8 – FORMATION OF THE FUNCTIONAL STRUCTURE OF ENZYMES: CO- AND POST-TRANSLATIONAL EVENTS

The acquisition of a three-dimensional structure is a defining event for the appearance of a functional active site in a protein. For this reason we feel it is important to describe succinctly here the mechanisms that ultimately generate a defined spatial architecture and thus enable the expression of activity. In view of the spectacular progress that has been made in understanding the mechanisms of genome expression, we might be tempted to believe that the events leading to the formation of an active protein are entirely elucidated. **The information contained in DNA, the genetic message, is one-dimensional information** which, during the diverse processes of biosynthesis – including transcription and translation – leads to the formation of a polypeptide chain having a well-defined sequence.

A number of key events also occurs in order that **the biological function, which requires a three-dimensional structure,** can be expressed. These events take place either throughout biosynthesis (co-translational events), or after the termination of the polypeptide chain (post-translational events). There are two groups: first, the covalent processes such as limited proteolysis and chemical modifications; second, the non-covalent processes, for instance, folding of the polypeptide chain and in certain cases the self-assembly of subunits to form a quaternary structure. However, all these events are still incompletely understood and poorly controlled, yet they are the subject of in-depth studies. Their importance is paramount since they generate the functional properties of proteins required for recognition, transport, catalysis and regulation. These functions generally only appear in very specific places within cells or organisms, permitting very precise and fine regulation to ensure the harmonious functioning of living beings.

Covalent and non-covalent processes will be considered in succession, although during protein formation they may arise chronologically in a different order.

8.1. COVALENT PROCESSES

8.1.1. LIMITED PROTEOLYSIS

During or after termination of the biosynthesis of polypeptide chains, or during their transport to particular regions of the cell or even the organism, some proteins

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_9, © Springer-Verlag Berlin Heidelberg 2010

undergo limited proteolysis. Essentially two types of proteolytic processing exist, which are important events for the acquisition of the functional structure of proteins; they are:

- cleavage of the signal peptide, and
- activation of precursors, zymogens or prohormones.

In 1975, BLOBEL and DOBBERSTEIN discovered that secreted proteins, particularly those that are synthesised on ribosomes linked to the rough endoplasmic reticulum (RER), when synthesised by cell-free systems not containing the RER membrane, possess a supplementary sequence of about twenty amino acids at their N-terminus, called the **signal peptide.** This sequence is largely hydrophobic and always begins at the N-terminus with a methionine residue. The addition of RER membranes to a cell-free preparation causes this sequence to disappear. The signal peptide, due to its hydrophobic character, enables the assisted deposition of the polypeptide chain into the intercisternal space. After traversing the membrane, it is digested by a membrane signal peptidase (Fig. 8.1a).

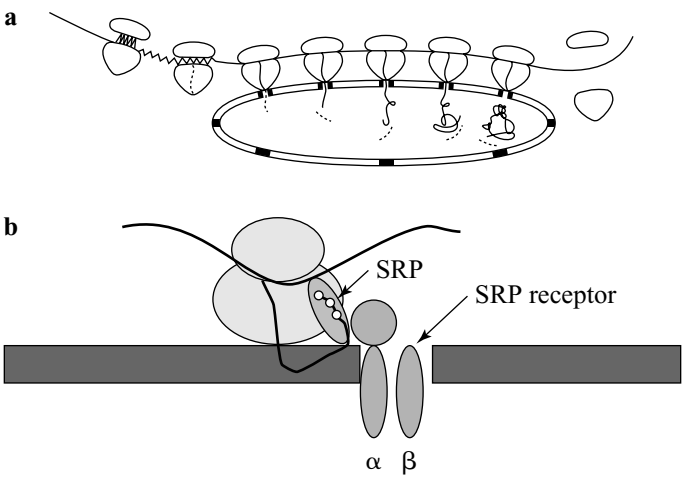


Fig. 8.1 Mechanism of enzymatic cleavage of the signal peptide during the transfer of the nascent polypeptide chain across the RER membrane

*(**a**) general scheme for the assisted transfer of the nascent protein across the membrane (**b**) interaction between the ribosome and signal recognition protein (SRP) associated to the anchoring protein.* (8.1a – © BLOBEL G. & DOBBERSTEIN B., 1975. Originally published in the *J. Cell Biol.*, **67**, 835–851)

- It is assumed that the signal peptide carries information for the association of the ribosome with the membrane via its interaction with particular proteins. These proteins were called ribophorins by KREIBICH et al. (1978). The mechanism was detailed later by MEYER et al. (1982) who identified the protein that recognises the signal sequence, signal recognition protein (SRP). BLOBEL and WALTER showed that SRP is composed of an assembly of 300 nucleotides and 6 different polypeptide chains forming a complex of 325 kDa. This complex is associated to another

integral membrane protein formed from one α subunit of 68 kDa and one β subunit of 30 kDa (Fig. 8.1b). The sequence of events optimises the co-translational process for secreted proteins. This process happens even when synthesis begins on the free ribosomes. Translation is blocked by the binding of SRP after a polypeptide chain comprising about 70–80 amino acids has been synthesised, when the signal sequence emerges from the large ribosomal subunit.

