Chapter 1

Review of cell and molecular biology

Gametogenesis, embryo development, implantation and in-vitro culture involve numerous complex pathways and interactions at the cellular and molecular level; a true understanding of their significance requires secure fundamental knowledge of the underlying principles. This chapter therefore provides a condensed overview and review of basic terminology and definitions, with particular emphasis on aspects relevant to reproductive biology and in-vitro fertilization.

Mammalian cell biology

In 1839, two German scientists, Matthias Jakob Schleiden and Theodor Schwann, introduced the "cell theory," the proposal that all higher organisms are made up of a single fundamental unit as a building block. In 1855, Rudolf Virchow extended this cell theory with a suggestion that was highly controversial at the time: "Omnis cellulae e celula" (all living cells arise from pre-existing cells). This statement has become known as the "biogenic law." The cell theory is now accepted to include a number of principles:

1. All known living things are made up of cells.
2. The cell is the structural and functional unit of all living things.
3. All cells come from pre-existing cells by division (spontaneous generation does not occur).
4. Cells contain hereditary information that is transmitted from cell to cell during cell division.
5. The chemical composition of all cells is basically the same.
6. The energy flow (metabolism and biochemistry) of life occurs within cells.

Although these features are common to all cells, the expression and repression of genes dictates individual variation, resulting in a large number of different types of variegated but highly organized cells, with convoluted intracellular structures and interconnected elements. The average size of a somatic cell is around 20 μm; the oocyte is the largest cell in the body, with a diameter of approximately 120 μm in its final stages of growth (Figure 1.1). The basic elements and organelles in an individual cell vary in distribution and number according to the cell type. Bacterial cells differ from mammalian cells in that they have no distinct nucleus, mitochondria or endoplasmic reticulum. Their cell membrane has numerous attachments, and their ribosomes are scattered throughout the cytoplasm.

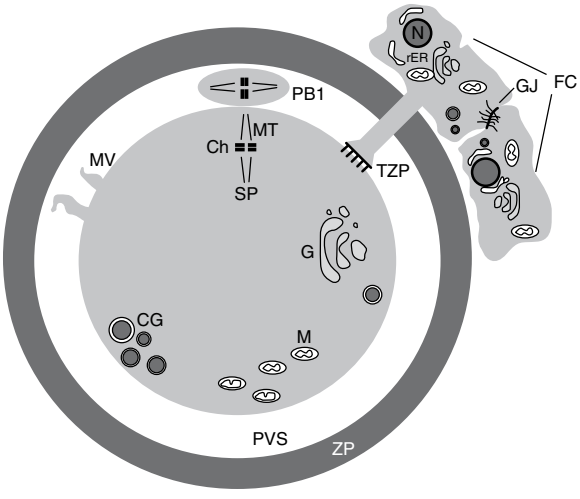

Figure 1.1 Schematic diagram of oocyte ultrastructure showing the zona pellucida (ZP) and the perivitelline space (PVS), first polar body (PB1), microvilli (MV), rough endoplasmic reticulum (rER), chromosomes (Ch) on the spindle (SP), Golgi complex (G), cortical granules (CG), two follicle cells (FC) attached to the oocyte and to each other via gap junctions (GJ). TZP = transzonal process, MT = Microtubules, M= Mitochondria.

Cell **membranes** are made up of a bimolecular layer of polar lipids, coated on both sides with protein films. Some proteins are buried in the matrix, others float independently of each other in or on the membrane surface,

In-Vitro Fertilization: Third Edition, ed. Kay Elder and Brian Dale. Published by Cambridge University Press. © K. Elder and B. Dale 2011.

forming a fluid mosaic of different functional units that are highly selective and specialized in different cells. Cells contain many different types of membrane, and each one encloses a space that defines an organelle, or a part of an organelle. The function of each organelle is determined largely by the types of protein in the membranes and the contents of the enclosed space. Membranes are important in the control of selective permeability, active and passive transport of ions and nutrients, contractile properties of the cell, and recognition of/association with other cells.

Cellular membranes always arise from pre-existing membranes, and the process of assembling new membranes is carried out by the endoplasmic reticulum (ER, see below). The synthesis and metabolism of fatty acids and cholesterol is important in membrane composition, and fatty acid oxidation (e.g., by the action of reactive oxygen species, ROS) can cause the membranes to lose their fluidity, as well as have an effect on transport mechanisms.

Microvilli are extensions of the plasma membrane that increase the cell surface area; they are abundant in cells with a highly absorptive capacity, such as the brush border of the intestinal lumen. Microvilli are present on the surface of oocytes, zygotes and early cleavage stage embryos in many species, and in some species (but not humans) their distribution is thought to be important in determining the site of sperm entry.

Cell **cytoplasm** is a fluid space, containing water, enzymes, nutrients and macromolecules; the cytoplasm is permeated by the cell's architectural support, the cytoskeleton.

Microtubules are hollow polymer tubes made up of alpha-beta dimers of the protein tubulin. They are part of the cytoskeletal structure, and are involved in intracellular transport, for example, the movement of mitochondria. Specialized structures such as centrioles, basal bodies, cilia and flagella are made up of microtubules. During prophase of mitosis or meiosis, microtubules form the **spindle** for chromosome attachment and movement.

Microfilaments are threads of actin protein, usually found in bundles just beneath the cell surface; they play a role in cell motility, and in endo- and exocytosis.

