggest fraction in the molecular function category. Within the proteins of catalytic activity, especially oxidoreductases and hydrolases are overrepresented. Therefore, corresponding processes seem to be important during early embryonic development. Especially proteins associated to redox processes are already known to play a role during early embryonic development [139] and are furthermore subjects to alterations during the morula to blastocyst transition as discussed in chapter 4.6.

To quantitatively compare the datasets from 2-cell-stage embryos and morulae they were subjected to a so called “spectral count” analysis. The method “spectral counting” is based on a connection between a protein’s abundance and the number of MS/MS spectra assigned to an identified protein [86, 87, 140]. This rather simple method facilitates a quantitative comparison between unlabelled samples which have been measured in two different LC-MS/MS runs (see chapter 1.5.2). As listed in Table 9, 28 proteins were altered in abundance by a “spectral count fold change “of < 0.5 or > 2 .

However, spectral count is a rather simple but fairly crude method of quantification, demanding a validation of the results by an independent approach [141].

For this reason, the abundance alterations of two of the differentially abundant proteins, being especially interesting with respect to early embryonic development, were validated by the so-called “SRM” technique. SRM stands for “selected reaction monitoring” and is a mass spectrometry based method for a precise and highly specific quantification of proteins. In contrast to protein quantification by the holistic spectral count approach, SRM is used for the targeted quantification of known candidate proteins. The abundance alterations of i) Y-box protein 2 (YBX2) and ii) the IGF 2 mRNA-binding protein 3 (IF2B3/IMP3) between 2-cell stage embryos and morulae were analysed in three biological replicates, each containing ten

embryos per stage. In Fig. 19 an example of the quantification of the two proteins in ten 2-cell stage embryos is shown, demonstrating high signal intensities sufficient for quantification.

4.5.1 The decrease of YBX2 in morulae could rely on its maternal RNA storage function

YBX2 is an mRNA binding protein, which is the major and functional component of “messenger ribonucleoprotein complexes” (mRNPs). It protects the maternal mRNA storage from translation and degradation in mouse and *Xenopus* oocytes [142-144]. YBX2 mRNA and protein expression levels had already been analysed in bovine oocytes and preimplantation embryos [145, 146]. Both in mouse and in *Bos taurus*, YBX2 is highly abundant in the cortex of early oocytes and early cleavage stages. It decreases after the MET, which occurs between the 8- and 16-cell stage in the cow. This expression pattern was explained by YBX2's function in storage maintenance of maternal mRNA, which is no longer required in a transcriptional active stage [145]. The results presented here do not only confirm those findings but moreover specify the relative abundance of YBX2 in morulae to 10 % of its abundance in 2-cell stage embryos.

4.5.2 The decrease of IMP3 in morulae could be involved in first cell determinations

The protein IMP3 (IGF 2 mRNA-binding protein 3) belongs to a family of foetal proteins, the IMP family, that are capable of strong and specific mRNA binding. So far the mRNAs of β -actin, VG1, C-MYC, H19, TAU and Insulin like growth factor 2 (IGF2) have been reported as IMP binding targets [147-152]. Similar to YBX2, members of this protein family can mainly be found in mRNP complexes [153] and take part in translational control as well as RNA stability and trafficking [151, 154]. Moreover, they are also involved in forming cell polarity, migration and proliferation [154] and are expressed in various cancer cell lines [155-157]. The IMP expression levels decrease soon after birth, and in adults they have been detected only in testis and placenta [158]. Taken together, these facts indicate an important role for IMPs during embryonic development, yet this role is controversially discussed. For example, IMPs specifically bind and repress translation of the IGF-2 - mRNA , which is known to enhance proliferation. However, IMP expression is also positively correlated with cell proliferation [158].

The results presented here are the first analyses of an IMP family member abundance alteration during bovine embryo development. In the morulae IMP3 is decreased to half of its abundance in 2-cell stage embryos. Like the expression pattern of YBX2, the decrease of IMP3 can be

explained with its function in storage maintenance of maternal mRNA, which is no longer required in a transcriptional active stage.

This role was more specified by Nielsen et al.[151]: *IMP* mRNA levels have been analysed during embryonic development in mice and showed a sharp increase until day E.12.5 followed by a decline towards birth. The increase of IMP1 was also confirmed on the protein level. However, Nielsen et al. did not analyse the abundance alterations of IMPs in embryonic stages occurring prior to E.8.5. Therefore, the decrease of IMP3 as in the results presented here is not contrary to the results of Nielsen et al. Moreover, a similar tissue-specific distribution of IMPs and IGF2 was observed which led the authors, especially on the background of their findings of the IMPs' ability to suppress IGF2, to the suggestion that IMPs participate in the physiological regulation of IGF2's foetal production. Therefore, the authors assumed that IMPs might prevent a too early initiation of the expansion or differentiation of a particular cell line [151]. When considering that *Igf2*^{-/-} mutants are indeed infertile and smaller than the normal phenotype, whereas mice with high IGF2 levels express severe malformations this becomes even more likely [159]. The decrease of IMP3 in morulae to the half of its abundance in 2-cell stage embryos as detected by the experiment performed for this thesis, is also coherent with the hypothesis of Nielsen et al.: A role in the prevention of too early induction of cell determinations between the 2-cell and the morula stage can be assumed for the bovine species as well. Hence, the decrease of IMP3 in morulae, where first cell determinations are definitely known to be already made (see chapter 1.3.2) would mean that prevention of cell determinations through IMP3 is no longer needed. If this was the case, further analyses about a more exact course of IMP3 alterations between the bovine 2-cell and the morula stage (which was not analysed by Nielsen et al.) could be helpful to determine the developmental stage of first cell differentiations. Moreover, this would be interesting because IMPs are known to be involved in the forming of cell polarities [154]. Cell polarities in turn are important for the induction of the first lineages decisions in the early embryo (see chapter 1.3.2).

4.6 Differential proteome analyses of morulae and blastocysts

The morula to blastocyst transition is a crucial step in early embryonic development. During this period the presence of two different cell populations in the embryo becomes morphologically obvious. One cell population consisting of a flat cell layer forms the trophectoderm (TE) and the other one forms the “inner cell mass” (ICM). TE cells are progenitor cells of the placenta, while the ICM cells are progenitor cells of the embryo. A fluid filled cavity, the blastocoel, is also formed by merging of intercellular spaces (for a more

detailed description see chapter 1.3.2). In ARTs, the morula to blastocyst transition is usually performed *in vitro*. However, on average only two thirds of all morulae develop into blastocysts [160-162]. To further deepen the knowledge about the underlying biochemical processes, quantitative differences in protein expression levels between morulae and blastocysts were analysed. To maximize the analytical depth, two different techniques were applied, based on either 2D gel electrophoresis and spot intensity quantification, or mass spectrometry and peptide ion intensity quantification.

4.6.1 Saturation DIGE analysis of morulae and blastocysts revealed several protein isoforms of different abundance alterations

For the saturation DIGE analysis of morulae and blastocysts, 12 2D gels containing 500 ng protein each (corresponding to five embryos) were prepared. In Fig. 20 representative Cy5 gel images from morulae and blastocysts demonstrate the high separation strength of these gels in both dimensions. Moreover, these gels are free from under- or overlabelling artefacts, demonstrating the use of a correct dye/protein ratio (see chapter 1.5.1).

In total, 2024 signals were detected and quantified in all gels and can therefore be considered as signals from real protein spots. Differentially abundant spots had to be detected in all gel images. The analysis of six biological replicates enabled a valid statistical evaluation of these spots [163]. In an analysis where thousands of analytes are quantified, false positive detections of abundance alterations usually occur due to statistical reasons. To avoid these detections of false positive proteome alterations, a FDR correction of p-values according to Benjamini and Hochberg [124] was applied. Therefore, the small p-values below 0.05 represent significant abundance alterations. The quantification of these protein spots was highly reproducible as demonstrated in Fig. 22. This figure contains representative 3D intensity shape plots and from differentially abundant proteins which are similar in all biological replicates.

For the identification of protein spots, preparative 2D gels containing several hundreds of micrograms protein (corresponding to several thousands of embryos) are needed. Since the preparation of such high numbers of embryos is rather time and cost intensive, 4444 GV oocytes were used instead. Although the 2D gel spot patterns from oocyte and blastocysts are similar, aberrations in the intensity of singular spots occur. Especially spots containing proteins which have very low expression levels in oocytes cannot be visualized and picked. In total, 40 spots, which could unambiguously be matched between analytical embryo and preparative oocyte gels, were picked, from which 34 were identified. The remaining six protein spots did not contain enough protein for a successful identification by LC-MS/MS. In 18 protein spots, a

single protein was unambiguously identified (Table 10), in 11 spots two proteins (Table 11) and in five spots three or more proteins (Table 12) were identified. These multiple identifications can be explained by the fact that not all proteins can be separated from each other by isoelectric focussing and SDS PAA gel electrophoresis and are therefore detected from the same position (see chapter 1.5.1). In these cases of non-unique identifications from a single spot, it is impossible to decide which of the proteins is responsible for the detected intensity alteration.

Eleven proteins could be identified in more than one spot, most probably representing different isoforms of the corresponding proteins. Isoforms which were individually altered by different abundance alterations were for example detected from “Aldo-keto reductase family 1 member B1” (AKR1B1), Peroxiredoxin 2 (PRDX2), Prohibitin (PHB), “Chloride intracellular channel protein 4” (CLIC4) and “Stomatin-like protein 2” (STOML2).

4.6.2 The iTRAQ analysis led to 141 quantified and identified proteins

For the nano LC-MS/MS analysis of morulae and blastocysts, three 2D nano LC-MS/MS run sets as described in chapter 4.3 were performed. This setup enabled the multiplexed analysis of two biological replicates per run set (Table 14) leading to 141 proteins identified and quantified in all biological replicates. Differentially abundant proteins had (i) to be identified and quantified in all biological replicates and (ii) to have an intensity alteration of more than 1.5. The analysis of six biological replicates enabled a valid statistical evaluation of these proteins [163]. Student's t test p-values lower than 0.05 represent significant abundance alterations. In total, 50 proteins were found to be differentially abundant (Table 15). The reproducibility of quantification of these proteins is demonstrated in Fig. 24 in which HCD MS/MS spectra of the peptide derived from the protein Transketolase are shown. The intensity ratios of iTRAQ reporter signals are similar in all biological replicates. Moreover, the intensity distributions within the entire spectra are highly reproducible. The distribution of these spots according to their p-values and fold changes is shown in the Volcano Plot in Fig. 23.

To ensure correct identifications only proteins which were identified with at least two peptides were accepted. Furthermore a decoy search was performed. As described in chapter 4.1, this procedure leads to valid datasets with a FDR below 1 %.

