

1

The Application of Science to Clinical Practice

Learning Points

- There are endocrine, paracrine and autocrine hormones
- They comprise three groups, those derived from tyrosine, protein and peptide hormones and steroids
- Two different types of cell synthesize either peptide and protein hormones or steroid hormones
- Hormones elicit effects on target cells by interacting with receptors either at the cell surface or within the cytoplasm and nucleus
- Both traditional and non-traditional inheritance play a major part in human health and disease

Hormones belong to a class of regulatory molecules synthesized in special cells. These cells may be collected into distinct glands or may be found as single cells within some other organ, such as the gastrointestinal tract. Endocrine hormones are released from the cells that make them into the adjacent extracellular space, whence they enter a local blood vessel and circulate to their target cells. Some cells secrete hormones that act on themselves (autocrine hormones); some (paracrine hormones) affect nearby cells without entering the bloodstream. Molecules secreted by neurones that excite or inhibit other neurones or muscle by means of synapses are called neurotransmitters. Sometimes, however, both neurotransmitters and hormones are secreted by neurones, thus forming the neuroendocrine system.

From the chemical standpoint, there are three groups of hormones: the first are derived from the amino acid tyrosine, the second are peptide and protein hormones and the third are steroid hormones.

Hormones derived from tyrosine include epinephrine, which is secreted by the adrenal medulla, and norepinephrine, which can be produced in the adrenal medulla but is also produced at sympathetic nerve endings, where it acts as a neurotransmitter. Dopamine, another derivative of tyrosine, is a neurotransmitter that can also act as a hormone when released from the median eminence to suppress the secretion of prolactin (PRL) from the anterior pituitary. The thyroid hormones (thyroxine and triiodothyronine) each have two fused molecules of tyrosine; thyroxine has four iodine atoms attached to the amino acid rings, while triiodothyronine has three iodine atoms substituted in the two aromatic rings (Fig. 1.1).

Protein and peptide hormones vary considerably in size. Thyrotrophin-releasing hormone (TRH) has only three amino acid residues while many of the hormones from the gastrointestinal tract, such as secretin from the duodenum and gastrin from the stomach, are larger, with up to 34 amino acids, while parathyroid hormone (PTH) is larger still, with 84. Ring structures linked by disulphide bridges are present in some hormones (Fig. 1.2). An intrachain disulphide bond to form a

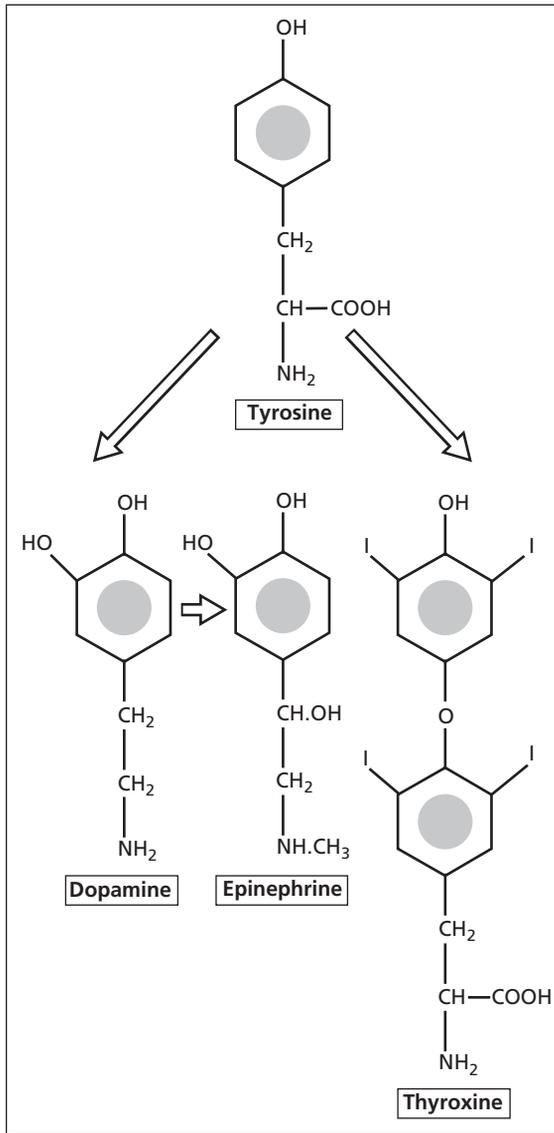


Figure 1.1 Hormones derived from tyrosine.

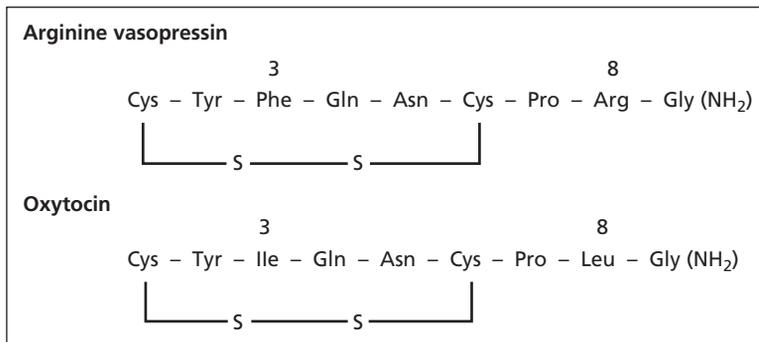


Figure 1.2 Structures of arginine vasopressin and oxytocin. Note the minor differences in chemical structure which confer major differences in action.

ring of seven amino acids at the amino-terminus is also found in calcitonin. Insulin may be regarded as a small protein or a large peptide; it consists of A-chains and B-chains linked by interchain disulphide bonds. Insulin is synthesized as a large precursor molecule (proinsulin) in a single chain of which a section, the connecting- (or C-) peptide, is subsequently removed by enzymatic hydrolysis after the disulphide bonds have been formed: the remaining two linked chains make up the insulin molecule. A number of other peptide hormones are also synthesized in larger precursor forms that are modified prior to secretion.

