

7

Biocatalysis, Enzyme Engineering and Biotechnology

G. A. Kotzia, D. Platis, I. A. Axarli, E. G. Chronopoulou, C. Karamitros, and N. E. Labrou

- Preface
- Enzyme Structure and Mechanism
 - Nomenclature and Classification of Enzymes
 - Basic Elements of Enzyme Structure
 - The Primary Structure of Enzyme
 - The Three-Dimensional Structure of Enzymes
 - Theory of Enzyme Catalysis and Mechanism
 - Coenzymes, Prosthetic Groups and Metal Ion Cofactors
 - Kinetics of Enzyme-Catalysed Reactions
 - Thermodynamic Analysis
 - Enzyme Dynamics During Catalysis
- Enzyme Production
 - Enzyme Heterologous Expression
 - The Choice of Expression System
 - Bacterial Cells
 - Mammalian Cells
 - Yeast
 - Filamentous Fungi
 - Insect Cells
 - Dictyostelium discoideum
 - Trypanosomatid Protozoa
 - Transgenic Plants
 - Transgenic Animals
 - Enzyme Purification
 - Ion-Exchange Chromatography
 - Affinity Chromatography
- Enzyme Engineering
 - Tailor-Made Enzymes by Protein Engineering
 - Rational Enzyme Design
 - Directed Enzyme Evolution
- Immobilised Enzymes
 - Methods for Immobilisation
 - Adsorption
 - Covalent Coupling
 - Cross-linking
 - Entrapment and Encapsulation
 - New Approaches for Oriented Enzyme Immobilisation:
 - The Development of Enzyme Arrays

- Enzyme Utilisation in Industry
- Enzymes Involved in Xenobiotic Metabolism and Biochemical Individuality
 - Phase I
 - Phase II
 - Phase III
- Acknowledgements
- References

Abstract: Enzymes are biocatalysts evolved in nature to achieve the speed and coordination of nearly all the chemical reactions that define cellular metabolism necessary to develop and maintain life. The application of biocatalysis is growing rapidly, since enzymes offer potential for many exciting applications in industry. The advent of whole genome sequencing projects enabled new approaches for biocatalyst development, based on specialised methods for enzyme heterologous expression and engineering. The engineering of enzymes with altered activity, specificity and stability, using site-directed mutagenesis and directed evolution techniques are now well established. Over the last decade, enzyme immobilisation has become important in industry. New methods and techniques for enzyme immobilisation allow for the reuse of the catalysts and the development of efficient biotechnological processes. This chapter reviews advances in enzyme technology as well as in the techniques and strategies used for enzyme production, engineering and immobilisation and discuss their advantages and disadvantages.

PREFACE

Enzymes are proteins with powerful catalytic functions. They increase reaction rates sometimes by as much as one million fold, but more typically by about one thousand fold. Catalytic activity can also be shown, to a limited extent, by biological molecules other than the 'classical' enzymes. For example, antibodies raised to stable analogues of the transition states of a number of enzyme-catalysed reactions can act as effective

catalysts for those reactions (Hsieh-Wilson et al. 1996). In addition, RNA molecules can also act as a catalyst for a number of different types of reactions (Lewin 1982). These antibodies and RNA catalysts are known as abzymes and ribozymes, respectively.

Enzymes have a number of distinct advantages over conventional chemical catalysts. Among these are their high productivity, catalytic efficiency, specificity and their ability to discriminate between similar parts of molecules (regiospecificity) or optical isomers (stereospecificity). Enzymes, in general, work under mild conditions of temperature, pressure and pH. This advantage decreases the energy requirements and therefore reduces the capital costs. However, there are some disadvantages in the use of enzymes, such as high cost and low stability. These shortcomings are currently being addressed mainly by employing protein engineering approaches using recombinant DNA technology (Stemmer 1994, Ke and Madison 1997). These approaches aim at improving various properties such as thermostability, specificity and catalytic efficiency. The advent of designer biocatalysts enables production of not only process-compatible enzymes, but also novel enzymes able to catalyse new or unexploited reactions (Schmidt-Dannert et al. 2000, Umeno and Arnold 2004). This is just the start of the enzyme technology era.

ENZYME STRUCTURE AND MECHANISM

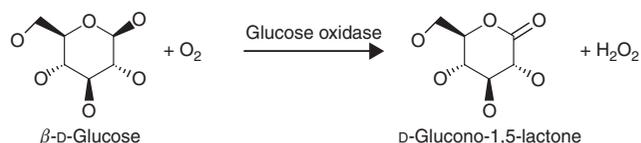
NOMENCLATURE AND CLASSIFICATION OF ENZYMES

Enzymes are classified according to the nature of the reaction they catalyse (e.g. oxidation/reduction, hydrolysis, synthesis, etc.) and sub-classified according to the exact identity of their substrates and products. This nomenclature system was established by the *Enzyme Commission* (a committee of the International Union of Biochemistry). According to this system, all enzymes are classified into six major classes:

1. *Oxidoreductases*, which catalyse oxidation–reduction reactions.
2. *Transferases*, which catalyse group transfer from one molecule to another.
3. *Hydrolases*, which catalyse hydrolytic cleavage of C–C, C–N, C–O, C–S or O–P bonds. These are group transfer reactions but the acceptor is always water.
4. *Lyases*, which catalyse elimination reactions, resulting in the cleavage of C–C, C–O, C–N, C–S bonds or the formation of a double bond, or conversely adding groups to double bonds.
5. *Isomerases*, which catalyse isomerisation reactions, e.g., racemisation, epimerisation, *cis-trans*-isomerisation, tautomerisation.
6. *Ligases*, which catalyse bond formation, coupled with the hydrolysis of a high-energy phosphate bond in ATP or a similar triphosphate.

The Enzyme Commission system consists of a numerical classification hierarchy of the form ‘E.C. a.b.c.d’ in which ‘a’ represents the class of reaction catalysed and can take values from 1 to 6 according to the classification of reaction types given

above. ‘b’ denotes the sub-class, which usually specifies more precisely the type of the substrate or the bond cleaved, e.g. by naming the electron donor of an oxidation–reduction reaction or by naming the functional group cleaved by a hydrolytic enzyme. ‘c’ denotes the sub-subclass, which allows an even more precise definition of the reaction catalysed. For example, sub-subclasses of oxidoreductases are denoted by naming the acceptor of the electron from its respective donor. ‘d’ is the serial number of the enzyme within its sub-subclass. An example will be analysed. The enzyme that oxidises D-glucose using molecular oxygen catalyses the following reaction:



Scheme 7.1.

Hence, its systematic name is D-glucose: oxygen oxidoreductase, and its systematic number is EC 1.1.3.4.

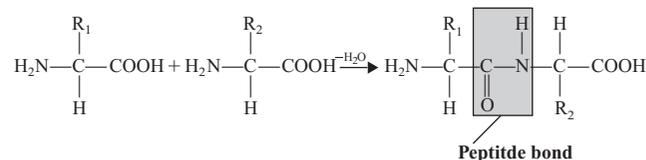
The systematic names are often quite long, and therefore, short trivial names along with systematic numbers are often more convenient for enzyme designation. These shorter names are known as *recommended names*. The recommended names consist of the suffix ‘-ase’ added to the substrate acted on. For example for the enzyme mentioned above, the recommended name is glucose oxidase.

It should be noted that the system of nomenclature and classification of enzymes is based only on the reaction catalysed and takes no account of the origin of the enzyme, that is from the species or tissue it derives.

BASIC ELEMENTS OF ENZYME STRUCTURE

The Primary Structure of Enzyme

Enzymes are composed of L-α-amino acids joined together by a *peptide bond* between the carboxylic acid group of one amino acid and the amino group of the next.



Scheme 7.2.

There are 20 common amino acids in proteins, which are specified by the genetic code; in rare cases, others occur as the products of enzymatic modifications after translation. A common

feature of all 20 amino acids is a central carbon atom (C_α) to which a hydrogen atom, an amino group ($-NH_2$) and a carboxyl group ($-COOH$) are attached. Free amino acids are usually zwitterionic at neutral pH, with the carboxyl group deprotonated and the amino group protonated. The structure of the most common amino acids found in proteins is shown in Table 7.1. Amino acids can be divided into four different classes depending on the structure of their side chains, which are called R groups: non-polar, polar uncharged, negatively charged (at neutral pH) and positively charged (at neutral pH; Richardson 1981). The properties of the amino acid side chains determine the properties of the proteins they constitute. The formation of a succession of peptide bonds generates the 'main chain' or 'backbone'.

The primary structure of a protein places several constraints on how it can fold to produce its three-dimensional structure (Cantor 1980, Fersht 1999). The backbone of a protein consists of a carbon atom C_α to which the side chain is attached, a NH group bound to C_α and a carbonyl group $C=O$, where the carbon atom C is attached to C_α (Fig. 7.1). The peptide bond is planar since it has partial ($\sim 40\%$) double bond character with π electrons shared between the C–O and C–N bonds (Fig. 7.1). The peptide bond has a *trans*-conformation, that is the oxygen of the carbonyl group and the hydrogen of the NH group are in the *trans* position; the *cis* conformation occurs only in exceptional cases (Richardson 1981).

Enzymes have several 'levels' of structure. The protein's sequence, that is the order of amino acids, is termed as its 'primary structure'. This is determined by the sequence of nucleotide bases in the gene that codes for the protein. Translation of the mRNA transcript produces a linear chain of amino acids linked together by a peptide bond between the carboxyl carbon of the

first amino acid and the free amino group of the second amino acid. The first amino acid in any polypeptide sequence has a free amino group and the terminal amino acid has a free carboxyl group. The primary structure is responsible for the higher levels of enzyme's structure and therefore for the enzymatic activity (Richardson 1981, Price and Stevens 1999).

The Three-Dimensional Structure of Enzymes

Enzymes are generally very closely packed globular structures with only a small number of internal cavities, which are normally filled by water molecules. The polypeptide chains of enzymes are organised into ordered, hydrogen bonded regions, known as *secondary structure* (Andersen and Rost 2003). In these ordered structures, any buried carbonyl oxygen forms a hydrogen bond with an amino NH group. This is done by forming α -helices and β -pleated sheets, as shown in Figure 7.2. The α -helix can be thought of as having a structure similar to a coil or spring (Surewicz and Mantsch 1988). The β -sheet can be visualised as a series of parallel strings laid on top of an accordion-folded piece of paper. These structures are determined by the protein's primary structure. Relatively small, uncharged, polar amino acids readily organise themselves into α -helices, while relatively small, non-polar amino acids form β -sheets. Proline is a special amino acid because of its unique structure (Table 7.1). Introduction of proline into the sequence creates a permanent bend at that position (Garnier et al. 1990). Therefore, the presence of proline in an α -helix or β -sheet disrupts the secondary structure at that point. The presence of a glycine residue confers greater than normal flexibility on a polypeptide chain. This is due to the absence of a bulky side chain, which reduces steric hindrance.

Another frequently observed structural unit is the β -turn (Fang and Shortle 2003). This occurs when the main chain sharply changes direction using a bend composed of four successive residues, often including proline and glycine. In these units, the C=O group of residue *i* is hydrogen bonded to the NH of residue *i*+3 instead of *i*+4 as in α -helix. Many different types of β -turn have been identified, which differ in terms of the number of amino acids and in conformation (e.g. Type I, Type II, Type III; Sibanda et al. 1989).

The three-dimensional structure of a protein composed of a single polypeptide chain is known as its *tertiary structure*. Tertiary structure is determined largely by the interaction of R groups on the surface of the protein with water and with other R groups on the protein's surface. The intermolecular non-covalent attractive forces that are involved in stabilising the enzyme's structure are usually classified into three types: ionic bonds, hydrogen bonds and van der Waals attractions (Matthews 1993). Hydrogen bonding results from the formation of hydrogen bridges between appropriate atoms; electrostatic bonds are due to the attraction of oppositely charged groups located on two amino acid side chains. Van der Waals bonds are generated by the interaction between electron clouds. Another important weak force is created by the three-dimensional structure of water, which tends to force hydrophobic groups together in order to minimise their disruptive effect on hydrogen-bonded

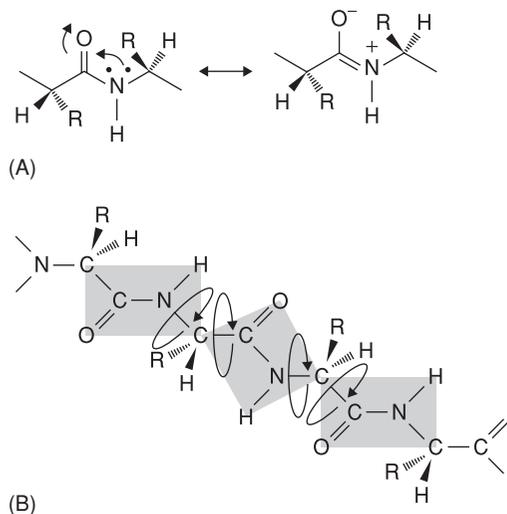
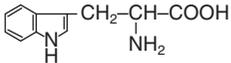
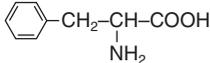
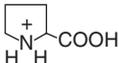
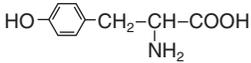
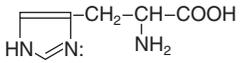


Figure 7.1. (A) The amide bond showing delocalisation of electrons. (B) A tripeptide unit of a polypeptide chain showing the planar amide units and the relevant angles of rotation about the bonds to the central α -carbon atom.

Table 7.1. Names, Symbols (One Letter and Three Letters Code) and Chemical Structures of the 20 Amino Acids Found in Proteins

<i>Non-polar amino acids</i>		
Alanine	Ala (A)	$\text{CH}_3\text{-CH-COOH}$ NH_2
Valine	Val (V)	H_3C H_3C \text{> } CH-CH-COOH NH_2
Leucine	Leu (L)	G_2B G_2B \text{> } $\text{CH-CH}_2\text{-CH-COOH}$ NH_2
Isoleucine	Ile (I)	$\text{H}_3\text{C-CH}_2$ \text{> } CH-CH-COOH H_3C NH_2
Methionine	Met (M)	$\text{CH}_3\text{-S-CH}_2\text{-CH-COOH}$ NH_2
Tryptophan	Trp (W)	 $\text{CH}_2\text{-CH-COOH}$ NH_2
Phenylalanine	Phe (P)	 $\text{CH}_2\text{-CH-COOH}$ NH_2
Proline	Pro (P)	
<i>Polar uncharged amino acids</i>		
Glycine	Gly (G)	H-CH-COOH NH_2
Asparagine	Asp (D)	$\text{H}_2\text{N-C-CH}_2\text{-CH-COOH}$ O NH_2
Glutamine	Gln (Q)	$\text{H}_2\text{N-C-CH}_2\text{-CH}_2\text{-CH-COOH}$ O NH_2
Serine	Ser (S)	$\text{HO-CH}_2\text{-CH-COOH}$ NH_2
Threonine	Thr (T)	CH_3 $\text{HO-CH}_2\text{-CH-COOH}$ NH_2
Cysteine	Cys (C)	$\text{HS-CH}_2\text{-CH-COOH}$ NH_2
Tyrosine	Tyr (Y)	HO-  $\text{CH}_2\text{-CH-COOH}$ NH_2
<i>Negatively charged amino acids</i>		
Glutamate	Glu (E)	$\text{HOOC-CH}_2\text{-CH}_2\text{-CH-COOH}$ NH_2
Aspartate	Asn (N)	$\text{HOOC-CH}_2\text{-CH-COOH}$ NH_2
<i>Positively charged amino acids</i>		
Arganine	Arg (R)	$\text{HN-CH}_2\text{-CH}_2\text{-CH}_2\text{-CH-COOH}$ C=NH NH_2 NH_2
Lysine	Lys (K)	$\text{H}_2\text{N-(CH}_2\text{)}_4\text{-CH-COOH}$ NH_2
Histidine	His (H)	 $\text{CH}_2\text{-CH-COOH}$ NH_2

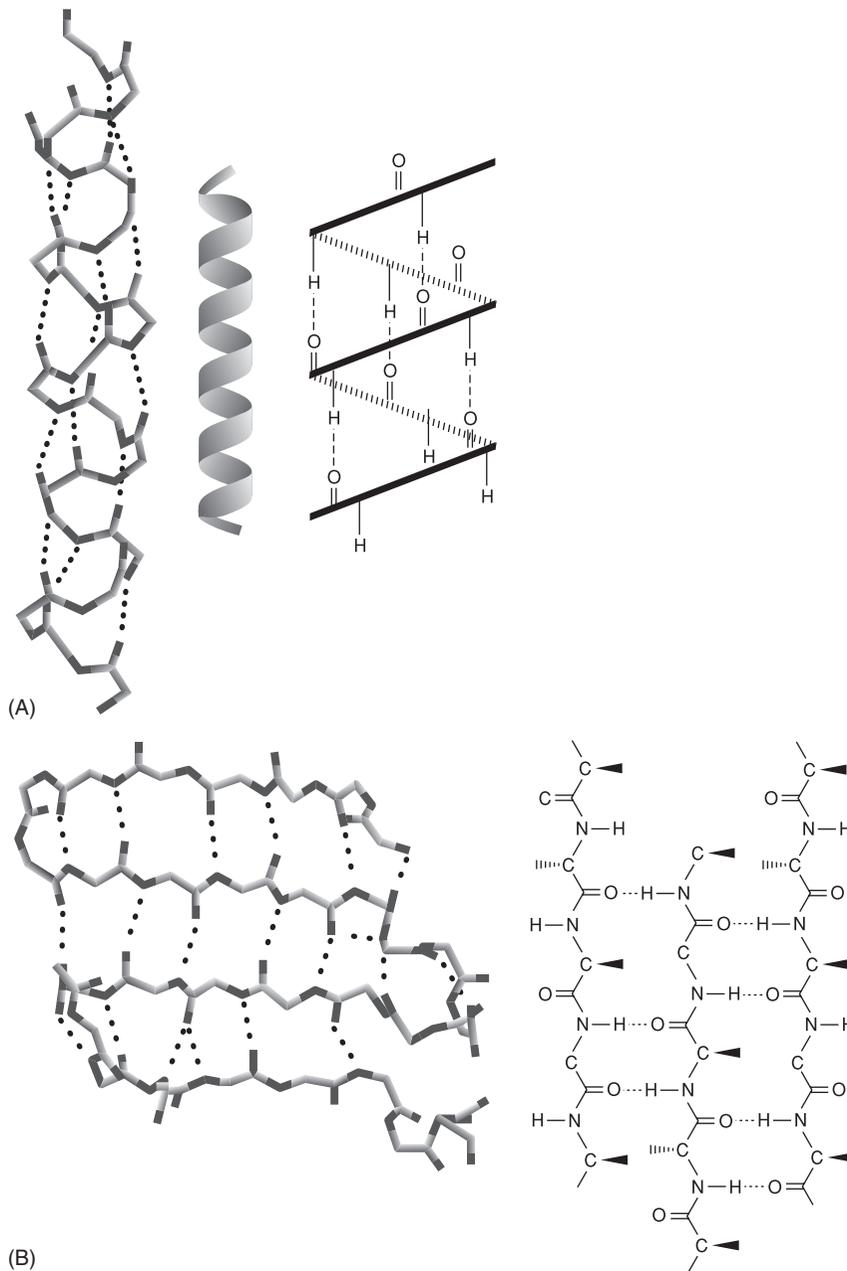


Figure 7.2. Representation of α -helix and β -sheet. Hydrogen bonds are depicted by dotted lines.

network of water molecules. Apart from the peptide bond, the only other type of covalent bond involved in linking amino-acids is the disulphide ($-S-S-$) bond, which can be formed between two cysteine side-chains under oxidising conditions. The disulphide bond contributes significantly to the structural stability of an enzyme and more precisely in tertiary structural stabilisation (see below for details; Matthews 1993, Estape et al. 1998).

Detailed studies have shown that certain combinations of α -helices or β -sheet together with turns occur in many proteins. These often-occurring structures have been termed *motifs*, or super-secondary structure (Kabsch and Sander 1983, Adamson et al. 1993). Examples of motifs found in several enzymes are shown in Figure 7.3. These protein folds represent highly stable structural units, and it is believed that they may form nucleating centres in protein folding (Richardson 1981).

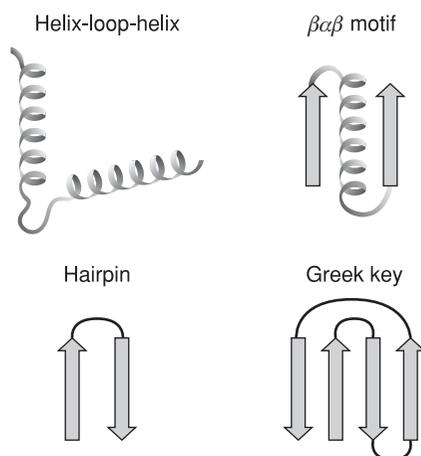


Figure 7.3. Common examples of motifs found in proteins.

Certain combinations of α -helices and β -sheets pack together to form compactly folded globular units, each of which is called protein domain with molecular mass of 15,000 to 20,000 Da (Orengo et al. 1997). Domains may be composed of secondary structure and identifiable motifs and therefore represent a higher level of structure than motifs. The most widely used classification scheme of domains has been the four-class system (Fig. 7.4; Murzin et al. 1995). The four classes of protein structure are as follows:

1. All- α -proteins, which have only α -helix structure.
2. All- β -proteins, which have β -sheet structure.
3. α/β -proteins, which have mixed or alternating segments of α -helix and β -sheet structure.
4. $\alpha+\beta$ proteins, which have α -helix and β -sheet structural segments that do not mix but are separated along the polypeptide chain.

While small proteins may contain only a single domain, larger enzymes contain a number of domains. In fact, most enzymes are dimers, tetramers or polymers of several polypeptide chains. Each polypeptide chain is termed 'subunit' and it may be identical or different to the others. The side chains on each polypeptide chain may interact with each other, as well as with water molecules to give the final enzyme structure. The overall organisation of the subunits is known as the *quaternary structure* and therefore the quaternary structure is a characteristic of multi-subunit enzymes. The four levels of enzyme structure are illustrated in Figure 7.5.

THEORY OF ENZYME CATALYSIS AND MECHANISM

In order for a reaction to occur, the reactant molecules must possess sufficient energy to cross a potential energy barrier, which is known as the *activation energy* (Fig. 7.6; Hackney 1990). All reactant molecules have different amounts of energy, but only a small proportion of them have sufficient energy to

cross the activation energy of the reaction. The lower the activation energy, the more substrate molecules are able to cross the activation energy. The result is that the reaction rate is increased.

Enzyme catalysis requires the formation of a specific reversible complex between the substrate and the enzyme. This complex is known as the *enzyme-substrate complex* (ES) and provides all the conditions that favour the catalytic event (Hackney 1990, Marti et al. 2004). Enzymes accelerate reactions by lowering the energy required for the formation of a complex of reactants that is competent to produce reaction products. This complex is known as the *transition state complex* of the reaction and is characterised by lower free energy than it would be found in the uncatalysed reaction.



Scheme 7.3.

The ES must pass to the transition state (ES^*). The transition state complex must advance to an *enzyme-product complex* (EP), which dissociates to free enzyme and product (P). This reaction's pathway goes through the transition states TS_1 , TS_2 and TS_3 . The amount of energy required to achieve the transition state is lowered; hence, a greater proportion of the molecules in the population can achieve the transition state and cross the activation energy (Benkovic and Hammes-Schiffer 2003, Wolfenden 2003). Enzymes speed up the forward and reverse reactions proportionately, so that they have no effect on the equilibrium constant of the reactions they catalyse (Hackney 1990).

Substrate is bound to the enzyme by relatively weak non-covalent forces. The free energy of interaction of the ES complex ranges between -12 to -36 kJ/mole. The intermolecular attractive forces between enzyme-substrate, in general, are of three types: ionic bonds, hydrogen bonds and van der Waals attractions.

Specific part of the protein structure that interacts with the substrate is known as the *substrate binding site* (Fig. 7.7). The substrate binding site is a three-dimensional entity suitably designed as a pocket or a cleft to accept the structure of the substrate in three-dimensional terms. The binding residues are defined as any residue with any atom within 4 \AA of a bound substrate. These binding residues that participate in the catalytic event are known as the *catalytic-residues* and form the *active-site*. According to Bartlett et al. (Bartlett et al. 2002), a residue is defined as catalytic if any of the following take place:

1. Direct involvement in the catalytic mechanism, for example as a nucleophile.
2. Exerting an effect, that aids catalysis, on another residue or water molecule, which is directly involved in the catalytic mechanism.
3. Stabilisation of a proposed transition-state intermediate.
4. Exerting an effect on a substrate or cofactor that aids catalysis, for example by polarising a bond that is to be broken.