4.6.3 The complementary characteristics of the saturation DIGE and the iTRAQ analyses are reflected by the comparison of morulae and blastocysts

Both the Saturation DIGE and the iTRAQ approach have been performed with six biological replicates containing 25 morulae and 25 blastocysts each. It was already demonstrated that more

than 2000 protein spots can be quantified from 2D gels containing only 500 ng protein by Berendt et al. [74]. In chapter 4.3 the method adaptation for the iTRAQ quantification of around 400 proteins from protein amounts obtained from 50 oocytes was already discussed. The results obtained by the application to the comparative proteome analysis of morulae and blastocysts of both techniques were compared:

Around 2024 protein spots have been quantified in each 2D gel, while 141 proteins have been quantified in each run set from the iTRAQ analysis. The obvious differences concerning the numbers of quantifications can be explained by the fact that quantification from 2D gels is performed on the protein level, while quantification in the iTRAQ analysis is performed on the peptide level. The 141 proteins correspond to a number of averagely 1000 quantified peptides, generated by tryptic hydrolysis of the proteins. Therefore, a number of 1000 quantified analytes (peptides in iTRAQ analysis) is the correct comparison to the number of 2024 analytes (protein spots in Saturation DIGE analysis). As a consequence of tryptic protein hydrolysis during the iTRAQ analysis, no information concerning the molecular weight and the isoelectric point (IP) of the protein can be acquired. Thus, the detection of protein isoforms is much less likely by iTRAQ based quantification. Yet, many isoforms are quantified as single spots on 2D gels what explains the double number of quantified analytes in the 2D DIGE approach. However, the main advantage of the iTRAQ approach is that the 141 quantified proteins were simultaneously quantified and identified, while the main disadvantage of the 2D DIGE analysis is the difficult protein identification requiring preparative gels. This was especially problematic in this analysis of early embryos, since a different developmental stage (GV oocytes) of sample material had to be used due to cost reasons. It led to the problem that some spots from proteins, which were too low abundant in oocytes, could not be assigned to their counterparts on the analytical gels and that spot identification was biased by the abundance alteration. Spots from proteins decreasing in abundance from oocytes to the blastocysts stage are more likely to be identified than spots from proteins increasing from oocytes to the blastocysts stage. Hence, the fact that 76 % of all identified protein spots showed decreased abundance alterations between morulae and blastocysts, cannot be interpreted as a biological phenomenon. In contrast, the fact that 66 % of all differentially abundant proteins from the iTRAQ analysis are higher abundant in blastocysts than in morulae allows a biological interpretation since protein identification is not biased by the abundance alteration. Furthermore, the problem of non-unique protein identifications from a single spot makes the interpretation of identified proteins even more difficult.

In total, 18 differentially abundant proteins were detected and unambiguously identified by the 2D DIGE analysis. This number could be enhanced by 47 proteins from the iTRAQ analysis. Seven proteins were detected in both approaches. They are listed together with their fold

changes in Table 16. Fig. 25 shows a comparison of HCD MS/MS spectra of Peroxiredoxin 2's peptide QVTINDLPVGR and corresponding 3D intensity shape plots from the 2D DIGE analysis, demonstrating the similarity of abundance ratios measured by both techniques. Furthermore, Fig. 25 demonstrates the reproducibility of quantification in both approaches. In contrast to six proteins which have similar abundance alterations in both approaches, one protein spot identified as "Glutathione S-transferase mu 3" (GSTM3) is increased by a factor 1.7 in the 2D DIGE analysis while the majority of the corresponding protein's peptides are decreased in abundance by a factor of 0.6 as reflected by the iTRAQ analysis. This again demonstrates the advantage of the 2D DIGE analysis in detecting abundance alterations of single protein isoforms. Opposing abundance alterations for Glutathione S-transferase mu 5 isoforms were also previously described in a 2D gel based analysis of GV and *in vitro* matured MII oocytes [74].

4.6.4 Gene ontology clustering of proteins indicates a switch from a catabolic to an anabolic stage as well as cell proliferation during the morula to blastocyst transition

David GO protein clustering shows in which functional categories proteins from an input list show up more than they are expected to appear from a randomized protein list. Thus the results indicate in which molecular and functional processes these proteins are involved [123].

Within this experiment, the David GO clustering revealed that nine proteins decreased in abundance in blastocysts are associated with catabolic processes, while only four proteins are associated with biosynthetic processes. In contrast, in the group of proteins increased in blastocysts, seven proteins are associated with biosynthetic and only two proteins are associated with catabolic processes. These findings indicate a switch from a catabolic to an anabolic state during the morula to blastocyst transition. This is also confirmed by the abundance alteration of single proteins. The two proteins "Suppressor of G2 allele of SKP1 homolog" (SUGP1) and the "UBFD1 protein", both involved in ubiquitination initiated protein degradation [164, 165], were identified from the same spot in the saturation DIGE analysis. The spot was altered in intensity by a factor 0.4 between morulae and blastocysts. The decrease of these proteins is a further hint to the termination of the degradation process of maternal proteins. These hypotheses can be strengthened by a study in which the overall protein content of bovine oocytes and pre-elongation stage embryos was measured. From the results it was concluded that the early cleavage stage bovine embryo has a higher rate of protein degradation than that of synthesis until initiation of compaction at the morula stage [166]. Besides this, especially ATP biosynthesis seems to increase between the morula to blastocyst transition since four of the

proteins increased in blastocysts are assigned to ATP biosynthetic processes. This is coherent with a finding of Thompson et al. that ATP production starts to increase with compaction of the morula and peaks at the blastocysts stage [167].

The increase of several other proteins in blastocysts can be considered as a result of the ongoing proliferation during blastocyst development. These are, for example, the four proteins which cluster into terms of chromosome or nucleosome organization. This can be explained by the rising number of cell nuclei during embryo development. Furthermore, the analysis also revealed that proteins involved in cytoskeletal organization are increasing. The three proteins from the analysis within this class are (i) Ezrin, which is involved in connections of major cytoskeletal structures to the plasma membrane [168], (ii) the “F-actin-capping protein subunit beta” (CAPZB) which binds to the fast growing ends of actin filaments [169] and (iii) “Cytoplasmic actin”. This could reflect the enhanced production of cytoskeletal structures.

4.6.5 Proteome analysis of morulae and blastocysts reflects increasing translation during the morula to blastocyst transition

Embryogenesis is characterized by a gradual degradation of maternally inherited messenger RNAs and activation of the embryonic genome, for example, initiation of embryonic transcription and translation of embryonic proteins (MET) [35-37]. The enhanced translation after the MET is also reflected in the proteome analysis between morulae and blastocysts on several levels: First, 66 % of all differentially abundant proteins from the iTRAQ analysis are increased in blastocysts. This is visualized in the Volcano Plot in Fig. 23. Second, in the GO clustering of proteins increased in blastocysts, eight proteins are assigned to the GO term translation. And third, proteins responsible for translation initiation like (i) Nucleophosmin (NPM1), (ii) the protein “Receptor of activated protein kinase C 1” (GNB2L1/RACK1) and (iii) the “Eukaryotic translation initiation factor 6” (EIF6) as well as (iv) the “Eukaryotic translation initiation factor 5A-1” (EIF5A) are increased in blastocysts in the analysis. NPM1, which directs the nuclear export of the ribosomal subunits leading to increased rates of protein synthesis [170] is increased by a factor 2.6 in blastocysts. RACK1 is increased in blastocysts by a factor 2.8. It is a ribosomal protein which recruits activated protein kinase C to the ribosome, leading to the stimulation of translation by phosphorylation of the EIF6 [171]. Interestingly, EIF6 was identified together with Clusterin from a mixed spot increased by a factor 2.4 in the saturation DIGE analysis. Since both proteins were unambiguously identified in the spot, it is not possible to decide which of the proteins is responsible for the intensity alteration. Yet, this is an example for a case in which regarding the biological context of such proteins can be a hint to

a likely responsibility for the abundance alteration to one of the proteins. The ongoing translation and the increase of RACK1 allow assuming, that EIF6 is at least involved in the increased abundance alteration of this spot between morulae and blastocysts.

4.6.6 Creatin kinase B (CKB) and Annexin A6 are promising candidates in the molecular studies of the morula to blastocyst transition

CKB is of special interest due to the fact that it was found to be by a factor 5.9 higher abundant in blastocysts than in morulae in the iTRAQ analysis. It belongs to the proteins which show the highest abundance alterations of all differentially abundant proteins. Interestingly, in other tissues, especially in the rat uterus, CKB mRNA expression is induced by oestrogen [172, 173]. Since the media surrounding the embryos was not changed during the embryo culture process, the CKB expression was not induced by additional external oestrogen. Hence, an early onset of oestrogen production by the developing embryo might have induced the rapid increase of CKB. Especially the trophoctodermal tissues from later embryonic stages are already known to produce placental estrogens, the role of which is still not understood in cattle [174]. To further evaluate this hypothesis, localisation studies which could reveal whether CKB expression is restricted either to the TE or to the ICM would be helpful. However, it is also possible that blastomeres become reactive to oestrogen, already contained in the medium, during the morula to blastocyst transition. This connection as well as the connection of CKB and oestrogen expression in general, could be studied by measuring the CKB expression at different developmental stages before and after oestrogen supplementation.

Another interesting candidate protein for further functional analyses during the ongoing blastocyst development is Annexin A6. It was found to be increased in blastocysts by a factor of 4.3 in the iTRAQ analysis. As reviewed by Grewal et al. [175], this protein has several functions which could play important roles in preimplantation embryonic development. It (i) provides a scaffold to form membrane-bound multifactorial signalling complexes, (ii) it is able to regulate transient membrane-actin cytoskeleton interactions during endocytosis, (iii) it stabilises the protein/lipid composition during membrane microdomain formation and (iv) it regulates secretory events. Especially interesting would be to further evaluate its role in spatio-temporal signalling during embryogenesis.

4.6.7 The increase of the LGALS3 protein, RACK1 and Ion transport associated proteins could reflect the embryo remodelling during blastocyst development

At the morula stage, the embryo starts a remodelling process: Cells change their shape with a remodelling of membrane lipids and become much more tightly apposed, a process known as compaction. Shortly afterwards, the embryo starts its first differentiation process by forming a flat layer of TE cells, a group of cells representing the ICM and a fluid-filled blastocoel by merging of intercellular spaces (see chapter 1.3.2).

The process of merging intercellular spaces happens by mechanisms that regulate intracellular osmolarity and cell volume through the controlled release of osmolytes from the cytoplasm via ion channels. Therefore, the increase of seven proteins, which were associated with ion transport (according to David GO clustering) in blastocysts, is coherent with the ongoing blastocyst development [38].

The “LGALS3 protein” is increased in blastocysts by a fold change of 2.8. It is a carbohydrate-binding-protein [176], named Galectin 3 in other species. It was shown to be involved in the prevention of apoptosis by being translocated to the mitochondria and blocking changes in the mitochondrial membrane potential [177]. Interestingly, Galectin 3 can especially prevent an apoptosis form induced by the loss of cell anchorage (anoikis). Such cells have been shown to undergo cell cycle arrest instead [176]. During the embryo remodelling process at the morula to blastocyst transition at least the transient loss of cell anchorages seems possible. High Galectin 3 levels might be essential to prevent so caused apoptosis, thereby keeping these cells alive until their function can be replaced by new cells. Especially the cellular environment of TE cells is affected by the remodelling process, since they form a flat single cell layer in the blastocyst, inwards facing the fluid filled blastocoel, instead of being surrounded by blastomeres from three sides in the compacted morula. The survival of the blastocysts depends on the correct forming of “tight junctions” between the TE cells which prevents water leakage from the blastocoel. Therefore apoptosis of one of the single layer TE cells would probably impair the embryo survival. Considering Galectin 3’s function in apoptosis, this could be the reason that galectin 3 mRNA expression in the mouse embryo is known to occur during the fourth day of gestation in TE cells of blastocysts [178]. To further evaluate a cellular restriction to TE cells in the bovine embryo localisation studies would be interesting.