Some hormones are quite large proteins, such as the glycoprotein hormones from the anterior pituitary, each of which has two peptide chains. Follicle-stimulating hormone (FSH), luteinizing hormone (LH) and thyrotrophin-stimulating hormone (TSH) each have two chains, referred to as the α - and β -subunits. The two subunits are synthesized quite separately; the α -subunit in each is very similar but the β -subunits are different and confer the biological specificities on the hormones. Within a given species, there may be considerable microheterogeneity of the structures of these hormones, so that a number of naturally occurring variants coexist. The variability of the glycoprotein hormones is largely due to differences in their carbohydrate composition, and the variants are referred to as isohormones or isoforms.

Steroids are a class of lipids derived from cholesterol and include cortisol, aldosterone, testosterone, progesterone and oestradiol. Small changes in the basic chemical structure (Fig. 1.3) cause dramatic changes in the physiological action of this group of hormones.

Biosynthesis of peptide hormones

Protein or peptide hormone synthesis starts with the transcription of a gene, proceeds through translation of a messenger ribonucleic acid (mRNA) and culminates in post-translational modification of the peptide or protein hormone (Fig. 1.4). The gene consists of double-stranded (ds) deoxyribonucleic acid (DNA).

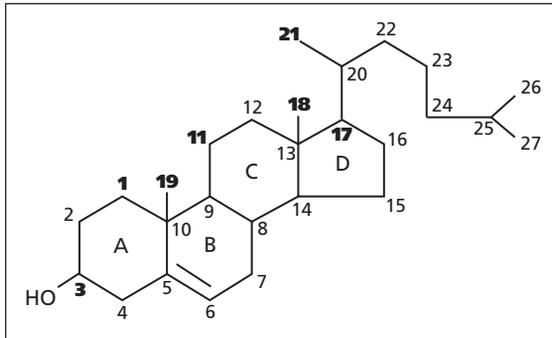


Figure 1.3 Structure of cholesterol, which is the base of the steroid nucleus. Cholesterol itself has a side chain on C17 (making it a C27 compound) which is removed in the biosynthesis of classic steroid hormones but retained in vitamin D.

At its 5' end (the 'upstream' side) is the regulatory region known as the promoter, which includes the TATA box (a sequence of seven thymidine–adenosine bases). This is followed by a variable number of exons (expressed sequence regions) and introns (intervening sequences), which make up the structural gene.

RNA polymerase produces an RNA transcript of the exons and introns in the form of pre-mRNA. Removal of the RNA sequences derived from the introns is followed by splicing together of the exon-derived sequences. Further post-transcriptional changes include the addition of a 7-methyl guanosine (7-meg) cap at the 5' end and a poly a (a series of adenosine residues) tail on the 3' end. When the mature mRNA is bound to a ribosome, translation occurs to give a peptide precursor that includes the signal peptide of the prehormone (or preprohormone). Post-translational processing is needed before the hormone is ready for secretion.

Storage

Protein- or peptide-hormone-secreting cells store the newly synthesized hormone in small vesicles or secretory granules scattered around the periphery of

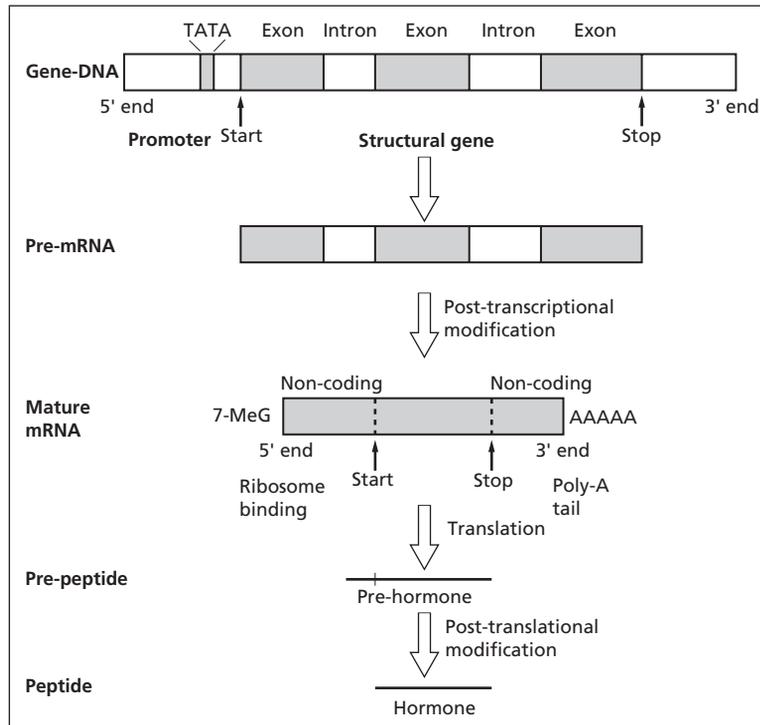


Figure 1.4 Schematic representation of a gene and of the events leading to the synthesis of a peptide hormone.

the cells just inside the cell membrane. Movement of the vesicles from the Golgi apparatus to a position near the cell membrane is influenced by two types of filamentous structure, called microtubules and microfilaments, which are found in all cells.