The pause in translation persists until contact is established with the anchoring protein. Translation then continues and translocation ensues. This safety-lock mechanism ensures that the protein is not terminated in the cytoplasm. The GTP-GDP cycle regulates the dissociation of the signal sequence from its receptor (RAPPOPORT, 1992). The anchoring protein is the first site of interaction between the nascent secreted protein and the endoplasmic reticulum membrane. However, the signal peptide is not always cleaved from membrane proteins, which are therefore not delivered to the intercisternal space, but are instead inserted into the membrane. The existence of signal peptides is not specific to proteins from eukaryotic cells; they have also been observed in proteins from prokaryotic organisms. Signal peptides also exist in proteins synthesised by free polysomes, such as mitochondrial proteins. ◢

The sequence of the signal peptide has been determined for a significant number of proteins. It is compatible with a helical structure, which would seem to be the preferential structure. The size of the helical segment formed would correspond approximately to the thickness of the RER membrane.

◤ To what extent is the proteolytic removal of the signal peptide necessary for the correct folding of the polypeptide chain and consequently for the emergence of a protein's function? There is no general answer to this question; indeed, pre-ribonuclease has an enzymatic activity whereas pre-amylase is inactive. Many proteins are synthesised as preproteins; the cleavage of the signal peptide leads, in some cases, to the gain of functional properties, but this situation really depends on the protein. ◢

As well as the removal of signal peptides, other controlled proteolytic processing is necessary in order for biological activity to manifest itself in certain proteins. Proteolysis of the precursor proteins triggers their activation and thus provides a means to regulate the formation of these active proteins in specific compartments within an organism. Many biological systems are regulated by proteases, for example, the formation of hormones and active peptides, the activation of zymogens to produce active enzymes, the blood coagulation cascade and the activation of complement, protein assembly in the heads of phages and the transformation of fibrinogen into fibrin. The role of proteases is also essential in fertilisation processes, development, inflammatory reactions or malignant cell migration and has been the subject of numerous studies.

◤ Highly specific proteolytic cleavage is used to generate certain peptide hormones. For instance, proinsulin is converted into insulin by the separation of a peptide of 33 amino acids, which gives rise to two protein chains, A and B, and is accompanied by a large change in conformation (Fig. 8.2 below).

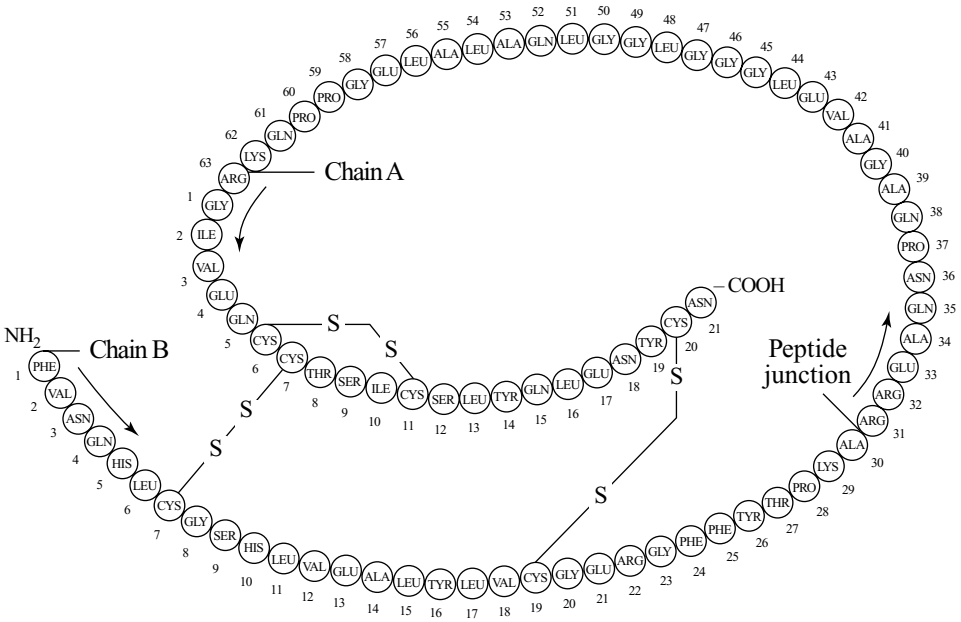


Fig. 8.2 Conversion of proinsulin, formed from a single chain, into insulin
The start sites of chains A and B are indicated by arrows

Multifunctional precursors also exist, such as pro-ACTH, which undergoes several types of proteolytic cleavage in a variety of places. It is not only the precursor of ACTH (adrenocorticotropic hormone), but also of α- and β-MSH (melanocyte-stimulating hormone), α- and β-LPH (lipotropic hormone), endorphins and enkephalins (Fig. 8.3).

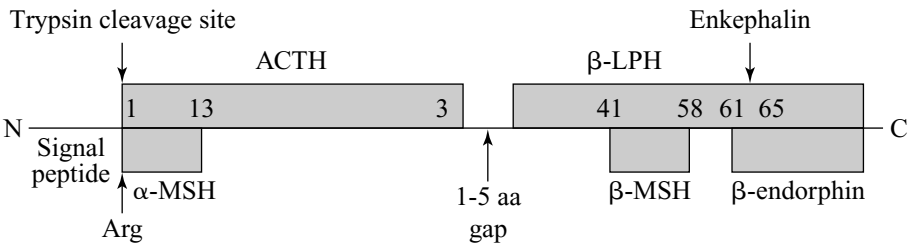


Fig. 8.3 Pro-ACTH: a multifunctional precursor of adrenocorticotropic hormone and other peptide hormones

The activation of zymogens to enzymes by limited proteolytic cleavage generally only leads to small variations in the polypeptide chain, but is sufficient for their ultimate conformational maturation. Thus with the serine proteases limited proteolysis of the zymogen induces, in appropriate conditions, the formation of a salt bridge between the amino group liberated by proteolytic digestion and a carboxylate neighbouring the reactive serine. This salt bridge stabilises the enzyme's functional structure.