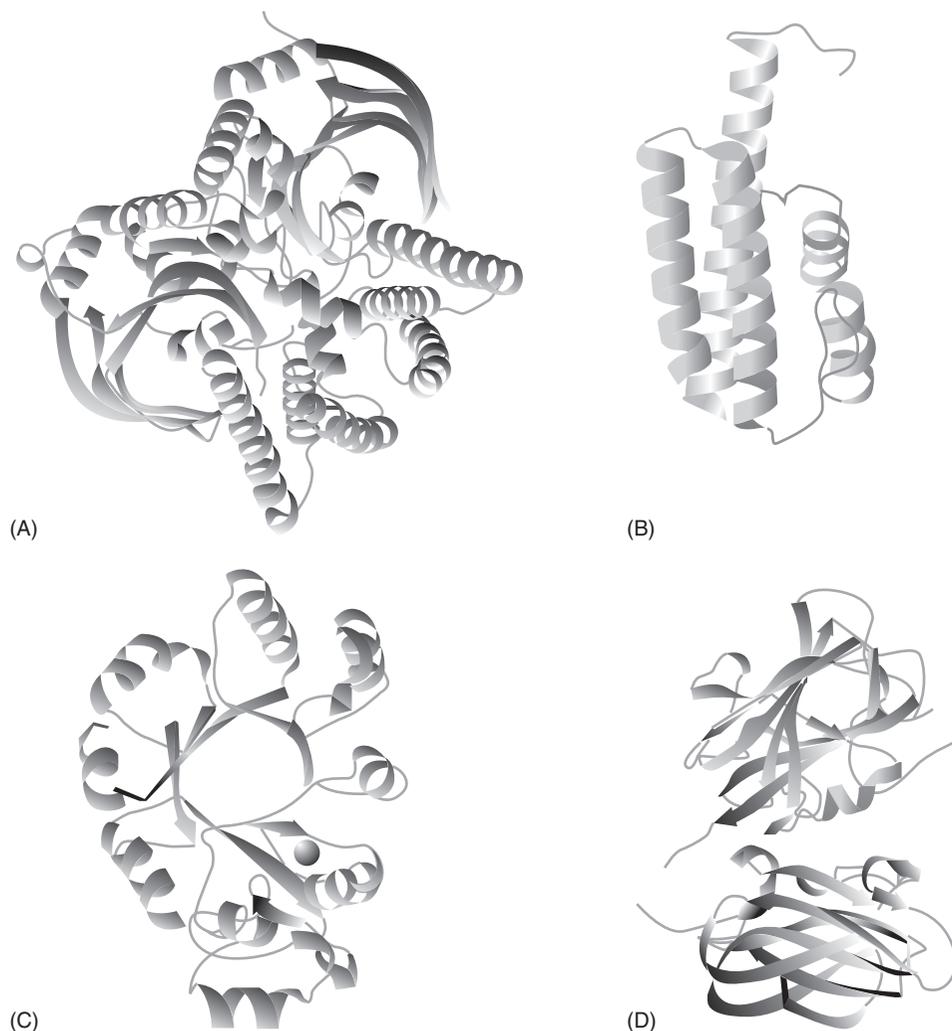


Figure 7.4. The four-class classification system of domains. **(A)** The $\alpha + \beta$ class (structure of glycyl-tRNA synthetase α -chain). **(B)** the all α class (structure of the hypothetical protein (Tm0613) from *Thermotoga maritima*). **(C)** The α/β class (structure of glycerophosphodiester phosphodiesterase). **(D)** The all β class (structure of allantoicase from *Saccharomyces cerevisiae*).

Despite the impression that the enzyme's structure is static and locked into a single conformation, several motions and conformational changes of the various regions always occur (Hammes 2002). The extent of these motions depends on many factors, including temperature, the properties of the solvating medium, the presence or absence of substrate and product (Hammes 2002). The conformational changes undergone by the enzyme play an important role in controlling the catalytic cycle. In some enzymes, there are significant movements of the binding residues, usually on surface loops, and in other cases, there are larger conformational changes. Catalysis takes place in the closed form and the enzyme opens again to release the product. This favoured model that explains enzyme catalysis and substrate interaction is the so-called *induced fit hypothesis* (Anderson et al. 1979, Joseph

et al. 1990). In this hypothesis, the initial interaction between enzyme and substrate rapidly induces conformational changes in the shape of the active site, which results in a new shape of the active site that brings catalytic residues close to substrate bonds to be altered (Fig. 7.8). When binding of the substrate to the enzyme takes place, the shape adjustment triggers catalysis by generating transition-state complexes. This hypothesis helps to explain why enzymes only catalyse specific reactions (Anderson et al. 1979, Joseph et al. 1990). This basic cycle has been seen in many different enzymes, including triosephosphate isomerase, which uses a small hinged loop to close the active site (Joseph et al. 1990) and kinases, which use two large lobes moving towards each other when the substrate binds (Anderson et al. 1979).

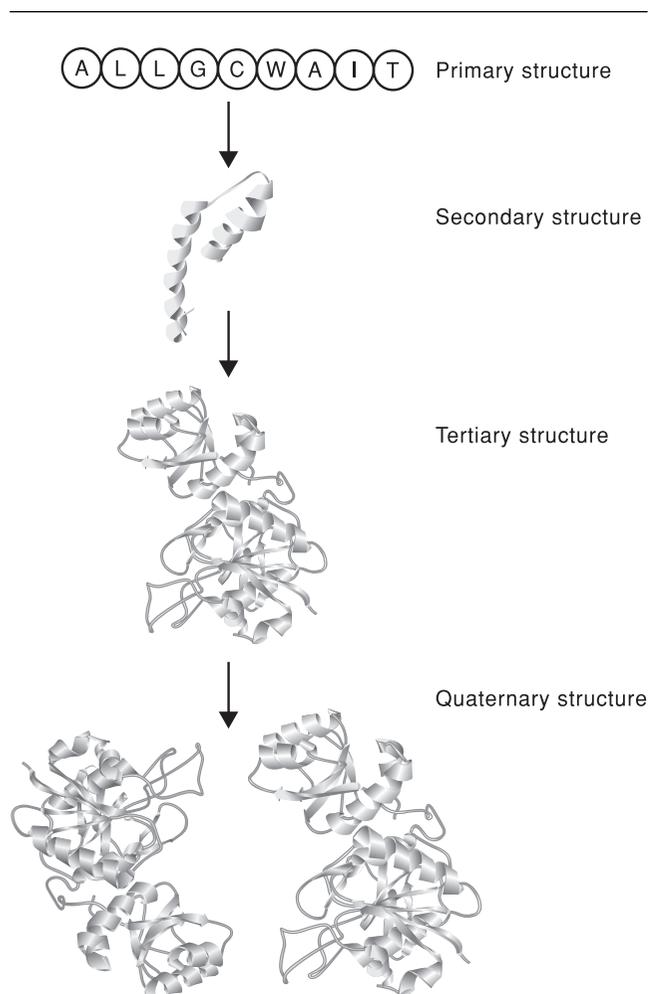


Figure 7.5. Schematic representation of the four levels of protein structure.

COENZYMES, PROSTHETIC GROUPS AND METAL ION COFACTORS

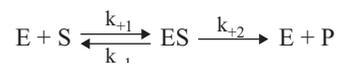
Non-protein groups can also be used by enzymes to affect catalysis. These groups, called *cofactors*, can be organic or inorganic and are divided into three classes: coenzymes, prosthetic groups and metal ion cofactors (McCormick 1975). Prosthetic groups are tightly bound to an enzyme through covalent bond. Coenzymes bind to enzyme reversibly and associate and dissociate from the enzyme during each catalytic cycle and therefore may be considered as co-substrates. An enzyme containing a cofactor or prosthetic group is termed as *holoenzyme*. Coenzymes can be broadly classified into three main groups: coenzymes that transfer groups onto substrate, coenzymes that accept and donate electrons and compounds that activate substrates (Table 7.2). Metal ions such as Ca^{+2} , Mg^{+2} , Zn^{+2} , Mn^{+2} , Fe^{+2} and Cu^{+2} may in some cases act as cofactors. These may be bound to the enzyme by simple coordination with electron-donating atoms of side chains (imidazole of His, $-\text{SH}$ group of Cys, carboxy-

late O^- of Asp and Glu). In some cases, metals, such as Mg^{+2} , are associated with the substrate rather than the enzyme. For example, Mg-ATP is the true substrate for kinases (Anderson et al. 1979). In other cases, metals may form part of a prosthetic group in which they are bound by coordinate bonds (e.g. heme; Table 7.2) in addition to side-chain groups. Usually, in this case, metal ions participate in electron transfer reactions.

KINETICS OF ENZYME-CATALYSED REACTIONS

The term enzyme kinetics implies a study of the velocity of an enzyme-catalysed reaction and of the various factors that may affect this (Moss 1988). An extensive discussion of enzyme kinetics would stay too far from the central theme of this chapter, but some general aspects will be briefly considered.

The concepts underlying the analysis of enzyme kinetics continue to provide significant information for understanding *in vivo* function and metabolism and for the development and clinical use of drugs aimed at selectively altering rate constants and interfering with the progress of disease states (Bauer et al. 2001). Central scope of any study of enzyme kinetics is knowledge of the way in which reaction velocity is altered by changes in the concentration of the enzyme's substrate and of the simple mathematics underlying this (Wharton 1983, Moss 1988, Watson and Dive 1994). As we have already discussed, the enzymatic reactions proceed through an intermediate ES in which each molecule of enzyme is combined, at any given instant during the reaction, with one substrate molecule. The reaction between enzyme and substrate to form the ES is reversible. Therefore, the overall enzymatic reaction can be shown as



Scheme 7.4.

where k_{+1} , k_{-1} and k_{+2} are the respective rate constants. The reverse reaction concerning the conversion of product to substrate is not included in this scheme. This is allowed at the beginning of the reaction when there is no, or little, product present. In 1913, biochemists Michaelis and Menten suggested that if the reverse reaction between E and S is sufficiently rapid, in comparison with the breakdown of ES complex to form product, the latter reaction will have a negligible effect on the concentration of the ES complex. Consequently, E, S and ES will be in equilibrium, and the rates of formation and breakdown of ES will be equal. On the basis of these assumptions Michaelis and Menten produced the following equation:

$$v = \frac{V_{\max} \cdot [\text{S}]}{K_m + [\text{S}]}$$

This equation is a quantitative description of the relationship between the rate of an enzyme-catalysed reaction (v) and the concentration of substrate $[\text{S}]$. The parameters V_{\max} and K_m are

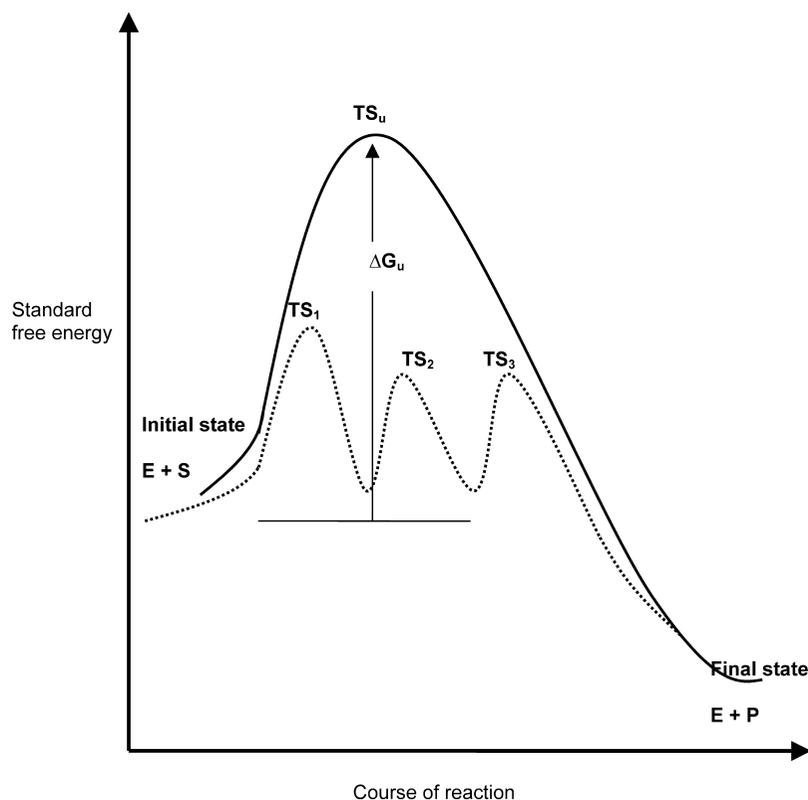


Figure 7.6. A schematic diagram showing the free energy profile of the course of an enzyme catalysed reaction involving the formation of enzyme–substrate (ES) and enzyme–product (EP) complexes. The catalysed reaction pathway goes through the transition states TS_1 , TS_2 , and TS_3 , with standard free energy of activation ΔG_c , whereas the uncatyalsed reaction goes through the transition state TS_u with standard free energy of activation ΔG_u .

constants at a given temperature and a given enzyme concentration. The K_m or Michaelis constant is the substrate concentration at which $v = V_{max}/2$ and its usual unit is M. The K_m provides us with information about the substrate *binding affinity* of the enzyme. A high K_m indicates a low affinity and vice versa (Moss 1988, Price and Stevens 1999).

The V_{max} is the maximum rate of the enzyme-catalysed reaction and it is observed at very high substrate concentrations where all the enzyme molecules are saturated with substrate, in the form of ES complex. Therefore

$$V_{max} = k_{cat}[E_t]$$

where $[E_t]$ is the total enzyme concentration and k_{cat} is the rate of breakdown of the ES complex (k_{-2} in the equation), which is known as the *turnover number*. k_{cat} represents the maximum number of substrate molecules that the enzyme can convert to product in a set time. The K_m depends on the particular enzyme and substrate being used and on the temperature, pH, ionic strength, etc. However, note that K_m is independent of the enzyme concentration, whereas V_{max} is proportional to enzyme concentration. A plot of the initial rate (v) against initial substrate concentration ($[S]$) for a reaction obeying the Michaelis–Menten

kinetics has the form of a rectangular hyperbola through the origin with asymptotes $v = V_{max}$ and $[S] = -K_m$ (Fig. 7.9A). The term *hyperbolic kinetics* is also sometimes used to characterise such kinetics.

There are several available methods for determining the parameters from the Michaelis–Menten equation. A better method for determining the values of V_{max} and K_m was formulated by Hans Lineweaver and Dean Burk and is termed the Lineweaver–Burk (LB) or *double reciprocal plot* (Fig. 7.9B). Specifically, it is a plot of $1/v$ versus $1/[S]$, according to the equation:

$$\frac{1}{v} = \frac{K_m}{V_{max}} \cdot \frac{1}{[S]} + \frac{1}{V_{max}}$$

Such a plot yields a straight line with a slope of K_m/V_{max} . The intercept on the $1/v$ axis is $1/V_{max}$ and the intercept on the $1/[S]$ axis is $-1/K_m$.

The rate of an enzymatic reaction is also affected by changes in pH and temperature (Fig. 7.10). When pH is varied, the velocity of reaction in the presence of a constant amount of enzyme is typically greatest over a relatively narrow range of pH. Since enzymes are proteins, they possess a large number of ionic groups, which are capable of existing in different ionic forms (Labrou et al. 2004a). The existence of a fairly narrow pH-optimum for

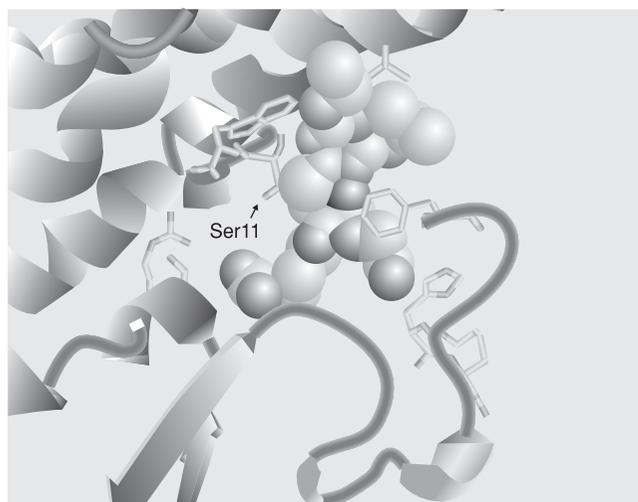


Figure 7.7. The substrate binding site of maize glutathione *S*-transferase. The binding residues are depicted as sticks, whereas the substrate is depicted in a space fill model. Only Ser 11 is involved directly in catalysis and is considered as catalytic residue.

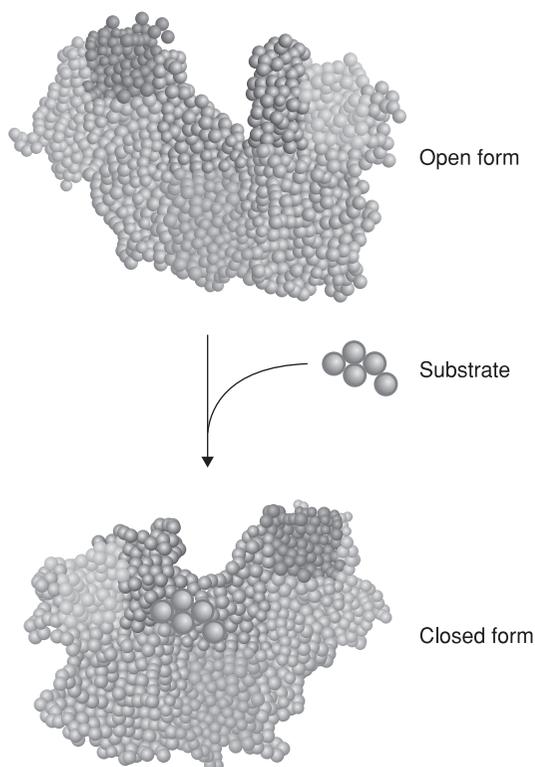


Figure 7.8. A schematic representation of the induced fit hypothesis.

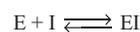
most enzymes suggests that one particular ionic form of the enzyme molecule, out of the many that it can potentially exist, is the catalytically active one. The effect of pH changes on v is reversible, except after exposure to extremes of pH at which denaturation of the enzyme may occur.

The rate of an enzymatic reaction increases with increasing temperature. Although there are significant variations from one enzyme to another, on average, for each 10°C rise in temperature, the enzymatic activity is increased by an order of two. After exposure of the enzyme to high temperatures (normally greater than 65°C), denaturation of the enzyme may occur and the enzyme activity decreased. The Arrhenius equation

$$\text{Log } V_{\text{max}} = \frac{-E_a}{2303RT} + A$$

provides a quantitative description of the relationship between the rate of an enzyme-catalysed reaction (V_{max}) and the temperature (T). Where E_a is the activation energy of the reaction, R is the gas constant, and A is a constant relevant to the nature of the reactant molecules.

The rates of enzymatic reactions are affected by changes in the concentrations of compounds other than the substrate. These modifiers may be activators, that is they increase the rate of reaction or their presence may inhibit the enzyme's activity. Activators and inhibitors are usually small molecules or even ions. Enzyme inhibitors fall into two broad classes: those causing irreversible inactivation of enzymes and those whose inhibitory effects can be reversed. Inhibitors of the first class bind covalently to the enzyme so that physical methods of separating the two are ineffective. Reversible inhibition is characterised by the existence of equilibrium between enzyme and inhibitor (I):



Scheme 7.5.

The equilibrium constant of the reaction, K_i , is given by the equation:

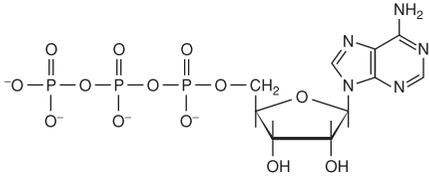
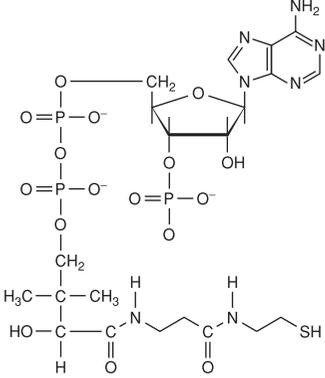
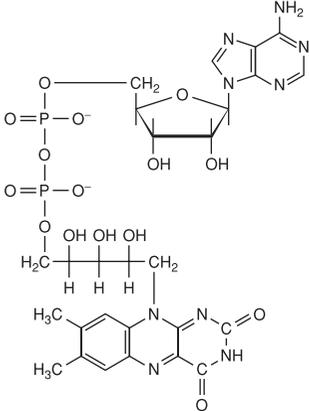
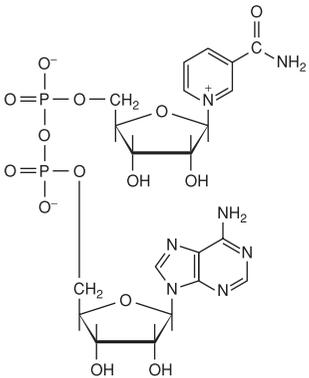
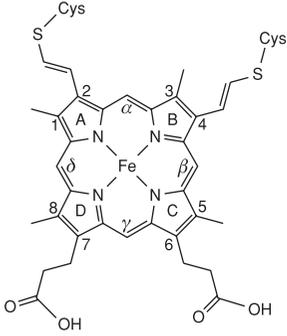
$$K_i = \frac{[ES]}{[E][I]}$$

K_i is a measure of the affinity of the inhibitor for the enzyme. Reversible inhibitors can be divided into three main categories: competitive inhibitors, non-competitive inhibitors and uncompetitive inhibitors. The characteristic of each type of inhibition and their effect on the kinetic parameters K_m and V_{max} are shown in Table 7.3.

THERMODYNAMIC ANALYSIS

Thermodynamics is the science that deals with energy. Chemical bonds store energy, as the subatomic particles are being attracted and chemical reactions are occurred with energy changes. Originally, thermodynamic laws were used to analyse and characterise mechanical systems, but the kinetic molecular theory

Table 7.2. The Structure of Some Common Coenzymes: Adenosine Triphosphate (ATP), Coenzyme A (CoA), Flavin Adenine Dinucleotide (FAD), Nicotinamide Adenine Dinucleotide (NAD⁺) and Heam c

Cofactor	Type of Reactions Catalysed	Structure
Adenosine triphosphate (ATP)	Phosphate transfer reactions (e.g. kinases)	
Coenzyme A (CoA)	Acyl transfer reactions (transferases)	
Flavin adenine dinucleotide (FAD)	Redox reactions (reductases)	
Nicotinamide adenine dinucleotide (NAD ⁺)	Redox reactions (e.g. dehydrogenases)	
Heam c	Activate substrates	

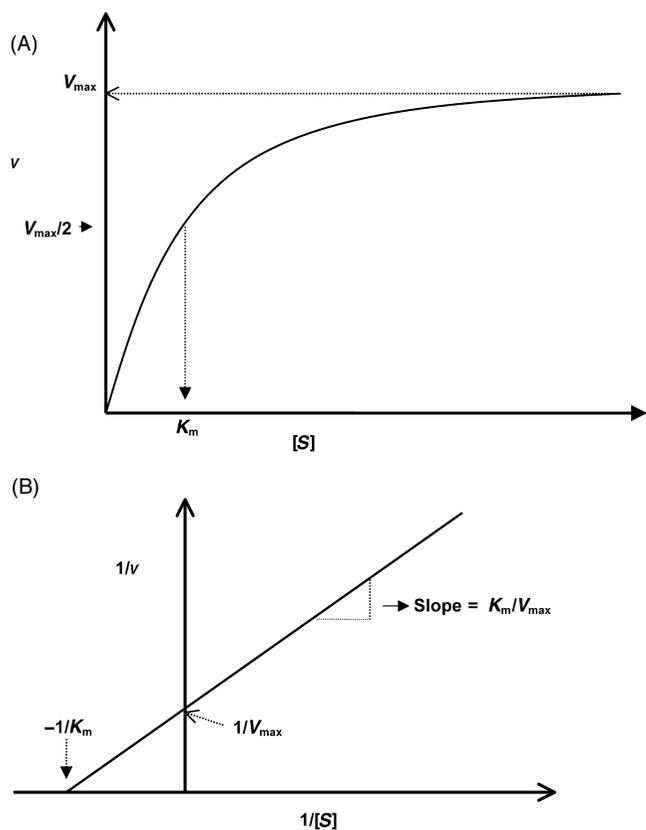


Figure 7.9. (A) A plot of the initial rate (v) against initial substrate concentration ($[S]$) for a reaction obeying the Michaelis-Menten kinetics. The substrate concentration, which gives a rate of half the maximum reaction velocity, is equal to the K_m . (B) The Lineweaver-Burk plot. The intercept on the $1/v$ axis is $1/V_{max}$, the intercept on the $1/[S]$ axis is $-1/K_m$ and the slope is K_m/V_{max} .