Secretion

The cell requires a stimulus before the stored prohormone is activated and released. The stimulation may be hormonal and usually involves a change in permeability of the cell to calcium ions which are required for interaction between the vesicle and plasma membranes and for the activation of enzymes, microfilaments and microtubules. Specific endopeptidases in the storage vesicle are activated during the secretory process and produce the active form of the hormone for release from the cell.

The mode of secretion in the cell is by exocytosis. The membrane of an intracellular storage granule fuses with the plasma membrane of the cell which parts near the point of fusion so that the content of the vesicle is secreted into the extracellular space surrounding the blood vessels. The membrane that originally surrounded the vesicle is recycled within the cell.

Steroid hormones

Cholesterol is the precursor of all steroid hormones. All the steroid-hormone-synthesizing cells of the body, the adrenal cortex, placenta, testis and ovary, contain intracellular fat droplets in the cytoplasm composed principally of cholesterol esters, the storage form of the hormone precursor. Cholesterol moves to the mitochondria to be converted to pregnenolone. This is transported to the surrounding smooth endoplasmic reticulum where it is transformed into the appropriate steroid hormone by a series of reactions (see Chapter 6, page 99). It is not known how the hormones get out of the cell but steroid-secreting cells, unlike protein- and peptide-producing cells, do not store hormone in a state ready for secretion but synthesize it for secretion as required.

Transport of hormones in the blood

Most peptide and protein hormones circulate in the bloodstream with little or no association with

serum proteins. There are specific transport proteins (thyroxine-, cortisol- and sex-hormone-binding globulins) in the circulation that bind thyroxine and many of the steroid hormones. The high specificity of these transport proteins is such that minor changes in the structure of hormones affect binding. For example, aldosterone is only weakly bound to cortisol-binding globulin. Many hormones may also loosely associate with other circulating proteins, especially albumin.

Protein-bound hormone is in equilibrium with 'free' (or unbound) hormone. The free hormone can diffuse to tissues more readily and so the physiological state usually corresponds more closely with the concentration of the free hormone. As a result of changes in the concentration of binding protein, the total and bound concentrations of hormone may alter quite markedly, even though this is accompanied by only a small change in the free hormone concentration with the result that the physiological status remains unaltered.

Hormone action

Hormones elicit their effects on target cell function by interacting with receptors either at the cell surface or within the cytoplasm and nucleus (Fig. 1.5). Receptors for the pituitary-derived proteins, insulin and the catecholamines are present at the plasma membrane; steroid and thyroid hormones access intracellular-binding sites.

The concentration of each receptor can vary and a cell may become more or less sensitive to a given extracellular concentration of ligand. Sensitization can occur by increasing the number of binding sites available through a combination of increased receptor synthesis and decreased degradation. Cells can become refractory (desensitized) by altering receptor localization (e.g. by internalizing cell-surface receptors), reducing receptor levels or recruiting molecules that deactivate intracellular signaling pathways.

Cell-surface receptors

Of the two major groups of cell-surface receptors, the first relies upon tyrosine kinase for the initiation of signaling and the second group tends to activate serine or threonine kinases by coupling to G-proteins.

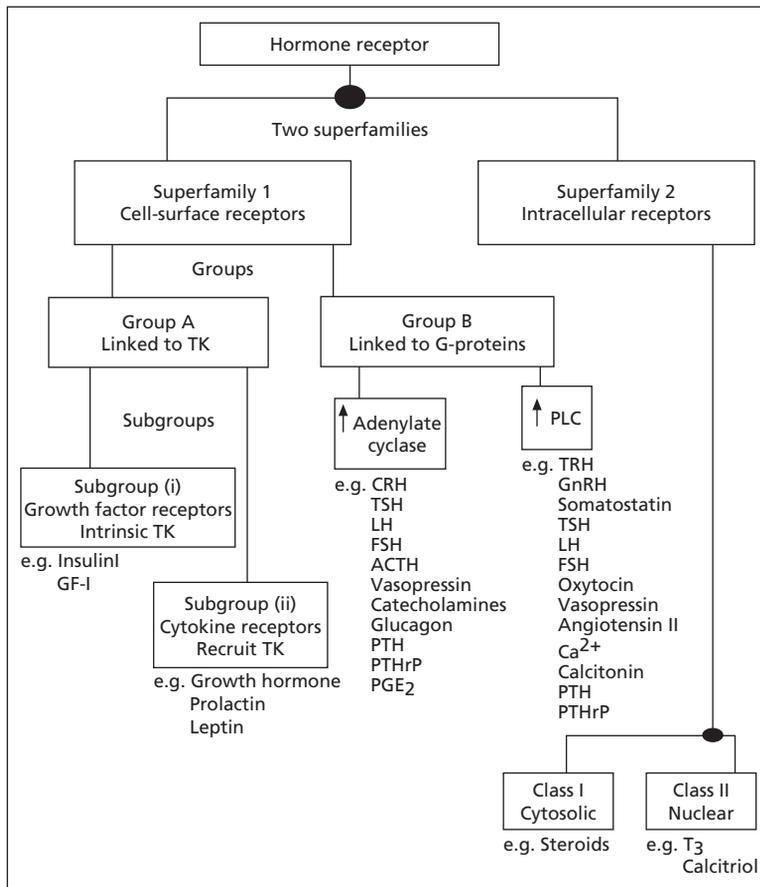


Figure 1.5 A composite diagram showing the different classes of hormone receptors.