The protein RACK1 is increased in blastocysts by a factor 2.8. Its involvement in translation initiation has been discussed above. Besides that it regulates several processes that involve contact with the extracellular matrix, such as cell spreading, the establishment of focal adhesions and cell–cell contacts [171]. Since it is known that first cell fate decisions in the

embryo partly depend on the location of the blastomeres (inside or outside the morula), and on the cellular connections [13, 39, 53, 179, 180], these processes seem to be important during embryo development. Therefore, it could well be that RACK1 is involved in blastocyst formation, which involves the loss of intercellular contacts as well as the establishment of contact to the fluid-filled blastocoel cavity.

4.6.8 The protein NPM1 could be involved in first lineage decisions

As mentioned above, NPM1 is increased in blastocysts by a factor of 2.6 and is involved in translational processes. Besides that, also its involvement in cell proliferation [181] is coherent with ongoing cell divisions during blastocyst development. Even more interesting is that NPM1 was recently found to be involved in embryonic stem cell (ESC) maintenance and ESC fate determination: It can form complexes with OCT4, SOX2 and NANOG [182] which are essential factors for maintenance of pluripotency [183] and are involved in the first lineage decisions (see chapter 1.3.2). NPM1 downregulation leads to mesoderm and ectoderm differentiation [182]. For these reasons, the increase of NPM1 in blastocysts could be both involved in the cell fate's decisions to become a TE or an ICM cell and prevent the preimplantation embryo's cells from too early differentiation into the different germ layers.

4.6.9 The reduction of redox enzymes may alter HIF dependent gene regulation

In the group of proteins decreased in blastocysts, 12 proteins associated with redox processes were found. It is already known that the embryo possesses several antioxidant systems which have two main functions: fine-tuning of "reactive oxygen species" (ROS) levels for signalling purposes and protection from oxidative stress [184].

Concerning the signalling aspects of fine tuning ROS levels, a hypothesis would be that this alters gene regulation over a transcription factor family, the so-called "hypoxia inducible factors" (HIFs). They regulate gene expression in dependence of the cell's redox state, yet can be also altered under normoxic conditions by ROS. Its gene activity leads to the up-regulation of a variety of genes which are, beside others, involved in energy metabolism and cell proliferation [139]. Since these are important processes in early embryonic development, it is possible that the reduction of redox enzymes in blastocysts is required for a HIF induced gene regulation over ROS species. This hypothesis is supported by Harvey et al. who already supposed that the activation of HIF might be important for bovine *in vitro* embryos to adapt to a reduced oxygen environment normally encountered *in vivo* [139].

Moreover, within the group of redox enzymes, the three PRDX isoforms 1, 2 and 6 are found to be decreased by a factor between 0.4 and 0.6 in morulae versus blastocysts. Peroxiredoxin enzymes are ubiquitously present in various cell types and contain a conserved cysteine residue in the N-terminal region that is the primary site of oxidation. Mammalian cells express six isoforms of PRDX (PRDX1 to 6) which are thought to remove low levels of peroxides produced as a result of normal cellular metabolism [185]. An abundance alteration of PRDX3 during *in vitro* maturation of bovine oocytes has previously been observed [74], already indicating the involvement of this enzyme class in developmental processes. Furthermore, *PRDX* mRNA expression was previously quantified in early bovine embryos of different developmental stages [184]. While this analysis revealed the existence of *PRDX1* transcripts in all analysed stages (from oocyte to blastocyst), *PRDX2* and *PRDX6* transcripts could not be detected in 9- to 16-cell stage embryos, but again at the morula stage. From this result, the authors concluded a complete degradation of stored maternal *PRDX2* and *PRDX6* transcripts prior to the MET which occurs between the 8-cell and the 16-cell stage (see chapter 1.3.2). In contrast to that, our analysis clearly shows a decrease of PRDX2 and PRDX6 up to the blastocysts stage, demonstrating that the increase of PRDX2 and PRDX6 mRNA is not reflected on the protein level at these stages. This could be due to a delay between transcription and translation as well as an enhanced degradation of PRDX2 and PRDX 6. It is imaginable that also the timing of this delay between transcription and translation is important for correct blastocyst development, possibly also for correct HIF regulation over ROS. Moreover, PRDX2 has been identified from three different protein spots, all decreasing but differing in their abundance alteration. These different spots most probably represent different isoform variants of this protein. Therefore, this finding could indicate that especially a correct concentration of each PRDX2 isoform has an impact for the morula to blastocyst transition.

Although the reduction of redox enzymes might be needed to increase ROS levels for HIF regulation, high ROS levels can induce lipid peroxidation-derived lipid aldehydes and their glutathione-conjugates which are in the following reduced by “Aldo-keto reductase” (AKR1B1) to corresponding alcohols mediating inflammatory signals [186]. Hence, low levels of AKR1B1 can prevent these signals from being transmitted, which might be the reason for its decrease in blastocysts as detected by the saturation DIGE analysis. This hypothesis is supported by a study in which *AKR1B1* mRNA levels were found to be of higher abundance in biopsies derived from blastocysts, leading to resorption and pregnancy failure than from biopsies leading to calf delivery [187]. Besides that, AKR1B1 has two different activities, namely metabolizing progesterone and synthesizing PGF_{2α} [188], both processes with the ability to terminate

pregnancy, so that high AKR1B1 concentrations seem not to be beneficial for a successful pregnancy in general.

5 Perspectives

During this work, several proteins were discovered to be altered in abundance between *in vivo* and *in vitro* matured oocytes and between different stages of the preimplantation embryo development. These proteins are considered to be interesting candidates for further studies, concerning their participation in molecular processes underlying the developmental competence of oocytes and the first cell fate decisions after fertilization. It is imaginable that further hints on an involvement in cell fate decisions could be gained through localisation studies by immunofluorescence microscopy in blastocysts. If any of the proteins, like for example “Creatin kinase B” or the “LGALS3 protein” are specifically localised to trophoctoderm cells, a further evaluation of their stage dependent localisation could lead to new insights into lineage decisions. Moreover, the protein “Bisphosphoglycerate mutase” (BPMG) was detected with higher abundance in *in vivo* than *in vitro* matured oocytes, which are known to have an impaired developmental competence compared to *in vivo* matured oocytes. Hence, the functional analysis of BPMG in embryo development could help to evaluate the developmental potential of oocytes. However the abundance alteration of BPMG will have to be validated by the selected reaction monitoring (SRM) technique, due to the fact that the three biological replicates available for analysis did not allow a valid statistic evaluation.

Furthermore, the SRM technique will be used to assign the abundance alterations of interesting candidate proteins not only to a developmental period, like for example between 2-cell stage embryos and morulae, but to shorter periods like day to day intervals. This could help to understand whether abundance alterations of proteins do occur as a consequence of certain biological events (for example the maternal embryo transition or cell fate decisions) or are involved in their induction. Moreover, the SRM technique will be employed for absolute protein quantification and could thus lead to the determination of physiological expression levels for certain proteins. Since targeted protein quantification by SRM needs lower protein amounts for valid quantifications than a holistic approach, it is possible to also analyse abundance alterations of proteins in human samples, which are available in very limited numbers only.

6 Summary

Assisted reproduction techniques (ARTs) become increasingly important both in veterinary and in human medicine. Since ARTs are still associated with low conception and birth rates, it is necessary to improve the understanding of underlying biochemical processes. The presented work addresses, on the proteome level, the oocyte maturation process as well as the early embryonic development up to the blastocyst stage, since both oocyte quality and a faultless development of the early embryo is crucial for the success of assisted reproduction.

As a model organism, the bovine system was used since it reflects crucial parameters of the female human reproductive system. Besides, the bovine reproductive system is of particular interest with respect to basic research as well as to aspects targeting the efficiency enhancement of ARTs for cattle production.

The correct storage of mRNAs and proteins is an important feature in determining oocyte quality. To find out which proteins are stored after their production during oogenesis, a comprehensive qualitative proteomic dataset was generated from 900 bovine GV oocytes. The dataset consisted of 829 proteins, covering a broad spectrum of different molecular functions and processes. For a total of 621 proteins within this dataset the evidence of protein existence was not demonstrated in the bovine system prior to this experiment. Moreover, it was revealed that proteins involved in oocyte maturation and such proteins which are stage dependently altered during the development up to the blastocyst stage can already be found in GV oocytes.

In addition, experimental data addressing protein abundance alterations during oocyte maturation and early embryogenesis are crucial to deepen the understanding of underlying biochemical processes. However, the quantitative analysis of proteomes from mammalian oocytes and early embryos is challenging due to the very small sample amounts available (one oocyte/early embryo contains only 90 ng protein). Therefore, techniques for proteome analyses were successfully downscaled to facilitate identification and relative quantification of up to 403 proteins from the protein amount corresponding to only 50 oocytes/early embryos. These methods were then applied to the quantitative proteome analyses of oocytes and early embryos:

Due to the negative effect of hormonal superstimulation to maternal health in humane medicine it is tried to replace traditional “in vitro fertilization” (IVF) based on *in vivo* matured oocytes of superstimulated donors, by the use of IVF with *in vitro* matured oocytes. However, low success rates suggest an impaired developmental competence of *in vitro* matured oocytes. Hence, quantitative differences between the proteomes of *in vivo* and *in vitro* matured oocytes were investigated. Amongst several differentially abundant proteins, “Bisphosphoglycerate mutase”, known as a “maternal house keeping protein“ was found to be of decreased abundance in *in*

vitro matured oocytes, while for example “Glyoxylase 1” and Ubiquitin were higher abundant in *in vitro* versus *in vivo* matured oocytes. These results suggest an altered metabolic status as well as an increased abundance of proteins, which were incorrectly processed or should be replaced due to other reasons.

A key process in a successful pregnancy is a correct pre-implantational embryonic development. For this reason, the proteome signatures of relevant stages of early embryonic development were quantitatively analysed:

The developmental interval between 2-cell stage embryos and morulae, which comprises the “maternal embryo transition”, was analysed by an LC-MS/MS spectral count approach. In total, 28 proteins were found to be differentially abundant. The abundance decrease of the proteins “Y-box protein 2” and the “IGF 2 mRNA-binding protein 3” between 2-cell stage embryos and morulae, being particularly interesting with respect to early embryonic development due to their roles in translational control, was validated by the “selected reaction monitoring” (SRM) technique.