(Eblin 1964) suggests that molecules are in constant motion, which obey the same laws of mechanics as macroscopic objects. Since enzymes are biomolecules, which catalyse chemical reactions, they contribute to the energetic background of a chemical system. As a result, enzymes are directly related to thermodynamic parameters.

In this section, we will focus on the most representative thermodynamic parameters that are useful, in order to describe experimental data. The fundamental thermodynamic parameters are the Helmholtz free energy (F), the entropy (S) and the enthalpy (H). Free energy (F) correlates the energy (E) of a system with its entropy (S): $F = E - TS$, where T is the temperature. The entropy (S) of a system is the range of its disorder and the thermodynamic definition is (Rakintzis 1994) (the symbol q is used to represent the amount of heat absorbed by a system):

$$S = \int \frac{q}{T} dT \Leftrightarrow S = q \ln(T)$$

Furthermore, enthalpy is a constitutive function of pressure (P) and temperature (T), and energy (E) is a constitutive func-

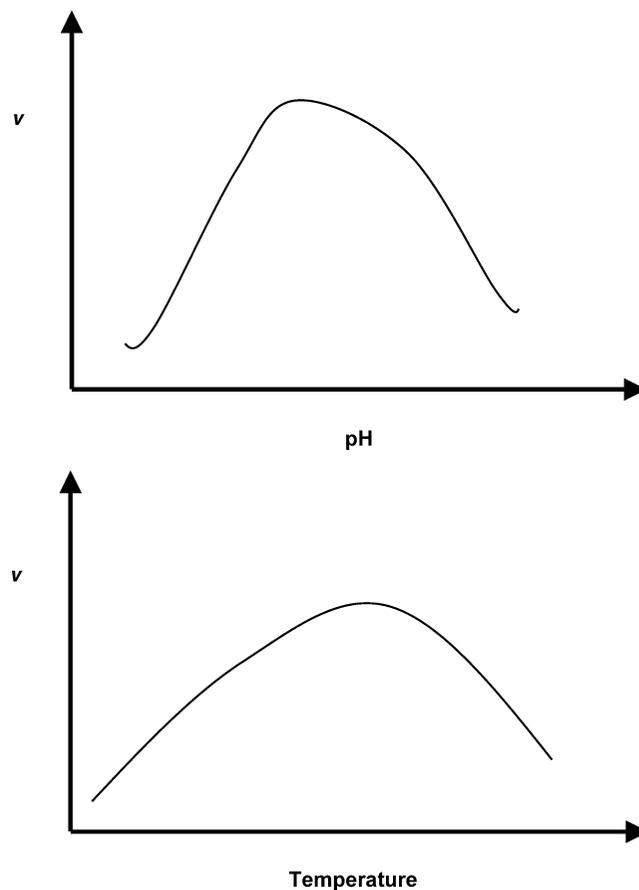


Figure 7.10. Relationship of the activity–pH and activity–temperature for a putative enzyme.

tion of volume (V) and temperature (T), respectively (Rakintzis 1994):

$$H = f(P, T)$$

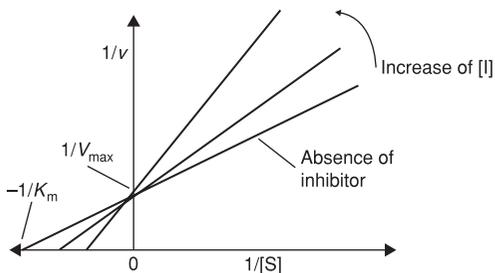
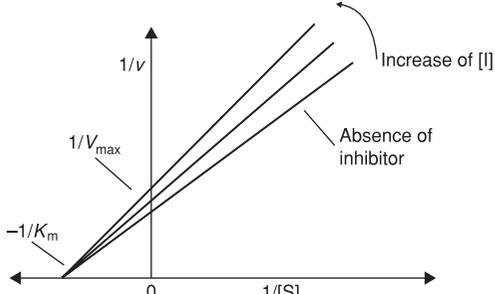
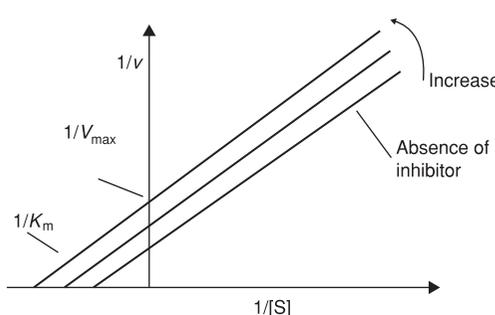
$$dH = \left[\frac{\partial H}{\partial P} \right]_T dP + \left[\frac{\partial H}{\partial T} \right]_P dT$$

$$E = f(V, T)$$

$$dE = \left[\frac{\partial E}{\partial V} \right]_T dV + \left[\frac{\partial E}{\partial T} \right]_V dT$$

At this point, we have to emphasise that in biological systems, a normal experiment deals with a constant pressure (atmospheric pressure) and not with a constant volume. This means, that we determine not a change of energy (E) of a studied body (e.g. enzyme), but a change in its enthalpy: $H = E + PV$. However, the quantity (PV) is considered insignificant in case that we study a protein for instance, because the volume per molecule is too small. Hence, there is no difference between H and E , and we can refer both of them simply as “energy” (Finkelstein and Ptitsyn 2002). Likewise, there is no difference between Helmholtz free energy ($F = E - TS$) and Gibb’s free energy ($G = H - TS$), since we consider that H is equivalent to E . As a result, we

Table 7.3. The Characteristic of Each Type of Inhibition and Their Effect on the Kinetic Parameters K_m and V_{max}

Inhibitor Type	Binding Site on Enzyme	Kinetic Effect
Competitive Inhibitor	<p>The inhibitor specifically binds at the enzyme's catalytic site, where it competes with substrate for binding:</p> $E \xrightleftharpoons[+I \downarrow -I]{+S} ES \rightarrow E + P$ <p>EI</p>	 <p>K_m is increased; V_{max} is unchanged. LB equation: $\frac{1}{v} = \frac{K_m}{V_{max}} \left(1 + \frac{[I]}{K_i} \right) \frac{1}{[S]} + \frac{1}{V_{max}}$</p>
Non-competitive Inhibitor	<p>The inhibitor binds to E or to the ES complex (may form an ESI complex) at a site other than the catalytic. Substrate binding is unchanged, whereas ESI complex cannot form products.</p> $E \xrightleftharpoons[+I \downarrow -I]{+S} ES \rightarrow E + P$ $EI \xrightleftharpoons[+I \downarrow -I]{+S} ESI \rightleftharpoons E + P$	 <p>K_m is unchanged; V_{max} is decreased proportionately to inhibitor concentration. LB equation: $\frac{1}{v} = \frac{K_m}{V_{max}} \cdot \left(1 + \frac{[I]}{K_i} \right) \cdot \frac{1}{[S]} + \frac{1}{V_{max}} \left(1 + \frac{[I]}{K_i} \right)$</p>
Uncompetitive Inhibitor	<p>Binds only to ES complexes at a site other than the catalytic site. Substrate binding alters enzyme structure, making inhibitor-binding site available:</p> $E \xrightleftharpoons[+I \downarrow -I]{+S} ES \rightarrow E + P$ <p>EIS</p>	 <p>K_m and V_{max} are decreased. LB equation: $\frac{1}{v} = \frac{K_m}{V_{max}} \cdot \frac{1}{[S]} + \frac{1}{V_{max}} \left(1 + \frac{[I]}{K_i} \right)$</p>

can refer both F and G as 'free energy' (Finkelstein and Ptitsyn 2002). Eventually, when we study the thermodynamic behaviour of an enzyme, regarding its structural stability or its functionality, we have to do with Gibb's free energy (G), entropy (S) and enthalpy (H).

As far as experimental procedures are concerned, someone can use two equations in order to measure the thermodynamic

parameters, applying temperature differentiation experimental protocols (Kotzia and Labrou 2005). These equations are the Arrhenius (Equation 4) and the Eyring equations, which basically describe the temperature dependence of reaction state. So, considering an enzymic reaction as a chemical reaction, which is characterised by a transition state, we can efficiently use the Eyring's equation. This equation is based on transition state

model, whose principles are (i) there is a thermodynamic equilibrium between the transition state and the state of reactants at the top of the energy barrier and (ii) the rate of chemical reaction is proportional to the concentration of the particles in the high-energy transition state.

Combining basic kinetic expressions, which describe a simple chemical reaction:



Scheme 7.6.

and thermodynamic equations, such as (Finkelstein and Ptitsyn 2002)

$$\Delta G = -RT \ln k$$

$$\Delta G = \Delta H - T \Delta S$$

Eyring's equation is formed:

$$k = \frac{k_B T}{h} e^{-\left(\frac{\Delta G}{RT}\right)} = \frac{k_B T}{h} e^{-\left(\frac{\Delta H}{RT} - \frac{\Delta S}{R}\right)}$$

where k is the reaction velocity, R is the universal gas constant ($8.3145 \text{ J} \cdot \text{mol}^{-1} \text{ K}$), k_B the Boltzmann's constant ($1.381 \cdot 10^{-23} \text{ J} \cdot \text{K}^{-1}$), h is the Plank constant ($6.626 \cdot 10^{-34} \text{ J} \cdot \text{s}$) and T is the absolute temperature in degrees Kelvin (K). If we compare Eyring's equation and Arrhenius equation (Laidler 1984), we can realise that E and ΔH are parallel quantities, as we concluded above. Eyring's equation is an appropriate and useful tool, which allows us to simplify the complicated meaning of thermodynamics and, simultaneously, to interpret experimental data to physical meanings.

ENZYME DYNAMICS DURING CATALYSIS

Multiple conformational changes and intramolecular motions appear to be a general feature of enzymes (Agarwal et al. 2002). The structures of proteins and other biomolecules are largely maintained by non-covalent forces and are therefore subject to thermal fluctuations ranging from local atomic displacements to complete unfolding. These changes are intimately connected to enzymatic catalysis and are believed to fulfil a number of roles in catalysis: enhanced binding of substrate, correct orientation of catalytic groups, removal of water from the active site and trapping of intermediates. Enzyme conformational changes may be classified into four types (Gutteridge and Thornton 2004): (i) domain motion, where two rigid domains, joined by a flexible hinge, move relative to each other; (ii) loop motion, where flexible surface loops (2–10 residues) adopt different conformations; (iii) side chain rotation, which alters the position of the functional atoms of the side chain and (iv) secondary structure changes.

Intramolecular motions in biomolecules are usually very fast (picosecond–nanosecond) local fluctuations. The flexibility associated with such motions provides entropic stabilisation of conformational states (Agarwal et al. 2002). In addition, there are

also slower- (microsecond–millisecond) and larger-scale, thermally activated, transitions. Large-scale conformational changes are usually key events in enzyme regulation.

ENZYME PRODUCTION

In the past, enzymes were isolated primarily from natural sources, and thus a relatively limited number of enzymes were available to the industry (Aehle 2007). For example, of the hundred or so enzymes being used industrially, over one half are from fungi and yeast and over a third are from bacteria, with the remainder divided between animal (8%) and plant (4%) sources (Panke and Wubbolts 2002, van Beilen and Li 2002, Bornscheuer 2005). Today, with the recent advances of molecular biology and genetic engineering, several expression systems were developed and exploited and used for the commercial production of several therapeutic (Mcgrath 2005), analytical or industrial enzymes (van Beilen and Li 2002, Kirk et al. 2002, Otero and Nielsen 2010). These systems have not only improved the efficiency, availability and cost with which enzymes can be produced, but they have also improved their quality (Labrou and Rigden 2001, Andreadeli et al. 2008, Kapoli et al. 2008, Kotzia and Labrou 2009, Labrou 2010).

ENZYME HETEROLOGOUS EXPRESSION

There are two basic steps involved in the assembly of every heterologous expression system:

1. The introduction of the DNA, encoding the gene of interest, into the host cells, which requires: (i) the identification and isolation of the gene of the protein we wish to express, (ii) insertion of the gene into a suitable expression vector and (iii) introduction of the expression vector into the selected cell system that will accommodate the heterologous protein.
2. The optimisation of protein expression by taking into account the effect of various factors such as growing medium, temperature, and induction period.

A variety of vectors able to carry the DNA into the host cells are available, ranging from plasmids, cosmids, phagemids, viruses as well as artificial chromosomes of bacterial, yeast or human origin (BAC, YAC or HAC, respectively; Grimes and Monaco 2005, Monaco and Moralli 2006, Takahashi and Ueda 2010). The vectors are either integrated into the host chromosomal DNA or remain in an episomal form. In general, expression vectors have the following characteristics (Fig. 7.11):

1. Polylinker: it contains multiple restriction sites that facilitate the insertion of the desired gene into the vector.
2. Selection marker: encodes for a selectable marker, allowing the vector to be maintained within the host cell under conditions of selective pressure (i.e. antibiotic).
3. Ori: a sequence that allows for the autonomous replication of the vector within the cells.
4. Promoter: inducible or constitutive, that regulates RNA transcription of the gene of interest.

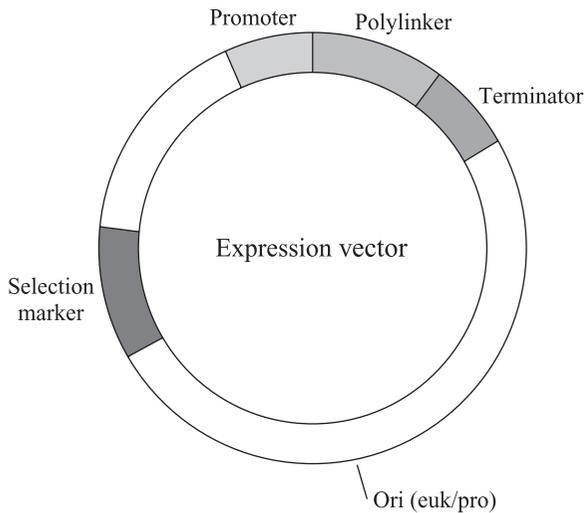


Figure 7.11. Generalised heterologous gene expression vector. Polylinker sequence (multiple cloning site (MCS)), promoter ($P^{\text{euk/pro}}$), terminator ($T^{\text{euk/pro}}$), selectable marker (prokaryotic and/or eukaryotic – $M^{\text{euk/pro}}$), and origin of replication (eukaryotic and/or prokaryotic – $Ori^{\text{euk/pro}}$).

5. Terminator: a strong terminator sequence that ensures that the RNA polymerase disengages and does not continue to transcribe other genes located downstream.

Vectors are usually designed with mixed characteristics for expression in both prokaryotic and eukaryotic host cells (Brondyk 2009, Demain and Vaishnav 2009). Artificial chromosomes are designed for cloning of very large segments of DNA (100 kb), usually for mapping purposes, and contain host-specific telomeric and centromeric sequences. These sequences permit the proper distribution of the vectors to the daughter cells during cell division and increase chromosome stability (Fig. 7.12).

THE CHOICE OF EXPRESSION SYSTEM

There are two main categories of expression systems: eukaryotic and prokaryotic (Demain and Vaishnav 2009). The choice of a suitable expression system involves the consideration of several important factors, such as protein yield, proper folding, post-translational modifications (e.g. phosphorylation, glycosylation), industrial applications of the expressed protein, as well as economic factors. For these reasons, there is no universally applied expression system. A comparison of the most commonly used expression systems is shown in Table 7.4.

Bacterial Cells

Expression of heterologous proteins in bacteria remains the most extensively used approach for the production of heterologous proteins (Zerbs et al. 2009), such as cytokines (Tang et al. 2006, Rabhi-Essafi et al. 2007), membrane proteins (Butzin

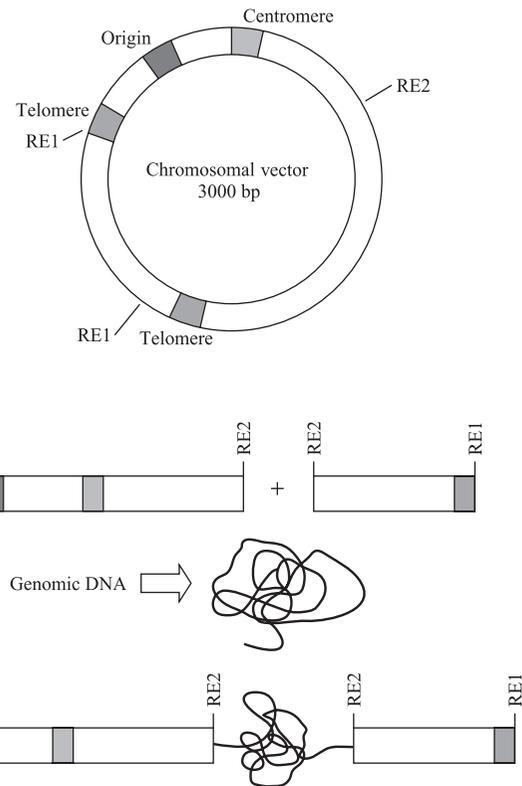


Figure 7.12. Artificial chromosome cloning system. Initially, the circular vector is digested with restriction endonucleases (RE) for linearisation and then ligated with size-fractionated DNA (≈ 100 kb). The vector contains centromeric and telomeric sequences, which assure chromosome-like propagation within the cell, as well as selection marker sequences for stable maintenance. Origin of replication (eukaryotic or prokaryotic – $Ori^{\text{euk/pro}}$), selectable marker (eukaryotic or prokaryotic – $M^{\text{euk/pro}}$), centromeric sequence (C), and telomeric sequence (T).

et al. 2009, Zoonens and Miroux 2010), enzymes (Melissis et al. 2006, Kotzia and Labrou 2007, Melissis et al. 2007, Andreadeli et al. 2008, Wu et al. 2009, Okino et al. 2010), antibodies (Kwong and Rader 2009, Xiong et al. 2009) and antigens (viral or non-viral; Donayre-Torres et al. 2009, Liu et al. 2009a, Vahedi et al. 2009, Ebrahimi et al. 2010) at both laboratory and industrial scale (Koehn and Hunt 2009, Sahdev et al. 2008, Peti and Page 2007, Jana and Deb 2005). Bacteria can be grown inexpensively and genetically manipulated very easily. They can reach very high densities rapidly and express high levels of recombinant proteins, reaching up to 50% of the total protein. However, in many cases, high-level expression correlates with poor quality. Often, the expressed protein is accumulated in the form of insoluble inclusion bodies (misfolded protein aggregate) and additional, sometimes labour-intensive, genetic manipulation or re-solubilisation/refolding steps are required (Burgess 2009). Bacterial cells do not possess the eukaryotes' extensive post-translational modification system (such as *N*- or

Table 7.4. Comparison of the Main Expression Systems

	Bacteria	Yeast	Fungi	Insect Cells	Mammalian Cells	Transgenic Plants	Transgenic Animals
Developing time	Short	Short	Short	Intermediate	Intermediate	Intermediate	Long
Costs for downstream processing	+++	++	++	++	++	++	++
Levels of expression	High	Intermediate	Intermediate	Intermediate	Low	Low	Low
Recombinant protein stability	+/-	+/-	+/-	+/-	+/-	+++	+/-
Production volume	Limited	Limited	Limited	Limited	Limited	Unlimited	Unlimited
Postranslational modifications (disulphide bond formation, glycosylation, etc.)	No	Yes	Yes	Yes	Yes	Yes	Yes
'Human-type' glycosylation	No	No	No	No	Yes	No	Yes
Folding capabilities	-	++	+++	+++	+++	+++	+++
Contamination level (pathogens, EPL, etc.)	++	-	-	-	++	-	++

O-glycosylation), a serious disadvantage, when post-translational modifications are essential to the protein's function (Zhang et al. 2004). However, they are capable of a surprisingly broad range of covalent modifications such as acetylation, amidation and deamidation, methylation, myristylation, biotinylation and phosphorylation.

Mammalian Cells

Mammalian cells are the ideal candidate for expression hosts (Engelhardt et al. 2009, Hacker et al. 2009) when post-translational modifications (*N*- and *O*-glycosylation, disulfide bond formation) are a critical factor for the efficacy of the expressed protein (Baldi et al. 2007, Werner et al. 2007, Durocher and Butler 2009, Geisse and Fux 2009, Hacker et al. 2009). Despite substantial limitations, such as high cost, low yield, instability of expression and time-consuming, a significant number of proteins (e.g. cytokines; Fox et al. 2004, Sunley et al. 2008, Suen et al. 2010), antibodies (Kim et al. 2008, Chusainow et al. 2009), enzymes (Zhuge et al. 2004), viral antigens (Holzer et al. 2003), blood factors and related proteins (Halabian et al. 2009, Su et al. 2010) are produced in this system because it offers very high product fidelity. However, oligosaccharide processing is species- and cell type-dependent among mammalian cells and differences in the glycosylation pattern have been reported in rodent cell lines and human tissues. The expressed proteins are usually recovered in a bioactive, properly folded form and secreted into the cell culture fluids.

Yeast

Yeast is a widely used expression system with many commercial, medical and basic research applications. The fact that the yeast is the most intensively studied eukaryote at the genetic/molecular level makes it an extremely advantageous expression system (Idiris et al. 2010). Being unicellular organism, it retains the

advantages of bacteria (low cultivation cost, high doubling rate, ease of genetic manipulation, ability to produce heterologous proteins in large-scale quantities) combined with the advantages of higher eukaryotic systems (post-translational modifications). The vast majority of yeast expression work has focused on the well-characterised baker's yeast *Saccharomyces cerevisiae* (Holz et al. 2003, Terpitz et al. 2008, Joubert et al. 2010), but a growing number of non-*Saccharomyces* yeasts are becoming available as hosts for recombinant polypeptide production, such as *Hansenula polymorpha*, *Candida boidinii*, *Kluyveromyces lactis*, *Pichia pastoris* (Cregg et al. 2000, Jahic et al. 2006, van Ooyen et al. 2006, Yurimoto and Sakai 2009), *Schizosaccharomyces pombe* (Alberti et al. 2007, Ahn et al. 2009, Takegawa et al. 2009), *Schwanniomyces occidentalis* and *Yarrowia lipolytica* (Madzak et al. 2004, 2005, Bankar et al. 2009). As in bacteria, expression in yeast relies on episomal or integrated multicopy plasmids with tightly regulated gene expression. Despite these advantages, expressed proteins are not always recovered in soluble form and may have to be purified from inclusion bodies. Post-translational modifications in yeast differ greatly from mammalian cells (Jacobs and Callewaert 2009, Hamilton and Gerngross 2007). This has sometimes proven to be a hindrance when high fidelity of complex carbohydrate modifications found in eukaryotic proteins appears to be important in many medical applications. Yeast cells do not add complex oligosaccharides and are limited to the high-mannose-type carbohydrates. These higher order oligosaccharides are possibly immunogenic and could potentially interfere with the biological activity of the protein.

Filamentous Fungi

Filamentous fungi have been extensively used for studies of eukaryotic gene organisation, regulation and cellular differentiation. Additionally, fungi belonging to the genus *Aspergillus* and *Penicillium* are of significant industrial importance because

of their applications in food fermentation and their ability to secrete a broad range of biopolymer degrading enzymes and to produce primary (organic acids) and secondary metabolites (antibiotics, vitamins). The extensive genetic knowledge as well as the already well-developed fermentation technology has allowed for the development of heterologous protein expression systems (Nevalainen et al. 2005, Wang et al. 2005) expressing fungal (e.g. glucoamylase (Verdoes et al. 1993), propyl aminopeptidase (Matsushita-Morita et al. 2009)) or mammalian (e.g. human interleukin-6 (Contreras et al. 1991), antitrypsin (Chill et al. 2009), lactoferrin (Ward et al. 1995), bovine chymosin (Ward et al. 1990, Cardoza et al. 2003)) proteins of industrial and clinical interest using filamentous fungi as hosts. However, the expression levels of mammalian proteins expressed in *Aspergillus* and *Trichoderma* species are low compared to homologous proteins. Significant advances in heterologous protein expression have dramatically improved the expression efficiency by fusing the heterologous gene to the 3'-end of a highly expressed homologous gene (mainly glucoamylase). Even so, limitations in protein folding, post-translational modifications, translocation and secretion as well as secretion of extracellular proteases could pose a significant hindrance for the production of bioactive proteins (Gouka et al. 1997, Nevalainen et al. 2005, Lubertozzi and Keasling 2009).