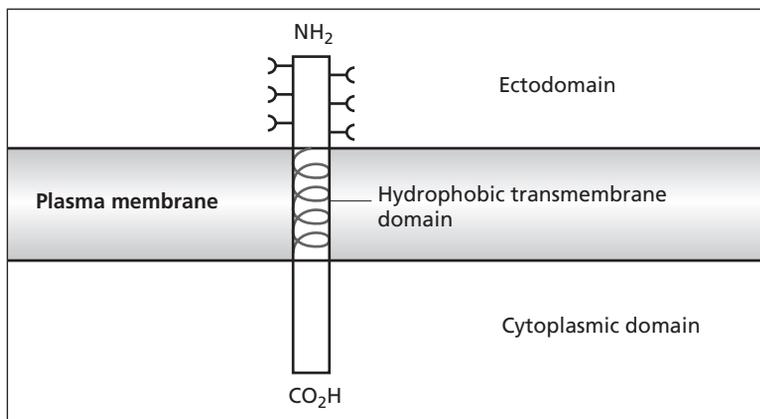


Figure 1.6 Schematic representation of a membrane-spanning cell-surface receptor with three clearly identifiable domains.

However, there is an underlying structural unity in all cell-surface receptors because each is made up of three segments, an extracellular domain, a transmembrane region and a cytoplasmic domain (Fig. 1.6).

The N-terminus of the protein forms the extracellular component of the receptor, which is responsible for hormone recognition and binding. The transmembrane region varies in structure from a simple linear stretch

Table 1.1 Examples of G-protein-coupled receptors (GPCRs).

| Family | Examples | G-protein | |
|---------------------|--------------------------|---------------------------------------|--------------------------|
| A | TRH receptor | Gq α | |
| | GnRH receptor | | |
| | Oxytocin | | |
| | Biogenic amine receptors | Gs α /Gq α | |
| | FSH receptor | | |
| | LH receptor | | |
| | TSH receptor | | |
| | Vasopressin | | |
| | B | Somatostatin | Gi α /Gq α |
| | | Melanocortin receptor | Gs α |
| Calcitonin receptor | | Gs α /Gi α /Gq α | |
| CRH receptor | | Gs α | |
| Glucagon receptor | | | |
| C | PTH receptor | Gs α /Gq α | |
| | PTHrP receptor | Gs α /Gq α | |
| | Calcium receptors | Gq α /Gi α | |
| | Glutamate receptors | Gs α /Gq α | |

CRH, corticotrophin-releasing hormone; GnRh, gonadotropin-releasing hormone; PTHrP, parathyroid hormone-related protein

of amino acids to a more complex arrangement that threads the plasma membrane. This segment is often regarded as a passive anchor but it can influence receptor function as, for example, mutations in the transmembrane region of the fibroblast growth factor (FGF) receptor are associated with achondroplasia. The cytoplasmic C-terminus of the receptor forms the effector region of the molecule because it initiates an intracellular signaling cascade that eventually results in the cellular response.

G-protein-coupled receptors (GPCRs) form a superfamily of more than 1000 membrane proteins. These receptors transduce hormonal signals and also mediate the cellular response to neurotransmitters, lipids, nucleotides, ions and sensory stimuli, such as light, smell and taste.

As their name suggests, activation of GPCRs generally leads to the recruitment of intracellular G (guanine)-proteins and then the generation of second messengers, for example cyclic adenosine monophosphate (cAMP) and inositol 1,4,5-triphosphate (IP₃). Some of these receptors can signal through G-protein-independent pathways.

Although GPCRs have the same basic design as the tyrosine kinase-linked receptors, in that they possess extracellular, transmembrane and intracellular domains. They can be grouped into three families, A, B and C (Table 1.1) on the basis of sequence similarity within the transmembrane region. There is little similarity between the groups, apart from the characteristic tertiary structure facilitated by the seven transmembrane helices.

Defects

Given their numerous and varied ligands, it is not surprising that mutations in GPCRs or their interacting G-proteins are associated with endocrine disease. Mutations that alter the extracellular (ligand-binding) domains of the receptor lead to hormone resistance (e.g. the TSH receptor), whereas aberrations in the transmembrane region of the receptor can result in altered receptor function.

Germline mutations in Xq28, which codes for the vasopressin V2 receptor, cause receptor misfolding and loss of receptor function so that circulating vasopressin, despite being present at very high levels, cannot

Table 1.2 Examples of defects in intracellular receptors that are associated with endocrine disease.

| Receptor | Clinical effects | Which are due to decreased |
|---------------------|---|-----------------------------------|
| ARs | Partial or complete AR insensitivity syndromes | Receptor number |
| | | AR binding |
| | | AR dimerization |
| | | Expanded CAG repeat in N-terminus |
| Glucocorticoid | Kennedy syndrome | Expanded CAG repeat in N-terminus |
| | Breast cancer | AR dimerization |
| | Prostate cancer | AR response to progesterone |
| ER | Generalized inherited glucocorticoid resistance | Hormone binding |
| | | GR number |
| | | DNA binding |
| T ₃ (TR) | Usually lethal | Hormone binding |
| | ER resistance | DNA binding |
| Calcitriol (VDR) | Resistance to thyroid hormone | TR β gene defects |
| | | T ₃ binding |
| | Calcitriol-resistant rickets | VDR dimerization |

Calcitriol (VDR) Calcitriol-resistant rickets VDR dimerization AR, androgen; ER, estrogen; GR, glucocorticoid receptor, VDR, vitamin D receptor.

increase urine concentration and nephrogenic diabetes insipidus results. Some cases of early onset severe obesity may be explained by functional defects in the melanocortin-4 receptor.