The most comprehensive experiment was performed on morulae and blastocysts. During the developmental interval between these stages, the first cell fate determinations become morphologically obvious. Morula and blastocyst proteomes were compared using two complementary methods, the 2D gel based saturation DIGE technique and the LC-MS/MS based iTRAQ technique. In six biological replicates, 61 spots from differentially abundant proteins were found by the 2D DIGE approach and 50 differentially abundant proteins by the iTRAQ analysis. In particular, proteins associated with translation and biosynthetic processes were increased during the morula to blastocyst transition. Several interesting candidate proteins which represent promising candidates for functional studies and localization experiments were discovered. For example, “Creatin kinase B” was increased in blastocysts by a factor of 5.9, which could indicate an early onset of oestrogen expression by trophoctodermal cells. Within the group of proteins being decreased in blastocysts, a notably high number of proteins associated to redox processes were found. These proteins could alter genes of energy metabolism and cell proliferation by the regulation of the “hypoxia inducible factor” by fine tuning of “reactive oxygen species”.

These experiments represent the most comprehensive holistic proteome analyses of bovine oocytes and preimplantation embryos available so far. The results facilitate now to analyse interesting protein candidates by localisation studies and targeted quantification by SRM to elucidate a possible involvement of these proteins in cell fate decisions and the induction of certain biological events (e.g., embryonic genome activation).

7 Zusammenfassung

Qualitative und quantitative Proteomanalyse von bovinen Oozyten und frühen Embryonalstadien

Die Bedeutung **Assistierter Reproduktionstechniken (ARTs)** nimmt sowohl in der Tier- als auch in der Humanmedizin mehr und mehr zu. Da diese jedoch immer noch mit niedrigen Konzeptions- und Geburtenraten verbunden sind, ist es notwendig, das Verständnis zugrunde liegender biochemischer Prozesse zu verbessern. Die hier präsentierte Doktorarbeit adressiert sowohl den Oozytenreifungsprozess als auch die frühe Embryonalentwicklung bis zum Stadium der Blastozyste, da sowohl die Oozytenqualität wie auch die fehlerfreie Embryonalentwicklung entscheidende Schritte für den Erfolg von ARTs darstellen.

Als Modell wurde das bovine System verwendet, da dieses entscheidende Parameter der weiblichen humanen Reproduktionsbiologie widerspiegelt. Darüber hinaus ist das bovine System von großem Interesse für die Grundlagenforschung sowie für angewandte Studien zur Verbesserung der Effizienz von ARTs.

Die korrekte Speicherung von Proteinen und mRNA ist ein wichtiges Charakteristikum der Oozytenqualität. Um herauszufinden, welche Proteine nach ihrer Produktion in der Oogenese gespeichert werden, wurde ein umfangreicher qualitativer proteomischer Datensatz aus 900 bovinen GV Oozyten generiert. Dieser Datensatz enthält 829 Proteine, die ein breites Spektrum verschiedener molekularer Prozesse und Funktionen abdecken. Für insgesamt 621 Proteine aus diesem Datensatz wurde zuvor kein Nachweis ihrer Existenz im bovinen System auf Proteinebene geführt. Darüber hinaus wurde gezeigt, dass sowohl Proteine, die an der Oozytenreifung beteiligt sind, als auch solche, die in der Entwicklung bis zur Blastozyste stadienabhängig in ihrer Abundanz verändert werden, bereits in der GV Oozyte vorhanden sind. Daneben sind vor allem Experimente, welche die Abundanzveränderungen von Proteinen während der Oozytenreifung und frühen Embryonalentwicklung erfassen, entscheidend, um das Verständnis zugrunde liegender biochemischer Prozesse zu vertiefen. Aufgrund der geringen Verfügbarkeit von Probenmaterial (eine Oozyte/ein früher Embryo enthält nur 90 ng Protein) stellt die quantitative Proteinanalyse von Säugeroozyten und frühen Embryonen jedoch eine Herausforderung dar. Deshalb wurden im ersten Teil dieser Arbeit Techniken zur Proteomanalyse erfolgreich an die Analyse von geringen Probenmengen angepasst, so dass bis zu 403 Proteine aus einer Proteinmenge, die der von 50 Oozyten/frühen Embryonen entsprach, identifiziert und quantifiziert werden konnten. Diese Techniken wurden dann zur quantitativen Proteomanalyse von Oozyten und frühen Embryonalstadien herangezogen.

Wegen des negativen Effekts von hormoneller Superstimulation auf die mütterliche Gesundheit wird versucht, die traditionelle „in vitro Fertilisation“ (IVF) mit *in vivo* gereiften Oozyten von superstimulierten Spendern durch IVF mit *in vitro* gereiften Oozyten zu ersetzen. Die niedrigen Erfolgsraten deuten jedoch auf eine beeinträchtigte Entwicklungskompetenz von *in vitro* gereiften Oozyten hin. Daher wurden die Unterschiede in der Proteinexpression zwischen *in vivo* und *in vitro* maturierten Oozyten untersucht. Innerhalb mehrerer differentiell abundanter Proteine wurde das Protein „Bisphosphoglycerat Mutase“, welches als „maternales house keeping Protein“ bekannt ist, mit erniedrigter Abundanz in *in vitro* gereiften Oozyten detektiert, während beispielsweise „Glyoxylase 1“ und Ubiquitin höher abundant in *in vitro* als in *in vivo* gereiften Oozyten waren. Diese Befunde deuten sowohl auf einen veränderten metabolischen Status als auch auf die erhöhte Präsenz von Proteinen, die entweder nicht korrekt prozessiert wurden oder aus anderen Gründen ersetzt werden sollten, in *in vitro* gereiften Oozyten hin.

Eine Schlüsselrolle für eine erfolgreiche Trächtigkeit spielt der korrekte Ablauf der Embryonalentwicklung während der Preimplantationsphase. Deshalb wurden die proteomischen Signaturen relevanter früher Embryonalstadien quantitativ analysiert.

Zunächst wurde das Entwicklungsintervall zwischen 2-Zell-Embryonen und Morulae, welches den Start der embryonalen Genomaktivierung beinhaltet, mittels eines LC-MS/MS basierten, so genannten „Spectral Count“ Ansatzes, verglichen. Insgesamt wurden 28 differentiell abundante Proteine zwischen 2-Zell-Embryonen und Morulae detektiert. Die abnehmende Abundanz zwischen 2-Zell-Embryonen und Morulae der zwei Proteine „Y-box protein 2“ und „IGF 2 mRNA-binding protein 3“, welche aufgrund ihrer Funktion in der Translationskontrolle im Kontext der frühen Embryonalentwicklung besonders interessant sind, wurde mit der „Selected Reaction Monitoring“ (SRM) Technik validiert.

Das umfangreichste Experiment wurde anhand von Morulae und Blastozysten durchgeführt. In dem Entwicklungsintervall zwischen diesen Stadien werden die ersten Zelldifferenzierungen morphologisch sichtbar. Die Proteome von Morulae und Blastozysten wurden mit zwei komplementäre Methoden, (i) der 2D Gel basierten „Saturation DIGE“ Technik und (ii) der LC-MS/MS basierten „iTRAQ“ Technik, vergleichend quantitativ analysiert. In sechs biologischen Replikaten wurden 61 Protein Spots von differentiell abundanten Proteinen mit der 2D DIGE Technik und 50 differentiell abundante Proteine durch den Einsatz der iTRAQ Technik detektiert. Insbesondere waren solche Proteine, die mit Translations- und biosynthetischen Prozessen assoziiert sind, in ihrer Abundanz in Blastozysten erhöht. Einige interessante Proteine, welche vielversprechende Kandidaten für Lokalisations- und funktionelle Studien sein könnten, wurden entdeckt. Zum Beispiel war die „Kreatin Kinase B“ in Blastozysten gegenüber

Morulae um einen Faktor von 5.9 erhöht, was auf einen frühen Beginn der Östrogenproduktion in den trophoctodermalen Zellen hindeuten könnte. Innerhalb der Gruppe von Proteinen, die in ihrer Abundanz in Blastozysten gegenüber Morulae erniedrigt sind, war eine große Anzahl mit Redoxprozessen assoziiert. Diese Proteine könnten die Expression von Genen, die für Metabolismus und Zellproliferation relevant sind, über eine Beeinflussung der „Hypoxie induzierten Faktoren“ durch eine Feinregulierung von reaktiven Sauerstoffspezies verändern.

Die in dieser Arbeit vorgestellten Experimente stellen die bisher umfangreichsten holistischen Proteomanalysen boviner Oozyten und früher Embryonen dar. Diese Ergebnisse ermöglichen nun die weitere Untersuchung von entdeckten Kandidatenproteinen in Lokalisationsstudien und durch gezielte Proteinquantifizierung mit der SRM Technologie, um herauszufinden, ob diese in die ersten Zelldifferenzierungen und in die Auslösung von bestimmten biologischen Ereignissen (z.B. embryonale Genomaktivierung) involviert sind.

8 Index of figures

Fig. 1: Stages of primordial germ cell migration	3
Fig. 2: Small antral/tertiary follicle	5
Fig. 3: A model for folliculogenesis initially generated and based on the ewe by Scaramuzzi et al.	7
Fig. 4: Illustration of the subsequent proteome analyses of bovine oocytes and early embryonic stages as performed in this doctoral thesis	22
Fig. 5: 1 D SDS PAA gels for the prefractionation of a protein lysate from 450 denuded GV oocytes.....	45
Fig. 6: Uniprot gene ontology analysis of 739 identified GV oocytes proteins.....	45
Fig. 7: KEGG oocytes meiosis pathway (bta04114).....	46
Fig. 8: Comparison of protein numbers identified from 1, 5, 10, 25 and 50 GV oocytes.....	48
Fig. 9: Base Peak Ion Chromatograms (BICs) of individual salt steps from the analysis of 25 GV oocytes by 2D LC-MS/MS	48
Fig. 10: HCD MS/MS spectrum of the Peptide LSDGVALVK derived from Heat shock 60 kDa Protein 1	49
Fig. 11: Numbers of proteins identified and quantified by iTRAQ 1D nano LC-MS/MS analysis from 25 and 50 GV oocytes	50
Fig. 12: Total numbers of identified and quantified proteins after each 1D LC-MS/MS run of a “run set”.....	51
Fig. 13: Numbers of identified and quantified proteins from the protein amount of 50 oocytes using different nano LC-MS/MS approaches.....	51
Fig. 14: Fold changes of the abundance of Ubiquitin, GLO1 and LAMP2 <i>in vivo</i> versus <i>in vitro</i> matured oocytes from different age groups.....	54
Fig. 15: Proportional Venn diagram of identified proteins from oocytes, 2-cell stage embryos and morulae	56
Fig. 16: Uniprot gene ontology analysis of the identified proteins from oocytes, two-cell-stage embryos and morulae	57
Fig. 17: Principle of SRM quantification	59
Fig. 18: Areas under the curve values from the SRM quantification of IMP3 and YBX2 in 2-cell stage embryos and morulae	60
Fig. 19: Chromatogram of a multiplexed SRM analysis.....	60
Fig. 20: 2D saturation DIGE analysis of morulae and blastocysts.....	62
Fig. 21: De Cyder 6.5 Cy5 readout image from a morula gel.....	63
Fig. 22: Graphic representations of five protein spots differing in intensity between morulae and blastocysts	63
Fig. 23: Volcano Plot of 141 proteins identified and quantified from morulae and blastocysts.	68
Fig. 24: Quality and reproducibility of HCD spectra.....	69
Fig. 25: Comparison of abundance alterations determined by DIGE spot intensity and iTRAQ reporter ion intensity	72