Figure 8.4 illustrates the different activation pathways of chymotrypsinogen.

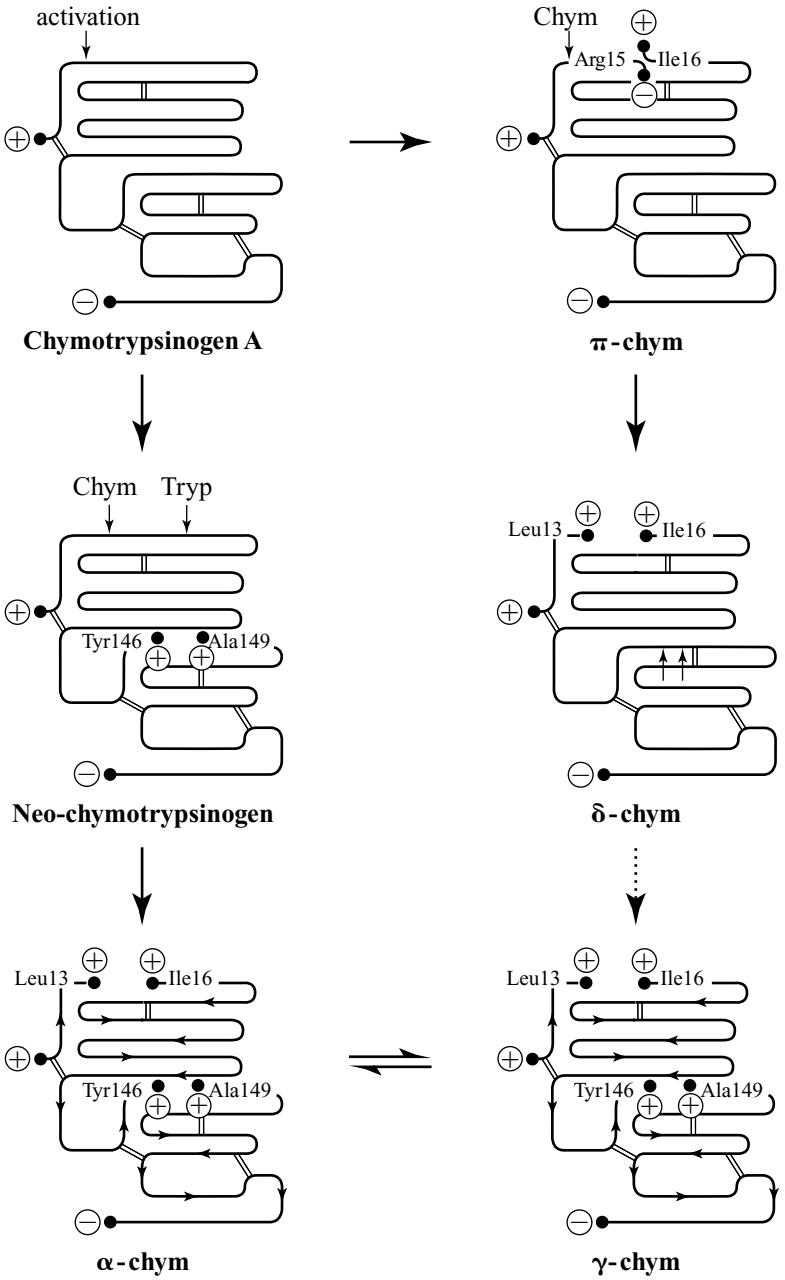


Fig. 8.4 The activation pathways of chymotrypsinogen

(Reprinted from *The Enzymes*, 3rd ed., Vol. III, BLOW D., The structure of chymotrypsin, 187.
© (1971) Academic Press, with permission from Elsevier)

Figure 8.5 shows, based on crystallographic data, the salt bridge formed between Ile 16 and Asp 194 adjacent to serine 195 in the active site of chymotrypsin. It is interesting to note that the active site is partially functional in the zymogen. Indeed, chymotrypsinogen has a very weak but significant catalytic activity; it is 10^6 to 10^7 times weaker than that of the enzyme (GERTLER et al., 1974). This system is studied in detail in Chap. 12. After limited proteolysis and creation of the salt bridge, the conformational rearrangements are only minor, yet decisive for augmenting the activity. Experimental evidence suggests that the zymogen is in a folded state that is close to the native state and yet different, probably a state that precedes the native one on the enzyme's folding pathway. The formation of the salt bridge must then promote the conformational coupling of the two structural domains that make up the enzyme: this coupling is necessary for the emergence of its activity (GHÉLIS & YON, 1979).