Activating mutations are also detrimental, presumably by altering crucial helix–helix interactions so that the receptor is active even in the absence of ligand. Familial male precocious puberty (testotoxicosis) is the result of such a mutation in the gene coding for the LH receptor (LHR), and activating mutations in the transmembrane domain of the TSH receptor have been reported in association with neonatal hyperthyroidism and toxic thyroid adenomas in adults.

Mutations resulting in the loss of Gs α function are linked to pseudohypoparathyroidism (Albright's hereditary osteodystrophy). If the mutation is maternally transmitted, resistance to the multiple hormones that activate Gs α in their target tissues occurs. Mutations resulting in the constitutive activation of Gs α cause McCune–Albright syndrome and some cases of acromegaly.

Intracellular receptors

Receptors for hormones such as the sex steroids, glucocorticoids, thyroxine and aldosterone are part

of a large family of receptors (>150 members) that are located inside the cell (Fig. 1.5). These receptors function as hormone-regulated transcription factors and control the expression of specific target genes by interacting with regions close to the gene promoters. Consequently, the cellular response to these hormones takes longer than the quickfire cell-surface receptor/second-messenger systems described above.

Defects

Mutations in the genes coding for intracellular receptors are responsible for numerous endocrinopathies as they can result in hormone resistance (Table 1.2). Studies of the glucocorticoid receptor suggest that defects resulting in abnormal interactions with co-activator molecules, and indeed problems with the co-activators themselves, may also be the cause of hormone resistance syndromes.

Target tissue metabolism

Some of the hormones that work through intracellular receptors are converted by enzymes expressed in their target cells to metabolites that are more potent because of their higher affinity for the receptor. For example, tissue-specific deiodinases convert T₄ to T₃, 5 α -reductase metabolizes testosterone to dihydrotestosterone and 1 α -hydroxylase in the mitochondria

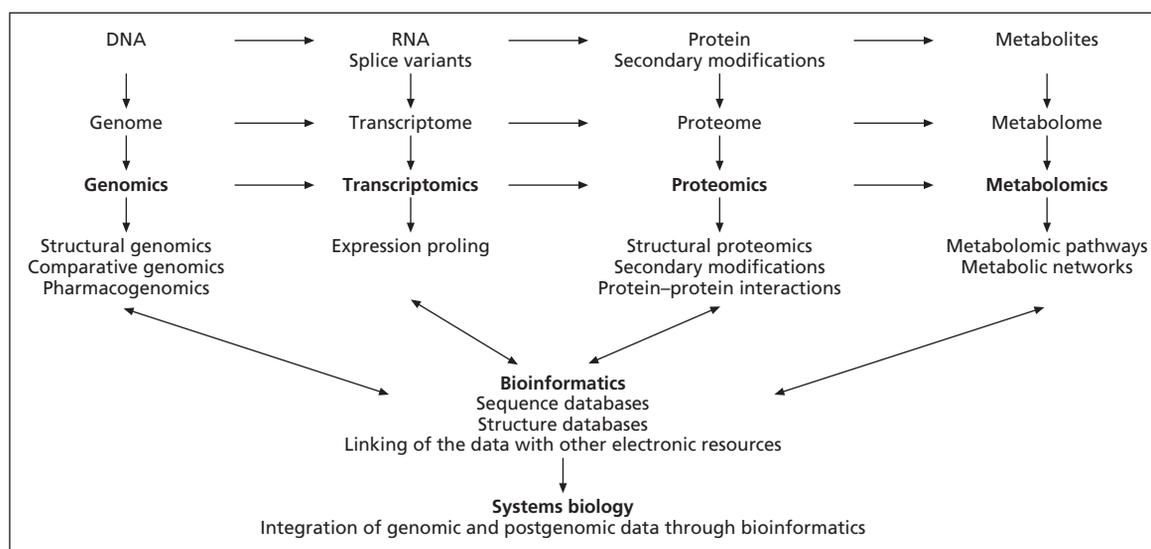


Figure 1.7 Overview of genomics, transcriptomics, proteomics, metabolomics, bioinformatics and systems biology.

of cells in the renal tubule converts 25-OH vitamin D to calcitriol. These ‘activation’ steps offer a way of achieving a range of effects, and various disorders can result from defects in target tissue metabolism. The best-known example is androgen insensitivity.

11 β -hydroxysteroiddehydrogenase (11 β HSD), which is expressed by aldosterone responsive cells in the kidney, converts cortisol to cortisone to prevent the overstimulation of the mineralocorticoid receptor that would otherwise occur as a result of the high concentration of cortisol in relation to the circulating level of aldosterone. Deficiency or impaired function of this enzyme leads to the hypertension and hypokalemia characteristic of the apparent mineralocorticoid excess (AME) syndrome.

Measurement of the concentrations of hormones in blood

Hormones are generally measured by immunoassays, although other techniques, particularly bioassays, play significant roles. All assays rely on a comparison between responses produced in the assay system by the sample and those produced by different known concentrations of a reference preparation. For immunoassays and bioassays, a calibration curve is generated with the reference preparation, and the

unknown concentration of the hormone in the sample can then be interpolated from this.

Bioassays, which measure the potency of a hormone by quantifying its biological effect, suffer many practical problems so their use is limited to research. Immunoassays, which rely on the recognition of a hormone by an antibody, are capable of high sample throughputs, which has resulted in their widespread use. The four main attributes that account for their successful application in diagnostic services are their sensitivity, specificity, precision and convenience.