Fig. 26: Comparative “David” GO analysis.....	73
Fig. 27: Composition scheme of the iTRAQ tag.....	78

9 Index of tables

Table 1: Scheme of superstimulation for the generation of <i>in vivo</i> matured oocytes	24
Table 2: Transitions for YBX2 and IMP3 Quantification.....	30
Table 3: Labelling and run set scheme for one biological replicate of <i>in vivo</i> and <i>in vitro</i> matured oocytes analysed by nano LC-MS/MS iTRAQ analysis.....	53
Table 4: Differentially abundant proteins between <i>in vivo</i> and <i>in vitro</i> matured oocytes from heifers	54
Table 5: Differentially abundant proteins and abundance ratios between <i>in vivo</i> and <i>in vitro</i> matured oocytes from young cows.....	54
Table 6: Differentially abundant proteins and abundance ratios between <i>in vivo</i> and <i>in vitro</i> matured oocytes from old cows	55
Table 7: Differentially abundant proteins and abundance ratios between <i>in vitro</i> matured oocytes of old cows and heifers	55
Table 8: Differentially abundant proteins and abundance ratios between <i>in vitro</i> matured oocytes of young cows and heifers	55
Table 9: Differentially abundant proteins between 2-cell stage embryos and morulae detected by nano LC-MS/MS spectral count analysis.....	58
Table 10: Unambiguously identified spots from differentially abundant proteins between morulae and blastocysts from the 2D saturation DIGE analysis.....	64
Table 11: Spots from differentially abundant proteins between morulae and blastocysts of the 2D saturation DIGE analysis, from which two proteins were identified	65
Table 12: Spots from differentially abundant proteins between morulae and blastocysts of the 2D saturation DIGE analysis, from which three or more proteins were identified.....	66
Table 13: Spots from differentially abundant proteins between morulae and blastocysts which were not identified.....	67
Table 14: Labelling and run set scheme of six biological replicates from morulae and blastocysts analysed by nano LC-MS/MS iTRAQ analysis.....	67
Table 15: Differentially abundant proteins between morulae and blastocysts detected by the LC-MS/MS iTRAQ analysis	70
Table 16: Differentially abundant proteins between morulae and blastocysts detected by both, the Saturation DIGE and the LC-MS/MS iTRAQ analysis	72

10 Index of abbreviations

ACN	acetonitril
BPB	bromphenolblau
BIC	base peak ion chromatogram
ART	assisted reproduction techniques
CID	collision-induced dissociation
COC	cumulus oocyte complexes
DF	dominant follicle
DIGE	difference gel electrophoresis
DTT	dithioerythritol
ESC	embryonic stem cells
ESI	electrospray ionisation
FA	formic acid
FAB	fast atom bombardment
FDR	false discovery rate
FSH	follicle stimulating hormone
GO	gene ontology
GV	germinal vesicle
GVBD	germinal vesicle break down
HIF	hypoxia inducible factor
HPLC	high performance liquid chromatography
ICAT	isotope-coded affinity tag”
ICM	inner cell mass
ICPL	isotope-coded protein labelling
ICR	imprinting control regions
IEF	isoelectric focussing
IGFBP	IGF binding protein
IP	ioselectric point
IPS	internal pooled standard
iTRAQ	isobaric tag for relative and absolute quantification
IVF	in vitro fertilization
IVM	in vitro maturation
IVP	in vitro production of embryos
LC	liquid chromatography
LH-R	luteinizing hormone receptor
LH-R	luteinizing hormone
LOS 1	arge offspring syndrome
MALDI	matrix-assisted laser desorption ionisation
MET	maternal embryo transition
mRNA	messenger RNA
MHKP	maternal house keeping proteins
MI	metaphase I
MII	metaphase II
MPF	M-phase promoting factor

mRNP	messenger ribonucleoprotein complexes
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MW	molecular weight
OHSS	ovarian hyperstimulation syndrome
OPU	ovum pick up
OSF	oocyte secreted factors
p.i.	post insemination
PAA	polyacrylamide
PBS	phosphate buffered saline
PCG	primordial germ cells
PMF	peptide mass fingerprinting
PVP	polyvinylpyrrolidone
RP	reversed phase
SCX	strong cation exchange
SELDI	surface enhanced laser desorption ionisation
SILAC	stable isotope labelling by amino acids in cell Culture
SOF	synthetic oviductal fluid
SRM	selected reaction monitoring
TCEP	tris (2-carboxyethyl) phosphine hydrochloride
TE	trophectoderm
TEMED	tetramethylethylenediamine
TOF	time-of-flight

11 Reference list

1. Farin, P.W., J.A. Piedrahita, and C.E. Farin, *Errors in development of fetuses and placentas from in vitro-produced bovine embryos*. Theriogenology, 2006. **65**(1): p. 178-91.
2. Fortunato, A. and E. Tosti, *The impact of in vitro fertilization on health of the children: an update*. Eur J Obstet Gynecol Reprod Biol, 2011. **154**(2): p. 125-9.
3. Budev, M.M., A.C. Arroliga, and T. Falcone, *Ovarian hyperstimulation syndrome*. Crit Care Med, 2005. **33**(10 Suppl): p. S301-6.
4. Brackett, B.G., et al., *Normal development following in vitro fertilization in the cow*. Biol Reprod, 1982. **27**(1): p. 147-58.
5. Thibier, M., *Data Retrieval Committee Statistics of Embryo Transfer- Year 2007*
6. Seli, E., C. Robert, and M.A. Sirard, *OMICS in assisted reproduction: possibilities and pitfalls*. Mol Hum Reprod, 2010. **16**(8): p. 513-30.
7. Pierson, G.P.A.a.R.A., *Bovine model for study of ovarian follicular dynamics in humans*. Theriogenology, 1995. **43**(1): p. 113-120.
8. Mihm, M. and A.C. Evans, *Mechanisms for dominant follicle selection in monovulatory species: a comparison of morphological, endocrine and intraovarian events in cows, mares and women*. Reprod Domest Anim, 2008. **43** Suppl 2: p. 48-56.
9. Baerwald, A.R., G.P. Adams, and R.A. Pierson, *A new model for ovarian follicular development during the human menstrual cycle*. Fertil Steril, 2003. **80**(1): p. 116-22.
10. Baerwald, A.R., G.P. Adams, and R.A. Pierson, *Characterization of ovarian follicular wave dynamics in women*. Biol Reprod, 2003. **69**(3): p. 1023-31.
11. Malhi, P.S., G.P. Adams, and J. Singh, *Bovine model for the study of reproductive aging in women: follicular, luteal, and endocrine characteristics*. Biol Reprod, 2005. **73**(1): p. 45-53.
12. Malhi, P.S., et al., *Superovulatory response in a bovine model of reproductive aging*. Anim Reprod Sci, 2008. **109**(1-4): p. 100-9.
13. Rüsse I, S.F., *Lehrbuch der Embryologie der Haustiere*. 1998.
14. Richardson, B.E. and R. Lehmann, *Mechanisms guiding primordial germ cell migration: strategies from different organisms*. Nat Rev Mol Cell Biol. **11**(1): p. 37-49.
15. Nagy, A., et al., *Manipulating the Mouse Embryo*. 2003: p. 38-42.
16. Hajkova, P., et al., *Epigenetic reprogramming in mouse primordial germ cells*. Mech Dev, 2002. **117**(1-2): p. 15-23.
17. Szabo, P.E. and J.R. Mann, *Biallelic expression of imprinted genes in the mouse germ line: implications for erasure, establishment, and mechanisms of genomic imprinting*. Genes Dev, 1995. **9**(15): p. 1857-68.
18. Weaver, J.R., M. Susiarjo, and M.S. Bartolomei, *Imprinting and epigenetic changes in the early embryo*. Mamm Genome, 2009. **20**(9-10): p. 532-43.
19. Heinzmann, J., et al., *Epigenetic profile of developmentally important genes in bovine oocytes*. Mol Reprod Dev, 2011. **78**(3): p. 188-201.
20. Fortune, J.E., M.Y. Yang, and W. Muruvi, *In vitro and in vivo regulation of follicular formation and activation in cattle*. Reprod Fertil Dev, 2011. **23**(1): p. 15-22.
21. Scaramuzzi, R.J., et al., *Regulation of folliculogenesis and the determination of ovulation rate in ruminants*. Reprod Fertil Dev, 2011. **23**(3): p. 444-67.
22. Krisher, R.L., *The effect of oocyte quality on development*. J Anim Sci, 2004. **82** E-Suppl: p. E14-23.
23. Sirard, M.A., et al., *Contribution of the oocyte to embryo quality*. Theriogenology, 2006. **65**(1): p. 126-36.
24. Forde, N., et al., *Oestrous cycles in Bos taurus cattle*. Anim Reprod Sci, 2010.