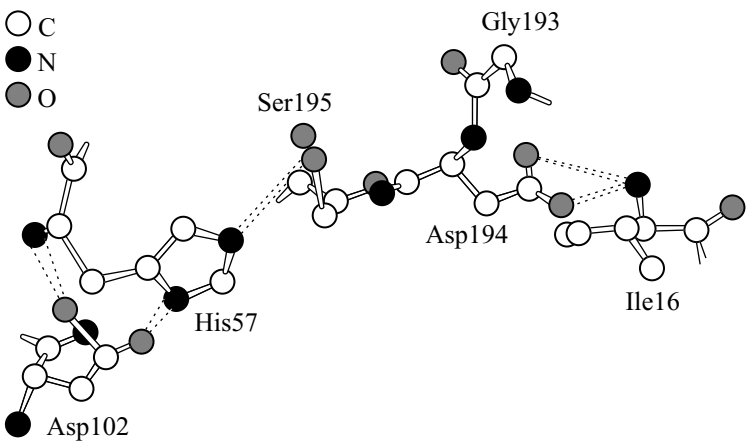


Fig. 8.5 The salt bridge between Asp 194 and Ile 16, which promotes chymotrypsin's activity (taken from crystallographic data)

(Reprinted from *The Enzymes*, 3rd ed., Vol. III, BLOW D., The structure of chymotrypsin, 187. © (1971) Academic Press, with permission from Elsevier)

The zymogen-activation cascades responsible for blood coagulation also come about from highly specific limited proteolyses. In this system, illustrated in Figure 8.6 opposite, each enzyme formed at a given step in the sequence activates a zymogen in the subsequent step, which in turn continues the cascade of events that culminates in the conversion of fibrinogen into fibrin.

Two different pathways exist to initiate the blood coagulation cascade involving either a contact factor, or a tissue factor. These processes engender a considerable amplification, which enables organisms to respond quickly to injury. Furthermore, they are modulated by non-covalent inhibitors at each step of the cascade. The cascade system represents an extremely efficient mechanism of covalent regulation, which we return to in Part V.

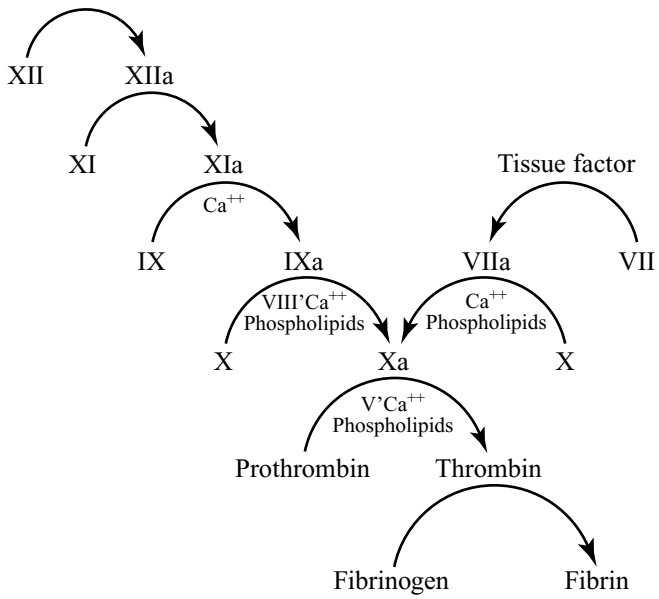


Fig. 8.6 Blood coagulation cascade

Several types of protein exist in circulation systems that are capable of generating, by limited proteolysis and in specific places, particular well-regulated enzymes in high concentrations. The fibronectins are an example. Apparently inert matrix proteins can also be activated in this way.

8.1.2. CHEMICAL MODIFICATIONS

Aside from proteolyses, diverse chemical events contribute to the genesis or the stabilisation of functional proteins. Their co- and post-translational chemical modifications are very varied. It is beyond the scope to detail them all here, but we mention briefly those which seem the most crucial from a structural and functional viewpoint. Many reviews deal with this problem, though we refer in particular to those that feature in *The enzymology of post-translational modifications of proteins*, edited by R.B. FREEMAN and H.C. HAWKINS (1980).

Let us firstly note that the formation of **disulphide bonds** results from the oxidation and condensation of sulphydryl groups. The formation of disulphide bonds is a major co- or post-translational modification in protein biosynthesis. Many free thiols are found in intracellular proteins that do not generally have disulphide bonds, yet on the other hand are frequently oligomeric. Disulphide bonds predominate in extracellular proteins. In fact the intracellular environment is more reducing than the circulating fluids in organisms. Furthermore, disulphide bonds confer greater stability on proteins helping them to maintain their functional structure despite environmental fluctuations. If the chemistry of sulphydryl groups and disulphide bonds is well known, disulphide-bond formation in vivo has not been completely elucidated.

◤ Thiol oxidation to give disulphide bonds is a spontaneous process that happens when oxidoreduction conditions are favourable. The existence of a microsomal enzyme, disulphide isomerase, capable of catalysing the conversion of thiols into disulphide bonds suggests that an enzymatic process occurs in vivo. For a few proteins, there is evidence that seems to indicate that disulphide bonds are formed on the nascent polypeptide chain after the completion of each structural domain (FREEDMAN & HILLSON, 1980). ◢

Glycosylation is a very widespread chemical modification, found in particular in membrane proteins, but also in many soluble proteins (PHELS, 1980). Only some amino acids are capable of binding to a sugar group, these include: serine, threonine, asparagine and hydroxylysine. The glycosidic motifs that link to these different residues are not the same, however. Glycosylation frequently arises on Asn residues in loops (β-turns) and therefore at the protein's surface in contact with solvent. This event confers on the protein precise structural properties in a given environment. The presence of sugar groups has the general effect of increasing the protein's stability and making it more resistant to proteases. Indeed, a glycosidic chain on the protein's surface creates an important screen between the protein and the solvent. In addition, the role of this carbohydrate extension in the specificity of cell-surface recognition is fundamental for cell adhesion, the immune response, hormone binding and differentiation.