Genetics (Fig. 1.7)

With the exception of simple trauma, almost every disease has a genetic component. In *monogenic disorders*, such as congenital adrenal hyperplasia (CAH), the genetic component is the major etiological factor. In *complex disorders*, multiple genes, in conjunction with environmental and lifestyle factors, contribute to the pathogenesis; hence their designation as *polygenic* or *multifactorial* disorders. In other instances, genetic factors influence the manifestation of disease indirectly by defining the host’s susceptibility and resistance as, for example, in infectious disease.

The term *genome*, introduced before the recognition that DNA is the genetic material, designates the

totality of all genes on all chromosomes in the nucleus of a cell. *Genomics* refers to the discipline of mapping, sequencing and analyzing genomes. Because of the rapidly growing list of mapped and sequenced genomes of numerous organisms, genomics is currently undergoing a transition with increasing emphasis on functional aspects.

Genome analysis can be divided into *structural genomics* and *functional genomics*. The analysis of differences among genomes of individuals of a given species is the focus of *comparative genomics*. The complement of mRNAs transcribed by the cellular genome is called the *transcriptome* and the generation of mRNA expression profiles is referred to as *transcriptomics*.

The term *proteome* has been coined to describe all the proteins expressed and modified following expression of the entire genome in the lifetime of a cell. *Proteomics* refers to the study of the proteome using techniques of protein separation and identification. The emerging field of *metabolomics* aims at determining the composition and alterations of the *metabolome*, the complement of low-molecular-weight molecules. The relevance of these analyses lies in the fact that proteins and metabolites function in *modular networks* rather than linear pathways. Hence, any physiological or pathological alteration may have many effects on the proteome and metabolome.

The growth of biological information has required computerized databases to store, organize, annotate and index the data. This has led to the development of *bioinformatics*, the application of informatics to (molecular) biology. Computational and mathematical tools are essential for the management of nucleotide and protein sequences, the prediction and modeling of secondary and tertiary structures, the analysis of gene and protein expression and the modeling of molecular pathways, interactions and networks. Numerous continuously evolving databases provide easy access to the expanding information about the genome of man and other species. The integration of data generated by transcriptomic, proteomic and metabolomic analyses through informatics, *systems biology*, is aimed at understanding phenotypic variations and creating comprehensive models of cellular organization and function. These efforts are based on the expectation that an understanding of the complex and dynamic changes in a biological system may provide insights into pathogenic processes and

the development of novel therapeutic strategies and compounds.

Mutations and human disease

Mutations are an important cause of genetic diversity as well as disease. A *mutation* is any change in the nucleotide sequence of DNA, regardless of its functional consequences. Mutations can affect one or a few nucleotides or consist of gross numerical or structural alterations in individual genes or chromosomes. Large deletions may affect a portion of a gene or an entire gene or, if several genes are involved, they may lead to a *contiguous gene syndrome*. Occasionally, mispairing of homologous sequences leads to *unequal cross-over*. This results in gene duplication on one of the chromosomes and gene deletion on the other chromosome. For example, a significant fraction of growth hormone (*GH*) gene deletions involves unequal crossing-over.

Mutations involving single nucleotides are referred to as *point mutations*. Substitutions are called *transitions* if a purine is replaced by another purine base (A to G) or if a pyrimidine is replaced by another pyrimidine (C to T). Changes from a purine to a pyrimidine or vice versa are referred to as *transversions*. Certain DNA sequences, such as successive pyrimidines or CG dinucleotides, are particularly susceptible to mutagenesis. Therefore, certain types of mutations (C to T or G to A) are relatively common. Moreover, the nature of the genetic code results in over-representation of certain amino acid substitutions. If the DNA sequence change occurs in a coding region and alters an amino acid, it is called a *missense mutation*. Depending on the functional consequences of such a missense mutation, amino acid substitutions in different regions of the protein can lead to distinct phenotypes. Small deletions and insertions alter the reading frame if they do not represent a multiple of three bases. Such *frameshift* mutations lead to an altered carboxy-terminus. Mutations may also be found in the regulatory sequences of genes and result in reduced gene transcription. Mutations in intronic sequences or in exon junctions may destroy or create splice donor or splice acceptor sites.

Some mutations are lethal, some have less deleterious yet recognizable consequences and some confer evolutionary advantage. Mutations in germ cells (*germline mutations*) can be transmitted to the progeny. Mutations also occur during embryogenesis

(*somatic mutations*). Mutations that occur during development lead to *mosaicism*, a situation in which tissues are composed of cells with different genetic constitutions, as illustrated by Turner or McCune–Albright syndromes. If the germline is mosaic, a mutation can be transmitted to some progeny but not others, which some-times leads to confusion in assessing the pattern of inheritance. Some somatic mutations are associated with neoplasia because they confer a growth advantage to cells by activating (proto)oncogenes or inactivating tumor suppressor genes.

Polymorphisms are sequence variations that have a frequency of at least 1% and do not usually result in an overt phenotype. Often they consist of single base pair substitutions that do not alter the protein coding sequence but some alter mRNA stability, translation or the amino acid sequence. Silent base substitutions and single-nucleotide polymorphisms (SNPs) are encountered frequently during genetic testing and must be distinguished from mutations that alter protein expression or function. Some SNPs or combinations of SNPs may play a pathogenic role in complex disorders by conferring susceptibility for the development of the disease.

Functional consequences of mutations

Mutations can broadly be classified as gain- and loss-of-function mutations. The consequences of an altered protein sequence often need experimental evaluation *in vitro* to determine that the mutation alters protein function.