25. Gilchrist, R.B., M. Lane, and J.G. Thompson, *Oocyte-secreted factors: regulators of cumulus cell function and oocyte quality*. Hum Reprod Update, 2008. **14**(2): p. 159-77.
26. Gilchrist, R.B., L.J. Ritter, and D.T. Armstrong, *Oocyte-somatic cell interactions during follicle development in mammals*. Anim Reprod Sci, 2004. **82-83**: p. 431-46.
27. Ginther, O.J., L. Knopf, and J.P. Kastelic, *Temporal associations among ovarian events in cattle during oestrous cycles with two and three follicular waves*. J Reprod Fertil, 1989. **87**(1): p. 223-30.
28. Rajakoski, E., *The ovarian follicular system in sexually mature heifers with special reference to seasonal, cyclical, and left-right variations*. Acta Endocrinol Suppl (Copenh), 1960. **34**(Suppl 52): p. 1-68.
29. Adams, G.P., et al., *Association between surges of follicle-stimulating hormone and the emergence of follicular waves in heifers*. J Reprod Fertil, 1992. **94**(1): p. 177-88.
30. Sunderland, S.J., et al., *Selection, dominance and atresia of follicles during the oestrous cycle of heifers*. J Reprod Fertil, 1994. **101**(3): p. 547-55.
31. Mihm, M., et al., *Identification of potential intrafollicular factors involved in selection of dominant follicles in heifers*. Biol Reprod, 2000. **63**(3): p. 811-9.
32. Ireland, J.J. and J.F. Roche, *Development of nonovulatory antral follicles in heifers: changes in steroids in follicular fluid and receptors for gonadotrophins*. Endocrinology, 1983. **112**(1): p. 150-6.
33. Ireland, J.J. and J.F. Roche, *Growth and differentiation of large antral follicles after spontaneous luteolysis in heifers: changes in concentration of hormones in follicular fluid and specific binding of gonadotrophins to follicles*. J Anim Sci, 1983. **57**(1): p. 157-67.
34. Scaramuzzi, R.J., et al., *A model for follicle selection and the determination of ovulation rate in the ewe*. Reprod Fertil Dev, 1993. **5**(5): p. 459-78.
35. Frei, R.E., G.A. Schultz, and R.B. Church, *Qualitative and quantitative changes in protein synthesis occur at the 8-16-cell stage of embryogenesis in the cow*. J Reprod Fertil, 1989. **86**(2): p. 637-41.
36. Meirelles, F.V., et al., *Genome activation and developmental block in bovine embryos*. Anim Reprod Sci, 2004. **82-83**: p. 13-20.
37. Memili, E. and N.L. First, *Developmental changes in RNA polymerase II in bovine oocytes, early embryos, and effect of alpha-amanitin on embryo development*. Mol Reprod Dev, 1998. **51**(4): p. 381-9.
38. Tosti, E., *Dynamic roles of ion currents in early development*. Mol Reprod Dev, 2010. **77**(10): p. 856-67.
39. Cockburn, K. and J. Rossant, *Making the blastocyst: lessons from the mouse*. J Clin Invest, 2010. **120**(4): p. 995-1003.
40. Vestweber, D., et al., *Expression and distribution of cell adhesion molecule uvomorulin in mouse preimplantation embryos*. Dev Biol, 1987. **124**(2): p. 451-6.
41. Fleming, T.P. and S.J. Pickering, *Maturation and polarization of the endocytotic system in outside blastomeres during mouse preimplantation development*. J Embryol Exp Morphol, 1985. **89**: p. 175-208.
42. Reeve, W.J. and F.P. Kelly, *Nuclear position in the cells of the mouse early embryo*. J Embryol Exp Morphol, 1983. **75**: p. 117-39.
43. Johnson, M.H. and B. Maro, *The distribution of cytoplasmic actin in mouse 8-cell blastomeres*. J Embryol Exp Morphol, 1984. **82**: p. 97-117.
44. Louvet, S., et al., *Ezrin becomes restricted to outer cells following asymmetrical division in the preimplantation mouse embryo*. Dev Biol, 1996. **177**(2): p. 568-79.
45. Handyside, A.H., *Distribution of antibody- and lectin-binding sites on dissociated blastomeres from mouse morulae: evidence for polarization at compaction*. J Embryol Exp Morphol, 1980. **60**: p. 99-116.

46. Ducibella, T., et al., *Changes in cell surface and cortical cytoplasmic organization during early embryogenesis in the preimplantation mouse embryo*. J Cell Biol, 1977. **74**(1): p. 153-67.
47. Guo, S. and K.J. Kemphues, *par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed*. Cell, 1995. **81**(4): p. 611-20.
48. Hung, T.J. and K.J. Kemphues, *PAR-6 is a conserved PDZ domain-containing protein that colocalizes with PAR-3 in Caenorhabditis elegans embryos*. Development, 1999. **126**(1): p. 127-35.
49. Ahringer, J., *Control of cell polarity and mitotic spindle positioning in animal cells*. Curr Opin Cell Biol, 2003. **15**(1): p. 73-81.
50. Plusa, B., et al., *Downregulation of Par3 and aPKC function directs cells towards the ICM in the preimplantation mouse embryo*. J Cell Sci, 2005. **118**(Pt 3): p. 505-15.
51. Vinot, S., et al., *Asymmetric distribution of PAR proteins in the mouse embryo begins at the 8-cell stage during compaction*. Dev Biol, 2005. **282**(2): p. 307-19.
52. Johnson, M.H. and C.A. Ziomek, *Induction of polarity in mouse 8-cell blastomeres: specificity, geometry, and stability*. J Cell Biol, 1981. **91**(1): p. 303-8.
53. Ziomek, C.A. and M.H. Johnson, *Cell surface interaction induces polarization of mouse 8-cell blastomeres at compaction*. Cell, 1980. **21**(3): p. 935-42.
54. Houliston, E., S.J. Pickering, and B. Maro, *Alternative routes for the establishment of surface polarity during compaction of the mouse embryo*. Dev Biol, 1989. **134**(2): p. 342-50.
55. Johnson, M.H. and C.A. Ziomek, *The foundation of two distinct cell lineages within the mouse morula*. Cell, 1981. **24**(1): p. 71-80.
56. Palmieri, S.L., et al., *Oct-4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation*. Dev Biol, 1994. **166**(1): p. 259-67.
57. Chambers, I., et al., *Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells*. Cell, 2003. **113**(5): p. 643-55.
58. Jedrusik, A., et al., *Role of Cdx2 and cell polarity in cell allocation and specification of trophoctoderm and inner cell mass in the mouse embryo*. Genes Dev, 2008. **22**(19): p. 2692-706.
59. Ralston, A. and J. Rossant, *Cdx2 acts downstream of cell polarization to cell-autonomously promote trophoctoderm fate in the early mouse embryo*. Dev Biol, 2008. **313**(2): p. 614-29.
60. Strumpf, D., et al., *Cdx2 is required for correct cell fate specification and differentiation of trophoctoderm in the mouse blastocyst*. Development, 2005. **132**(9): p. 2093-102.
61. Nishioka, N., et al., *Tead4 is required for specification of trophoctoderm in pre-implantation mouse embryos*. Mech Dev, 2008. **125**(3-4): p. 270-83.
62. Yagi, R., et al., *Transcription factor TEAD4 specifies the trophoctoderm lineage at the beginning of mammalian development*. Development, 2007. **134**(21): p. 3827-36.
63. Nishioka, N., et al., *The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophoctoderm from inner cell mass*. Dev Cell, 2009. **16**(3): p. 398-410.
64. Klose, J., *Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals*. Humangenetik, 1975. **26**(3): p. 231-43.
65. O'Farrell, P.H., *High resolution two-dimensional electrophoresis of proteins*. J Biol Chem, 1975. **250**(10): p. 4007-21.
66. Arnold, G.J. and T. Frohlich, *Dynamic proteome signatures in gametes, embryos and their maternal environment*. Reprod Fertil Dev, 2011. **23**(1): p. 81-93.

67. Merrill, C.R., R.C. Switzer, and M.L. Van Keuren, *Trace polypeptides in cellular extracts and human body fluids detected by two-dimensional electrophoresis and a highly sensitive silver stain*. Proc Natl Acad Sci U S A, 1979. **76**(9): p. 4335-9.
68. Fazekas de St Groth, S., R.G. Webster, and A. Datyner, *Two new staining procedures for quantitative estimation of proteins on electrophoretic strips*. Biochim Biophys Acta, 1963. **71**: p. 377-91.
69. Neuhoff, V., et al., *Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250*. Electrophoresis, 1988. **9**(6): p. 255-62.
70. Berggren, K., et al., *A luminescent ruthenium complex for ultrasensitive detection of proteins immobilized on membrane supports*. Anal Biochem, 1999. **276**(2): p. 129-43.
71. Lopez, M.F., et al., *A comparison of silver stain and SYPRO Ruby Protein Gel Stain with respect to protein detection in two-dimensional gels and identification by peptide mass profiling*. Electrophoresis, 2000. **21**(17): p. 3673-83.
72. Unlu, M., M.E. Morgan, and J.S. Minden, *Difference gel electrophoresis: a single gel method for detecting changes in protein extracts*. Electrophoresis, 1997. **18**(11): p. 2071-7.
73. Kondo, T., et al., *Application of sensitive fluorescent dyes in linkage of laser microdissection and two-dimensional gel electrophoresis as a cancer proteomic study tool*. Proteomics, 2003. **3**(9): p. 1758-66.
74. Berendt, F.J., et al., *Highly sensitive saturation labeling reveals changes in abundance of cell cycle-associated proteins and redox enzyme variants during oocyte maturation in vitro*. Proteomics, 2009. **9**(3): p. 550-64.
75. Edman, P. and G. Begg, *A protein sequenator*. Eur J Biochem, 1967. **1**(1): p. 80-91.
76. Edman, P., *Sequence determination*. Mol Biol Biochem Biophys, 1970. **8**: p. 211-55.
77. Frohlich, T. and G.J. Arnold, *Proteome research based on modern liquid chromatography--tandem mass spectrometry: separation, identification and quantification*. J Neural Transm, 2006. **113**(8): p. 973-94.
78. Barber, M., et al., *Fast-atom-bombardment mass spectra of enkephalins*. Biochem J, 1981. **197**(2): p. 401-4.
79. Morris, H.R., et al., *Fast atom bombardment: a new mass spectrometric method for peptide sequence analysis*. Biochem Biophys Res Commun, 1981. **101**(2): p. 623-31.
80. Williams, D.H., et al., *Fast-atom-bombardment mass spectrometry. A new technique for the determination of molecular weights and amino acid sequences of peptides*. Biochem J, 1982. **201**(1): p. 105-17.
81. Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules*. Science, 1989. **246**(4926): p. 64-71.
82. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons*. Anal Chem, 1988. **60**(20): p. 2299-301.
83. Kelleher, N.L., *Top-down proteomics*. Anal Chem, 2004. **76**(11): p. 197A-203A.
84. Henzel, W.J., et al., *Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases*. Proc Natl Acad Sci U S A, 1993. **90**(11): p. 5011-5.
85. Henzel, W.J., C. Watanabe, and J.T. Stults, *Protein identification: the origins of peptide mass fingerprinting*. J Am Soc Mass Spectrom, 2003. **14**(9): p. 931-42.
86. Gao, J., et al., *Changes in the protein expression of yeast as a function of carbon source*. J Proteome Res, 2003. **2**(6): p. 643-9.
87. Liu, H., R.G. Sadygov, and J.R. Yates, 3rd, *A model for random sampling and estimation of relative protein abundance in shotgun proteomics*. Anal Chem, 2004. **76**(14): p. 4193-201.