Insect Cells

Recombinant baculoviruses are widely used as a vector for the expression of recombinant proteins in insect cells (Hitchman et al. 2009, Jarvis 2009, Trometer and Falson 2010), such as enzymes (Zhao et al. 2010), immunoglobulins (Iizuka et al. 2009), viral antigens (Takahashi et al. 2010) and transcription factors (Fabian et al. 1998). The recombinant genes are usually expressed under the control of the polyhedrin or *p10* promoter of the *Autographa californica* nuclear polyhedrosis virus (AcNPV) in cultured insect cells of *Spodoptera frugiperda* (Sf9 cells) or in insect larvae of *Lepidopteran* species infected with the recombinant baculovirus containing the gene of interest. The polyhedron and *p10* genes possess very strong promoters and are highly transcribed during the late stages of the viral cycle. Usually, the recombinant proteins are recovered from the infected insect cells in soluble form and targeted in the proper cellular environment (membrane, nucleus and cytoplasm). Insect cells have many post-translational modification, transport and processing mechanisms found in higher eukaryotic cells (Durocher and Butler 2009), although their glycosylation efficiency is limited and they are not able to process complex-type oligosaccharides containing fucose, galactose and sialic acid.

Dictyostelium discoideum

Recently, the cellular slime mould *Dictyostelium discoideum*, a well-studied single-celled organism, has emerged as a promising eukaryotic alternative system (Arya et al. 2008b) for the expression of recombinant proteins (e.g. human antithrombin III; Dingermann et al. 1991, Tiltscher and Storr 1993, Dittrich et al. 1994) and enzymes (e.g. phosphodiesterase; Arya et al.

2008a). Its advantage over other expression systems lies in its extensive post-translational modification system (glycosylation, phosphorylation, acylation), which resembles that of higher eukaryotes (Jung and Williams 1997, Jung et al. 1997, Slade et al. 1997). It is a simple organism with a haploid genome of 5×10^7 bp and a life cycle that alternates between single-celled and multicellular stages. Recombinant proteins are expressed from extrachromosomal plasmids (*Dictyostelium discoideum* is one of a few eukaryotes that have circular nuclear plasmids) rather than being integrated in the genome (Ahern et al. 1988). The nuclear plasmids can be easily genetically manipulated and isolated in a one-step procedure, as in bacteria. This system is ideally suited for the expression of complex glycoproteins and although it retains many of the advantages of the bacterial (low cultivation cost) and mammalian systems (establishment of stable cell lines, glycosylation), the development of this system at an industrial scale is hampered by the relatively low productivity, compared to bacterial systems.

Trypanosomatid Protozoa

A newly developed eukaryotic expression system is based on the protozoan lizard parasites of the *Leishmania* and *Trypanosoma* species (Basile and Peticca 2009). Its gene and protein regulation and editing mechanisms are remarkably similar to those of higher eukaryotes and include the capability of 'mammalian-like' glycosylation. It has a very rapid doubling time and can be grown to high densities in relatively inexpensive medium. The recombinant gene is integrated into the small ribosomal subunit rRNA gene and can be expressed to high levels. Increased expression levels and additional promoter control can be achieved in T7 polymerase-expressing strains. Being a lizard parasite, it is not pathogenic to humans, which makes this system invaluable and highly versatile. Proteins and enzymes of significant interest, such as human tissue plasminogen activator (Soleimani et al. 2007), EPO (Breitling et al. 2002), IFN γ (Tobin et al. 1993) and IL-2 (La Flamme 1995), have been successfully expressed in this system.

Transgenic Plants

The current protein therapeutics market is clearly an area of enormous interest from a medical and economic point of view. Recent advances in human genomics and biotechnology have made it possible to identify a plethora of potentially important drugs or drug targets. Transgenic technology has provided an alternative, more cost-effective bioproduction system than the previously used (*E. coli*, yeast, mammalian cells; Larrick and Thomas 2001, Demain and Vaishnav 2009). The accumulated knowledge on plant genetic manipulation has been recently applied to the development of plant bioproduction systems (Twyman et al. 2005, Boehm 2007, Lienard et al. 2007, Faye and Champey 2008, Sourrouille et al. 2009). Expression in plants could be either constitutive or transient and directed to a specific tissue of the plant (depending on the type of promoter used). Expression of heterologous proteins in plants offers significant advantages, such as low production cost, high biomass

production, unlimited supply and ease of expandability. Plants also have high-fidelity expression, folding and post-translational modification mechanisms, which could produce human proteins of substantial structural and functional equivalency compared to proteins from mammalian expression systems (Gomord and Faye 2004, Joshi and Lopez 2005). Additionally, plant-made human proteins of clinical interest (Schillberg et al. 2005, Twyman et al. 2005, Vitale and Pedrazzini 2005, Tiwari et al. 2009), such as antibodies (Hassan et al. 2008, Ko et al. 2009, De Muynck et al. 2010, Lai et al. 2010), vaccines (Hooper 2009, Alvarez and Cardineau 2010), cytokines (Elias-Lopez et al. 2008) and enzymes (Kermode et al. 2007, Stein et al. 2009), are free of potentially hazardous human diseases, viruses or bacterial toxins. However, there is considerable concern regarding the potential hazards of contamination of the natural gene pool by the transgenes and possible additional safety precautions will raise the production cost.

Transgenic Animals

Besides plants, transgenic technology has also been applied to many different species of animals (mice, cows, rabbits, sheep, goats and pigs; Niemann and Kues 2007, Houdebine 2009). The DNA containing the gene of interest is microinjected into the pronucleus of a single-cell fertilised zygote and integrated into the genome of the recipient; therefore, it can be faithfully passed on from generation to generation. The gene of interest is coupled with a signal targeting protein expression towards specific tissues, mainly the mammary gland, and the protein can therefore be harvested and purified from milk. The proteins produced by transgenic animals are almost identical to human proteins, greatly expanding the applications of transgenic animals in medicine and biotechnology. Several human protein of pharmaceutical value have been produced in transgenic animals, such as haemoglobin (Swanson 1992, Logan and Martin 1994), lactoferrin (Han et al. 2008, Yang et al. 2008), antithrombin III (Yang et al. 2009), protein C (Velander 1992) and fibrinogen (Prunkard 1996), and there is enormous interest for the generation of transgenic tissues suitable for transplantation in humans (only recently overshadowed by primary blastocyte technology (Klimanskaya et al. 2008)). Despite the initial technological expertise required to produce a transgenic animal, the subsequent operational costs are low and subsequent inbreeding ensures that the ability to produce the transgenic protein will be passed on to its offspring. However, certain safety issues have arisen concerning the potential contamination of transgenically produced proteins by animal viruses or prions, which could possibly be passed on to the human population. Extensive testing required by the FDA substantially raises downstream costs.

ENZYME PURIFICATION

Once a suitable enzyme source has been identified, it becomes necessary to design an appropriate purification procedure to isolate the desired protein. The extent of purification required for an enzyme depends on several factors, the most important of which being the degree of enzyme purity required as well as

Table 7.5. Protein Properties Used During Purification

Protein Property	Technique
Solubility	Precipitation
Size	Gel filtration
Charge	Ion exchange
Hydrophobicity	Hydrophobic interaction chromatography
Biorecognition	Affinity chromatography

the starting material, for example the quantity of the desired enzyme present in the initial preparation (Lesley 2001, Labrou and Clonis 2002). For example, industrial enzymes are usually produced as relatively crude preparations. Enzymes used for therapeutic or diagnostic purposes are generally subjected to the most stringent purification procedures, as the presence of compounds other than the intended product may have an adverse clinical impact (Berthold and Walter 1994).

Purification of an enzyme usually occurs by a series of independent steps in which the various physicochemical properties of the enzyme of interest are utilised to separate it progressively from other unwanted constituents (Labrou and Clonis 2002, Labrou et al. 2004b). The characteristics of proteins that are utilised in purification include solubility, ionic charge, molecular size, adsorption properties and binding affinity to other biological molecules. Several methods that exploit differences in these properties are listed in Table 7.5.

Precipitation methods (usually employing $(\text{NH}_4)_2\text{SO}_4$, polyethyleneglycol or organic solvents) are not very efficient method of purification (Labrou and Clonis 2002). They typically give only a few fold purification. However, with these methods, the protein may be removed from the growth medium or from cell debris where harmful proteases and other detrimental compounds may affect protein stability. On the other hand, *chromatography* is a highly selective separation technique (Regnier 1987, Fausnaugh 1990). A wide range of chromatographic techniques has been used for enzyme purification: size-exclusion chromatography, ion-exchange, hydroxyapatite, hydrophobic interaction chromatography, reverse-phase chromatography and affinity chromatography (Labrou 2003). Of these, ion-exchange and affinity chromatography are the most common and probably the most important (Labrou and Clonis 1994).

Ion-Exchange Chromatography

Ion-exchange resins selectively bind proteins of opposite charge; that is, a negatively charged resin will bind proteins with a net positive charge and vice versa (Fig. 7.13). Charged groups are classified according to type (cationic and anionic) and strength (strong or weak); the charge characteristics of strong ion exchange media do not change with pH, whereas with weak ion-exchange media they do. The most commonly used charged groups include diethylaminoethyl, a weakly anionic exchanger; carboxymethyl, a weakly cationic exchanger; quaternary ammonium, a strongly anionic exchanger; and methyl sulfonate,

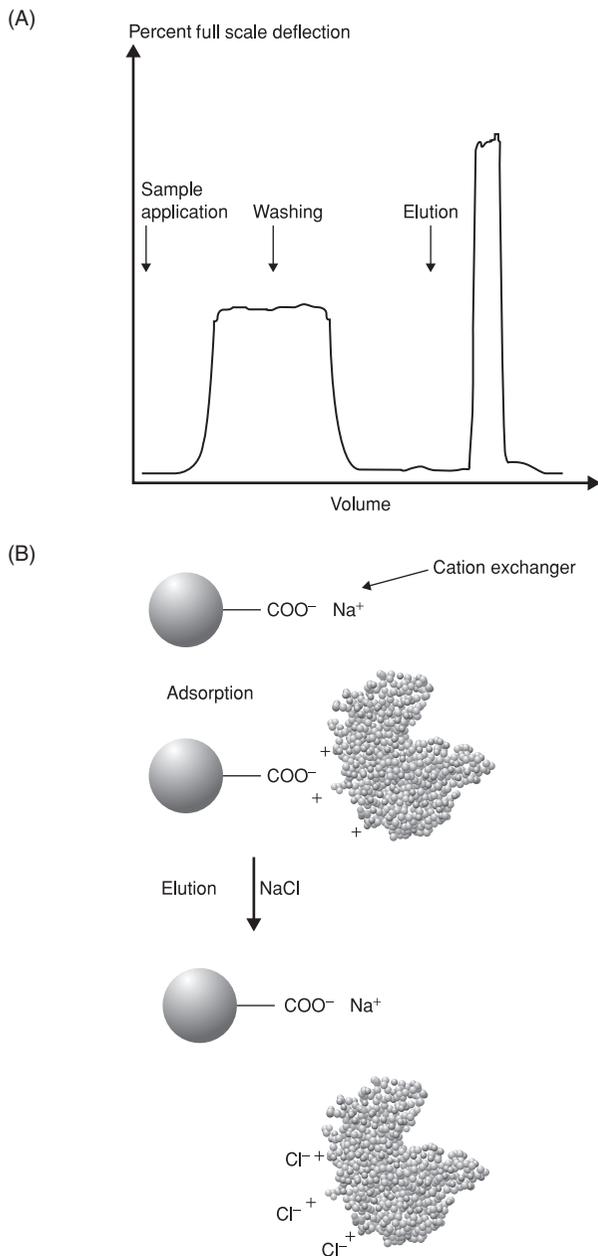


Figure 7.13. (A) Schematic diagram of a chromatogram showing the steps for a putative purification. (B) Schematic diagram depicting the principle of ion-exchange chromatography.

a strongly cationic exchanger (Table 7.6; Levison 2003). The matrix material for the column is usually formed from beads of carbohydrate polymers, such as agarose, cellulose or dextrans (Levison 2003).

The technique takes place in five steps (Labrou 2000; Fig. 7.13): *equilibration* of the column to pH and ionic strength conditions suitable for target protein binding; protein *sample application* to the column and reversible adsorption through

counter-ion displacement; *washing* of the unbound contaminating proteins, enzymes, nucleic acids and other compounds; introduction of *elution* conditions in order to displace bound proteins; and *regeneration* and *re-equilibration* of the adsorbent for subsequent purifications. Elution may be achieved either by increasing the salt concentration or by changing the pH of the irrigating buffer. Both methods are used in industry, but raising the salt concentration is by far the most common because it is easier to control (Levison 2003). Most protein purifications are done on anion exchange columns because most proteins are negatively charged at physiological pH values (pH 6–8).

Affinity Chromatography

Affinity chromatography is potentially the most powerful and selective method for protein purification (Fig. 7.14; Labrou and Clonis 1994, Labrou 2003). According to the International Union of Pure and Applied Chemistry, affinity chromatography is defined as a liquid chromatographic technique that makes use of a ‘biological interaction’ for the separation and analysis of specific analytes within a sample. Examples of these interactions include the binding of an enzyme with a substrate/inhibitor or of an antibody with an antigen or in general the interaction of a protein with a binding agent, known as the ‘affinity ligand’ (Fig. 7.14; Labrou 2002, 2003, Labrou et al. 2004b). The development of an affinity chromatography-based purification step involves the consideration of the following factors: (i) selection of an appropriate ligand and (ii) immobilisation of the ligand onto a suitable support matrix to make an *affinity adsorbent*. The selection of the immobilised ligand for affinity chromatography is the most challenging aspect of preparing an affinity adsorbent. Certain factors need to be considered when selecting a ligand (Labrou and Clonis 1995, 1996): (i) the specificity of the ligand for the protein of interest, (ii) the reversibility of the interaction with the protein, (iii) its stability against the biological and chemical operation conditions and (iv) the affinity of the ligand for the protein of interest. The binding site of a protein is often located deep within the molecule and adsorbents prepared by coupling the ligands directly to the support exhibit low binding capacities. This is due to steric interference between the support matrix and the protein’s binding site. In these circumstances, a ‘spacer arm’ is inserted between the matrix and ligand to facilitate effective binding (Fig. 7.14). A hexyl spacer is usually inserted between ligand and support by substitution of 1,6-diaminohexane (Lowe 2001).

The ideal matrix should be hydrophilic, chemically and biologically stable and have sufficient modifiable groups to permit an appropriate degree of substitution with the enzyme. Sepharose is the most commonly used matrix for affinity chromatography on the research scale. Sepharose is a commercially available beaded polymer, which is highly hydrophilic and generally inert to microbiological attack (Labrou and Clonis 2002). Chemically, it is an agarose (poly- $\{\beta\text{-1,3-D-galactose-}\alpha\text{-1,4-(3,6-anhydro)-L-galactose}\}$) derivative.

The selection of conditions for an optimum affinity chromatographic purification involves the study of the following factors: (1) choice of adsorption conditions (e.g. buffer composition,

Table 7.6. Functional Groups Used on Ion Exchangers

	Functional Group
<i>Anion exchangers</i>	
Diethylaminoethyl (DEAE)	$-\text{O}-\text{CH}_2-\text{CH}_2-\text{N}^+\text{H}(\text{CH}_2\text{CH}_3)_2$
Quaternary aminoethyl (QAE)	$-\text{O}-\text{CH}_2-\text{CH}_2-\text{N}^+(\text{C}_2\text{H}_5)_2-\text{CH}_2-\text{CHOH}-\text{CH}_3$
Quaternary ammonium (Q)	$-\text{O}-\text{CH}_2-\text{CHOH}-\text{CH}_2-\text{O}-\text{CH}_2-\text{CHOH}-\text{CH}_2\text{N}^+(\text{CH}_3)_2$
<i>Cation exchangers</i>	
Carboxymethyl (CM)	$-\text{O}-\text{CH}_2-\text{COO}^-$
Sulphopropyl (SP)	$-\text{O}-\text{CH}_2-\text{CHOH}-\text{CH}_2-\text{O}-\text{CH}_2-\text{CH}_2-\text{CH}_2\text{SO}_3^-$

pH, ionic strength) to maximise the conditions required for the formation of strong complex between the ligand and the protein to be purified, (2) choice of washing conditions to desorb non-specifically bound proteins and (3) choice of elution conditions to maximise purification (Labrou and Clonis 1995). The elution conditions of the bound macromolecule should be both

tolerated by the affinity adsorbent and effective in desorbing the biomolecule in good yield and in the native state. Elution of bound proteins is performed in a non-specific or biospecific manner. Non-specific elution usually involves (1) changing the ionic strength (usually by increasing the buffer's molarity or including salt, e.g. KCl or NaCl) and the pH (adsorption generally weakens with increasing pH), (2) altering the polarity of the irrigating buffer by employing, for example ethylene glycol or other organic solvents, if the hydrophobic contribution in the protein-ligand complex is large. Biospecific elution is achieved by inclusion in the equilibration buffer of a suitable ligand, which usually competes with the immobilised ligand for the same binding site on the enzyme/protein (Labrou 2000). Any competing ligand may be used. For example, substrates, products, cofactors, inhibitors or allosteric effectors are all potential candidates as long as they have higher affinity for the macromolecule than the immobilised ligand.

Dye-ligand affinity chromatography represents a powerful affinity-based technique for enzyme and protein purification (Clonis et al. 2000, Labrou 2002, Labrou et al. 2004b). The technique has gained broad popularity due to its simplicity and wide applicability to purify a variety of proteins. The employed dyes as affinity ligands are commercial textile chlorotriazine polysulfonated aromatic molecules, which are usually termed as triazine dyes (Fig. 7.15). Such dye-ligands have found wide applications over the past 20 years as general affinity ligands in the research market to purify enzymes, such as oxidoreductases, decarboxylases, glycolytic enzymes, nucleases, hydrolases, lyases, synthetases and transferases (Scopes 1987). Anthraquinone triazine dyes are probably the most widely used dye-ligands in enzyme and protein purification. Especially the triazine dye Cibacron Blue F3GA (Fig. 7.18) has been widely exploited as an affinity chromatographic tool to separate and purify a variety of proteins (Scopes 1987). With the aim of increasing the specificity of dye-ligands, the biomimetic dye-ligand concept was introduced. According to this concept, new dyes that mimic natural ligands of the targeted proteins are designed by substituting the terminal 2-aminobenzene sulfonate moiety of the dye Cibacron Blue 3GA (CB3GA) for (η) with a substrate-mimetic moiety (Clonis et al. 2000, Labrou 2002, 2003, Labrou et al. 2004b). These biomimetic dyes exhibit increased purification ability and specificity and provide useful tools for designing simple and effective purification protocols.

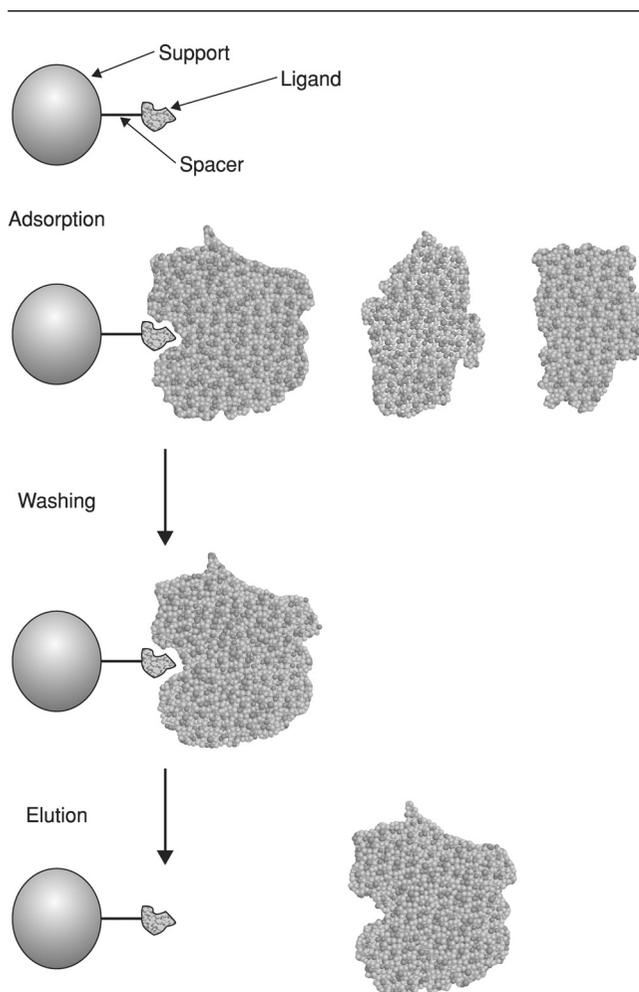


Figure 7.14. Schematic diagram depicting the principle of affinity chromatography.

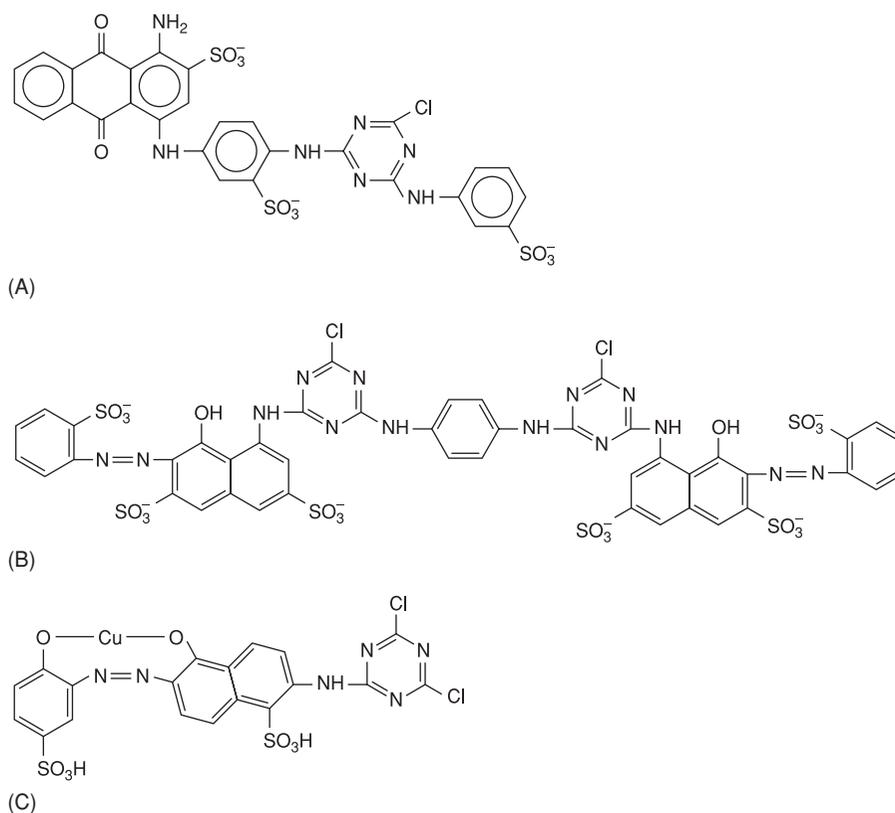


Figure 7.15. Structure of several representative triazine dyes: **(A)** Cibacron Blue 3GA, **(B)** Procion Red HE-3B and **(C)** Procion Rubine MX-B.

The rapid development of recombinant DNA technology since the early 1980s has changed the emphasis of a classical enzyme purification work. For example, epitope tagging is a recombinant DNA method for inserting a specific protein sequence (affinity tag) into the protein of interest (Terpe 2003). This allows the expressed tagged protein to be purified by affinity interactions with a ligand that selectively binds to the affinity tag. Examples of affinity tags and their respective ligands used for protein and enzyme purification are shown in Table 7.7.