Centrioles are a pair of hollow tubes at right angles to each other, just outside the nucleus. These structures organize the nuclear spindle in preparation for the separation of chromatids during nuclear division. When the cell is about to divide, one of the centrioles migrates to the other side of the nucleus so that one lies at each end. The microtubule fibers in the spindle are contractile, and they pull the chromosomes apart during cell division.

The **nucleus** of each cell is surrounded by a layered membrane, with a thickness of 7.5 nm. The outer layer of this membrane is connected to the ER, and the outer and inner layers are connected by "press studs," creating pores in the nuclear membrane that allow the passage of ions, RNA and other macromolecules between the nucleus and the cell cytoplasm. These pores have an active role in the regulation of DNA synthesis, since they control the passage of DNA precursors and thus allow only a single duplication of the pre-existing DNA during each cell cycle. The inner surface of the membrane has nuclear lamina, a regular network of three proteins that separate the membrane from peripheral chromatin. DNA is distributed throughout the nucleoplasm wound around spherical clusters of histones to form nucleosomes, which are strung along the DNA like beads. These are then further aggregated into the chromatin fibers of approximately 30 nm diameter. The nucleosomes are supercoiled within the fibers in a cylindrical or solenoidal structure to form chromatin, and the nuclear lamina provide anchoring points for chromosomes during interphase (Figure 1.2):

- Active chromatin = euchromatin – less condensed
- Inactive (turned off) = heterochromatin – more condensed
- Before and during cell division, chromatin becomes organized into chromosomes.

Three types of cell lose their nuclei as part of normal differentiation, and their nuclear contents are broken down and recycled:

- Red blood cells (RBCs)
- Squamous epithelial cells
- Platelets.

Other cells may be multinuclear: syncytia in muscle and giant cells (macrophages), syncytiotrophoblast.

Nuclear RNA is concentrated in **nucleoli**, which form dense, spherical particles within the nucleoplasm (Figure 1.3); these are the sites where ribosome subunits, ribosomal RNA and transfer RNA are manufactured. RNA polymerase I rapidly transcribes the genes for ribosomal RNA from large loops of DNA, and the product is packed in situ with ribosomal proteins to generate new ribosomes (RNP: ribonucleoprotein particles).

Mitochondria are the site of aerobic respiration. Each cell contains 40–1000 mitochondria, and they are

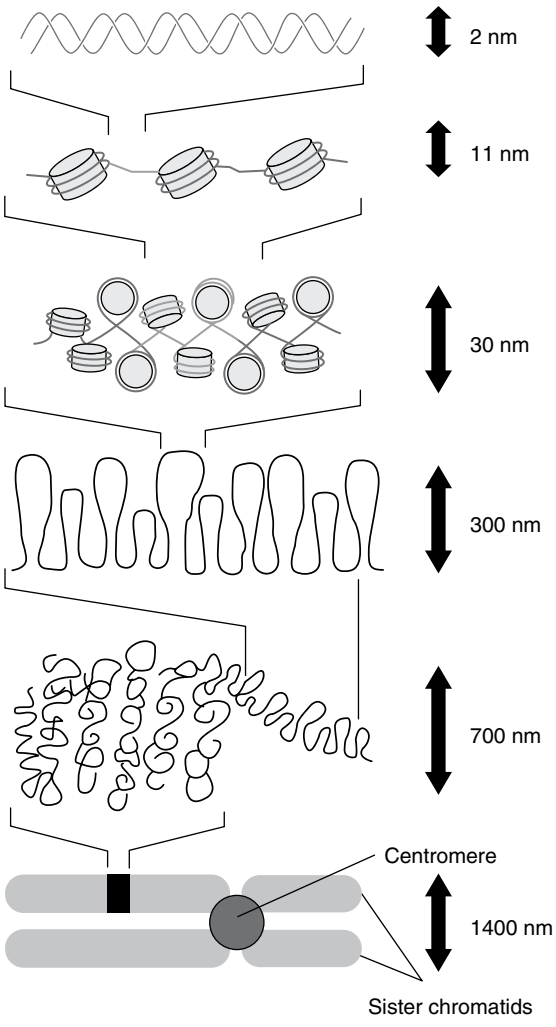

Figure 1.2 Levels of chromatin packaging. From the top: DNA double helix, nucleosome "beads on a string," chromatin fiber of packed nucleosomes, section of extended chromosome, condensed chromosome and finally the entire chromosome.

most abundant in cells that are physically and metabolically active. They are elliptical, 0.5–1 μm in size, with a smooth outer membrane, an intermembranous space, and a highly organized inner membrane which forms cristae (crests) with elementary particles attached to them, "F1-F0 lollipops," which act as molecular dynamos. The cristae are packed with proteins, some in large complexes: the more active the tissue, the more cristae in the mitochondria. Cristae are the site of intracellular energy production and transduction, via the Krebs (TCA) cycle, as well as processes of oxidation, dehydrogenation, fatty acid oxidation, peroxidation, electron transport chains and oxidative phosphorylation. They