88. Ong, S.E., et al., *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics*. Mol Cell Proteomics, 2002. **1**(5): p. 376-86.
89. Gygi, S.P., et al., *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags*. Nat Biotechnol, 1999. **17**(10): p. 994-9.
90. Schmidt, A., J. Kellermann, and F. Lottspeich, *A novel strategy for quantitative proteomics using isotope-coded protein labels*. Proteomics, 2005. **5**(1): p. 4-15.
91. Ross, P.L., et al., *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents*. Mol Cell Proteomics, 2004. **3**(12): p. 1154-69.
92. Duncan, M.W., A.L. Yergey, and S.D. Patterson, *Quantifying proteins by mass spectrometry: the selectivity of SRM is only part of the problem*. Proteomics, 2009. **9**(5): p. 1124-7.
93. Zhang, Y., et al., *Absolute quantification of semicarbazide-sensitive amine oxidase in human umbilical artery by single-reaction monitoring with electrospray tandem mass spectrometry*. Anal Bioanal Chem, 2010. **397**(2): p. 709-15.
94. Barnidge, D.R., et al., *Absolute quantification of the G protein-coupled receptor rhodopsin by LC/MS/MS using proteolysis product peptides and synthetic peptide standards*. Anal Chem, 2003. **75**(3): p. 445-51.
95. Schultz, R.M. and P.M. Wassarman, *Biochemical studies of mammalian oogenesis: Protein synthesis during oocyte growth and meiotic maturation in the mouse*. J Cell Sci, 1977. **24**: p. 167-94.
96. Calvert, M.E., et al., *Oolemmal proteomics--identification of highly abundant heat shock proteins and molecular chaperones in the mature mouse egg and their localization on the plasma membrane*. Reprod Biol Endocrinol, 2003. **1**: p. 27.
97. Tutuncu, L., et al., *Calreticulin on the mouse egg surface mediates transmembrane signaling linked to cell cycle resumption*. Dev Biol, 2004. **270**(1): p. 246-60.
98. Yurttas, P., E. Morency, and S.A. Coonrod, *Use of proteomics to identify highly abundant maternal factors that drive the egg-to-embryo transition*. Reproduction, 2010. **139**(5): p. 809-23.
99. Vitale, A.M., et al., *Proteomic profiling of murine oocyte maturation*. Mol Reprod Dev, 2007. **74**(5): p. 608-16.
100. Albee, A.J. and C. Wiese, *Xenopus TACC3/maskin is not required for microtubule stability but is required for anchoring microtubules at the centrosome*. Mol Biol Cell, 2008. **19**(8): p. 3347-56.
101. Schneider, L., et al., *The transforming acidic coiled coil 3 protein is essential for spindle-dependent chromosome alignment and mitotic survival*. J Biol Chem, 2007. **282**(40): p. 29273-83.
102. Ma, M., et al., *Protein expression profile of the mouse metaphase-II oocyte*. J Proteome Res, 2008. **7**(11): p. 4821-30.
103. Zhang, P., et al., *Proteomic-based identification of maternal proteins in mature mouse oocytes*. BMC Genomics, 2009. **10**: p. 348.
104. Wang, S., et al., *Proteome of mouse oocytes at different developmental stages*. Proc Natl Acad Sci U S A, 2010. **107**(41): p. 17639-44.
105. Pfeiffer, M.J., et al., *Proteomic Analysis of Mouse Oocytes Reveals 28 Candidate Factors of the "Reprogrammome"*. J Proteome Res, 2011.
106. Peddinti, D., E. Memili, and S.C. Burgess, *Proteomics-based systems biology modeling of bovine germinal vesicle stage oocyte and cumulus cell interaction*. PLoS One, 2010. **5**(6): p. e11240.
107. Coenen, K., L. Massicotte, and M.A. Sirard, *Study of newly synthesized proteins during bovine oocyte maturation in vitro using image analysis of two-dimensional gel electrophoresis*. Mol Reprod Dev, 2004. **67**(3): p. 313-22.

108. Ellederova, Z., et al., *Protein patterns of pig oocytes during in vitro maturation*. Biol Reprod, 2004. **71**(5): p. 1533-9.
109. Susor, A., et al., *Proteomic analysis of porcine oocytes during in vitro maturation reveals essential role for the ubiquitin C-terminal hydrolase-L1*. Reproduction, 2007. **134**(4): p. 559-68.
110. Powell, M.D., et al., *Discovery of putative oocyte quality markers by comparative ExacTag proteomics*. Proteomics Clin Appl, 2010. **4**(3): p. 337-51.
111. Katz-Jaffe, M.G., et al., *A proteomic analysis of mammalian preimplantation embryonic development*. Reproduction, 2005. **130**(6): p. 899-905.
112. Katz-Jaffe, M.G., D.K. Gardner, and W.B. Schoolcraft, *Proteomic analysis of individual human embryos to identify novel biomarkers of development and viability*. Fertil Steril, 2006. **85**(1): p. 101-7.
113. Massicotte, L., et al., *Maternal housekeeping proteins translated during bovine oocyte maturation and early embryo development*. Proteomics, 2006. **6**(13): p. 3811-20.
114. Gupta, M.K., et al., *Proteomic analysis of parthenogenetic and in vitro fertilized porcine embryos*. Proteomics, 2009. **9**(10): p. 2846-60.
115. Gilchrist, R.B., *Recent insights into oocyte-follicle cell interactions provide opportunities for the development of new approaches to in vitro maturation*. Reprod Fertil Dev, 2011. **23**(1): p. 23-31.
116. Berg, U.B., G., *In-vitro production of bovine blastocysts by in-vitro maturation and fertilization of oocytes and subsequent in-vitro culture*. Zuchthygiene, 1989. **24** **134-139**: p. 134-139.
117. Parrish, J.J., et al., *Bovine in vitro fertilization with frozen-thawed semen*. Theriogenology, 1986. **25**(4): p. 591-600.
118. Machado, S.A., et al., *The variability of ovum pick-up response and in vitro embryo production from monozygotic twin cows*. Theriogenology, 2006. **65**(3): p. 573-83.
119. Keller, A., et al., *Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search*. Anal Chem, 2002. **74**(20): p. 5383-92.
120. Nesvizhskii, A.I., et al., *A statistical model for identifying proteins by tandem mass spectrometry*. Anal Chem, 2003. **75**(17): p. 4646-58.
121. Ashburner, M., et al., *Gene ontology: tool for the unification of biology*. The Gene Ontology Consortium. Nat Genet, 2000. **25**(1): p. 25-9.
122. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources*. Nat Protoc, 2009. **4**(1): p. 44-57.
123. Dennis, G., Jr., et al., *DAVID: Database for Annotation, Visualization, and Integrated Discovery*. Genome Biol, 2003. **4**(5): p. P3.
124. Benjamini, Y. and Y. Hochberg, *On the adaptive control of the false discovery rate in multiple testing with independent statistics*. Journal of Educational and Behavioral Statistics, 2000. **25**(1): p. 60-83.
125. Elias, J.E., et al., *Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations*. Nat Methods, 2005. **2**(9): p. 667-75.
126. Memili, E., et al., *Bovine germinal vesicle oocyte and cumulus cell proteomics*. Reproduction, 2007. **133**(6): p. 1107-20.
127. Armstrong, D.T., *Effects of maternal age on oocyte developmental competence*. Theriogenology, 2001. **55**(6): p. 1303-22.
128. Matthiesen, M.M., *Effect of donor age on the developmental capacity of bovine cumulus oocyte complexes obtained by repeated OPU from nonstimulated and FSH-superstimulated German Simmental heifers and cows at different life cycle stages*, doctoral thesis, Ludwig-Maximilians-Universität München, 2011

129. Komander, D., *The emerging complexity of protein ubiquitination*. *Biochem Soc Trans*, 2009. **37**(Pt 5): p. 937-53.
130. Min, T., et al., *Critical role of proteostasis-imbalance in pathogenesis of COPD and severe emphysema*. *J Mol Med*, 2011.
131. Thornalley, P.J., *Glyoxalase I-structure, function and a critical role in the enzymatic defence against glycation*. *Biochem Soc Trans*, 2003. **31**(Pt 6): p. 1343-8.
132. Sutton-McDowall, M.L., R.B. Gilchrist, and J.G. Thompson, *The pivotal role of glucose metabolism in determining oocyte developmental competence*. *Reproduction*, 2010. **139**(4): p. 685-95.
133. Rieger, D., et al., *Comparison of the effects of oviductal cell co-culture and oviductal cell-conditioned medium on the development and metabolic activity of cattle embryos*. *J Reprod Fertil*, 1995. **105**(1): p. 91-8.
134. Rose, Z.B., *The enzymology of 2,3-bisphosphoglycerate*. *Adv Enzymol Relat Areas Mol Biol*, 1980. **51**: p. 211-53.
135. Arnone, A., *X-ray diffraction study of binding of 2,3-diphosphoglycerate to human deoxyhaemoglobin*. *Nature*, 1972. **237**(5351): p. 146-9.
136. Wang, Y., et al., *Crystal structure of human bisphosphoglycerate mutase*. *J Biol Chem*, 2004. **279**(37): p. 39132-8.
137. Pritlove, D.C., et al., *Novel placental expression of 2,3-bisphosphoglycerate mutase*. *Placenta*, 2006. **27**(8): p. 924-7.
138. Tomek, W., H. Torner, and W. Kanitz, *Comparative analysis of protein synthesis, transcription and cytoplasmic polyadenylation of mRNA during maturation of bovine oocytes in vitro*. *Reprod Domest Anim*, 2002. **37**(2): p. 86-91.
139. Harvey, A.J., *The role of oxygen in ruminant preimplantation embryo development and metabolism*. *Anim Reprod Sci*, 2007. **98**(1-2): p. 113-28.
140. Washburn, M.P., D. Wolters, and J.R. Yates, 3rd, *Large-scale analysis of the yeast proteome by multidimensional protein identification technology*. *Nat Biotechnol*, 2001. **19**(3): p. 242-7.
141. Lundgren, D.H., et al., *Role of spectral counting in quantitative proteomics*. *Expert Rev Proteomics*, 2010. **7**(1): p. 39-53.
142. Weston, A. and J. Sommerville, *Xp54 and related (DDX6-like) RNA helicases: roles in messenger RNP assembly, translation regulation and RNA degradation*. *Nucleic Acids Res*, 2006. **34**(10): p. 3082-94.
143. Yu, J., N.B. Hecht, and R.M. Schultz, *RNA-binding properties and translation repression in vitro by germ cell-specific MSY2 protein*. *Biol Reprod*, 2002. **67**(4): p. 1093-8.
144. Yu, J., N.B. Hecht, and R.M. Schultz, *Requirement for RNA-binding activity of MSY2 for cytoplasmic localization and retention in mouse oocytes*. *Dev Biol*, 2003. **255**(2): p. 249-62.
145. Vigneault, C., S. McGraw, and M.A. Sirard, *Spatiotemporal expression of transcriptional regulators in concert with the maternal-to-embryonic transition during bovine in vitro embryogenesis*. *Reproduction*, 2009. **137**(1): p. 13-21.
146. Vigneault, C., et al., *Transcription factor expression patterns in bovine in vitro-derived embryos prior to maternal-zygotic transition*. *Biol Reprod*, 2004. **70**(6): p. 1701-9.
147. Ross, A.F., et al., *Characterization of a beta-actin mRNA zipcode-binding protein*. *Mol Cell Biol*, 1997. **17**(4): p. 2158-65.
148. Deshler, J.O., et al., *A highly conserved RNA-binding protein for cytoplasmic mRNA localization in vertebrates*. *Curr Biol*, 1998. **8**(9): p. 489-96.
149. Doyle, G.A., et al., *The c-myc coding region determinant-binding protein: a member of a family of KH domain RNA-binding proteins*. *Nucleic Acids Res*, 1998. **26**(22): p. 5036-44.