ENZYME ENGINEERING

Another extremely promising area of enzyme technology is enzyme engineering. New enzyme structures may be designed and produced in order to improve existing ones or create new activities. Over the past two decades, with the advent of protein engineering, molecular biotechnology has permitted not only the improvement of the properties of these isolated proteins, but also the construction of 'altered versions' of these 'naturally occurring' proteins with novel or 'tailor-made' properties (Ryu

Table 7.7. Adsorbents and Elution Conditions of Affinity Tags

Affinity Tag	Matrix	Elution Condition
Poly-His	Ni ²⁺ -NTA	Imidazole 20–250 mM or low pH
FLAG	Anti-FLAG monoclonal antibody	pH 3.0 or 2–5 mM EDTA
Strep-tag II	Strep-Tactin (modified streptavidin)	2.5 mM desthiobiotin
c-myc	Monoclonal antibody	Low pH
S	S-fragment of RNaseA	3 M guanidine thiocyanate, 0.2 M citrate pH 2, 3 M magnesium chloride
Calmodulin-binding peptide	Calmodulin	EGTA or EGTA with 1 M NaCl
Cellulose-binding domain	Cellulose	Guanidine HCl or urea > 4 M
Glutathione S-transferase	Glutathione	5–10 mM reduced glutathione

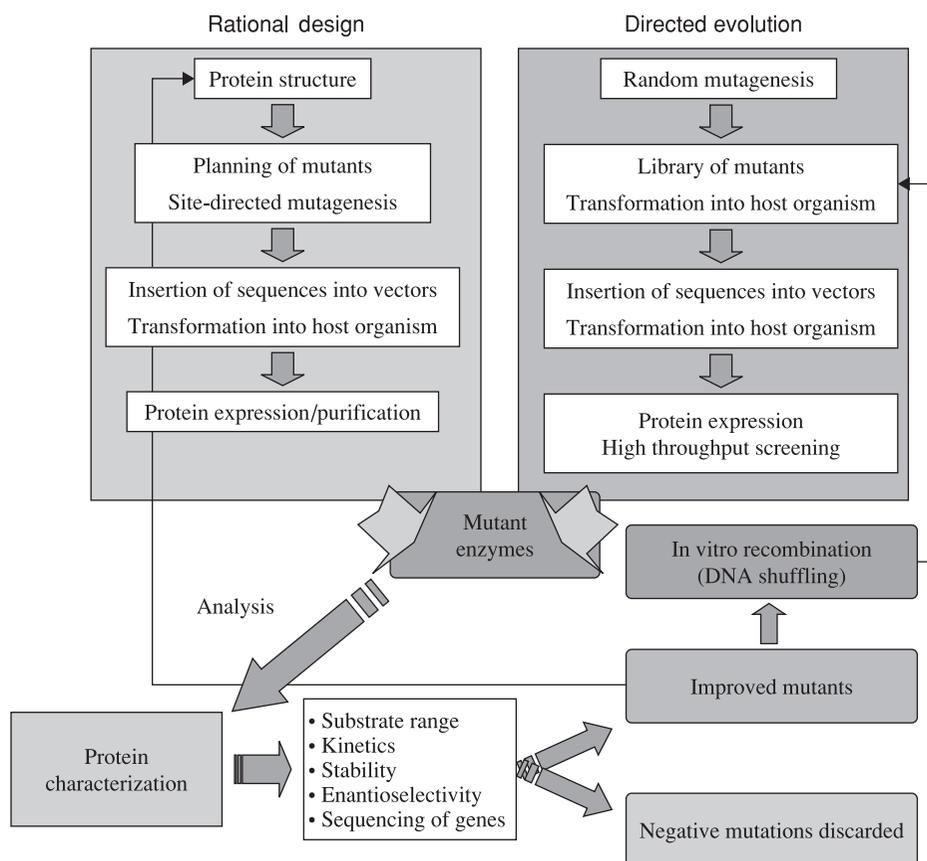


Figure 7.16. Comparison of rational design and directed evolution.

and Nam 2000, Gerlt and Babbitt 2009, Tracewell and Arnold 2009).

TAILOR-MADE ENZYMES BY PROTEIN ENGINEERING

There are two main intervention approaches for the construction of tailor-made enzymes: rational design and directed evolution (Chen 2001, Schmidt et al. 2009; Fig. 7.16).

Rational design takes advantage of knowledge of the three-dimensional structure of the enzyme, as well as structure/function and sequence information to predict, in a ‘rational/logical’ way, sites on the enzyme that when altered would endow the enzyme with the desired properties (Craik et al. 1985, Wells et al. 1987, Carter et al. 1989, Scrutton et al. 1990, Cedrone et al. 2000). Once the crucial amino acids are identified, site-directed mutagenesis is applied and the expressed mutants are screened for the desired properties. It is clear that protein engineering by rational design requires prior knowledge of the ‘hot spots’ on the enzyme. Directed evolution (or molecular evolution) does not require such prior sequence or three-dimensional structure knowledge, as it usually employs random-mutagenesis protocols to engineer enzymes that are subsequently screened for the desired properties (Tao and Cornish 2002, Dalby 2003,

Jaeger and Eggert 2004, Jestin and Kaminski 2004, Williams et al. 2004). However, both approaches require efficient expression as well as sensitive detection systems for the protein of interest (Kotzia et al. 2006). During the selection process, the mutations that have a positive effect are selected and identified. Usually, repeated rounds of mutagenesis are applied until enzymes with the desired properties are constructed. For example, it took four rounds of random mutagenesis and DNA shuffling of *Drosophila melanogaster* 2'-deoxynucleoside kinase, followed by FACS analysis, in order to yield an orthogonal ddT kinase with a 6-fold higher activity for the nucleoside analogue and a 20-fold k_{cat}/K_m preference for ddT over thymidine, an overall 10,000-fold change in substrate specificity (Liu et al. 2009b).

Usually, a combination of both methods is employed by the construction of combinatorial libraries of variants, using random mutagenesis on selected (by rational design) areas of the parental ‘wild-type’ protein (typically, binding surfaces or specific amino acids; Altamirano et al. 2000, Arnold 2001, Saven 2002, Johannes and Zhao 2006). For example, Park et al. rationally manipulated several active site loops in the ab/ba metallohydrolase scaffold of glyoxalase II through amino acid insertion, deletion, and substitution, and then used directed evolution to introduce random point-mutations to fine-tune the enzyme

activity (Park et al. 2006). The resulting enzyme completely lost its original activity and instead showed β -lactamase activity.

The industrial applications of enzymes as biocatalysts are numerous. Recent advances in genetic engineering have made possible the construction of enzymes with enhanced or altered properties (change of enzyme/cofactor specificity and enantioselectivity, altered thermostability, increased activity) to satisfy the ever-increasing needs of the industry for more efficient catalysts (Bornscheuer and Pohl 2001, Zaks 2001, Jaeger and Eggert 2004, Chaput et al. 2008, Zeng et al. 2009).

RATIONAL ENZYME DESIGN

The rational protein design approach is mainly used for the identification and evaluation of functionally important residues or sites in proteins. Although the protein sequence contains all the information required for protein folding and functions, today's state of technology does not allow for efficient protein design by simple knowledge of the amino acid sequence alone. For example, there are 10^{325} ways of rearranging amino acids in a 250-amino-acid-long protein, and prediction of the number of changes required to achieve a desired effect is an obstacle that initially appears impossible. For this reason, a successful rational design cycle requires substantial planning and could be repeated several times before the desired result is achieved. A rational protein design cycle requires the following:

1. Knowledge of the amino acid sequence of the enzyme of interest and availability of an expression system that allows for the production of active enzyme. Isolation and characterisation (annotation) of cDNAs encoding proteins with novel or pre-observed properties has been significantly facilitated by advances in genomics (Skena et al. 1995, Zweiger and Scott 1997, Skena et al. 1998, Carbone 2009) and proteomics (Anderson and Anderson 1998, Anderson et al. 2000, Steiner and Anderson 2000, Xie et al. 2009) and is increasing rapidly. These cDNA sequences are stored in gene (NCBI) and protein databanks (UniProtKB/Swiss-Prot; Release 57.12 of 15 Dec 09 of UniProtKB/Swiss-Prot contains 513,77 protein sequence entries; Apweiler et al. 2004, The UniProt Consortium 2008). However, before the protein design cycle begins, a protein expression system has to be established. Introduction of the cDNA encoding the protein of interest into a suitable expression vector/host cell system is nowadays a standard procedure (see above).
2. Structure/function analysis of the initial protein sequence and determination of the required amino acid changes. As mentioned before, the enzyme engineering process could be repeated several times until the desired result is obtained. Therefore, each cycle ends where the next begins. Although, we cannot accurately predict the conformation of a given protein by knowledge of its amino acid sequence, the amino acid sequence can provide significant information. Initial screening should therefore involve sequence comparison analysis of the original protein sequence to other sequence homologous proteins with

potentially similar functions by utilising current bioinformatics tools (Andrade and Sander 1997, Fenyo and Beavis 2002, Nam et al. 2009, Yen et al. 2009, Zhang et al. 2009a, 2009b). Areas of conserved or non-conserved amino acid residues can be located within the protein and could possibly provide valuable information, concerning the identification of binding and catalytic residues. Additionally, such methods could also reveal information pertinent to the three-dimensional structure of the protein.

3. Availability of functional assays for identification of changes in the properties of the protein. This is probably the most basic requirement for efficient rational protein design. The expressed protein has to be produced in a bioactive form and characterised for size, function and stability in order to build a baseline comparison platform for the ensuing protein mutants. The functional assays should have the required sensitivity and accuracy to detect the desired changes in the protein's properties.
4. Availability of the three-dimensional structure of the protein or capability of producing a reasonably accurate three-dimensional model by computer modelling techniques. The structures of thousands of proteins have been solved by various crystallographic techniques (X-ray diffraction, NMR spectroscopy) and are available in protein structure databanks. Current bioinformatics tools and elaborate molecular modeling software (Wilkins et al. 1999, Gasteiger et al. 2003, Guex et al. 2009) permit the accurate depiction of these structures and allow the manipulation of the amino acid sequence. For example, they are able to predict, with significant accuracy, the consequences of a single amino acid substitution on the conformation, electrostatic or hydrophobic potential of the protein (Guex and Peitsch 1997, Gasteiger et al. 2003, Schwede et al. 2003). Additionally, protein–ligand interactions can, in some cases, be successfully simulated, which is especially important in the identification of functionally important residues in enzyme–cofactor/substrate interactions (Saxena et al. 2009). Finally, in allosteric regulation, the induced conformational changes are very difficult to predict. In last few years, studies on the computational modelling of allostery have also begun (Kidd et al. 2009).

Where the three-dimensional structure of the protein of interest is not available, computer modelling methods (homology modelling, fold recognition using threading and *ab initio* prediction) allow for the construction of putative models based on known structures of homologous proteins (Schwede et al. 2003, Kopp and Schwede 2004, Jaroszewski 2009, Qu et al. 2009). Additionally, comparison with proteins having homologous three-dimensional structure or structural motifs could provide clues as to the function of the protein and the location of functionally important sites. Even if the protein of interest shows no homology to any other known protein, current amino acid sequence analysis software could provide putative tertiary structural models. A generalised approach to predict protein structure is shown in Figure 7.17.

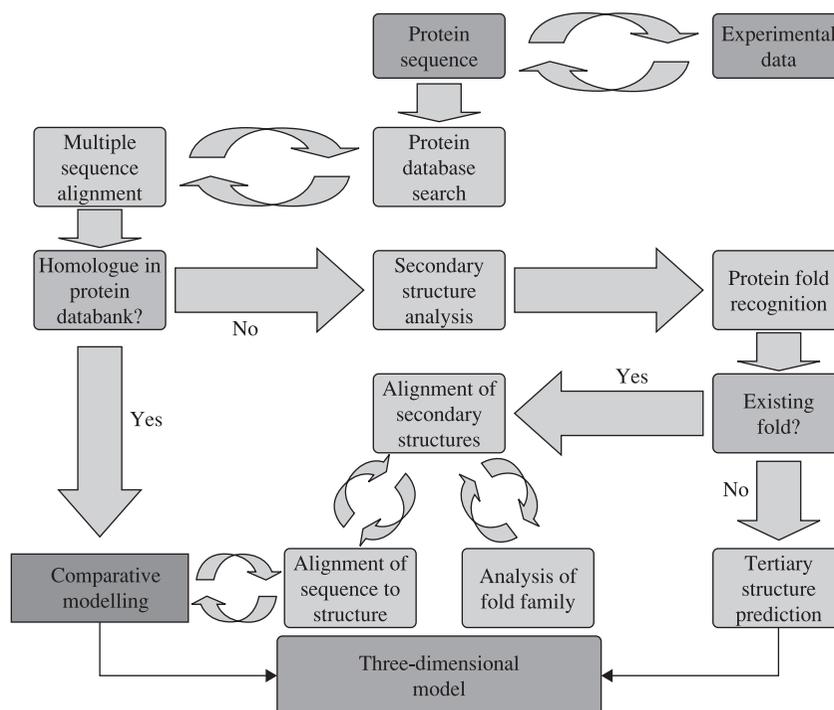


Figure 7.17. A generalised schematic for the prediction of protein three-dimensional structure.

5. Genetic manipulation of the wild-type nucleotide sequence. A combination of previously published experimental literature and sequence/structure analysis information is usually necessary for the identification of functionally important sites in the protein. Once an adequate three-dimensional structural model of the protein of interest has been constructed, manipulation of the gene of interest is necessary for the construction of mutants. Polymerase chain reaction (PCR) mutagenesis is the basic tool for the genetic manipulation of the nucleotide sequences. The genetically redesigned proteins are engineered by the following:

- a. Site-directed mutagenesis: alteration of specific amino acid residues. There are a number of experimental approaches designed for this purpose. The basic principle involves the use of synthetic oligonucleotides (oligonucleotide-directed mutagenesis) that are complementary to the cloned gene of interest but contain a single (or sometimes multiple) mismatched base(s) (Balland et al. 1985, Garvey and Matthews 1990, Wagner and Benkovic 1990). The cloned gene is either carried by a single-stranded vector (M13 oligonucleotide-directed mutagenesis) or a plasmid that is later denatured by alkali (plasmid DNA oligonucleotide-directed mutagenesis) or heat (PCR-amplified oligonucleotide-directed mutagenesis) in order for the mismatched oligonucleotide to anneal.

The latter then serves as a primer for DNA synthesis catalysed by externally added DNA polymerase for the creation of a copy of the entire vector, carrying, however, a mutated base. PCR mutagenesis is the most frequently used mutagenesis method (Fig. 7.18). For example, substitution of specific amino acid positions by site-directed mutagenesis (S67D/H68D) successfully converted the coenzyme specificity of the short-chain carbonyl reductase from NADP(H) to NAD(H) as well as the product enantioselectivity without disturbing enzyme stability (Zhang et al. 2009). In another example, engineering of the maize GSTF1-1 by mutating selected G-site residues resulted in substantial changes in the pH-dependence of kinetic parameters of the enzyme (Labrou et al. 2004a). Mutation of a key residue in the H-site of the same enzyme (Ile118Phe) led to a fourfold improved specificity of the enzyme towards the herbicide alachlor (Labrou et al. 2005).

So far, substitution of a specific amino acid by another has been limited by the availability of only 20 naturally occurring amino acids. However, it is chemically possible to construct hundreds of designer-made amino acids. Incorporation of these novel protein building blocks could help shed new light into the cellular and protein functions (Wang and Schultz 2002, Chin et al. 2003, Deiters et al. 2003, Arnold 2009).

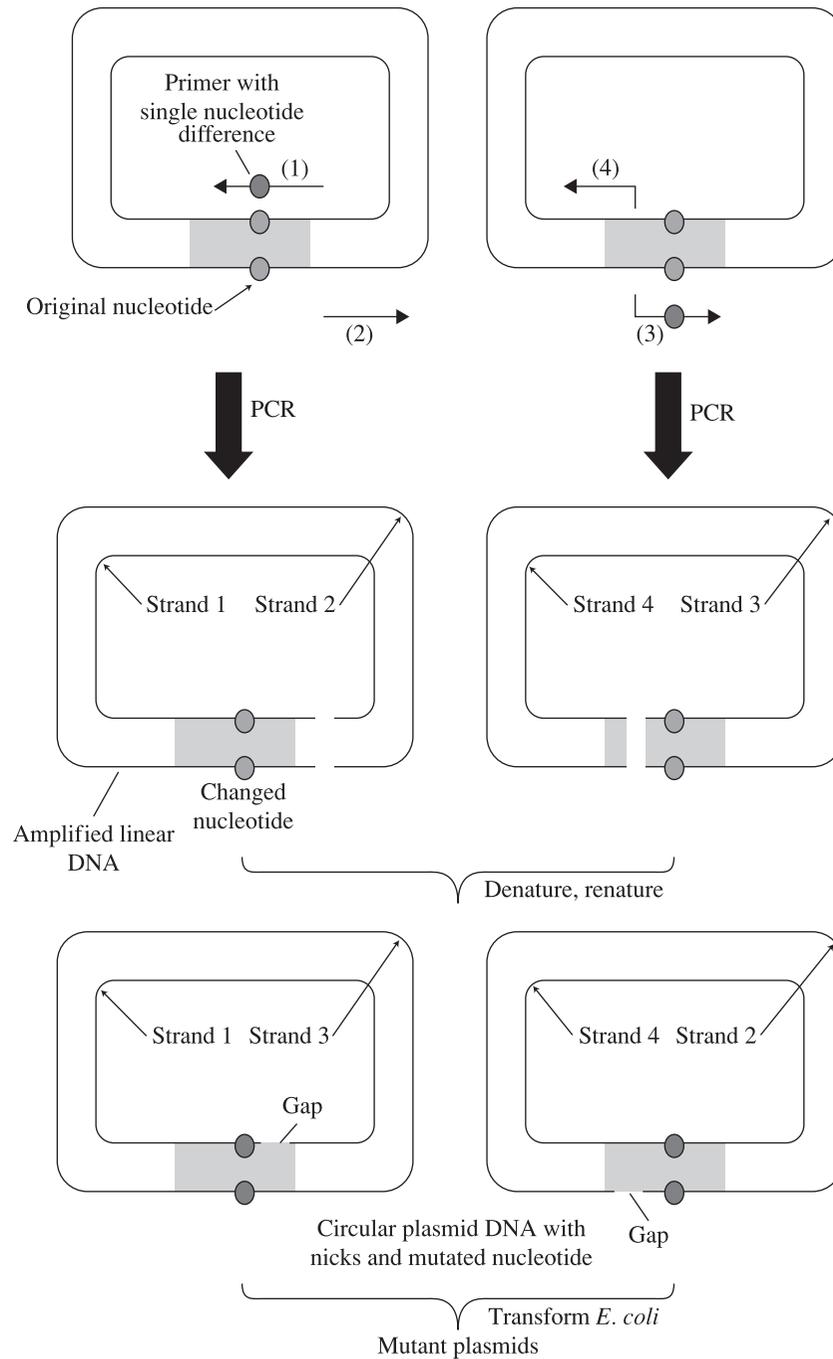


Figure 7.18. PCR oligonucleotide-directed mutagenesis. Two sets of primers are used for the amplification of the double-stranded plasmid DNA. The primers are positioned as shown and only one contains the desired base change. After the initial PCR step, the amplified PCR products are mixed together, denatured and renatured to form, along with the original amplified linear DNA, nicked circular plasmids containing the mutations. Upon transformation into *E. coli*, the nicked are repaired by host cell enzymes and the circular plasmids can be maintained.

- b. Construction of deletion mutants: deletion of specified areas within or at the 5'/3' ends (truncation mutants) of the gene.
- c. Construction of insertion/fusion mutants: insertion of a functionally/structurally important epitope or fusion to another protein fragment.

There are numerous examples of fusion proteins designed to facilitate protein expression and purification, display of proteins on surfaces of cells or phages, cellular localisation, metabolic engineering as well as protein–protein interaction studies (Nixon et al. 1998).

- d. Domain swapping: exchanging of protein domains between homologous or heterologous proteins. For example, exchange of a homologous region between *Agrobacterium tumefaciens* β -glucosidase (optimum at pH 7.2–7.4 and 60°C) and *Cellvibrio gilvus* β -glucosidase (optimum at pH 6.2–6.4 and 35°C) resulted in a hybrid enzyme with optimal activity at pH 6.6–7.0 and 45–50°C (Singh et al. 1995). Also, domain swapping was used to clarify the control of electron transfer in nitric-oxide synthases (Nishida and Ortiz de Montellano 2001). In another example, domain swapping was observed in the structurally unrelated capsid of a rice yellow mottle virus, a member of the plant icosahedral virus group, where it was demonstrated to increase stability of the viral particle (Qu et al. 2000).

Although site-directed mutagenesis is widely used, it is not always feasible due to the limited knowledge of protein structure–function relationship and the approximate nature of computer-graphic modelling. In addition, rational design approaches can fail due to unexpected influences exerted by the substitution of one or more amino acid residues (Cherry and Fidantsef 2003, Johannes and Zhao 2006). Irrational approaches can therefore be preferable alternatives to direct the evolution of enzymes with highly specialised traits (Hibbert and Dalby 2005, Chatterjee and Yuan 2006, Johannes and Zhao 2006).

DIRECTED ENZYME EVOLUTION

Directed evolution by DNA recombination can be described as a mature technology for accelerating protein evolution. Evolution is a powerful algorithm with proven ability to alter enzyme function and especially to ‘tune’ enzyme properties (Cherry and Fidantsef 2003, Williams et al. 2004, Hibbert and Dalby 2005, Roodveldt et al. 2005, Chatterjee and Yuan 2006). The methods of directed evolution use the process of natural selection but in a directed way (Altreuter and Clark 1999, Kaur and Sharma 2006, Wong et al. 2006, Glasner et al. 2007, Gerlt and Babbitt 2009, Turner 2009). The major step in a typical directed enzyme evolution experiment is first to make a set of mutants and then to find the best variants through a high-throughput selection or screening procedure (Kotzia et al. 2006). The process can be iterative, so that a ‘generation’ of molecules can be created in a few weeks or even in a few days, with large numbers of progeny subjected to selective pressures not encountered in nature (Arnold 2001, Williams et al. 2004).

There are many methods to create combinatorial libraries, using directed evolution (Labrou 2010). Some of these are random mutagenesis using mainly error-prone PCR (Ke and Madison 1997, Cirino et al. 2003), DNA shuffling (Stemmer 1994, Cramer et al. 1998, Baik et al. 2003, Bessler et al. 2003, Dixon et al. 2003, Wada et al. 2003), StEP (staggered extension process; Zhao et al. 1998, Aguinaldo and Arnold 2003), RPR (random-priming in vitro recombination; Shao et al. 1998, Aguinaldo and Arnold 2003), incremental truncation for the creation of hybrid enzymes (ITCHY; Lutz et al. 2001), RACHITT (random chimeragenesis on transient templates; Coco et al. 2001, Coco 2003), ISM (iterative saturation mutagenesis; Reetz 2007), GSSM (gene site saturation mutagenesis; DeSantis et al. 2003, Dumon et al. 2008), PDLGO (protein domain library generation by overlap extension; Gratz and Jose 2008) and DuARChEM (dual approach to random chemical mutagenesis; Mohan and Banerjee 2008). The most frequently used methods for DNA shuffling are shown in Figure 7.19.

Currently, directed evolution has gained considerable attention as a commercially important strategy for rapid design of molecules with properties tailored for the biotechnological and pharmaceutical market. Over the past four years, DNA family shuffling has been successfully used to improve enzymes of industrial and therapeutic interest (Kurtzman et al. 2001, Chiang 2004, Dai and Copley 2004, Yuan et al. 2005). For example, by applying the DNA family shuffling approach, the catalytic properties of cytochrome P450 enzymes were further extended in the chimeric progeny to include a new range of blue colour formations. Therefore, it may be possible to direct the new enzymes towards the production of new dyes (Rosic 2009).

IMMOBILISED ENZYMES

The term ‘immobilised enzymes’ describes enzymes physically confined, localised in a certain region of space or attached on a support matrix (Abdul 1993). The main advantages of enzyme immobilisation are listed in Table 7.8.

There are at least four main areas in which immobilised enzymes may find applications, that is industrial, environmental, analytical and chemotherapeutic (Powell 1984, Liang et al. 2000). Environmental applications include waste water treatment and the degradation of chemical pollutants of industrial and agricultural origin (Dravis et al. 2001). Analytical applications include biosensors. Biosensors are analytical devices, which have a biological recognition mechanism (most commonly enzyme) that transduce it into a signal, usually electrical, and can be detected by using a suitable detector (Phadke 1992). Immobilised enzymes, usually encapsulated, are also being used for their possible chemotherapeutic applications in replacing enzymes that are absent from individuals with certain genetic disorders (DeYoung 1989).