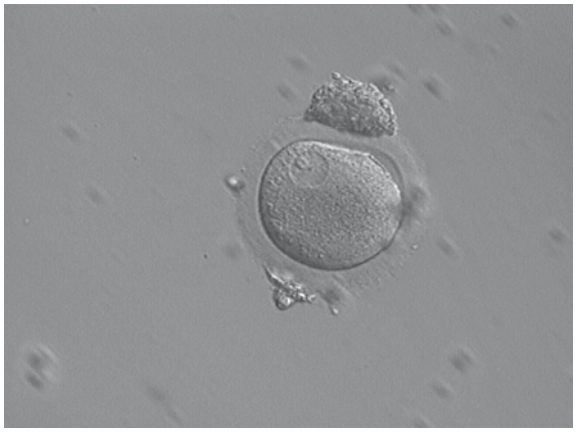

Figure 1.3 Human oocyte at germinal vesicle stage, showing prominent nucleolus.

also act as a Ca^{2+} store, and are important in calcium regulation. Mitochondria contain their own double-stranded DNA that can replicate independently of the cell, but the information for their assembly is coded for by nuclear genes that direct the synthesis of mitochondrial constituents in the cytoplasm. These are transported into the mitochondria for integration into its structures.

A number of rare diseases are caused by mutations in mitochondrial DNA, and the tissues primarily affected are those that most rely on respiration, i.e., the brain and nervous system, muscles, kidneys and the liver. All the mitochondria in the developing human embryo come from the oocyte, and therefore all mitochondrial diseases are maternally inherited, transmitted exclusively from mother to child. In the sperm, mitochondria are located in the midpiece, providing the metabolic energy required for motility; there are no mitochondria in the sperm head.

- Oocytes contain 100 000–1 000 000 mitochondria.
- Sperm contain 70–100 mitochondria, in the midpiece of each sperm. These are incorporated into the oocyte cytoplasm, but do not contribute to the zygote mitochondrial population – they are eliminated at the four- to eight-cell stage.
- All of the mitochondria of an individual are descendants of the mitochondria of the zygote, which contains mainly oocyte mitochondria.

The human mitochondrial genome

The sequence of human mitochondrial DNA was published by Fred Sanger in 1981, who shared the 1980 Nobel Prize in Chemistry with Paul Berg and Walter

Gilbert, "for their contributions concerning the determination of base sequences in nucleic acids." The mitochondrial genome has:

- Small double-stranded circular DNA molecule (mtDNA) 16 568 bp in length
- 37 genes that code for:
  2 ribosomal RNAs
  22–23 tRNAs
  10–13 proteins associated with the inner mitochondrial membrane, involved in energy production
- Other mitochondrial proteins are encoded by nuclear DNA and specifically transported to the mitochondria.
- Mitochondrial DNA is much less tightly packed and protected than nuclear DNA, and is therefore more susceptible to ROS damage that can cause mutations.
- As it is inherited only through the maternal line, mutations can be clearly followed through generations and are used as "markers" in forensic science and archaeology, as well as in tracking different human populations and ethnic groups.

Mitochondria can be seen in different distributions during early development (Figure 1.4); they do not begin to replicate until the blastocyst stage, and therefore an adequate store of active mitochondria in the mature oocyte is a prerequisite for early development.

- Germinal vesicle oocyte: homogeneous clusters associated with endoplasmic reticulum (ER)
- Metaphase I oocyte: polarized towards the spindle
- Metaphase II oocyte: perinuclear ring and polar body
- Embryos at 1c, 2c, 4c stages: perinuclear ring
- Cytoplasmic fragments in cleavage stage embryos contain large amounts of active mitochondria

The **endoplasmic reticulum (ER)** is an interconnected lipoprotein membrane network of tubules, vesicles and flattened sacs that extends from the nuclear membrane outwards to the plasma membrane, held together by the cytoskeleton. The ER itself is a membrane-enclosed organelle that carries out complex biosynthetic processes, producing proteins, lipids and polysaccharides. As new lipids and proteins are made, they are inserted into the existing ER membrane and the space enclosed by it. **Smooth ER (sER)** is involved in metabolic processes, including synthesis and metabolism of lipids, steroids and carbohydrates, as well as regulation of calcium levels. The surface of **rough ER (rER)** is studded with ribosomes, the units of protein synthesis machinery. Membrane-bound vesicles shuttle proteins between the rER and the Golgi apparatus, another part of the membrane system. The **Golgi apparatus** is important in modifying, sorting and packaging macromolecules for secretion from the cell; it is also involved in transporting lipids around the cell, and in making lysosomes.