Among other modifications, the **hydroxylation of prolines and lysines** leads to the formation, respectively, of 4-hydroxyproline, 3-hydroxyproline and hydroxylysine (Fig. 8.7).

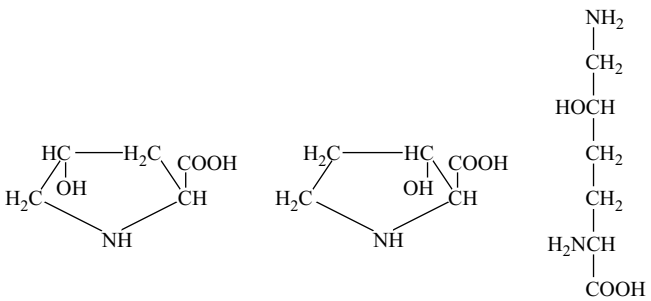


Fig. 8.7 Structures of 4-hydroxyproline, 3-hydroxyproline and 5-hydroxylysine

◤ These modifications are present in collagen and a few proteins having similar sequences such as acetyl cholinesterase, protein Clq from complement, elastin and some plasma proteins (KIVIRIKKO & MILLYLÄ, 1980). They are catalysed by three different enzymes: prolyl-4-hydroxylase, prolyl-3-hydroxylase and lysyl hydroxylase, enzymes which are found in higher organisms, and also in lower organisms and plants. These enzymes have been identified and characterised. ◢

Phosphorylations are important modifications for the regulation and control of a great number of cellular processes (HUNTER, 1987). They require the action of specific kinases, which transfer a phosphate group to the protein substrate. Nearly a hundred protein kinases have been identified to date. The reversibility of this covalent modification by the activity of phosphatases is a means to control certain cellular processes. Phosphorylation generally affects serine, threonine or tyrosine residues and can considerably modify a protein's properties. For enzymes, this may mean a change in activity either causing inhibition or, conversely, activation. Many enzymes are activated by phosphorylation, for instance, glycogen phosphorylase, phosphofructokinase, triacyl glycerol lipase, tyrosine hydroxylase and DNA-dependent RNA polymerase to name but a few. Others are inhibited by phosphorylation; a few examples include glycogen synthase, pyruvate dehydrogenase and glycerophosphate acyltransferase. A certain number of proteins having no enzymatic activity are phosphorylated in vivo, among which, troponin I, ribosomal proteins and perilipin, a protein from adipose tissue. The activity of certain kinases is dependent on cyclic AMP (cAMP) or cyclic GMP (cGMP). The cAMP-dependent kinases are in general composed of two types of subunit: catalytic (C) subunits, which catalyse the phosphorylation reaction and regulatory (R) subunits, which bind to cAMP. They are frequently tetramers having the composition R_2C_2. These enzymes are very specific for their protein substrates, more specifically the phosphorylation sites. Specific sequences are found proximal to the phosphorylated serine, namely, Lys-Arg-X-X-Ser(P) or Arg-Arg-X-Ser(P), where X denotes any amino acid. The phosphorylations catalysed by this type of enzyme are integral to metabolic control, muscle contraction, as well as transcription and translation. These aspects are discussed in Part V, since phosphorylation participates in the covalent regulation of the activity of several enzymes.

The **carboxylation** of glutamyl residues to form γ-carboxyglutamyl residues (γ-carboxyglutamic acid or Gla) plays an important role in the activity of proteins required for blood coagulation (SUTTIE, 1980). This modification has been observed in several plasma proteins, in particular prothrombin where the exact Gla residues have been identified; they are localised to the molecule's N-terminus. This modification is also present in the clotting factors IX and X and probably in factor VII, which possesses a similar N-terminal sequence to that of prothrombin; the Gla residues are found in this region. Other plasma proteins such as bovine protein C and human protein S contain γ-carboxyglutamic acid. These also have close homology to the sequences of the above-mentioned proteins. Let us also briefly mention protein Z from bovine plasma and osteocalcin from chicken bones. The carboxylase that introduces the additional carboxylate group on these proteins requires vitamin K.

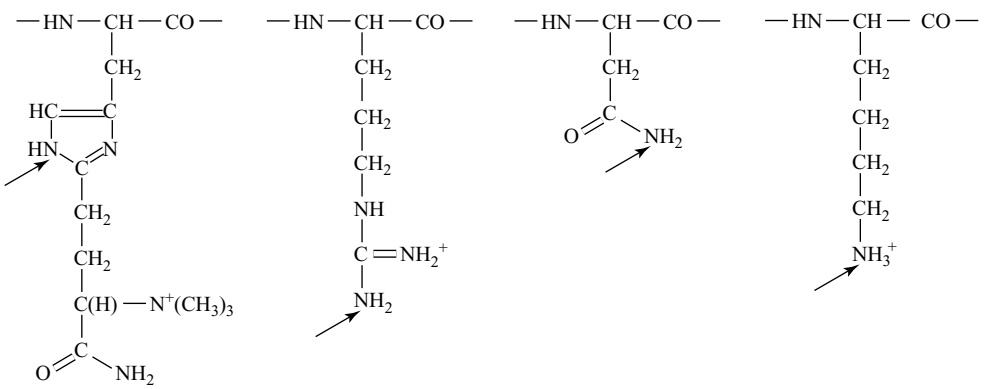
Among the other post-translational chemical modifications, **methylations** take part in various processes. O-methyl aspartate, O-methyl glutamate and N-methyl glutamine are present in a variety of proteins, e.g. N-methyl histidine in actin, α-N-trimethyl alanine in ribosomal protein from *E. coli* and myosin light chain, and α-N-dimethyl proline in bacterial cytochrome c. The ε-amino group of lysine