METHODS FOR IMMOBILISATION

There are a number of ways in which an enzyme may be immobilised: adsorption, covalent coupling, cross-linking, matrix

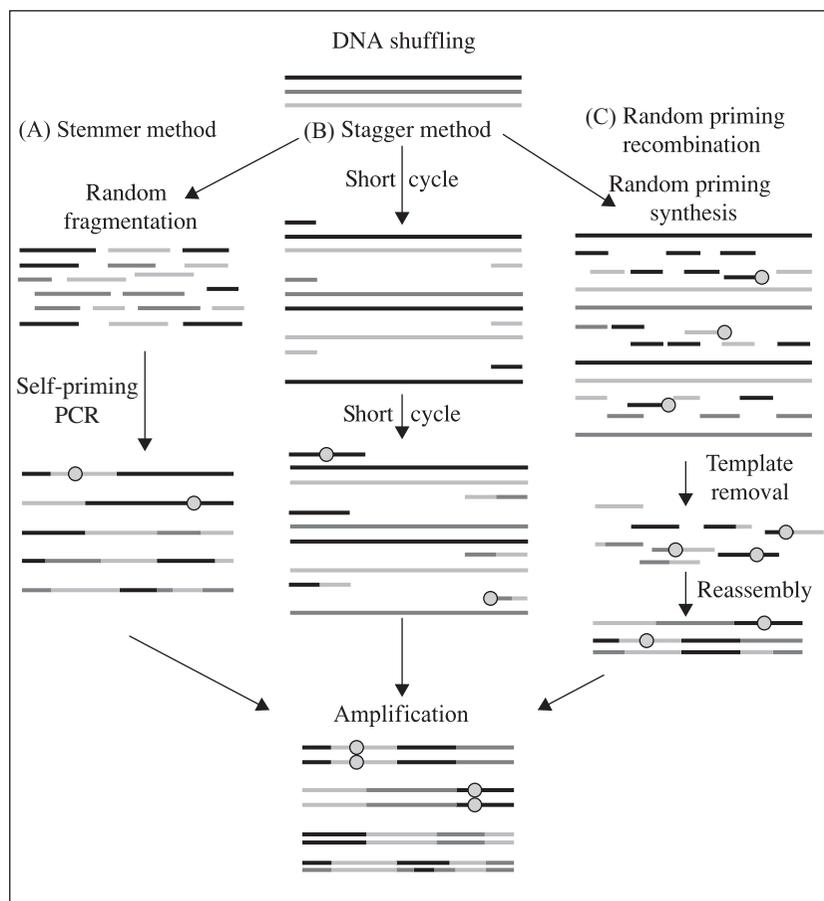


Figure 7.19. A schematic representation of the most frequently used methods for DNA shuffling.

entrapment or encapsulation (Podgornik and Tennikova 2002; Fig. 7.20). These methods will be discussed in the following sections.

Adsorption

Adsorption is the simplest method and involves reversible interactions between the enzyme and the support material (Fig. 7.20A). The driving force causing adsorption is usually the formation of several non-covalent bonds such as salt links,

van der Waals, hydrophobic and hydrogen bonding (Calleri et al. 2004). The methodology is easy to carry out and can be applied to a wide range of support matrices such as alumina, bentonite, cellulose, anion and cation exchange resins, glass, hydroxyapatite, kaolinite, etc. The procedure consists of mixing together the enzyme and a support under suitable conditions of pH, ionic strength, temperature, etc. The most significant advantages of this method are (i) absence of chemicals resulting to a little damage to enzyme and (ii) reversibility, which allows regeneration with fresh enzyme. The main disadvantage of the method is the leakage of the enzyme from the support under many conditions of changes in the pH, temperature and ionic strength. Another disadvantage is the non-specific adsorption of other proteins or other substances to the support. This may modify the properties of the support or of the immobilised enzyme.

Covalent Coupling

The covalent coupling method is achieved by the formation of a covalent bond between the enzyme and the support

Table 7.8. Advantages of Immobilised Enzymes

1. Repetitive use of a single batch of enzymes.
2. Immobilisation can improve enzyme's stability by restricting the unfolding of the protein.
3. Product is not contaminated with the enzyme. This is very important in the food and pharmaceutical industries.
4. The reaction is controlled rapidly by removing the enzyme from the reaction solution (or vice versa).

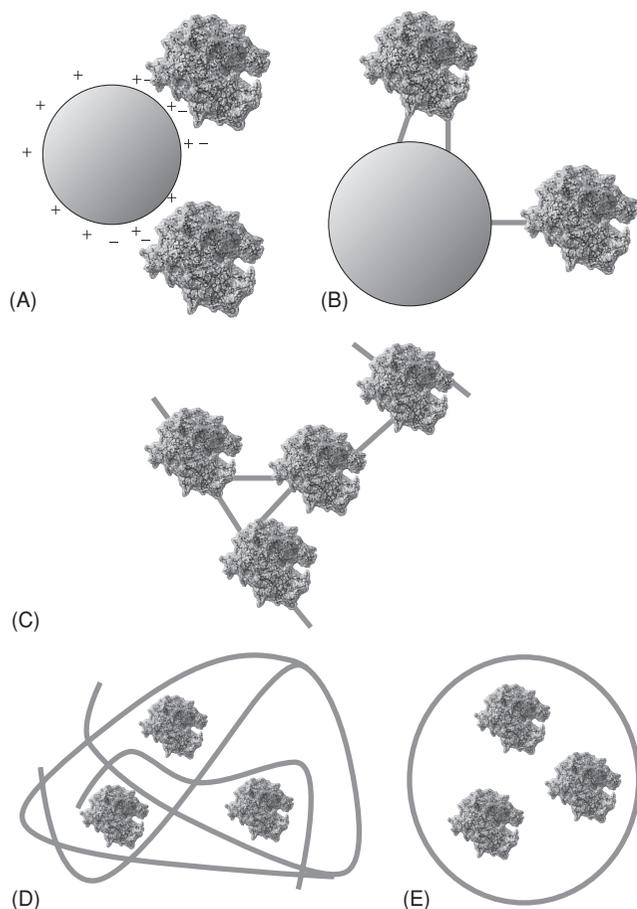


Figure 7.20. Representation of the methods by which an enzyme may be immobilised: adsorption, covalent coupling, cross-linking, matrix entrapment and encapsulation.

(Fig. 7.20B). The binding is very strong and therefore little leakage of enzyme from the support occurs (Calleri et al. 2004). The bond is formed between reactive electrophile groups present on the support and nucleophile side chains on the surface of the enzyme. These side-chains are usually the amino group ($-\text{NH}_2$) of lysine, the imidazole group of histidine, the hydroxyl group ($-\text{OH}$) of serine and threonine, and the sulfhydryl group ($-\text{SH}$) of cysteine. Lysine residues are found to be the most generally useful groups for covalent bonding of enzymes to insoluble supports due to their widespread surface exposure and high reactivity, especially in slightly alkaline solutions.

It is important that the amino acids essential to the catalytic activity of the enzyme are not involved in the covalent linkage to the support (Dravis et al. 2001). This may be difficult to achieve, and enzymes immobilised in this fashion generally lose activity upon immobilisation. This problem may be prevented if the enzyme is immobilised in the presence of saturating concentrations of substrate, product or a competitive inhibitor to protect active site residues. This ensures that the active site remains 'unre-

acted' during the covalent coupling and reduces the occurrence of binding in unproductive conformations.

Various types of beaded supports have been used successfully as for example, natural polymers (e.g. agarose, dextran and cellulose), synthetic polymers (e.g. polyacrylamide, polyacryloyl trihydroxymethylacrylamide, polymethacrylate), inorganic (e.g. silica, metal oxides and controlled pore glass) and microporous flat membrane (Calleri et al. 2004).

The immobilisation procedure consists of three steps (Calleri et al. 2004): (i) activation of the support, (ii) coupling of ligand and (iii) blocking of residual functional groups in the matrix. The choice of coupling chemistry depends on the enzyme to be immobilised and its stability. A number of methods are available in the literature for efficient immobilisation of enzyme through a chosen particular functional side chain's group by employing glutaraldehyde, oxirane, cyanogen bromide, 1,1-carbonyldiimidazole, cyanuric chloride, trialkoxysilane to derivatise glass, etc. Some of them are illustrated in Figure 7.21.

Cross-linking

This type of immobilisation is achieved by cross-linking the enzymes to each other to form complex structures as shown in Figure 7.20C. It is therefore a support-free method and less costly than covalent linkage. Methods of cross-linking involve covalent bond formation between the enzymes using bi- or multi-functional reagent. Cross-linking is frequently carried out using glutaraldehyde, which is of low cost and available in industrial quantities. To minimise close proximity problems associated with the cross-linking of a single enzyme, albumin and gelatin are usually used to provide additional protein molecules as spacers (Podgornik and Tennikova 2002).

Entrapment and Encapsulation

In the immobilisation by entrapment, the enzyme molecules are free in solution, but restricted in movement by the lattice structure of the gel (Fig. 7.20D; Balabushevich et al. 2004). The entrapment method of immobilisation is based on the localisation of an enzyme within the lattice of a polymer matrix or membrane (Podgornik and Tennikova 2002). It is done in such a way as to retain protein while allowing penetration of substrate. Entrapment can be achieved by mixing an enzyme with chemical monomers that are then polymerised to form a cross-linked polymeric network, trapping the enzyme in the interstitial spaces of lattice. Many materials have been used, such as alginate, agarose, gelatin, polystyrene and polyacrylamide. As an example of this latter method, the enzymes' surface lysine residues may be derivatised by reaction with acryloyl chloride ($\text{CH}_2 = \text{CH}-\text{CO}-\text{Cl}$) to give the acryloyl amides. This product may then be copolymerised and cross-linked with acrylamide ($\text{CH}_2 = \text{CH}-\text{CO}-\text{NH}_2$) and bisacrylamide ($\text{H}_2\text{N}-\text{CO}-\text{CH} = \text{CH}-\text{CH} = \text{CH}-\text{CO}-\text{NH}_2$) to form a gel.

Encapsulation of enzymes can be achieved by enveloping the biological components within various forms of semipermeable

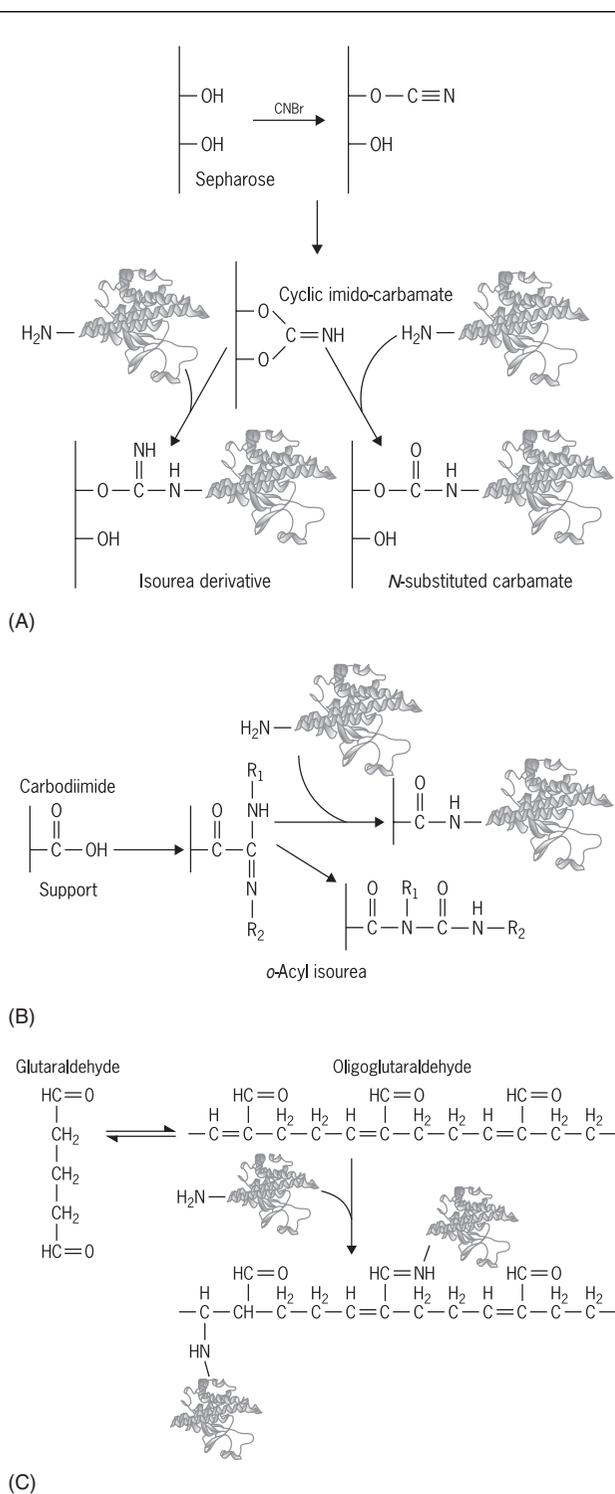


Figure 7.21. Commonly used methods for the covalent immobilisation of enzymes. **(A)** Activation of hydroxyl support by cyanogen bromide. **(B)** Carbodiimides may be used to attach amino groups on the enzyme to carboxylate groups on the support or carboxylate groups on the enzyme to amino groups on the support. **(C)** Glutaraldehyde is used to cross-link enzymes or link them to supports. The product of the condensation of enzyme and glutaraldehyde may be stabilised against dissociation by reduction with sodium borohydride.

membranes as shown in Figure 7.20E. Encapsulation is most frequently carried out using nylon and cellulose nitrate to construct microcapsules varying from 10 to 100 μM . In general, entrapment methods have found more application on the immobilisation of cells.

NEW APPROACHES FOR ORIENTED ENZYME IMMOBILISATION: THE DEVELOPMENT OF ENZYME ARRAYS

With the completion of several genome projects, attention has turned to the elucidation of functional activities of the encoded proteins. Because of the enormous number of newly discovered open reading frames, progress in the analysis of the corresponding proteins depends on the ability to perform characterisation in a parallel and high throughput format (Cahill and Nordhoff 2003). This typically involves construction of protein arrays based on recombinant proteins. Such arrays are then analysed for their enzymatic activities and the ability to interact with other proteins or small molecules, etc. The development of enzyme array technology is hindered by the complexity of protein molecules. The tremendous variability in the nature of enzymes and consequently in the requirement for their detection and identification makes the development of protein chips a particularly challenging task. Additionally, enzyme molecules must be immobilised on a matrix in a way that they preserve their native structures and are accessible to their targets (Cutler 2003). The immobilisation chemistry must be compatible with preserving enzyme molecules in native states. This requires good control of local molecular environments of the immobilised enzyme molecule (Yeo et al. 2004). There is one major barrier in enzyme microarray development: the immobilisation chemistry has to be such that it preserves the enzyme in native state and with optimal orientation for substrate interaction. This problem may be solved by the recently developed *in vitro* protein ligation methodology. Central to this method is the ability of certain protein domains (inteins) to excise themselves from a precursor protein (Lue et al. 2004). In a simplified intein expression system, a thiol reagent induces cleavage of the intein–extein bond, leaving a reactive thioester group on the C-terminus of the protein of interest. This group can then be used to couple essentially any polypeptide with an N-terminal cysteine to the thioester tagged protein by restoring the peptide bond. In another methodology, optimal orientation is based on the unique ability of protein prenyl-transferases to recognise short but highly specific C-terminal protein sequences (Cys–A–A–X–), as shown in Figure 7.22. The enzyme accepts a spectrum of phosphoisoprenoid analogues while displaying a very strict specificity for the protein substrate. This feature is explored for protein derivatisation. Several types of pyrophosphates (biotin analogues, photoreactive aside and benzophenone analogues; Fig. 7.22) can be covalently attached to the protein tagged with the Cys–A–A–X motif. After modification, the protein can be immobilised directly either reversibly through biotin–avidin interaction on avidin modified support or covalently through the photoreactive group on several supports.

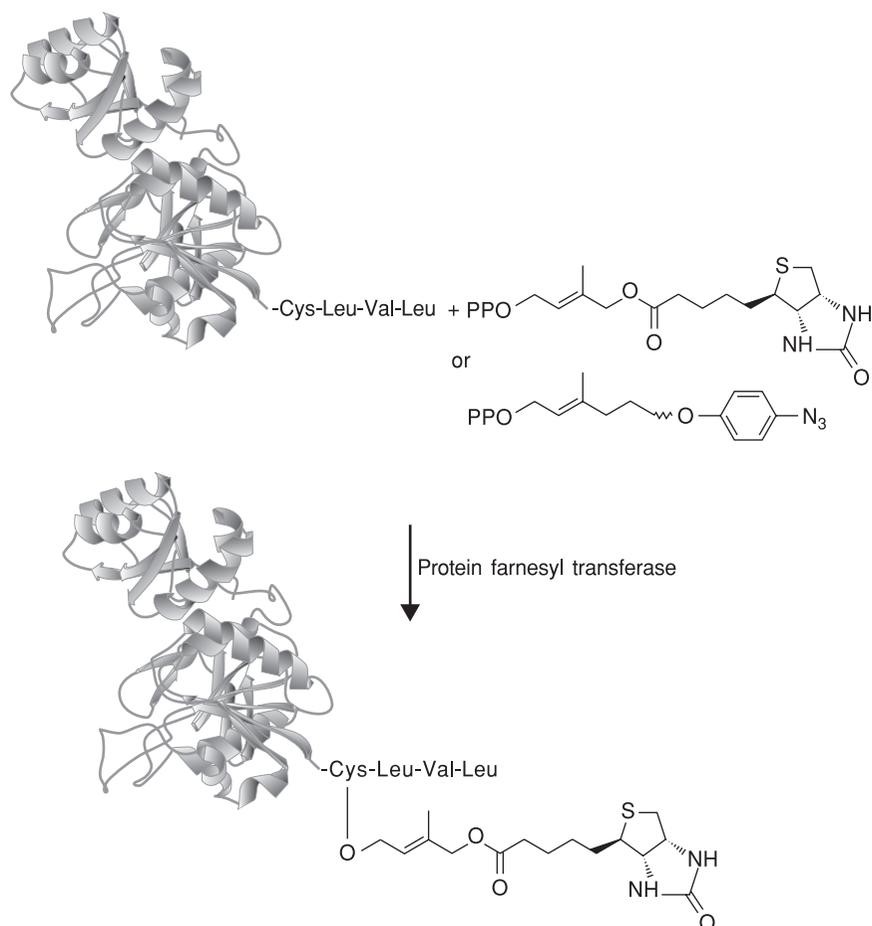


Figure 7.22. Principal scheme of using CAAX-tagged proteins for covalent modification with prenyl transferases.

ENZYME UTILISATION IN INDUSTRY

Enzymes offer potential for many exciting applications in industry. Some important industrial enzymes and their sources are listed in Table 7.9. In addition to the industrial enzymes listed above, a number of enzyme products have been approved for therapeutic use. Examples include tissue plasminogen activator and streptokinase for cardiovascular disease, adenosine deaminase for the rare severe combined immunodeficiency disease, β -glucocerebrosidase for Type 1 Gaucher disease, L-asparaginase for the treatment of acute lymphoblastic leukemia, DNase for the treatment of cystic fibrosis and neuraminidase which is being targeted for the treatment of influenza (Cutler 2003).

There are also thousands of enzyme products used in small amounts for research and development in routine laboratory practice and others that are used in clinical laboratory assays. This group also includes a number of DNA- and RNA-modifying enzymes (DNA and RNA polymerase, DNA ligase, restriction endonucleases, reverse transcriptase, etc.), which led to the de-

velopment of molecular biology methods and were a foundation for the biotechnology industry (Yeo et al. 2004). The clever application of one thermostable DNA polymerase led to the PCR and this has since blossomed into numerous clinical, forensic and academic embodiments. Along with the commercial success of these enzyme products, other enzyme products are currently in commercial development.

Another important field of application of enzymes is in metabolic engineering. Metabolic engineering is a new approach involving the targeted and purposeful manipulation of the metabolic pathways of an organism, aiming at improving the quality and yields of commercially important compounds. It typically involves alteration of cellular activities by manipulation of the enzymatic functions of the cell using recombinant DNA and other genetic techniques. For example, the combination of rational pathway engineering and directed evolution has been successfully applied to optimise the pathways for the production of isoprenoids such as carotenoids (Schmidt-Dannert et al. 2000, Umeno and Arnold 2004).

Table 7.9. Some Important Industrial Enzymes and Their Sources

Enzyme	EC Number	Source	Industrial Use
Rennet	3.4.23.4	Abomasum	Cheese
α -Amylase	3.2.1.1	Malted barley, <i>Bacillus</i> , <i>Aspergillus</i>	Brewing, baking
α -Amylase	3.2.1.2	Malted barley, <i>Bacillus</i>	Brewing
Bromelain	3.4.22.4	Pineapple latex	Brewing
Catalase	1.11.1.6	Liver, <i>Aspergillus</i>	Food
Penicillin amidase	3.5.1.11	<i>Bacillus</i>	Pharmaceutical
Lipoxygenase	1.13.11.12	Soybeans	Food
Ficin	3.4.22.3	Fig latex	Food
Pectinase	3.2.1.15	<i>Aspergillus</i>	Drinks
Invertase	3.2.1.26	<i>Saccharomyces</i>	Confectionery
Pectin lyase	4.2.2.10	<i>Aspergillus</i>	Drinks
Cellulase	3.2.1.4	<i>Trichoderma</i>	Waste
Chymotrypsin	3.4.21.1	Pancreas	Leather
Lipase	3.1.1.3	Pancreas, <i>Rhizopus</i> , <i>Candida</i>	Food
Trypsin	3.4.21.4	Pancreas	Leather
α -Glucanase	3.2.1.6	Malted barley	Brewing
Papain	3.4.22.2	Pawpaw latex	Meat
Asparaginase	3.5.1.1	<i>Erwinia chrysanthemy</i> , <i>Erwinia carotovora</i> , <i>Escherichia coli</i>	Human health
Glucose isomerase	5.3.1.5	<i>Bacillus</i>	Fructose syrup
Protease	3.4.21.14	<i>Bacillus</i>	Detergent
Aminoacylase	3.5.1.14	<i>Aspergillus</i>	Pharmaceutical
Raffinase	3.2.1.22	<i>Saccharomyces</i>	Food
Glucose oxidase	1.1.3.4	<i>Aspergillus</i>	Food
Dextranase	3.2.1.11	<i>Penicillium</i>	Food
Lactase	3.2.1.23	<i>Aspergillus</i>	Dairy
Glucoamylase	3.2.1.3	<i>Aspergillus</i>	Starch
Pullulanase	3.2.1.41	<i>Klebsiella</i>	Starch
Raffinase	3.2.1.22	<i>Mortierella</i>	Food
Lactase	3.2.1.23	<i>Kluyveromyces</i>	Dairy

ENZYMES INVOLVED IN XENOBIOTIC METABOLISM AND BIOCHEMICAL INDIVIDUALITY

The term xenobiotic metabolism refers to the set of metabolic pathways that chemically modify xenobiotics, which are compounds foreign to an organism's normal biochemistry, such as drugs and poisons. The term biochemical individuality of xenobiotic metabolism refers to variability in xenobiotic metabolism and drug responsiveness among different people. Biochemical individuality is a significant factor that can improve public health, drug therapy, nutrition and health impacts such as cancer, diabetes 2 and cardiovascular disease.

Most xenobiotics are lipophilic and able to bind to lipid membranes and be transported in the blood (Hodgson 2004). The enzymes that are involved in xenobiotic metabolism (Table 7.10) comprise one of the first defense mechanism against environmental carcinogens and xenobiotic compounds (Zhang et al. 2009a). Xenobiotic metabolism follows mainly three phases (I, II, III). In Phase I, the original compound obtain increased hydrophilicity and constitute an adequate substrate for phase II enzymes, by the introduction of a polar reactive group ($-\text{OH}$,

$-\text{NH}_2$, $-\text{SH}$ or $-\text{COOH}$). In Phase II, the products of Phase I can be conjugated to substrates such as GSH, which result in a significant increase of water solubility of xenobiotic, promoting its excretion (Hodgson 2004). The ATP-dependent transporters that facilitate the movement of the polar conjugates (by phase I and II) across biological membranes and their excretion from the cell constitute Phase III proteins (Josephy and Mannervik 2006). In general, the enzymes that are involved in xenobiotic metabolism are genetically polymorphic, affecting the individual delicacy to environmental pollutants (Zhang et al. 2009a).