Gain-of-function mutations are typically dominant and result in phenotypic alterations when a single allele is affected. *Loss-of-function* (inactivating) mutations are usually recessive, and an affected individual is homozygous or compound heterozygous (i.e. carrying two different mutant alleles) for the disease-causing mutations. Mutation in a single allele can result in *haploinsufficiency*, a situation in which one normal allele is not sufficient to maintain a normal phenotype. Haploinsufficiency is a commonly observed mechanism in diseases associated with mutations in transcription factors. For example, monoallelic mutations in the transcription factor TTF1 are associated with transient congenital hypothyroidism, respiratory distress and ataxia.

The clinical features among patients with an identical mutation in a transcription factor often vary

significantly. One mechanism underlying this variability consists of the influence of modifying genes. Haploinsufficiency can affect the expression of rate-limiting enzymes. For example, in MODY 2 (maturity onset diabetes of the young 2), heterozygous glucokinase mutations result in haploinsufficiency with a higher threshold for glucose-dependent insulin release and mild hyperglycemia.

Mutation of a single allele can result in loss-of-function due to a dominant-negative effect. In this case, the mutated allele interferes with the function of the normal gene product by several different mechanisms. The mutant protein may interfere with the function of a multimeric protein complex, as illustrated by Liddle syndrome, which is caused by mutations in the β - or γ -subunit of the renal sodium channel. In thyroid hormone resistance, mutations in the thyroid hormone receptor β lead to impaired T_3 binding; the receptors cannot release co-repressors and they silence transcription of target genes. The mutant protein can be cytotoxic, as in autosomal-dominant neurohypophyseal diabetes insipidus, in which abnormal folding leads to retention in the endoplasmic reticulum and degeneration of the arginine vasopressin (AVP)-secreting neurons.

An increase in dosage of a gene product may also result in disease. For example, duplication of the *DAX1* gene results in dosage-sensitive sex reversal.

Genotype and phenotype

An observed trait is referred to as a *phenotype*. The genetic information defining the phenotype is called the *genotype*. Alternative forms of a gene or a genetic marker are referred to as alleles, which may be polymorphic variants of nucleic acids that have no apparent effect on gene expression or function. In other instances, these variants may have subtle effects on gene expression, thereby conferring adaptive advantages or increased susceptibility. Commonly occurring allelic variants may reflect mutations in a gene that clearly alter its function, as illustrated, for example, by the DF508 deletion in the cystic fibrosis conductance regulator.

Because each individual has two copies of each chromosome, an individual can have only two alleles at a given locus. However, there can be many different alleles in the population. The normal or common allele is usually referred to as *wild type*. When alleles

at a given locus are identical, the individual is *homozygous*. Inheriting such identical copies of a mutant allele occurs in many autosomal-recessive disorders, particularly in circumstances of consanguinity. If the alleles are different, the individual is *heterozygous* at this locus. If two different mutant alleles are inherited at a given locus, the individual is referred to as a *compound heterozygote*. *Hemizygous* is used to describe males with a mutation in an X-chromosomal gene or a female with a loss of one X-chromosomal locus.

A *haplotype* refers to a group of alleles that are closely linked together at a genetic locus. Haplotypes are useful for tracking the transmission of genomic segments within families and for detecting evidence of genetic recombination, if the cross-over event occurs between the alleles.

Allelic and phenotypic heterogeneity

Allelic heterogeneity refers to the fact that different mutations in the same genetic locus can cause an identical or similar phenotype. *Phenotypic heterogeneity* occurs when more than one phenotype is caused by allelic mutations. For example, different mutations in the androgen receptor can result in a wide phenotypic spectrum. In some cases, the receptor is deleted or mutated in a manner that inactivates it completely which leads to complete androgen insensitivity syndrome in a karyotypic male. By contrast, the phenotype may be milder if the androgen receptor is only partially inactivated. In these patients, the phenotype may include infertility, gynecomastia or epispadias. Allelic heterogeneity is explained by the fact that many different mutations are capable of altering protein structure and function. Allelic heterogeneity creates a significant problem for genetic testing because the entire genetic locus must be examined for mutations, as these can differ in each patient.

Locus or non-allelic heterogeneity and phenocopies

Non-allelic or locus heterogeneity refers to the situation in which a similar disease phenotype results from mutations at different genetic loci. This occurs when more than one gene product produces different subunits of an interacting complex or when different genes are involved in the same genetic cascade or physiological pathway. For example, congenital hypothyroidism associated with dysmorphogenesis can arise from

mutations in several genes located on different chromosomes. The effects of inactivating mutations in these genes are similar because the protein products are all required for normal hormone synthesis. Similarly, the genetic forms of diabetes insipidus can be caused by mutations in several genes. Mutations in the AVP-NP2 gene cause autosomal-dominant or -recessive forms of neurohypophyseal diabetes insipidus. The nephrogenic forms can be caused by mutations in the X-chromosomal AVPR2 receptor gene, whereas mutations in the aquaporin-2 (AQP-2) gene cause either autosomal-recessive or -dominant nephrogenic diabetes insipidus.

Recognition of non-allelic heterogeneity is important because the ability to identify disease loci in linkage studies is reduced by including patients with similar phenotypes but different genetic disorders. Genetic testing is more complex because several different genes need to be considered along with the possibility of different mutations in each of the candidate genes.

Phenocopies designate a phenotype that is identical or similar but results from non-genetic or other genetic causes. For example, obesity may be due to several Mendelian defects or have a primarily behavioral origin. As in non-allelic heterogeneity, the presence of phenocopies has the potential to confound linkage studies and genetic testing. Patient history, subtle differences in clinical presentation, and rigorous testing are key in assigning the correct phenotype.