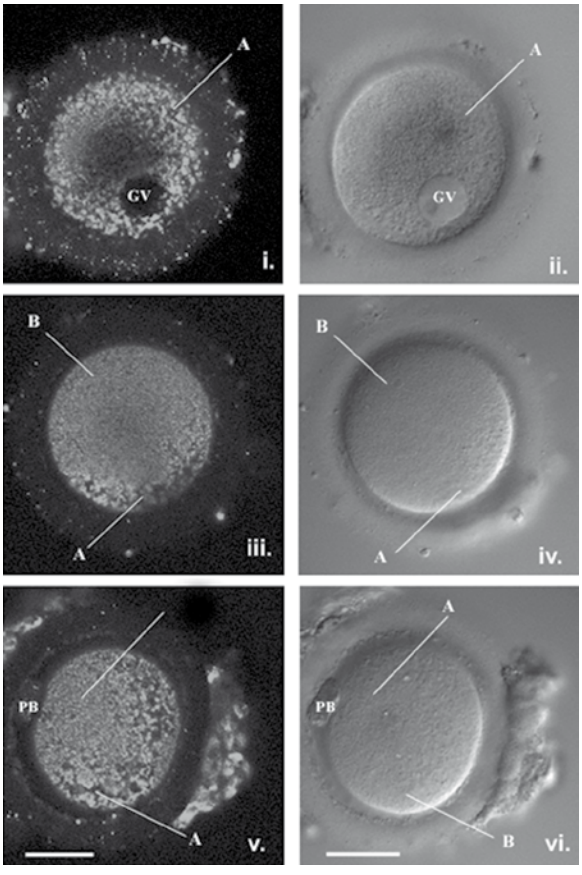

Figure 1.4 Mitochondrial aggregation patterns in a germinal vesicle (GV) oocyte (top), a metaphase I oocyte (center) and a metaphase II oocyte (bottom). Frames to the left are in fluorescence using the potential sensitive dye JC-1 to show the mitochondria, frames on the right are transmitted light images. The two mitochondrial patterns: A (granular-clumped) and B (smooth) are shown. PB = polar body. Scale bars = 50 µm. (With permission from Wilding *et al.*, 2001, *Human Reproduction* 16, pp. 909–917.) See color plate section.

rER

- Has attached 80 s ribonucleoprotein particles, the ribosomes (bacterial ribosomes are 70 s), which

are made in the nucleus and then travel out to the cytoplasm through nuclear pores.
- Ribosomes are composed of two subunits: 40 s and 60 s (bacteria: 30 s and 50 s); the association between the subunits is controlled by Mg^{2+} concentration.
- Polysomes = several ribosomes which move along a single strand of mRNA creating several copies of the same protein.

sER

- A series of flattened sacs and sheets, site of lipid and steroid synthesis.
- Cells that make large amounts of steroids have extensive sER.

The **Golgi apparatus** was first observed by Camillo Golgi in 1898, using a novel silver staining technique to observe cellular structures under the light microscope; he was awarded the 1906 Nobel Prize in Physiology or Medicine for his studies on the structure of the nervous system. The Golgi apparatus consists of a fine, compact network of tubules near the cell nucleus, a collection of closely associated compartments with stacked arrays of smooth sacs and variable numbers of cisternae, vesicles or vacuoles. It is connected to rER, linked to vacuoles that can develop into secretory granules, which contain and store the proteins produced by the rER. All of the proteins exported from the ER are funneled through the Golgi apparatus, and every protein passes in a strict sequence through each of the compartments (cis, tubules, trans). This process consists of three stages:

1. "Misdirected mail" – sends back misdirected proteins (cis).
2. "Addressing" – stacks of cisternae that modify lipid and sugar moieties, giving them "tags" for subsequent sorting.
3. "Sorting and delivering" (trans): proteins and lipids are identified, sorted and sent to their proper destination.

Transport occurs via vesicles, which bud from one compartment and fuse with the next. The Golgi apparatus will move to different parts of the cell according to the ongoing metabolic processes at the time – it is very well developed in secretory cells (e.g., in the pancreas).

The Golgi apparatus also makes **lysosomes**, which contain hydrolytic enzymes that digest worn-out organelles and foreign particles, acting as "rubbish bins" and providing a recycling apparatus for intracellular digestion; they contain at least 50 different enzymes, and "leaky" lysosomes can cause damage and kill cells. Macromolecules inside the cell are transported to lysosomes, those from outside the cell reach them by pinocytosis or phagocytosis; phagocytosis only occurs in specialized cells (e.g., white blood cells).

Peroxisomes are microbody vesicles that contain oxidative enzymes such as catalase; they dispose of toxic hydrogen peroxide, and are important in cell aging.

Metabolism in the mammalian cell

Four basic factors influence the metabolic activity of a cell:

1. Spatial: compartmentation, permeability, transport, interactions.
2. Temporal: products become substrates, positive and negative feedback.
3. Intensity/concentrations: precise amounts of reactants/substrates/products.
4. Determinants that specify the structure of enzymes and direct their formation/activation.

Molecules that are important in the biology/metabolism of the cell include carbohydrates, fats and lipids, and proteins.

Carbohydrates

Carbohydrates are made up of carbon (C), hydrogen (H) and oxygen (O), with the molecular ratio $C_x(H_2O)_y$.

- Monosaccharides: pentose – 5 C's (ribose, deoxyribose); hexose – 6 C's (glucose, fructose)
- Disaccharides: two monosaccharides (sucrose, maltose, lactose)
- Oligosaccharides: combine with proteins and lipids to form glycoproteins and glycolipids, important in cell–cell recognition and the immune response
- Polysaccharides: polymers, insoluble, normally contain 12 to 10 000 monosaccharides (starch, cellulose, glycogen)
  - Also form complexes with lipids and phosphate.

Fats and lipids

Fatty acids (FAs) have a long hydrocarbon chain ending in a carboxyl group:

- Saturated FAs have single bonds between carbon atoms.

- Unsaturated FAs have some double bonds between carbon atoms.

Lipids are made up of FAs plus water:

- Phospholipids are important in membranes.
- Glycolipids are important in receptors.

Proteins

The *primary structure* of a protein is a sequence of amino acids with peptide bonds:

–CONH–

Amino acids have at least one amino and one carboxyl group; they are amphoteric, and form dipolar zwitterions in solution. Proteins have *secondary structures*; they can be folded into a helix, or form beta sheets that are held together by hydrogen bonds:

- Alpha helix – tends to be soluble (most enzymes).
- Beta sheets – insoluble – fibrous tissue.