150. Havin, L., et al., *RNA-binding protein conserved in both microtubule- and microfilament-based RNA localization*. Genes Dev, 1998. **12**(11): p. 1593-8.
151. Nielsen, J., et al., *A family of insulin-like growth factor II mRNA-binding proteins represses translation in late development*. Mol Cell Biol, 1999. **19**(2): p. 1262-70.
152. Runge, S., et al., *H19 RNA binds four molecules of insulin-like growth factor II mRNA-binding protein*. J Biol Chem, 2000. **275**(38): p. 29562-9.
153. Vikesaa, J., et al., *RNA-binding IMPs promote cell adhesion and invadopodia formation*. EMBO J, 2006. **25**(7): p. 1456-68.
154. Yisraeli, J.K., *VICKZ proteins: a multi-talented family of regulatory RNA-binding proteins*. Biol Cell, 2005. **97**(1): p. 87-96.
155. Mueller-Pillasch, F., et al., *Cloning of a gene highly overexpressed in cancer coding for a novel KH-domain containing protein*. Oncogene, 1997. **14**(22): p. 2729-33.
156. Ioannidis, P., et al., *C-MYC and IGF-II mRNA-binding protein (CRD-BP/IMP-1) in benign and malignant mesenchymal tumors*. Int J Cancer, 2001. **94**(4): p. 480-4.
157. Ioannidis, P., et al., *CRD-BP/IMP1 expression characterizes cord blood CD34+ stem cells and affects c-myc and IGF-II expression in MCF-7 cancer cells*. J Biol Chem, 2005. **280**(20): p. 20086-93.
158. Yaniv, K. and J.K. Yisraeli, *The involvement of a conserved family of RNA binding proteins in embryonic development and carcinogenesis*. Gene, 2002. **287**(1-2): p. 49-54.
159. Baker, J., et al., *Role of insulin-like growth factors in embryonic and postnatal growth*. Cell, 1993. **75**(1): p. 73-82.
160. Sagirkaya, H., et al., *Developmental and molecular correlates of bovine preimplantation embryos*. Reproduction, 2006. **131**(5): p. 895-904.
161. Lim, J.M., et al., *Development of bovine IVF oocytes cultured in medium supplemented with a nitric oxide scavenger or inhibitor in a co-culture system*. Theriogenology, 1999. **51**(5): p. 941-9.
162. Yuan, Y.Q., et al., *Influence of oxygen tension on apoptosis and hatching in bovine embryos cultured in vitro*. Theriogenology, 2003. **59**(7): p. 1585-96.
163. Karp, N.A. and K.S. Lilley, *Design and analysis issues in quantitative proteomics studies*. Proteomics, 2007. **7 Suppl 1**: p. 42-50.
164. Kitagawa, K., et al., *SGT1 encodes an essential component of the yeast kinetochore assembly pathway and a novel subunit of the SCF ubiquitin ligase complex*. Mol Cell, 1999. **4**(1): p. 21-33.
165. Fenner, B.J., M. Scannell, and J.H. Prehn, *Identification of polyubiquitin binding proteins involved in NF-kappaB signaling using protein arrays*. Biochim Biophys Acta, 2009. **1794**(7): p. 1010-6.
166. Thompson, J.G., et al., *Total protein content and protein synthesis within pre-elongation stage bovine embryos*. Mol Reprod Dev, 1998. **50**(2): p. 139-45.
167. Thompson, J.G., et al., *Oxygen uptake and carbohydrate metabolism by in vitro derived bovine embryos*. J Reprod Fertil, 1996. **106**(2): p. 299-306.
168. Mangeat, P., C. Roy, and M. Martin, *ERM proteins in cell adhesion and membrane dynamics*. Trends Cell Biol, 1999. **9**(5): p. 187-92.
169. Amatruda, J.F. and J.A. Cooper, *Purification, characterization, and immunofluorescence localization of Saccharomyces cerevisiae capping protein*. J Cell Biol, 1992. **117**(5): p. 1067-76.
170. Maggi, L.B., Jr., et al., *Nucleophosmin serves as a rate-limiting nuclear export chaperone for the Mammalian ribosome*. Mol Cell Biol, 2008. **28**(23): p. 7050-65.
171. Nilsson, J., et al., *Regulation of eukaryotic translation by the RACK1 protein: a platform for signalling molecules on the ribosome*. EMBO Rep, 2004. **5**(12): p. 1137-41.
172. Pentecost, B.T., et al., *Estrogen regulation of creatine kinase-B in the rat uterus*. Mol Endocrinol, 1990. **4**(7): p. 1000-10.

173. Bergen, H.T., et al., *In situ hybridization for creatine kinase-B messenger RNA in rat uterus and brain*. Mol Cell Endocrinol, 1993. **92**(1): p. 111-9.
174. Schuler, G., et al., *Placental steroids in cattle: hormones, placental growth factors or by-products of trophoblast giant cell differentiation?* Exp Clin Endocrinol Diabetes, 2008. **116**(7): p. 429-36.
175. Grewal, T., et al., *Annexin A6-regulator of the EGFR/Ras signalling pathway and cholesterol homeostasis*. Int J Biochem Cell Biol, 2010. **42**(5): p. 580-4.
176. Liu, F.T., R.J. Patterson, and J.L. Wang, *Intracellular functions of galectins*. Biochim Biophys Acta, 2002. **1572**(2-3): p. 263-73.
177. Nangia-Makker, P., et al., *Galectin-3 in apoptosis, a novel therapeutic target*. J Bioenerg Biomembr, 2007. **39**(1): p. 79-84.
178. Domic, J., S. Dabelic, and M. Flogel, *Galectin-3: an open-ended story*. Biochim Biophys Acta, 2006. **1760**(4): p. 616-35.
179. Fleming, T.P., *A quantitative analysis of cell allocation to trophectoderm and inner cell mass in the mouse blastocyst*. Dev Biol, 1987. **119**(2): p. 520-31.
180. Marikawa, Y. and V.B. Alarcon, *Establishment of trophectoderm and inner cell mass lineages in the mouse embryo*. Mol Reprod Dev, 2009. **76**(11): p. 1019-32.
181. Okuwaki, M., *The structure and functions of NPM1/Nucleophsmin/B23, a multifunctional nucleolar acidic protein*. J Biochem, 2008. **143**(4): p. 441-8.
182. Johansson, H. and S. Simonsson, *Core transcription factors, Oct4, Sox2 and Nanog, individually form complexes with nucleophosmin (Npm1) to control embryonic stem (ES) cell fate determination*. Aging (Albany NY), 2010. **2**(11): p. 815-22.
183. Chen, L. and G.Q. Daley, *Molecular basis of pluripotency*. Hum Mol Genet, 2008. **17**(R1): p. R23-7.
184. Leyens, G., B. Knoops, and I. Donnay, *Expression of peroxiredoxins in bovine oocytes and embryos produced in vitro*. Mol Reprod Dev, 2004. **69**(3): p. 243-51.
185. Rhee, S.G., H.Z. Chae, and K. Kim, *Peroxiredoxins: a historical overview and speculative preview of novel mechanisms and emerging concepts in cell signaling*. Free Radic Biol Med, 2005. **38**(12): p. 1543-52.
186. Srivastava, S.K., et al., *Aldose Reductase Inhibition Suppresses Oxidative Stress-Induced Inflammatory Disorders*. Chem Biol Interact, 2011.
187. El-Sayed, A., et al., *Large-scale transcriptional analysis of bovine embryo biopsies in relation to pregnancy success after transfer to recipients*. Physiol Genomics, 2006. **28**(1): p. 84-96.
188. Madore, E., et al., *An aldose reductase with 20 alpha-hydroxysteroid dehydrogenase activity is most likely the enzyme responsible for the production of prostaglandin f2 alpha in the bovine endometrium*. J Biol Chem, 2003. **278**(13): p. 11205-12.

12 Acknowledgements

First of all I would like to express my deep and sincere gratitude to my supervisor Dr. Georg J. Arnold who gave me the possibility to carry out this work at his laboratory. Moreover, he gave me the freedom to develop my skills by always being open and supportive to new ideas and providing input through interesting and helpful discussions. Besides, I deeply enjoyed that he gave me the possibility to present my work to a broad audience all over the world and made me so growing into the scientific community.

I am deeply grateful to Univ.-Prof. Dr. Eckhard Wolf for being my doctoral advisor from the veterinarian department of the Ludwig-Maximilians-University. He provided some of the main recommendations for the aims of my research and gave me the possibility to do parts of my work at the “Chair for Molecular Animal Breeding”. In particular I would like to thank him for the careful reading of this doctoral thesis!

I also owe deep gratitude to Dr. Thomas Fröhlich. This thesis would not have been possible without his knowledge in mass spectrometry and friendly advice. His enthusiasm for mass spectrometry was inspiring and infected me with excitement for mass spectrometry myself. I thank him for mentoring me and spending immense time and patience during this doctorate and preparation of this work.

I thank Miwako Kösters for her help and support especially for spending the time to supervise my work on the extremely rare samples. I wish her and her family all the best for their life in their newly built own house.

I would also like to thank Tuna Güngör and Myriam Reichenbach for taking the time to teach me the *in vitro* production of embryos.

I wish to express my warm and sincere thanks to Dr. Julia Katharina Korte. After first meeting her as a collaborator when starting this thesis I quickly got to know her intelligence as well as her great supporting personality during this doctoral thesis. Without Julia my PhD time would not have been the same. She enjoyed and suffered all the ups and downs of my PhD time both scientific and private. I appreciated her being around me and hope that our great friendship will last on!

I am especially and warmly grateful to Rebekka Kubisch one of the most supporting persons I met during my life. I could always trust on her well-thought-out advices in particular when she helped me writing this manuscript. Even when she had a lack of time herself she read and improved this manuscript again and again. I experienced her generous personality together with her spirit and intelligence through many essential private and scientific events and hope that I will have the possibility to give some of that back during our friendship.

My deep thanks also go to Maria-Dorothee Faust. She is the best and honest friend I can imagine to have. Even before I started my PhD thesis she became a supportive and generous person in my life. Moreover she contributed to this work by actively helping me on the collection of oocytes and the *in vitro* production of embryos as well as taking care of my horse anytime I was stucked in the lab.

I also want to thank Daniela Deutsch, Stefan Kempf and Kathrin Otte. Especially the last phase of my PhD thesis was influenced by their imaginativeness and cheerfulness and made it to a great time. In particular I want to thank Dani for her input in scientific English writing and carefully reading this manuscript.

I am also deeply grateful to Nadine Wahner and Nilima Chowdhury who suffered being told my problems and joys during my PhD time even from far away. I cannot express with words how pleased I am that our long-lasting friendship is still going on and that I can always rely on the fact that they both are somewhere there for me in the world!

I also want to thank the members of the “Kommune 41” Barbara Rinser and Monika Alter for our friendship and all the extremely funny and relaxing Kommunabend which made me thinking of other stuff than this thesis and so resetted my mind. I hope the tradition of Kommunabend will last through our lives.

My thanks also go to my short term flatmates Attila Klier and Sophie Fehrenbacher. I got to know these guys in a very difficult time of my life both private and concerning my thesis. I benefited from the comfort I took from them as well as from the interesting and often funny conversations we had.

I owe my deepest gratitude to my parents Bernd and Brigitte Demant. Without their emotional and financial support as well as their love during my whole education this PhD thesis would not have been possible. Therefore, I dedicate this thesis to them.