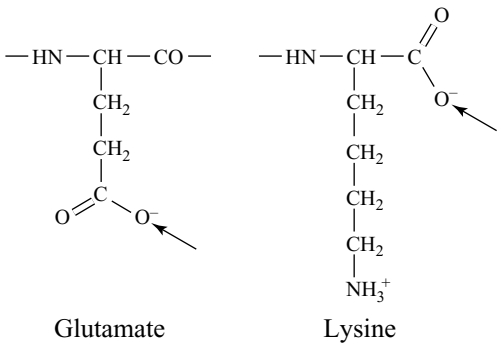
can be methylated giving rise to three derivatives: ε-N-methyl lysine, ε-N-dimethyl lysine, ε-N-trimethyl lysine. These derivatives have been found in a great variety of eukaryotic and prokaryotic proteins. The enzymes catalysing this reaction are protein-lysine methylases which have been isolated from different organisms and characterised.

Acetylation of the N-terminus (N-α-acetylation) of polypeptide chains is amino acid-dependent, occurring preferentially with Ala or Ser, and sometimes Met, Gly and Asp. N-α-acetylation has also been observed for Asn, Ile, Thr and Val, but none of the 11 remaining amino acids undergo this modification. It would seem that N-α-acetylation follows a phylogenetic scheme. For instance, cytochrome c is always N-α-acetylated in vertebrates and higher plants, but never in yeast or bacteria. The same is true for other enzymes such as enolase and several pancreatic enzymes. However, α-amylase is found to be acetylated in all species studied. N-α-acetylation of proteins is catalysed by the enzyme N-α-acetylase, which should be distinguished from N-ε-acetylase, the enzyme that acetylates the ε-amino group of lysines. Acetylation seems to take place on the nascent polypeptide chain after cleavage of the signal peptide.

The **iodination of tyrosines** in thyroglobulin yielding 3-iodotyrosine and 3,5-diiodotyrosine is an important chemical modification in the biosynthesis of thyroxine.

ADP ribosylation refers to a whole group of post-translational modifications on proteins. The ADP-ribosylation reactions are classed into two main groups: mono-ADP ribosylations and poly-ADP ribosylations. These two groups are distinct not only because of the ADP-ribose chain, but also due to the chemical nature of the bond between the ADP ribosyl and the protein, which is an N-glycoside bond in the first group and an O-glycoside bond in the second. The amino acid acceptors in the first group are: lysine, arginine, asparagine and diphtamide (a histidine derivative); and in the second group: glutamate and lysine (COO^- at the C-terminus) (Fig. 8.8).





Glutamate Lysine

Fig. 8.8 Amino acid acceptors in ADP-ribosylation reactions
The arrows *indicate the atoms involved in covalent-bond formation*

Many diverse proteins undergo these post-translational modifications, which are catalysed by specific enzymes. For instance, elongation factor EF2 is ADP-ribosylated, with diphatamide acting as the acceptor. Other proteins such as transducin, protein Ns from the adenylate cyclase complex, microtubule proteins and some *E. coli* and eukaryotic proteins are adenylated on their arginine or asparagine residues. The role of ADP-ribosylation in enzyme regulation is elaborated upon in Part V.

◤ Among other post-translational modifications, we also point out modification by the addition of fatty acids, namely, **N-myristylation and S-palmitoylation.** The myristate is linked by an amide bond on the N-terminal side of a glycine. The palmitate typically forms a thioester bond with a cysteine. Palmitoylated proteins are synthesised on free ribosomes and transported to the plasma membrane. **Isoprenylation,** well-known for G proteins, also occurs on Ras proteins. This modification is produced at a protein's C-terminus having the sequence CAAX by modifying the cysteine residue. The **glycolipid bond** is found in proteins anchored on the external side of cell membranes. ◢

8.2. NON-COVALENT PROCESSES

Although covalent modifications are not common to all proteins, the first non-covalent event, i.e. folding of the polypeptide chain, does apply to all proteins. For oligomeric proteins, the self-assembly of subunits constitutes an additional process.

8.2.1. PROTEIN FOLDING

The polypeptide chain has to fold in order to adopt a compact, globular three-dimensional structure. Protein folding is a vital process as it transforms one-dimensional information into three-dimensional information and only then can biological activity emerge, as has been previously emphasised. In this sense, protein folding can be

considered as the first act of morphogenesis. Studies into protein folding are however not detailed in this book; we refer the interested reader instead to other general books and reviews on the topic including: *Protein folding* by GHÉLIS and YON (1982), and the reviews by DOBSON and KARPLUS (1999), DILL and CHAN (1997), WOLYNES et al., (1995) and YON (2001). We shall nonetheless briefly outline the principal aspects of this research field.