Proteins also have a three-dimensional *tertiary structure*, which is formed by folding of the secondary structure, held in place by different types of bond to form a more rigid structure: disulfide bonds, ionic bonds, intermolecular bonds (van der Waals – non-polar side chains attracted to each other).

High temperatures and extremes of pH denature proteins, destroying their tertiary structure and their functional activity.

Some proteins have a *quaternary structure*, with several tertiary structures fitted together; e.g., collagen consists of a triple stranded helix.

Enzymes are proteins that catalyze a large number of biologically important actions, including anabolic and catabolic processes, and transfer of groups (e.g., methylase, kinase, hydroxylase, dehydrogenase). Some enzymes are isolated in organelles, others are free in the cytoplasm; there are more than 5000 enzymes in a typical mammalian cell.

- Kinases: add a phosphate group, key enzymes in many activation pathways.
- Methylases: add a methyl group. DNA methylation is important in modifications that are involved in imprinting, lipid methylation is important for membrane stability, and proteins are also stabilized by methylation.

Most enzymes are conjugated proteins, with an active site that has a definite shape; a substrate fits into the active site, or may induce a change of shape so that it can fit.

- The rate of an enzymatic reaction is affected by temperature, pH, substrate concentration, enzyme concentration.
- Enzymes can be activated by removal of a blocking peptide, maintaining the S-H groups, or by the presence of a cofactor.
- The active site of an enzyme is often linked to the presence of an amino acid OH^- group (serine, threonine). Mutations at this level render the enzymes inactive.

Enzyme inhibitors can be:

- Competitive – structurally similar
- Noncompetitive – no similarity, form an enzyme/inhibitor complex that changes the shape of the protein so that the active site is distorted
- Irreversible: heavy metal ions combine with -SH causing the protein to precipitate. Lead (Pb^{2+}) and cadmium (Cd^{2+}) are the most hazardous; these cations can also replace zinc (Zn^{2+}), which is usually a stabilizer of tertiary structures.

Allosteric enzymes are regulated by compounds that are not their substrate, but which bind to the enzyme away from the active site in order to modify activity. The compounds can be activators or inhibitors, increasing or decreasing the affinity of the enzyme for the substrate. These interactions help to regulate metabolism by end-product inhibition/feedback mechanisms.

For example, phosphofructokinase (PFK): high ATP inhibits, low ATP activates.

Km is the substrate concentration that sustains half the maximum rate of reaction. Two or more enzymes may catalyze the same substrate, but in different reactions; if the reserves of substrate are low, then the enzyme with the lowest Km will claim more of the substrate.

Cytokines

- Cytokines are proteins, peptides or peptidoglycan molecules that are involved in signaling pathways. They represent a large and diverse family of regulatory molecules that are produced by many different types of cell, and are used extensively in cellular communication:
  - Colony stimulating factors
  - Growth and differentiation factors

- Immunoregulatory and proinflammatory cytokines function in the immune system (interferon, interleukins, tumor necrosis factors).

- Each cytokine has a unique cell surface receptor that conducts a cascade of intracellular signaling that may include upregulation and/or downregulation of genes and their transcription factors.
- They can amplify or inhibit their own expression via feedback mechanisms:
  - Type 1 cytokines enhance cellular immune responses:
    - Interleukin-2 (IL-2), gamma interferon (IFN-γ), TGF-β, TNF-β, etc.
  - Type 2 favor antibody responses:
    - IL-4, IL-5, IL-6, IL-10, IL-13, etc.
  - Type 1 and type 2 cytokines can regulate each other.

Metabolic pathways

Each metabolic pathway is a series of reactions, organized such that the products of one reaction become substrates for the next (Figure 1.5). The reactants in a pathway may be modified in a series of small steps, so that energy is released in controlled amounts, or minor adjustments can be made to the structure of molecules.

Anabolic pathways require energy to synthesize complex molecules from smaller units.

Catabolic pathways break molecules up into smaller units which can then be used to generate energy.

Each step in a pathway is catalyzed by a specific enzyme, and each enzyme represents a point for control of the overall pathway. The steps of the pathway