Variable expressivity and incomplete penetrance

Penetrance and *expressivity* are two different yet related concepts that are often confused. Penetrance is a qualitative notion designating whether a phenotype is expressed for a particular genotype. Expressivity is a quantitative concept describing the degree to which a phenotype is expressed. It is used to describe the phenotypic spectrum in individuals with a particular disorder. Thus, expressivity is dependent on penetrance.

Penetrance is *complete* if all carriers of a mutant express the phenotype; it is *incomplete* if some individuals do not have any features of the phenotype. Dominant conditions with incomplete penetrance are characterized by skipping generations with unaffected carriers transmitting the mutant gene. For example, hypertrophic obstructive cardiomyopathy (HOCM) caused by mutations in the *myosin-binding protein C* gene is a dominant disorder with clinical features

in only a subset of patients who carry the mutation. Incomplete penetrance in some individuals can confound pedigree analysis. In many conditions with post-natal onset, the proportion of gene carriers affected varies with age so it is important to specify age when describing penetrance. Variable expressivity is used to describe the phenotypic spectrum in individuals with a particular disorder.

Some of the mechanisms underlying expressivity and penetrance include modifier genes, gender and environmental factors. Thus, variable expressivity and penetrance illustrate that genetic and/or environmental factors do not influence only complex disorders, but also 'simple' Mendelian traits. This has to be considered in genetic counseling, because one cannot always predict the course of disease, even when the mutation is known.

Sex-influenced phenotypes

Certain mutations affect males and females quite differently. In some instances, this is because the gene resides on the X or Y sex chromosomes. As a result, the phenotype of mutated X-linked genes will usually be expressed fully in males but variably in heterozygous females, depending on the degree of X inactivation and the function of the gene. Because only males have a Y chromosome, mutations in genes such as *SRY* (which causes male-to-female sex reversal) or *DAZ* (*deleted in azoospermia*, which causes abnormalities of spermatogenesis) are unique to males.

Other diseases are expressed in a sex-limited manner because of the differential function of the gene product in males and females. Activating mutations in the LHR cause dominant male-limited precocious puberty in boys. The phenotype is unique to males because activation of the receptor induces testosterone production in the testis, whereas it is functionally silent in the ovary. Homozygous inactivating mutations of the FSH receptor cause primary ovarian failure in females because the follicles do not develop in the absence of FSH action but, because testosterone production is preserved, sexual maturation occurs in affected males in whom spermatogenesis is impaired.

Chromosomal disorders

Chromosomal (cytogenetic) disorders are caused by numerical or structural aberrations in chromosomes.

Molecular cytogenetics has led to the identification of more subtle chromosome abnormalities referred to as *microdeletion* and *imprinting* syndromes.

Errors in meiosis and early cleavage divisions occur frequently. Some 10–25% of all conceptions harbor chromosomal abnormalities, which often lead to spontaneous abortion. Numerical abnormalities are much more common than structural defects, especially trisomy, which is found in about 25% of spontaneous abortions and 0.3% of newborns.

Numerical abnormalities in sex chromosomes are relatively common. Males with a 47XXY karyotype have Klinefelter syndrome and females with trisomy 47XXX may be subfertile. Autosomal monosomies are usually incompatible with life; 45XO is present in 1–2% of all conceptuses but leads to spontaneous abortion in 99% of all cases. Mosaicism (e.g. 45XO/45XX, 45XO/45XXX), partial deletions, isochromosomes and ring chromosomes all cause Turner syndrome. Sex chromosome monosomy usually results from loss of the paternal sex chromosome. The 47XXY can result from paternal or maternal non-disjunction, while the autosomal trisomies are most commonly caused by maternal non-disjunction during meiosis, a defect that increases with maternal age. Trisomies are typically associated with alterations in genetic recombination.

Structural rearrangements involve breakage and reunion of chromosomes. Rearrangements between different chromosomes (*translocations*) can be *reciprocal* or *Robertsonian*. Reciprocal translocations involve exchanges between any of the chromosomes; Robertsonian rearrangements designate the fusion of the long arms of two acrocentric chromosomes. Other structural defects include deletions, duplications, inversions and the formation of rings and isochromosomes. Deletions affecting several tightly clustered genes result in *contiguous gene syndromes*, disorders that mimic a combination of single gene defects. They have been useful for identifying the location of new disease-causing genes. Structural chromosome defects can be present in a 'balanced' form without an abnormal phenotype. They can, however, be transmitted in an 'unbalanced' form to offspring and thus cause a hereditary form of chromosome abnormality.

Paternal deletions of chromosome 15q11-13 cause Prader–Willi syndrome (PWS), while maternal deletions are associated with the Angelman syndrome.

The difference in phenotype results from the fact that this chromosomal region is imprinted, that is, differentially expressed on the maternal and paternal chromosomes.

Acquired somatic abnormalities in chromosome structure are often associated with malignancies and are important for diagnosis, classification and prognosis. Deletions can lead to loss of tumor suppressor genes or DNA repair genes. Duplications,

amplifications and rearrangements, in which a gene is put under the control of another promoter, can result in gain-of-function of genes controlling cell proliferation. For example, rearrangement of the 5' regulatory region of the *PTH* gene located on chromosome 11q15 with the *cyclin D1* gene from 11q13 creates the *PRAD1* oncogene, resulting in overexpression of *cyclin D1* and the development of parathyroid adenomas.