It has been assumed for a long time that all the information necessary for a protein to acquire its three-dimensional structure is contained in its sequence, and thus in the genome. In 1958, F. CRICK wrote: "[Protein] folding is simply a function of the order of the amino acids." In 1973, following his remarkable work on the folding of ribonuclease, ANFINSEN clarified: "The three-dimensional structure of [...] the native [protein] conformation is determined by the totality of the inter-atomic interactions and hence by the amino acid sequence, in a given environment." The role of the environment, especially the solvent, is fundamental for the formation and stabilisation of the functional structure of proteins. A corollary of this well-known concept, designated the ANFINSEN postulate, was the assumption that the native structure of a protein represents the most energetically stable structure, corresponding to a minimum in GIBBS energy in defined conditions. The thermodynamic control of protein folding was called into question in 1968 by LEVINTHAL, and then by WETLAUFER in 1973 regarding temporal considerations. These authors pointed out that a random search for the native conformation amongst all possible conformations of a polypeptide chain would require an astronomical amount of time (10^{26} years for a chain only 150 amino acids in length!), which is of course incompatible with the folding times observed both in vitro and in vivo. LEVINTHAL proposed a kinetic rather than a thermodynamic control of the folding process. This familiar problem, termed the LEVINTHAL paradox, had dominated discussions for nearly thirty years. It is clear that a random search for the most stable structure of a protein along the entire length of its polypeptide chain is not a plausible hypothesis. It became obvious that evolution had found an efficient solution to solve this combinatorial problem.

Diverse models have been proposed to solve the LEVINTHAL paradox, which were experimented with in order to find and characterise the folding intermediates. A sequential and hierarchical model, in which elementary structures would form and interact with each other to generate the native protein, was supported by numerous authors until relatively recently (KIM & BALDWIN, 1990). The diffusion-collision model developed by KARPLUS and WEAVER (1976, 1994) implicated the formation of nuclei of elementary structures in different parts of the polypeptide chain. These nuclei would diffuse and self-associate. Several centres would be formed at the same time and self-assemble if they had the correct structure, thus enabling a polypeptide chain of 100–200 amino acids to acquire its native structure in less than a second. In 1985, HARRISON and DURBIN introduced the "jigsaw puzzle" model. Using the metaphor of puzzle building, they assumed that folding followed multiple

routes before reaching a unique solution. This model, which implied the existence of heterogenous species in the course of folding, became quite controversial at the time.

With the convergence of theoretical and experimental studies in 1995, a new vision of protein folding, the *new view*, emerged. Presented by WOLYNES et al. in terms of an energy landscape using the metaphor of a folding funnel, this model described the kinetic and thermodynamic behaviour of an ensemble of unfolded molecules. The number of conformations to be explored gradually diminishing as the process advanced, until an energy minimum representing the native structure was reached. The model implied the existence of several folding pathways and the heterogeneity of its intermediate species, and was in accordance with the "jigsaw puzzle" model. This vision of the folding process has now been confirmed experimentally. The advantage of this model is that it takes into account the possibility of incorrect folding which might result from the existence of local energetic minima in which the molecules could become "trapped", thus delaying their folding or forcing them towards aggregration. Two types of aggregate exist: amorphic aggregates, which appear as inclusion bodies when the gene coding for a protein is overexpressed in a foreign host, and aggregates organised into amyloid structures. Amyloid structures are the basis of serious animal and human pathologies among which bovine spongiform encephalopathy (BSE, or mad cow disease), CREUTZFELD-JACOB disease and ALZHEIMER's disease (see the review by YON, 2004).

Another question is raised: how does folding proceed in a cellular context? The assumption is that the same mechanisms are involved in folding processes *in vivo* and in vitro. The discovery of molecular chaperones (ELLIS, 1987) has led to a reconsideration of this question. The rapid development of research in this field, the resolution of the structures of chaperonins and molecular chaperones, as well as the characterisation of their interactions with partially folded proteins has permitted the elucidation of the roles of these molecules. Molecular chaperones, through their transient association with a nascent or unstable protein when under conditions of stress (heat shock, for example), prevent its potentially incorrect folding and subsequent aggregation. They also interact with proteins that undergo translocation. Their association is predominantly mediated by hydrophobic interactions. Molecular chaperones do not interact with native proteins. They do not have any information capable of directing the protein towards a conformation different from that determined by the sequence. Furthermore, they increase the folding yield but do not alter the folding rate; **molecular chaperones are not folding catalysts.** Lastly, molecular chaperones assist the folding of a limited number of proteins. The incorrect folding of a protein and its subsequent aggregation arises from kinetic competition between the correct folding pathway and an alternative side pathway. When the formation of the correct structure is kinetically favoured, the presence of a chaperone is unnecessary.

Taken together, the most recent data enable us to conclude that the principles governing protein folding, established by studies *in vitro* and *in silico*, also govern the folding of the nascent polypeptide chain in a cellular context.

Currently, protein folding represents a very active research field comprising different aspects of biology, physics, chemistry and computing. The fundamental principles have applications for the use of the information contained in genomic sequences, in the development of new therapeutics, understanding certain pathologies and the design of non-natural proteins harbouring specific functions.

8.2.2. ASSEMBLY OF SUBUNITS

Oligomeric proteins result from the association of identical (α_n homo-oligomers) or non-identical ($\alpha_n\beta_n$ hetero-oligomers) subunits. In the nomenclature proposed by MONOD, WYMAN and CHANGEUX, the term oligomer signified that the number of subunits is limited. It is in contrast to the term polymer for which the association consists of a large, indeed, an unlimited number of subunits. When the subunits making up the oligomer are identical, the authors employed the term protomer, a monomer being a molecule in which a single subunit exists. Thus, a haemoglobin molecule ($\alpha_2\beta_2$) has four subunits and two protomers.