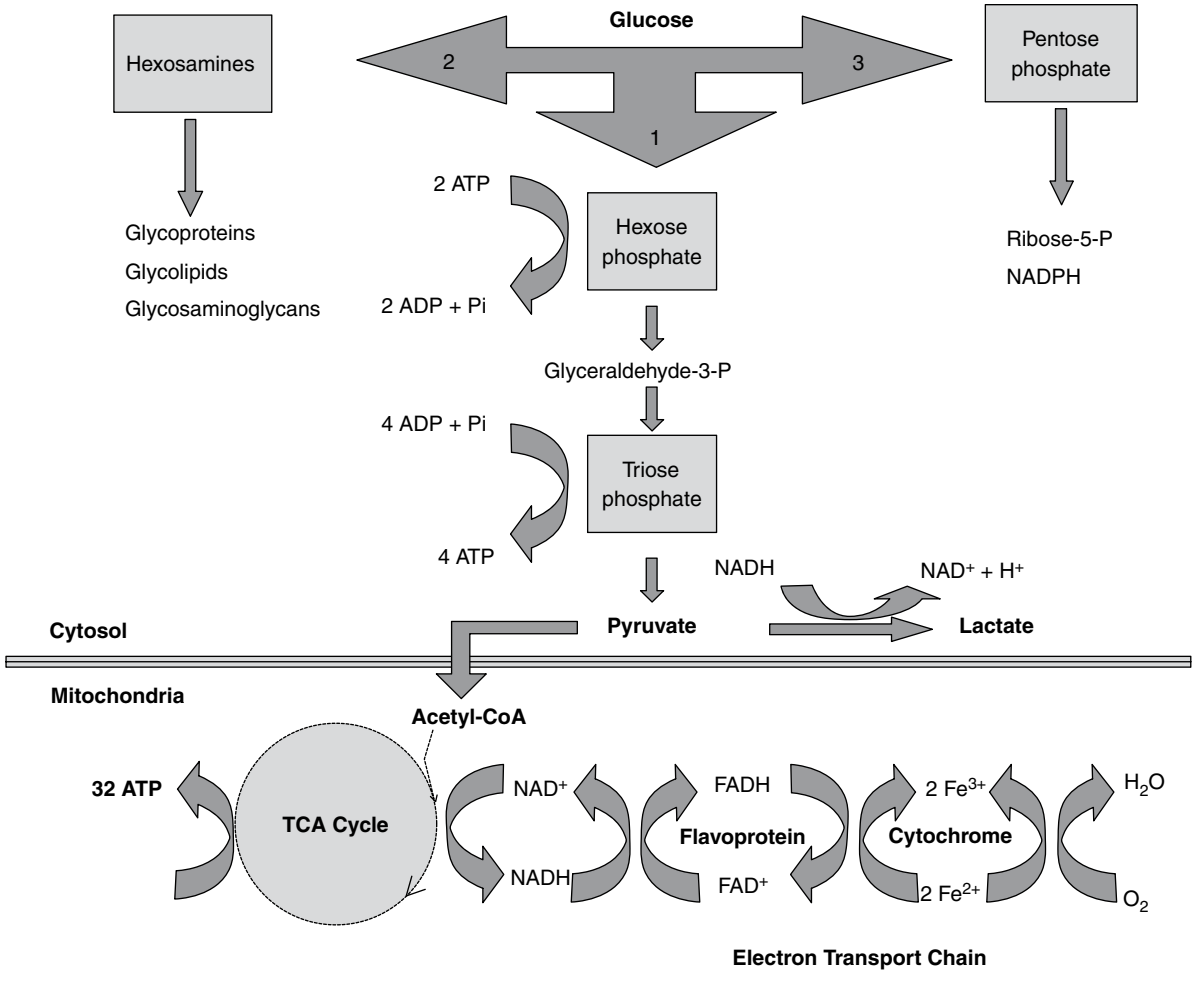

Figure 1.5 Pathways that metabolize glucose in a mammalian cell.

may be spatially arranged, so that the product of one reaction is in the right place to become the substrate of the next enzyme. This allows high local concentrations of substrate molecules to build up, and biochemical reactions to proceed rapidly. A pathway arranged in this manner may be catalyzed by a multienzyme complex.

- **Glycolysis**, which breaks glucose down into pyruvate, takes place in the cell cytoplasm; pyruvate enters mitochondria to be further metabolized.
- **Fatty acid oxidation** and the **Krebs cycle** (TCA or Citric acid cycle) take place in the mitochondrial matrix.

The Krebs cycle is part of a metabolic pathway that converts carbohydrates, fats and proteins into CO_2 and ATP, which is generated by a process of oxidative phosphorylation. ATP is exported from the mitochondria for use in protein synthesis, DNA replication, etc.: all energy-requiring processes of life are coupled to the cleavage of ATP:

$$ATP \Leftrightarrow ADP + Phosphate + Energy$$

ATP is exported from the mitochondria in exchange for ADP arising from the ATP that has been broken down to drive cellular metabolism.

Redox reactions: oxidation and reduction are electron-transfer processes, involving NAD-NADH.

- NADP(H) is generally used for anabolic reactions.
- NAD(H) is used for catabolic reactions.

These reactions need ubiquinone and cytochrome C, cytochrome oxidase (inhibited by cyanide).

- Oxidation: loss of electrons; reduction: gain of electrons.
- An oxidizing agent removes electrons and is itself reduced.
- A reducing agent gains electrons and is itself oxidized.

Reactive oxygen species (ROS, oxygen radicals)

ROS are molecules that contain the oxygen ion or peroxide; the presence of unpaired valence electrons makes them highly reactive. They are formed as a by-product of oxygen metabolism, and have an important role in cell signaling mechanisms. However, high levels of ROS (i.e., oxidative stress) can cause oxidative damage to nucleic acids, proteins and lipids, as well as inactivate enzymes by oxidation of cofactors.

Antioxidants such as ascorbic acid (vitamin C), tocopherol (vitamin E), glutathione, hypotaurine, pyruvic acid, uric acid and albumin are important in cellular defense mechanisms against ROS damage (Figure 1.6).