PHASE I

Human cytochrome P450, is one of the most important enzymes that takes part in xenobiotic metabolism; therefore, its genetic polymorphisms have been studied in depth. For example, P450 2A6 (CYP2A6) catalyses nicotine oxidation, and it has been found to have inter-individual and inter-ethnic variability. Genetic polymorphisms of this gene impact smoking behaviour (Xu et al. 2002). Another example of genetic polymorphism's impact of this enzyme came from Siraj et al., who suggested that CYP1A1 phenotype AA showed association with increased

Table 7.10. The Main Enzymes Involved in Xenobiotic Metabolism

	Reactions
<i>Phase I enzymes</i>	
Cytochrome P450 (CYPs)	Epoxidation/hydroxylation <i>N</i> -, <i>O</i> -, <i>S</i> -dealkylation <i>N</i> -, <i>S</i> -, <i>P</i> -oxidation Desulfuration Dehalogenation Azo reduction Nitro reduction
Flavin-containing monooxygenases (FMOs)	<i>N</i> -, <i>S</i> -, <i>P</i> -oxidation Desulfuration
Alcohol dehydrogenase	Oxidation Reductions
Aldehyde dehydrogenase	Oxidation
Prostaglandin synthetase co-oxidation	Dehydrogenation <i>N</i> -dealkylation Epoxidation/hydroxylation Oxidation
Molybdenum hydroxylases	Oxidation Reductions
Esterases and amidases	Hydrolysis
Epoxide hydrolase	Hydrolysis
<i>Phase II enzymes</i>	
UDP (uridine diphospho) glucuronosyl transferase (UGT)	Glucuronide conjugation
Sulfotransferases	Sulfation reaction
Sulfatases	Hydrolysis of sulfate esters
Methyltransferases	<i>N</i> -, <i>O</i> -, <i>S</i> -methylation,
Glutathione <i>S</i> -transferase	Alkyltransferase, aryltransferase, aralkyltransferase, alkenetransferase, epoxidetransferase
γ -Glutamyltranspeptidase	Hydrolysis Transpeptidation
<i>N</i> -acetyltransferase	Acetylation
Aminopeptidases	Hydrolysis of peptides
<i>N</i> , <i>O</i> -Acyltransferase	Acylation
Cysteine conjugate β -lyase	Methylation
<i>Phase III enzymes</i>	
MRP (multi-drug resistance – associated protein)	Transport and excretion of soluble products from phase I and II metabolic pathways
MDR (multi-drug resistance family/ <i>P</i> -glycoprotein)	
MXR (mitoxantrone-resistance protein) efflux	

Source: Rommel and Richard 2002, Hodgson 2004, Josephy and Mannervik 2006.

risk of developing papillary thyroid cancer in Middle Eastern population (Siraj et al. 2008).

Another studied genetic polymorphic enzyme is superoxide dismutase-2 (SOD). The heterozygosity 9Val-allele of MnSOD is associated with a higher risk and severity of orofaciolingual dyskinesias (TDof) in Russian psychiatric inpatients from Siberia (Al Hadithy et al. 2010). In addition, there are several studies that have reported and showed that the 9Val/9Val genotype confer great susceptibility to tardive dyskinesia (Zhang et al. 2003, Akyol et al. 2005, Galecki et al. 2006, Hitzeroth et al. 2007). In addition, human MnSOD gene encoding alanine

(A) or valine at codon 16 (Shimoda-Matsubayashi et al. 1996) can be a risk factor for several malignancies (Iguchi et al. 2009).

PHASE II

Polymorphisms of human UGT correlate with diseases and side effects of drugs, for example the isoform UGT1A1 is associated with diseases of bilirubin metabolism (Hodgson 2004). *N*-Acetyltransferases (NAT) are important enzymes that participate in metabolic activation of carcinogenic aromatic and heterocyclic amines that are present in cigarette smoke (Wang

et al. 1999). An association was found between smokers with MnSOD AA genotype and prostate cancer risk, especially in case of rapid NAT1 subjects (Iguchi et al. 2009). Tamini et al. found that the AA genotype of MnSOD in women smokers have an elevated risk for breast cancer (Tamini et al. 2004).

Tobacco smoke (Lioy and Greenberg 1990) as well as smoked foods, cereals, leafy green vegetables and fossil fuels combustion by-products (Waldman et al. 1991) are the sources of exposure to polycyclic aromatic hydrocarbons (PAHs). PAHs have been considered as potential carcinogens for human (Shimada 2006). According to McCarty et al., the lack of dose–response relationship of PAHs and breast cancer may be due to genetic differences in metabolic activation and detoxification of PAHs (McCarty et al. 2009).

PHASE III

The proteins of phase III are membrane transporters. These proteins seem to be significantly implicated in the absorption, distribution and discard of drugs (Rommel and Richard 2002). There are genetic polymorphisms of drug transporters, which appear to have clinical impact and have been detected in multiple clinical and in vitro studies (Maeda and Sugiyama 2008).

ACKNOWLEDGEMENTS

This work was partially supported by the following grants: HRAKLEITOS II and THALIS ((fall under the Operational Programme “Education and Lifelong Learning”) co-funded by the European Union – European Social Fund & National Resources), AquaPhage (funded by the European Union), BioExplore (co-funded by the European Union – European Social Fund & National Resources).

REFERENCES

- Abdul MM. 1993. Biocatalysis and immobilized enzyme/cell bioreactors. Promising techniques in bioreactor technology. *Biotechnol Bioeng (NY)* 11: 690–695.
- Adamson JG et al. 1993. Structure, function and application of the coiled-coil protein folding motif. *Curr Opin Biotechnol* 4: 428–437.
- Aehle W. 2007. *Enzymes in Industry: Production and Applications*, 3rd edn. Wiley-VCH, Weinheim.
- Agarwal PK et al. 2002. Network of coupled promoting motions in enzyme catalysis. *Proc Natl Acad Sci USA* 99: 2794–2799.
- Aguinaldo AM, Arnold FH. 2003. Staggered extension process (StEP) in vitro recombination. *Methods Mol Biol* 231: 105–110.
- Ahern KG et al. 1988. Identification of regions essential for extrachromosomal replication and maintenance of an endogenous plasmid in *Dictyostelium*. *Nucleic Acids Res* 16(14B): 6825–6837.
- Ahn J et al. 2009. Generation of expression vectors for high-throughput functional analysis of target genes in *Schizosaccharomyces pombe*. *J Microbiol* 47(6): 789–795.
- Akyol O et al. 2005. Association between Ala-9Val polymorphism of Mn-SOD gene and schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 29: 123–131.
- Al Hadithy AFY et al. 2010. Missense polymorphisms in three oxidative-stress enzymes (GSTP1, SOD2, and GPX1) and dyskinesias in Russian psychiatric inpatients from Siberia. *Hum Psychopharmacol Clin Exp* 25: 84–91.
- Alberti S et al. 2007. A suite of Gateway cloning vectors for high-throughput genetic analysis in *Saccharomyces cerevisiae*. *Yeast* 24(10): 913–919.
- Altamirano MM et al. 2000. Directed evolution of new catalytic activity using the α/β -barrel scaffold. *Nature* 403: 617–622.
- Altmann F et al. 1999. Insect cells as hosts for the expression of recombinant glycoproteins. *Glycoconj J* 16: 109–123.
- Altreuter DH, Clark DS. 1999. Combinatorial biocatalysis: taking the lead from nature. *Curr Opin Biotechnol* 10: 130–136.
- Alvarez ML, Cardineau GA. 2010. Prevention of bubonic and pneumonic plague using plant-derived vaccines. *Biotechnol Adv* 28(1): 184–196.
- Andersen CA, Rost B. 2003. Secondary structure assignment. *Methods Biochem Anal* 44: 341–363.
- Anderson CM et al. 1979. Space-filling models of kinase clefts and conformation changes. *Science* 204: 375–380.
- Anderson NL et al. 2000. Proteomics: applications in basic and applied biology. *Curr Opin Biotechnol* 11: 408–412.
- Anderson NL, Anderson NG. 1998. Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis* 19: 1853–1861.
- Andrade MA, Sander C. 1997. Bioinformatics: from genome data to biological knowledge. *Curr Opin Biotechnol* 8: 675–683.
- Andreaddi A et al. 2008. Structure-guided alteration of coenzyme specificity of formate dehydrogenase by saturation mutagenesis to enable efficient utilization of NADP⁺. *FEBS J* 275(15): 3859–3869.
- Apweiler R. 2001. Functional information in SWISS-PROT: the basis for large-scale characterisation of protein sequences. *Brief Bioinform* 2: 9–18.
- Apweiler R et al. 2004. Protein sequence databases. *Curr Opin Chem Biol* 8: 76–80.
- Archer DB. 1994. Enzyme production by recombinant *Aspergillus*. *Bioprocess Technol* 19: 373–393.
- Archer DB, Peberdy JF. 1997. The molecular biology of secreted production by fungi. *Crit Rev Biotechnol* 17: 273–306.
- Arnold FH. 2001. Combinatorial and computational challenges for biocatalyst design. *Nature* 409: 253–257.
- Arnold U. 2009. Incorporation of non-natural modules into proteins: structural features beyond the genetic code. *Biotechnol Lett* 31(8): 1129–1139.
- Arya R et al. 2008a. Production and characterization of pharmacologically active recombinant human phosphodiesterase 4B in *Dictyostelium discoideum*. *Biotechnol J* 3(7): 938–947.
- Arya R et al. 2008b. *Dictyostelium discoideum* – a promising expression system for the production of eukaryotic proteins. *FASEB J* 22(12): 4055–4066.
- Baik SH et al. 2003. Significantly enhanced stability of glucose dehydrogenase by directed evolution. *Appl Microbiol Biotechnol* 61: 329–335.
- Balabushevich NG et al. 2004. Encapsulation of catalase in polyelectrolyte microspheres composed of melamine formaldehyde, dextran sulfate, and protamine. *Biochemistry (Mosc.)* 69: 763–769.

- Balbas P. 2001. Understanding the art of producing protein and nonprotein molecules in *Escherichia coli*. *Mol Biotechnol* 19: 251–267.
- Baldi L et al. 2007. Recombinant protein production by large-scale transient gene expression in mammalian cells: state-of-the-art and future perspectives. *Biotechnol Lett* 29(5): 677–684.
- Balland A et al. 1985. Use of synthetic oligonucleotides in gene isolation and manipulation. *Biochimie* 67: 725–736.
- Bankar AV et al. 2009. Environmental and industrial applications of *Yarrowia lipolytica*. *Appl Microbiol Biotechnol* 84(5): 847–865.
- Bartlett GJ et al. 2002. Analysis of catalytic residues in enzyme active sites. *J Mol Biol* 324: 105–121.
- Basile G, Peticca M. 2009. Recombinant protein expression in *Leishmania tarentolae*. *Mol Biotechnol* 43(3): 273–278.
- Bauer C et al. 2001. A unified theory of enzyme kinetics based upon the systematic analysis of the variations of $k(\text{cat})$, $K(M)$, and $k(\text{cat})/K(M)$ and the relevant $\Delta G(0 \text{ not equal})$ values-possible implications in chemotherapy and biotechnology. *Biochem Pharmacol* 61: 1049–1055.
- Baumert TF et al. 1998. Hepatitis C virus structural proteins assemble into virus-like particles in insect cells. *J Virol* 72: 3827–3836.
- Bendig MM. 1988. The production of foreign proteins in mammalian cells. *Genet Eng* 91–127.
- Benkovic SJ, Hammes-Schiffer S. 2003. A perspective on enzyme catalysis. *Science* 301: 1196–1202.
- Berka RM, Barnett CC. 1989. The development of gene expression systems for filamentous fungi. *Biotechnol Adv* 7: 127–154.
- Berthold W, Walter J. 1994. Protein purification: aspects of processes for pharmaceutical products. *Biologicals* 22: 135–150.
- Bessler C et al. 2003. Directed evolution of a bacterial alpha-amylase: toward enhanced pH-performance and higher specific activity. *Protein Sci* 12: 2141–2149.
- Boehm R. 2007. Bioproduction of therapeutic proteins in the 21st century and the role of plants and plant cells as production platforms. *Ann NY Acad Sci* 1102: 121–134.
- Bornscheuer UT. 2005. Trends and challenges in enzyme technology. *Adv Biochem Eng Biotechnol* 100: 181–203.
- Bornscheuer UT, Pohl M. 2001. Improved biocatalysts by directed evolution and rational protein design. *Curr Opin Chem Biol* 5: 137–143.
- Breitling R et al. 2002. Non-pathogenic trypanosomatid protozoa as a platform for protein research and production. *Protein Expr Purif* 25(2): 209–218.
- Brondyk WH. 2009. Selecting an appropriate method for expressing a recombinant protein. *Methods Enzymol* 463: 131–147.
- Burgess RR. 2009. Refolding solubilized inclusion body proteins. *Methods Enzymol* 463: 259–282.
- Butzin NC et al. 2009. A new system for heterologous expression of membrane proteins: *Rhodospirillum rubrum*. *Protein Expr Purif* 70: 88–94.
- Cahill DJ, Nordhoff E. 2003. Protein arrays and their role in proteomics. *Adv Biochem Eng Biotechnol* 83: 177–187.
- Calleri E et al. 2004. Penicillin G acylase-based stationary phases: analytical applications. *J Pharm Biomed Anal* 35: 243–258.
- Cantor CR. 1980. *The Conformation of Biological Macromolecules*. W.H. Freeman, San Francisco, CA.
- Carbone R. 2009. An advanced application of protein microarrays: cell-based assays for functional genomics. *Methods Mol Biol* 570: 339–352.
- Cardoza RE et al. 2003. Expression of a synthetic copy of the bovine chymosin gene in *Aspergillus awamori* from constitutive and pH-regulated promoters and secretion using two different pre-pro sequences. *Biotechnol Bioeng* 83(3): 249–259.
- Carter P et al. 1989. Engineering subtilisin BPN' for site-specific proteolysis. *Proteins* 6: 240–248.
- Cedrone F et al. 2000. Tailoring new enzyme functions by rational redesign. *Curr Opin Struct Biol* 10: 405–410.
- Cereghino GP et al. 2002. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. *Curr Opin Biotechnol* 13: 329–332.
- Chaput JC et al. 2008. Creating protein biocatalysts as tools for future industrial applications. *Expert Opin Biol Ther* 8(8): 1087–1098.
- Chatterjee R, Yuan L. 2006. Directed evolution of metabolic pathways. *Trends Biotechnol* 24: 28–38.
- Chen R. 2001. Enzyme engineering: rational redesign versus directed evolution. *Trends Biotechnol* 19: 13–14.
- Chen W et al. 1999. Engineering of improved microbes and enzymes for bioremediation. *Curr Opin Biotechnol* 10: 137–141.
- Cherry JR, Fidantsef AL. 2003. Directed evolution of industrial enzymes: an update. *Curr Opin Biotechnol* 14: 438–443.
- Chiang SJ. 2004. Strain improvement for fermentation and biocatalysis processes by genetic engineering technology. *J Ind Microbiol Biotechnol* 31: 99–108.
- Chica RA et al. 2005. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. *Curr Opin Biotechnol* 16: 378–384.
- Chill L et al. 2009. Production, purification, and characterization of human alpha1 proteinase inhibitor from *Aspergillus niger*. *Biotechnol Bioeng* 102(3): 828–844.
- Chin JW et al. 2003. An expanded eukaryotic genetic code. *Science* 301: 964–967.
- Choi JH, Lee SY. 2004. Secretory and extracellular production of recombinant proteins using *Escherichia coli*. *Appl Microbiol Biotechnol* 64: 625–635.
- Chusainow J et al. 2009. A study of monoclonal antibody-producing CHO cell lines: what makes a stable high producer? *Biotechnol Bioeng* 102(4): 1182–1196.
- Cirino PC et al. 2003. Generating mutant libraries using error-prone PCR. *Methods Mol Biol* 231: 3–9.
- Cloniy YD et al. 2000. Biomimetic dyes as affinity chromatography tools in enzyme purification. *J Chromatogr A* 891: 33–44.
- Coco WM. 2003. RACHITT: gene family shuffling by random chimeragenesis on transient templates. *Methods Mol Biol* 231: 111–127.
- Coco WM et al. 2001. DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nat Biotechnol* 19: 354–359.
- Contreras R et al. 1991. Efficient KEX2-like processing of a glucoamylase-interleukin-6 fusion protein by *Aspergillus nidulans* and secretion of mature interleukin-6. *Biotechnology (NY)* 9(4): 378–381.
- Craik CS et al. 1985. Redesigning trypsin: alteration of substrate specificity. *Science* 228: 291–297.
- Cramer CL et al. 1996. Bioproduction of human enzymes in transgenic tobacco. *Ann NY Acad Sci* 792: 62–71.

- Cramer A et al. 1998. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* 391: 288–291.
- Cregg JM et al. 2000. Recombinant protein expression in *Pichia pastoris*. *Mol Biotechnol* 16(1): 23–52.
- Cutler P. 2003. Protein arrays: the current state-of-the-art. *Proteomics* 3: 3–18.
- Dai M, Copley SD. 2004. Genome shuffling improves degradation of the anthropogenic pesticide pentachlorophenol by *Sphingobium chlorophenolicum* ATCC 39723. *Appl Environ Microbiol* 70: 2391–2397.
- Dalby PA. 2003. Optimising enzyme function by directed evolution. *Curr Opin Struct Biol* 13: 500–505.
- Daniell H et al. 2001. Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. *Trends Plant Sci* 6: 219–226.
- De Muynck B et al. 2010. Production of antibodies in plants: status after twenty years. *Plant Biotechnol J* (in press).
- Deiters A et al. 2003. Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*. *J Am Chem Soc* 125: 11782–11783.
- Demain AL, Vaishnav P. 2009. Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv* 27(3): 297–306.
- DeSantis G et al. 2003. Creation of a productive, highly enantioselective nitrilase through gene site saturation mutagenesis (GSSM). *J Am Chem Soc* 125(38): 11476–11477.
- DeYoung JL. 1989. Development of pancreatic enzyme microsphere technology and US findings with pancreas in the treatment of chronic pancreatitis. *Int J Pancreatol* 5 Suppl: 31–36.
- Dingermann T et al. 1991. Expression of human antithrombin III in the cellular slime mould *Dictyostelium discoideum*. *Appl Microbiol Biotechnol* 35(4): 496–503.
- Dittrich W et al. 1994. Production and secretion of recombinant proteins in *Dictyostelium discoideum*. *Biotechnology (NY)* 12: 614–618.
- Dixon DP et al. 2003. Forced evolution of a herbicide detoxifying glutathione transferase. *J Biol Chem* 278: 23930–23935.
- Donayre-Torres AJ et al. 2009. Production and purification of immunologically active core protein p24 from HIV-1 fused to ricin toxin B subunit in *E. coli*. *Virol J* 6: 17.
- Dracheva S et al. 1995. Expression of soluble human interleukin-2 receptor alpha-chain in *Escherichia coli*. *Protein Expr Purif* 6: 737–747.
- Dravis BC et al. 2001. Haloalkane hydrolysis with an immobilized haloalkane dehalogenase. *Biotechnol Bioeng* 75: 416–423.
- Dumon C et al. 2008. Engineering hyperthermostability into a GH11 xylanase is mediated by subtle changes to protein structure. *J Biol Chem* 283(33): 22557–22564.
- Durocher Y, Butler M. 2009. Expression systems for therapeutic glycoprotein production. *Curr Opin Biotechnol* 20(6): 700–707.
- Eblin LP. 1964. An introduction to molecular kinetic theory (Hildebrand, Joel H.). *J Chem Educ* 41(3): 171.
- Ebrahimi F et al. 2010. Production and characterization of a recombinant chimeric antigen consisting botulinum neurotoxin serotypes A, B and E binding subdomains. *J Toxicol Sci* 35(1): 9–19.
- Edmunds T et al. 1998. Transgenically produced human antithrombin: structural and functional comparison to human plasma-derived antithrombin. *Blood* 91: 4561–4571.
- Eisenmesser EZ et al. 2002. Enzyme dynamics during catalysis. *Science* 295: 1520–1523.
- Elias-Lopez AL et al. 2008. Transgenic tomato expressing interleukin-12 has a therapeutic effect in a murine model of progressive pulmonary tuberculosis. *Clin Exp Immunol* 154(1): 123–133.
- Engelhardt EM et al. 2009. Suspension-adapted Chinese hamster ovary-derived cells expressing green fluorescent protein as a screening tool for biomaterials. *Biotechnol Lett* 31(8): 1143–1149.
- Estape D et al. 1998. Susceptibility towards intramolecular disulphide-bond formation affects conformational stability and folding of human basic fibroblast growth factor. *Biochem J* 335(Pt 2): 343–349.
- Fabian JR et al. 1998. Reconstitution and purification of eukaryotic initiation factor 2B (eIF2B) expressed in Sf21 insect cells. *Protein Expr Purif* 13(1): 16–22.
- Fang Q, Shortle D. 2003. Prediction of protein structure by emphasizing local side-chain/backbone interactions in ensembles of turn fragments. *Proteins* 53(Suppl 6): 486–490.
- Fausnaugh JL. 1990. Protein purification and analysis by liquid chromatography and electrophoresis. *Bioprocess Technol* 7: 57–84.
- Faye L, Champey Y. 2008. Plants, medicine and genetics, which applications for tomorrow? *Med Sci (Paris)* 24(11): 939–945.
- Fenyo D, Beavis RC. 2002. Informatics and data management in proteomics. *Trends Biotechnol* 20: S35–S38.
- Fersht A. 1999 *Structure and Mechanism in Protein Science: a Guide to Enzyme Catalysis and Protein Folding*. W.H. Freeman, New York.
- Finkelstein AV, Ptitsyn OB. 2002. Elements of thermodynamics. In: *Protein Physics*. Academic Press, New York, pp. 43–55.
- Fischer R, Emans N. 2000. Molecular farming of pharmaceutical proteins. *Transgenic Res* 9: 279–299.
- Fischer R et al. 1999d. Towards molecular farming in the future: transient protein expression in plants. *Biotechnol Appl Biochem* 30(Pt 2): 113–116.
- Fischer R et al. 1999b. Towards molecular farming in the future: *Pichia pastoris*-based production of single-chain antibody fragments. *Biotechnol Appl Biochem* 30(Pt 2): 117–120.
- Fischer R et al. 1999a. Towards molecular farming in the future: moving from diagnostic protein and antibody production in microbes to plants. *Biotechnol Appl Biochem* 30(Pt 2): 101–108.
- Fischer R et al. 1999c. Molecular farming of recombinant antibodies in plants. *Biol Chem* 380: 825–839.
- Fischer R et al. 2000. Antibody production by molecular farming in plants. *J Biol Regul Homeost Agents* 14: 83–92.
- Fox SR et al. 2004. Maximizing interferon-gamma production by Chinese hamster ovary cells through temperature shift optimization: experimental and modeling. *Biotechnol Bioeng* 85(2): 177–184.
- Galecki P et al. 2006. Manganese superoxide dismutase gene (MnSOD) polymorphism in schizophrenics with tardive dyskinesia from central Poland. *Psychiatr Pol* 40: 937–948.