The self-assembly of subunits takes place in the final stage of the folding process and is typically accompanied by specific conformational rearrangements. These may only be of a very subtle nature, but ensure the correct positioning of the functional sites (e.g. active site) in an enzyme. There are few known cases of oligomeric enzymes whose isolated subunits are functional. This is linked to the fact that, in many oligomeric enzymes, the active site is localised to the interface between two subunits and includes amino acid side chains belonging to each of them. Only the aldolases and transaldolases provide examples of oligomeric enzymes whose isolated subunits have a significant enzymatic activity. It is however important to emphasise that, even in these cases, the activity of the isolated protomer is lower than that of the oligomeric structure.

Thus, *the different levels of protein structure are essential for the formation of an enzyme's active site and for modulating its properties*.

BIBLIOGRAPHY

BOOKS

KIVIRIKKO K.I. & MILLYLA R. –1980– in *The Enzymology of post translational modifications of proteins* R.B. FREDMAN & H.C. HAWKINS eds, Academic Press, London p.53.

PHELPS C.F. –1980– in *The Enzymology of post translational modifications of proteins* R.B. FREDMAN & H.C. HAWKINS eds, Academic Press, London, p.112.

GENERAL REVIEWS

ANFINSEN C.B. –1973– Principles that govern protein folding, *Science* **181**, 223–230

CROSS G.A.M. –1990– Glycophorin anchoring of plasma membrane proteins, *Annu. Rev. Cell Biol.* **6**, 1–39.

DILL K.A. & CHAN H.S. –1997– From Levinthal paradox to funnel, *Nat. Struct. Biol.* **4**, 10–19.

DOBSON C.M. & KARPLUS M. –1999– The fundamentals of protein folding: bringing together theory and experiments, *Curr. Opin. Struct. Biol.* **9**, 92–101.

ELLIS R.J. & HARTL F.U. –1999– Principles of protein folding in the cellular environment, in *Curr. Opin. Struct. Biol.* **9**, 102–110.

GIBBS J.M. –1991– Ras C-terminal processing enzymes – New drug targets? *Cell* **65**, 1–4.

HUNTER T. –1987– A thousand and one protein kinases, *Cell* **50**, 823–829.

JAENICKE R. –1987– Folding and association of proteins, *Prog. Biophys. Biol. Mol.* **49**, 117–237.

KIM P.S. & BALDWIN R.L. –1990– Intermediates in protein folding reactions of small proteins, *Annu. Rev. Biochem.* **59**, 631–660.

TANFORD C. –1968– Protein denaturation, *Adv. Prot. Chem.* **23**, 121–182.

UEDA K. & HAYAISHI O. –1985– ADP ribosylation, *Annu. Rev. Biochem.* **54**, 73–100.

YON J.M. & BETTON J.M. –1991– Protein folding in vitro and in the cellular environment, *Biol. Cell* **71**, 17–23.

YON J.M. –2001– Protein folding: a perspective for biology, medicine and biotechnology, *Braz. J. Med. Biol. Res.* **34**, 419–435.

YON J.M. –2002– Protein folding in the post-geneomic era, *J. Cell. Mol. Biol.* **6**, 307–327.

YON J.M. –2004– Protein aggregation, in *Encyclopedia of Molecular Biology and Molecular Medicine*, Vol. I, 23–52, ed. by R.A. Meyers, Willey VCH, New York.

SPECIALISED ARTICLES

ANFINSEN C.B., HABER E., SELA M. & WHITE F.H. –1961– *Proc. Natl Acad. Sci. USA* **47**, 1309.

BLOBEL G. & DOBBERSTEIN B. –1975– *J. Cell Biol.* **67**, 835–862.

BLOW D.W. –1971– in *The Enzymes*, Vol. III, 3rd ed., P.D. BOYER ed., Acad. Press, New York.

CRICK F.H. –1958– *Symp. Soc. Exp. Biol.* **13**, 138.

ELLIS J.R. –1987– *Nature* **328**, 378.

GERTLER A., WALSH & NEURATH H. –1974– *Biochemistry* **13**, 1302.

GHÉLIS C. & YON J.M. –1979– *C. R. Acd. Sci.* **282**, 197.

HILLSON D.A. & FREEDMAN R.B. –1980– *Biochem. J.* **191**, 373.

JAMES G. & OLSEN E.N. –1990– *Biochemistry* **29**, 2623.

KARPLUS M. & WEAVER D.L. –1976– *Nature* **260**, 404.

KARPLUS M. & WEAVER D.L. –1994– *Protein Sci.* **3**, 650.

KREIBICH G., FREINSTEIN C.M., PEREYRA B.N., ULRECH B.N. & SABATINI D.D. –1978– *J. Cell Biol.* **77**, 464 and 488.

LEVINTHAL C. –1968– *J. Chim. Phys.* **65**, 44.

MEYER D.L., KRAUS E. & DOBBERSTEIN B. –1982– *Nature* **297**, 647.

RAPOPORT T.A. –1992– *Science* **258**, 931.

SUTTIE, J.W. –1980– *CRC Crit. Rev. Biochem.* **8**, 191.

WETLAUFER D.B. & RISTOW S. –1973– *Annu. Rev. Biochem.* **42**, 135.

WOLYNES P.G., ONUCHIC J.N. & THIRUMALAI D. –1995– *Science* **267**, 1619.