Superoxide dismutase (SOD)

SOD enzymes catalyze the dismutation of superoxide into oxygen and hydrogen peroxide, an important defense against potential ROS damage. Three SOD enzymes are present in mammalian cells:

- SOD1: dimer, present in the cytoplasm, contains Cu^{2+} and Zn^{2+}
- SOD2: tetramer, mitochondrial enzyme, contains Mn^{2+}
- SOD3: tetramer, extracellular, contains Cu^{2+} and Zn^{2+}

ROS can cause damage to DNA in oocytes, sperm and embryos, with important consequences for fertilization and embryo development (Guerin *et al.*, 2001). Oocytes are particularly susceptible during the final stages of follicular growth, and ROS damage to sperm DNA has been strongly linked to male infertility (Sakkas *et al.*, 1998).

Fundamental principles of molecular biology

The **nucleic acids, DNA** (deoxyribose nucleic acid; Figure 1.7) **and RNA** (ribose nucleic acid), are made up of:

1. Nucleotides: organic compounds containing a nitrogenous base
2. Sugar: deoxyribose in DNA, ribose in RNA
3. Phosphate group.

Nucleotides are purines and pyrimidines, determined by the structure of the nitrogenous base.

	DNA	RNA
Purines	Adenine (A)	Adenine (A)
(double ring)	Guanine (G)	Guanine (G)
Pyrimidines	Cytosine (C)	Cytosine (C)
(single ring)	Thymine (T = methylated U)	Uracil (U)

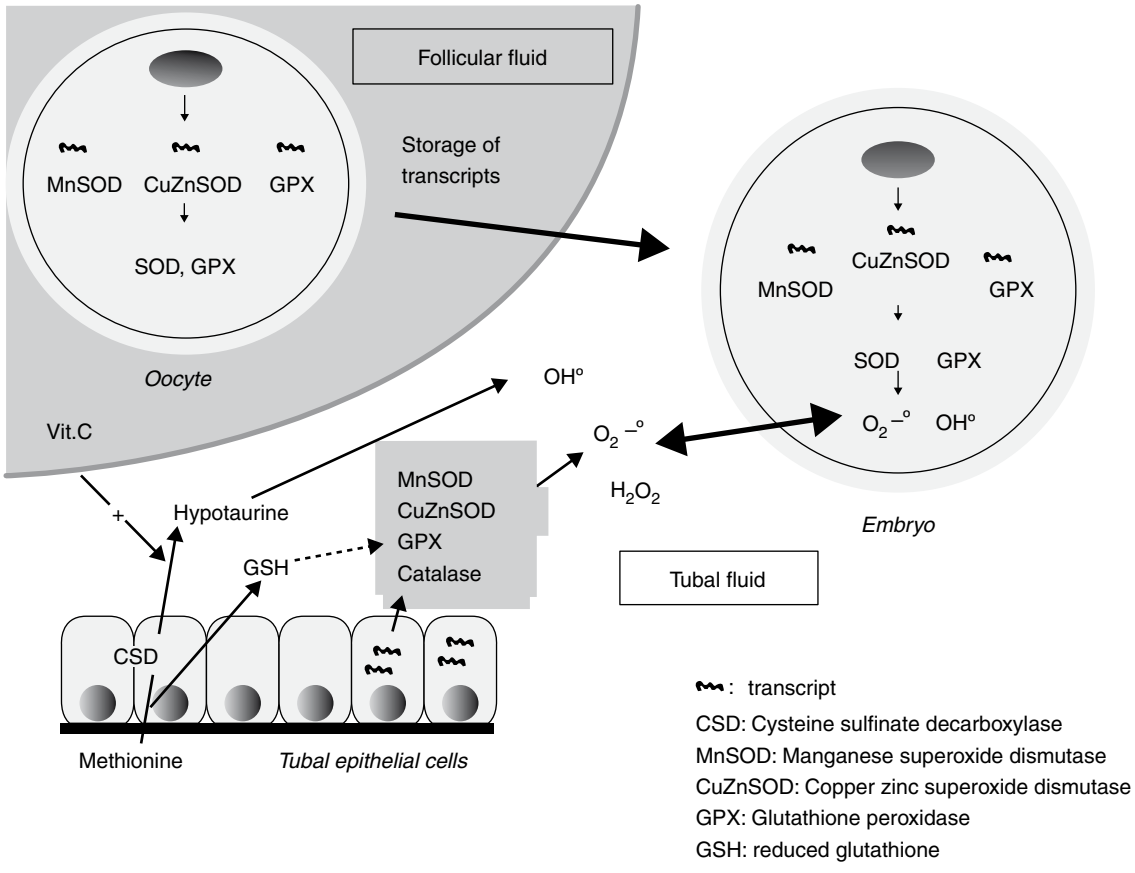

Figure 1.6 Mechanisms that protect oocytes and embryos from ROS damage (with thanks to Y. Ménézo).

Methylation of cytosine is important in gene silencing and imprinting processes.

Nucleotides also function as important cofactors in cell signaling and metabolism: coenzyme A (CoA), flavin adenine dinucleotide (FAD), flavin mononucleotide, adenosine triphosphate (ATP), nicotinamide adenine dinucleotide phosphate (NADP).

DNA

- Double-stranded helix with paired bases to form complementary strands
- G=C or A=T
- Pentose deoxyribose – phosphate backbone
- Stabilized by H bonds between purines and pyrimidines, on the inside of the helix
- Each pitch of the double helix has 10 base pairs.

DNA replication

DNA copies itself by semi-conservative replication: each strand acts as a template for synthesis of a complementary strand.

1. Free nucleotides are made in the cytoplasm, and are present in the nucleoplasm before replication begins.
2. The double helix unwinds, and hydrogen bonds, holding the two DNA strands together, break. This leaves unpaired bases exposed on each strand.
3. The sequence of unpaired bases serves as a template on which to arrange the free nucleotides from the nucleoplasm.
4. DNA polymerase moves along the unwound parts of the DNA, pairing complementary nucleotides from the nucleoplasm with each exposed base.

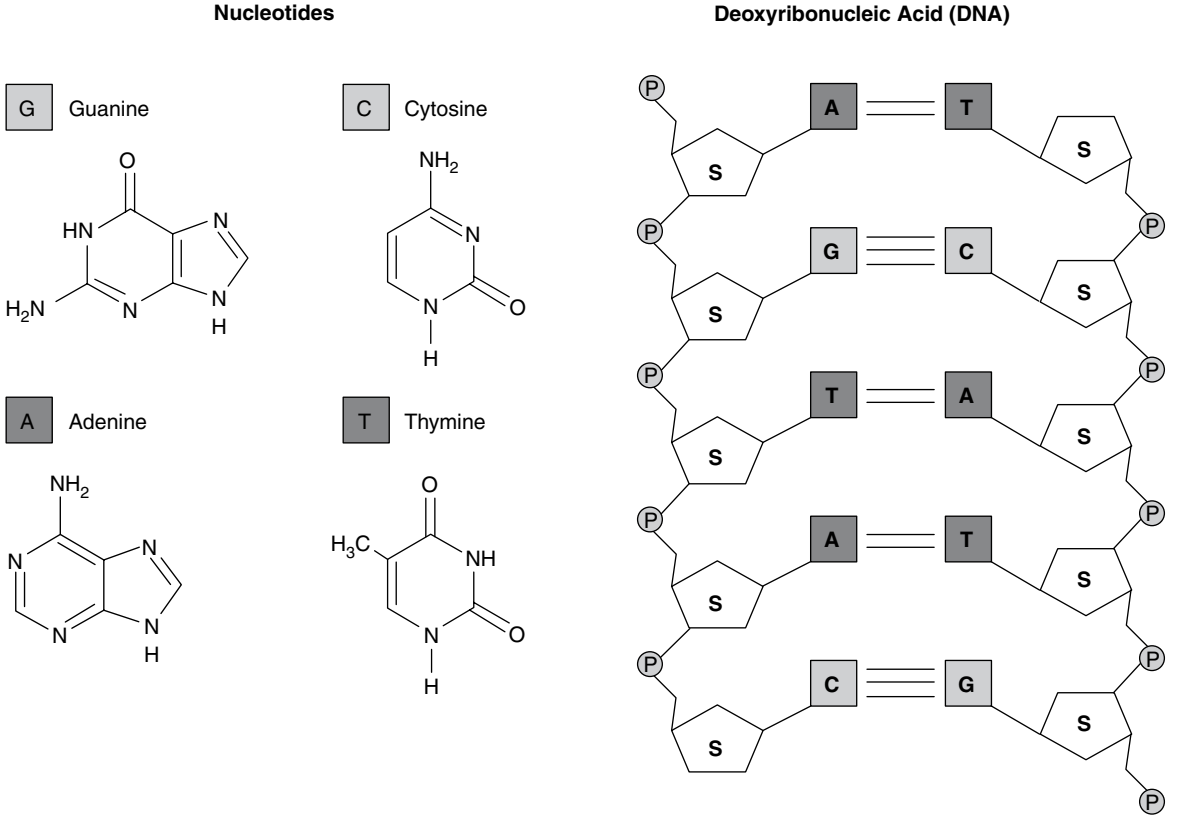

Figure 1.7 Structure of DNA. Complementary base pairs form the DNA double helix; two hydrogen bonds form between A and T, three hydrogen bonds form between G and C. The two polynucleotide chains must be antiparallel to each other to allow pairing. S = sugar, P = phosphate group.

5. The same enzyme connects the nucleotides together to form a new strand of DNA, hydrogen bonded to the old strand:
   - DNA polymerase forms new hydrogen bonds on the 5′3′ strand
   - DNA ligase acts on the 3′5′ strand
   - Several replication points appear along the strand, which eventually join.

6. DNA is then mounted on "scaffolding proteins," histones – and this is then wrapped around non-histones to form chromatin. Histones are basic proteins that bind to nuclear DNA and package it into nucleosomes; the regulation of gene expression involves histone acetylation and deacetylation. There are two ATP-dependent remodeling complexes and acetyltransferases that preferentially bind activated states and fix chromatin configurations:
   - Histone acetyltransferase coactivator complex
   - Histone deacetylase corepressor complex.

Methylation of protamines and histones is a crucial component of imprinting processes: an association has been found between Beckwith–Wiedemann syndrome and epigenetic alterations of LitI and H19 during in-vitro fertilization (DeBaum *et al.*, 2003).

Each mammalian cell contains around 1.8 m of DNA, of which only 10% is converted into specific proteins; the noncoding part of the DNA still carries genetic information, and probably functions in regulatory control mechanisms.

- Genes = chief functional unit of DNA
- Exons – contain information for the amino acid sequence of a protein (coding sequence)
- Introns = non-coding regions in between exons
- Codon = a group of three nucleotide bases which code for one amino acid.