- Garnier J et al. 1990. Secondary structure prediction and protein design. *Biochem Soc Symp* 57: 11–24.
- Garvey EP, Matthews CR. 1990. Site-directed mutagenesis and its application to protein folding. *Biotechnology* 14: 37–63.
- Gasteiger E et al. 2001. SWISS-PROT: connecting biomolecular knowledge via a protein database. *Curr Issues Mol Biol* 3: 47–55.
- Gasteiger E et al. 2003. ExpASY: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 31: 3784–3788.
- Geisse S, Fux C. 2009. Recombinant protein production by transient gene transfer into mammalian cells. *Methods Enzymol* 463: 223–238.
- Gerlt JA, Babbitt PC. 2009. Enzyme (re)design: lessons from natural evolution and computation. *Curr Opin Chem Biol* 13(1): 10–18.
- Giga-Hama Y, Kumagai H. 1999. Expression system for foreign genes using the fission yeast *Schizosaccharomyces pombe*. *Biotechnol Appl Biochem* 30(Pt 3): 235–244.
- Glasner ME et al. 2007. Mechanisms of protein evolution and their application to protein engineering. *Adv Enzymol Relat Area Mol Biol* 75: 193–239.
- Gomord V, Faye L. 2004. Posttranslational modification of therapeutic proteins in plants. *Curr Opin Plant Biol* 7(2): 171–181.
- Gouka RJ et al. 1997. Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects. *Appl Microbiol Biotechnol* 47: 1–11.
- Gratz A, Jose J. 2008. Protein domain library generation by overlap extension (PDLGO): a tool for enzyme engineering. *Anal Biochem* 378(2): 171–176.
- Grimes BR, Monaco ZL. 2005. Artificial and engineered chromosomes: developments and prospects for gene therapy. *Chromosoma* 114(4): 230–241.
- Guex N, Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18: 2714–2723.
- Guex N et al. 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis* (Suppl 1): S162–S173.
- Gutteridge A, Thornton J. 2004. Conformational change in substrate binding, catalysis and product release: an open and shut case? *FEBS Lett* 567: 67–73.
- Hacker DL et al. 2009. 25 years of recombinant proteins from reactor-grown cells – where do we go from here? *Biotechnol Adv* 27(6): 1023–1027.
- Hackney D. 1990. Binding energy and catalysis. In: *The Enzymes*, vol XIX. Academic Press, p. 136.
- Halabian R et al. 2009. Expression and purification of recombinant human coagulation factor VII fused to a histidine tag using Gateway technology. *Blood Transfus* 7(4): 305–312.
- Hamilton SR, Gerngross TU. 2007. Glycosylation engineering in yeast: the advent of fully humanized yeast. *Curr Opin Biotechnol* 18(5): 387–392.
- Hamilton SR et al. 2003. Production of complex human glycoproteins in yeast. *Science* 301: 1244–1246.
- Hammes GG. 2002. Multiple conformational changes in enzyme catalysis. *Biochemistry* 41: 8221–8228.
- Han ZS et al. 2008. Adenoviral vector mediates high expression levels of human lactoferrin in the milk of rabbits. *J Microbiol Biotechnol* 18(1): 153–159.
- Hasemann CA, Capra JD. 1990. High-level production of a functional immunoglobulin heterodimer in a baculovirus expression system. *Proc Natl Acad Sci USA* 87: 3942–3946.
- Hassan S et al. 2008. Considerations for extraction of monoclonal antibodies targeted to different subcellular compartments in transgenic tobacco plants. *Plant Biotechnol J* 6(7): 733–748.
- Hibbert EG, Dalby PA. 2005. Directed evolution strategies for improved enzymatic performance. *Microb Cell Fact* 4: 29.
- Hitchman RB et al. 2009. Baculovirus expression systems for recombinant protein production in insect cells. *Recent Pat Biotechnol* 3(1): 46–54.
- Hitzeroth A et al. 2007. Association between the MnSOD Ala-9Val polymorphism and development of schizophrenia and abnormal involuntary movements in the Xhosa population. *Prog Neuropsychopharmacol Biol Psychiatry* 31: 664–672.
- Hodgson E. 2004. *A Textbook of Modern Toxicology*, 3rd edn. Wiley-Interscience, Hoboken, NJ, Chapter 7, pp. 111–148.
- Holz C et al. 2003. Establishing the yeast *Saccharomyces cerevisiae* as a system for expression of human proteins on a proteome-scale. *J Struct Funct Genomics* 4: 97–108.
- Holzer GW et al. 2003. Overexpression of hepatitis B virus surface antigens including the preS1 region in a serum-free Chinese hamster ovary cell line. *Protein Expr Purif* 29: 58–69.
- Hooper DC. 2009. Plant vaccines: an immunological perspective. *Curr Top Microbiol Immunol* 332: 1–11.
- Houdebine LM. 2009a. Genetically modified animals. Introduction. *Comp Immunol Microbiol Infect Dis* 32(2): 45–46.
- Hsieh-Wilson LC et al. 1996. Insights into antibody catalysis: structure of an oxygenation catalyst at 1.9-angstrom resolution. *Proc Natl Acad Sci USA* 93: 5363–5367.
- Humphreys DP. 2003. Production of antibodies and antibody fragments in *Escherichia coli* and a comparison of their functions, uses and modification. *Curr Opin Drug Discov Devel* 6: 188–196.
- Idiris A et al. 2010. Engineering of protein secretion in yeast: strategies and impact on protein production. *Appl Microbiol Biotechnol* (in press).
- Iguchi T et al. 2009. MnSOD genotype and prostate cancer risk as a function of NAT genotype and smoking status. *In Vivo* 23: 7–12.
- Iizuka M et al. 2009. Production of a recombinant mouse monoclonal antibody in transgenic silkworm cocoons. *FEBS J* 276(20): 5806–5820.
- Ikeno M et al. 1998. Construction of YAC-based mammalian artificial chromosomes. *Nat Biotechnol* 16: 431–439.
- Jacobs PP, Callewaert N. 2009. N-glycosylation engineering of biopharmaceutical expression systems. *Curr Mol Med* 9(7): 774–800.
- Jaeger KE, Eggert T. 2004. Enantioselective biocatalysis optimized by directed evolution. *Curr Opin Biotechnol* 15: 305–313.
- Jahic M et al. 2006. Process technology for production and recovery of heterologous proteins with *Pichia pastoris*. *Biotechnol Prog* 22(6): 1465–1473.
- Jana S, Deb JK. 2005. Strategies for efficient production of heterologous proteins in *Escherichia coli*. *Appl Microbiol Biotechnol* 67(3): 289–298.
- Janne J et al. 1998. Transgenic bioreactors. *Biotechnol Annu Rev* 4: 55–74.
- Jaroszewski L. 2009. Protein structure prediction based on sequence similarity. *Methods Mol Biol* 569: 129–156.

- Jarvis DL. 2009. Baculovirus-insect cell expression systems. *Methods Enzymol* 463: 191–222.
- Jestin JL, Kaminski PA. 2004. Directed enzyme evolution and selections for catalysis based on product formation. *J Biotechnol* 113: 85–103.
- Johannes TW, Zhao H. 2006. Directed evolution of enzymes and biosynthetic pathways. *Current Opinion in Microbiology* 9: 261–267.
- Joseph D et al. 1990. Anatomy of a conformational change: hinged “lid” motion of the triosephosphate isomerase loop. *Science* 249: 1425–1428.
- Joseph PD, Mannervik B. 2006. *Molecular Toxicology*, 2nd edn. Oxford University Press, New York, Chapter 9, pp. 303–332.
- Joshi L, Lopez LC. 2005. Bioprospecting in plants for engineered proteins. *Curr Opin Plant Biol* 8(2): 223–226.
- Joubert O et al. 2010. Heterologous expression of human membrane receptors in the yeast *Saccharomyces cerevisiae*. *Methods Mol Biol* 601: 87–103.
- Jung E, Williams KL. 1997. The production of recombinant glycoproteins with special reference to simple eukaryotes including *Dictyostelium discoideum*. *Biotechnol Appl Biochem* 25(Pt 1): 3–8.
- Jung E et al. 1997. An in vivo approach for the identification of acceptor sites for O-glycosyltransferases: motifs for the addition of O-GlcNAc in *Dictyostelium discoideum*. *Biochemistry* 36(13): 4034–4040.
- Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22: 2577–2637.
- Kakkis ED et al. 1994. Overexpression of the human lysosomal enzyme alpha-L-iduronidase in Chinese hamster ovary cells. *Protein Expr Purif* 5: 225–232.
- Kapoli P et al. 2008. Engineering sensitive glutathione transferase for the detection of xenobiotics. *Biosens Bioelectron* 24(3): 498–503.
- Kaszubska W et al. 2000. Expression, purification, and characterization of human recombinant thrombopoietin in Chinese hamster ovary cells. *Protein Expr Purif* 18: 213–220.
- Kaur J, Sharma R. 2006. Directed evolution: an approach to engineer enzymes. *Crit Rev Biotechnol* 26(3): 165–199.
- Ke SH, Madison EL. 1997. Rapid and efficient site-directed mutagenesis by single-tube ‘megaprimer’ PCR method. *Nucleic Acids Res* 25: 3371–3372.
- Kermode AR et al. 2007. Ectopic expression of a conifer abscisic acid insensitive 3 transcription factor induces high-level synthesis of recombinant human alpha-L-iduronidase in transgenic tobacco leaves. *Plant Mol Biol* 63(6): 763–776.
- Kidd BA et al. 2009. Computation of conformational coupling in allosteric proteins. *PLoS Comput Biol* 5(8): e1000484.
- Kim SH et al. 2008. Neutralization of hepatitis B virus (HBV) by human monoclonal antibody against HBV surface antigen (HBsAg) in chimpanzees. *Antiviral Res* 79(3): 188–191.
- Kirk O et al. 2002. Industrial enzyme applications. *Curr Opin Biotechnol* 13(4): 345–351.
- Klimanskaya I et al. 2008. Derive and conquer: sourcing and differentiating stem cells for therapeutic applications. *Nat Rev Drug Discov* 7(2): 131–142.
- Ko K et al. 2009. Production of antibodies in plants: approaches and perspectives. *Curr Top Microbiol Immunol* 332: 55–78.
- Koehn J, Hunt I. 2009. High-throughput protein production (HTPP): a review of enabling technologies to expedite protein production. *Methods Mol Biol* 498: 1–18.
- Kopp J, Schwede T. 2004. Automated protein structure homology modeling: a progress report. *Pharmacogenomics* 5: 405–416.
- Kost TA, Condreay JP. 1999. Recombinant baculoviruses as expression vectors for insect and mammalian cells. *Curr Opin Biotechnol* 10: 428–433.
- Kost TA, Condreay JP. 2002. Recombinant baculoviruses as mammalian cell gene-delivery vectors. *Trends Biotechnol* 20: 173–180.
- Kotzia GA, Labrou NE. 2005. Cloning, expression and characterization of *Erwinia caratovora* L-asparaginase. *Journal of Biotechnology* 119(2005): 309–323.
- Kotzia GA, Labrou NE. 2007. L-Asparaginase from *Erwinia chrysanthemi* 3937: cloning, expression and characterization. *J Biotechnol* 127(4): 657–669.
- Kotzia GA, Labrou NE. 2009. Engineering thermal stability of L-asparaginase by *in vitro* directed evolution. *FEBS J* 276(6): 1750–1761.
- Kotzia GA et al. 2006. Evolutionary methods in enzyme technology: high-throughput screening and selection of new enzyme variants. In: SG Pandalai (ed.) *Recent Research Developments in Biotechnology & Bioengineering*, vol. 7. Research Signpost, Kerala, pp. 85–104.
- Kukuruzinska MA et al. 1987. Protein glycosylation in yeast. *Annu Rev Biochem* 56: 915–944.
- Kurtzman AL et al. 2001. Advances in directed protein evolution by recursive genetic recombination: applications to therapeutic proteins. *Curr Opin Biotechnol* 12: 361–370.
- Kwong KY, Rader C. 2009. *E. coli* expression and purification of Fab antibody fragments. *Curr Protoc Protein Sci* Chapter 6: Unit 6.10.
- La Flamme AC et al. 1995. Expression of mammalian cytokines by *Trypanosoma cruzi* indicates unique signal sequence requirements and processing. *Mol Biochem Parasitol* 75(1): 25–31.
- Labrou NE. 2000. Dye-ligand affinity chromatography for protein separation and purification. *Methods Mol Biol* 147: 129–139.
- Labrou NE. 2002. Affinity chromatography. In: MN Gupta (ed.) *Methods for Affinity-Based Separations of Enzymes and Proteins*. Birkhuser Verlag AG, Switzerland, pp. 16–18.
- Labrou NE. 2003. Design and selection of ligands for affinity chromatography. *J Chromatogr B Analyt Technol Biomed Life Sci* 790: 67–78.
- Labrou NE. 2010. Random mutagenesis methods for *in vitro* directed enzyme evolution. *Curr Protein Pept* 11: 91–100.
- Labrou NE, Rigden DJ. 2001. Active-site characterization of *Candida boidinii* formate dehydrogenase. *Biochem J* 354(Pt 2): 455–463.
- Labrou N, Clonis YD. 1994. The affinity technology in downstream processing. *J Biotechnol* 36: 95–119.
- Labrou NE, Clonis YD. 1995. Biomimetic dye affinity chromatography for the purification of bovine heart lactate dehydrogenase. *J Chromatogr A* 718: 35–44.
- Labrou NE, Clonis YD. 1996. Biomimetic-dye affinity chromatography for the purification of mitochondrial L-malate dehydrogenase from bovine heart. *J Biotechnol* 45: 185–194.

- Labrou NE, Clonis YD. 2002. Immobilised synthetic dyes in affinity chromatography. In: M.A. Vijayalakshmi (ed.) *Biochromatography – Theory and Practice*. Taylor and Francis Publishers, London, Chapter 8, pp. 235–251.
- Labrou NE, Rigden DJ. 2001. Active-site characterization of *Candida boidinii* formate dehydrogenase. *Biochem J* 354: 455–463.
- Labrou NE et al. 2004a. Engineering the pH-dependence of kinetic parameters of maize glutathione S-transferase I by site-directed mutagenesis. *Biomol Eng* 21: 61–66.
- Labrou NE et al. 2001. The conserved Asn49 of maize glutathione S-transferase I modulates substrate binding, catalysis and inter-subunit communication. *Eur J Biochem* 268: 3950–3957.
- Labrou NE et al. 2005. Kinetic analysis of maize glutathione S-transferase I catalyzing the detoxification from chloroacetanilide herbicides. *Planta* 222: 91–97.
- Labrou NE et al. 2004b. Dye-ligand and biomimetic affinity chromatography. In: DS Hage (ed.) *Handbook of Affinity Chromatography*. Marcel Dekker, New York.
- Lai H et al. 2010. Monoclonal antibody produced in plants efficiently treats West Nile virus infection in mice. *Proc Natl Acad Sci USA* 107(6): 2419–2424.
- Laidler KJ. 1984. The development of the Arrhenius equation. *J Chem Educ* 61(6): 494.
- Larrick JW, Thomas DW. 2001. Producing proteins in transgenic plants and animals. *Curr Opin Biotechnol* 12(4): 411–418.
- Lesley SA. 2001. High-throughput proteomics: protein expression and purification in the postgenomic world. *Protein Expr Purif* 22: 159–164.
- Lesley SA et al. 2002. Gene expression response to misfolded protein as a screen for soluble recombinant protein. *Protein Eng* 15: 153–160.
- Levison PR. 2003. Large-scale ion-exchange column chromatography of proteins. Comparison of different formats. *J Chromatogr B Anal Technol Biomed Life Sci* 790: 17–33.
- Lewin R. 1982. RNA can be a catalyst. *Science* 218: 872–874.
- Liang JF et al. 2000. Biomedical application of immobilized enzymes. *J Pharm Sci* 89: 979–990.
- Lienard D et al. 2007. Pharming and transgenic plants. *Biotechnol Annu Rev* 13: 115–147.
- Lioy PJ, Greenberg A. 1990. Factors associated with human exposures to polycyclic aromatic hydrocarbons. *Toxicol Ind Health* 6: 209–223.
- Liu D et al. 2009a. High-level expression and large-scale preparation of soluble HBx antigen from *Escherichia coli*. *Biotechnol Appl Biochem* 54(3): 141–147.
- Liu L et al. 2009b. Directed evolution of an orthogonal nucleoside analog kinase via fluorescence-activated cell sorting. *Nucleic Acids Res* 37(13): 4472–4481.
- Logan JS, Martin MJ. 1994. Transgenic swine as a recombinant production system for human hemoglobin. *Methods Enzymol* 231: 435–445.
- Lowe CR. 2001. Combinatorial approaches to affinity chromatography. *Curr Opin Chem Biol* 5: 248–256.
- Lubertozzi D, Keasling JD. 2009. Developing *Aspergillus* as a host for heterologous expression. *Biotechnol Adv* 27(1): 53–75.
- Lue RY et al. 2004. Versatile protein biotinylation strategies for potential high-throughput proteomics. *J Am Chem Soc* 126: 1055–1062.
- Lutz S et al. 2001. Rapid generation of incremental truncation libraries for protein engineering using alpha-phosphothioate nucleotides. *Nucleic Acids Res* 29: E16.
- Madzak C et al. 2004. Heterologous protein expression and secretion in the non-conventional yeast *Yarrowia lipolytica*: a review. *J Biotechnol* 109(1–2): 63–81.
- Madzak C et al. 2005. Heterologous production of a laccase from the basidiomycete *Pycnoporus cinnabarinus* in the dimorphic yeast *Yarrowia lipolytica*. *FEMS Yeast Res* 5(6–7): 635–646.
- Maeda K, Sugiyama Y. 2008. Impact of genetic polymorphisms of transporters on the pharmacokinetic, pharmacodynamic and toxicological properties of anionic drugs. *Drug Metab Pharmacokin* 23: 223–235.
- Marti S et al. 2004. Theoretical insights in enzyme catalysis. *Chem Soc Rev* 33: 98–107.
- Mason HS, Arntzen CJ. 1995. Transgenic plants as vaccine production systems. *Trends Biotechnol* 13: 388–392.
- Matsushita-Morita M et al. 2009. Characterization of recombinant prolyl aminopeptidase from *Aspergillus oryzae*. *J Appl Microbiol* 109(10): 156–165.
- Matsuura Y et al. 1987. Baculovirus expression vectors: the requirements for high level expression of proteins, including glycoproteins. *J Gen Virol* 68(Pt 5): 1233–1250.
- Matthews BW. 1993. Structural and genetic analysis of protein stability. *Annu Rev Biochem* 62: 139–160.
- McCarty KM et al. 2009. PAH–DNA adducts, cigarette smoking, GST polymorphisms, and breast cancer risk. *Environ Health Perspect* 117: 552–558.
- McCormick DB. 1997. *Methods in Enzymology. Vitamins and coenzymes, Part L*. Academic Press, San Diego, p. 282.
- Mcgrath BM. 2005. *Directory of Therapeutic Enzymes*, 1st edn. CRC Press.
- Melissis S et al. 2006. Nucleotide-mimetic synthetic ligands for DNA-recognizing enzymes one-step purification of Pfu DNA polymerase. *J Chromatogr A* 1122(1–2): 63–75.
- Melissis S et al. 2007. One-step purification of Taq DNA polymerase using nucleotide-mimetic affinity chromatography. *Biotechnol J* 2(1): 121–132.
- Mohan U, Banerjee UC. 2008. Molecular evolution of a defined DNA sequence with accumulation of mutations in a single round by a dual approach to random chemical mutagenesis (DuARChEM). *ChemBiochem* 9(14): 2238–2243.
- Monaco ZL, Moralli D. 2006. Progress in artificial chromosome technology. *Biochem Soc Trans* 34(Pt 2): 324–327.
- Moss DW. 1988. Theoretical enzymology: enzyme kinetics and enzyme inhibition. In: DL Williams, V Marks (eds.) *Principles of Clinical Biochemistry*, 2nd edn. Bath Press, Bath, UK, pp. 423–440.
- Murzin AG et al. 1995. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* 247: 536–540.
- Nam HJ et al. 2009. Bioinformatic approaches for the structure and function of membrane proteins. *BMB Rep* 42(11): 697–704.
- Nevalainen KM et al. 2005. Heterologous protein expression in filamentous fungi. *Trends Biotechnol* 23(9): 468–474.
- Niemann H, Kues WA. 2007. Transgenic farm animals: an update. *Reprod Fertil Dev* 19(6): 762–770.
- Nishida CR, Ortiz de Montellano PR. 2001. Control of electron transfer in nitric-oxide synthases. Swapping of autoinhibitory

- elements among nitric-oxide synthase isoforms. *J Biol Chem* 276(23): 20116–20124.
- Nixon AE et al. 1998. Hybrid enzymes: manipulating enzyme design. *Trends Biotechnol* 16: 258–264.
- Okino N et al. 2010. Expression, purification, and characterization of a recombinant neutral ceramidase from *Mycobacterium tuberculosis*. *Biosci Biotechnol Biochem* 74(2): 316–321.
- Orengo CA et al. 1997. CATH – a hierarchic classification of protein domain structures. *Structure* 5: 1093–1108.
- Otero JM, Nielsen J. 2010. Industrial systems biology. *Biotechnol Bioeng* 105(3): 439–460.
- Ozturk DH, Erickson-Viitanen S. 1998. Expression and purification of HIV-I p15NC protein in *Escherichia coli*. *Protein Expr Purif* 14: 54–64.
- Panda AK. 2003. Bioprocessing of therapeutic proteins from the inclusion bodies of *Escherichia coli*. *Adv Biochem Eng Biotechnol* 85: 43–93.
- Panke S, Wubbolts MG. 2002. Enzyme technology and bioprocess engineering. *Curr Opin Biotechnol* 13(2): 111–116.
- Park HS et al. 2006. Design and evolution of new catalytic activity with an existing protein scaffold. *Science* 311: 535–538.
- Peti W, Page R. 2007. Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost. *Protein Expr Purif* 51(1): 1–10.
- Phadke RS. 1992. Biosensors and enzyme immobilized electrodes. *Biosystems* 27: 203–206.
- Piefer AJ, Jonsson CB. 2002. A comparative study of the human T-cell leukemia virus type 2 integrase expressed in and purified from *Escherichia coli* and *Pichia pastoris*. *Protein Expr Purif* 25: 291–299.
- Platis D, Foster GR. 2003. High yield expression, refolding, and characterization of recombinant interferon alpha2/alpha8 hybrids in *Escherichia coli*. *Protein Expr Purif* 31: 222–230.
- Podgornik A, Tennikova TB. 2002. Chromatographic reactors based on biological activity. *Adv Biochem Eng Biotechnol* 76: 165–210.
- Powell LW. 1984. Developments in immobilized-enzyme technology. *Biotechnol Genet Eng Rev* 2: 409–438.
- Price N, Stevens L. 1999. *Fundamentals of Enzymology*, 3rd edn. Oxford University Press, New York, pp. 79–92.
- Prunkard D et al. 1996. High-level expression of recombinant human fibrinogen in the milk of transgenic mice. *Nat Biotechnol* 14(7): 867–871.
- Qu C et al. 2000. 3D domain swapping modulates the stability of members of an icosahedral virus group. *Struct Fold Des* 8: 1095–1103.
- Qu X et al. 2009. A guide to template based structure prediction. *Curr Protein Pept Sci* 10(3): 270–285.
- Rabhi-Essafi I et al. 2007. A strategy for high-level expression of soluble and functional human interferon alpha as a GST-fusion protein in *E. coli*. *Protein Eng Des Sel* 20(5): 201–209.
- Rakintzis TN. 1994. *Physical Chemistry, First Thermodynamic Law*. Chapter C, pp. 71–103.
- Reetz MT. 2007. Controlling the selectivity and stability of proteins by new strategies in directed evolution: the case of organocatalytic enzymes. *Ernst Schering Found Symp Proc* 2: 321–340.
- Regnier FE. 1987. Chromatography of complex protein mixtures. *J Chromatogr* 418: 115–143.
- Richardson JS. 1981. The anatomy and taxonomy of protein structure. *Adv Protein Chem* 34: 167–339.
- Rommel GT, Richard BK. 2002. Pharmacogenomics of drug transporters. In: J Licinio, M Wong (eds.) *Pharmacogenomics: The Search for Individualized Therapies*. Wiley VCH, Weinheim, Chapter 9, pp. 179–213.
- Roodveldt C et al. 2005. Directed evolution of proteins for heterologous expression and stability. *Curr Opin Struct Biol* 15: 50–56.
- Rosic NN. 2009. Versatile capacity of shuffled cytochrome P450s for dye production. *Appl Microbiol Biotechnol* 82(2): 203–210.
- Roy P et al. 1994. Long-lasting protection of sheep against blue-tongue challenge after vaccination with virus-like particles: evidence for homologous and partial heterologous protection. *Vaccine* 12: 805–811.
- Rudolph NS. 1999. Biopharmaceutical production in transgenic livestock. *Trends Biotechnol* 17: 367–374.
- Russell DA. 1999. Feasibility of antibody production in plants for human therapeutic use. *Curr Top Microbiol Immunol* 240: 119–138.
- Ryu DD, Nam DH. 2000. Recent progress in biomolecular engineering. *Biotechnol Prog* 16: 2–16.
- Sahdev S et al. 2008. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Mol Cell Biochem* 307(1–2): 249–264.
- Sala F et al. 2003. Vaccine antigen production in transgenic plants: strategies, gene constructs and perspectives. *Vaccine* 21: 803–808.
- Saven JG. 2002. Combinatorial protein design. *Curr Opin Struct Biol* 12(4): 453–458.
- Saxena A et al. 2009. The basic concepts of molecular modeling. *Methods Enzymol* 467: 307–334.
- Schatz SM et al. 2003. Higher expression of Fab antibody fragments in a CHO cell line at reduced temperature. *Biotechnol Bioeng* 84: 433–438.
- Schena M et al. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270: 467–470.
- Schena M et al. 1998. Microarrays: biotechnology's discovery platform for functional genomics. *Trends Biotechnol* 16: 301–306.
- Schillberg S et al. 2003a. 'Molecular farming' of antibodies in plants. *Naturwissenschaften* 90: 145–155.
- Schillberg S et al. 2003b. Molecular farming of recombinant antibodies in plants. *Cell Mol Life Sci* 60: 433–445.
- Schillberg S et al. 2005. Opportunities for recombinant antigen and antibody expression in transgenic plants—technology assessment. *Vaccine* 23(15): 1764–1769.
- Schmidt M et al. 2009. Protein engineering of carboxyl esterases by rational design and directed evolution. *Protein Pept Lett* 16(10): 1162–1171.
- Schmidt-Dannert C et al. 2000. Molecular breeding of carotenoid biosynthetic pathways. *Nat Biotechnol* 18: 750–753.
- Schwede T et al. 2003. SWISS-MODEL: an automated protein homology–modeling server. *Nucleic Acids Res* 31: 3381–3385.
- Scopes RK. 1987. Dye-ligands and multifunctional adsorbents: an empirical approach to affinity chromatography. *Anal Biochem* 165: 235–246.
- Scrutton NS et al. 1990. Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature* 343: 38–43.
- Sgaramella V, Eridani S. 2004. From natural to artificial chromosomes: an overview. *Methods Mol Biol* 240: 1–12.

- Shao Z et al. 1998. Random-priming in vitro recombination: an effective tool for directed evolution. *Nucleic Acids Res* 26: 681–683.
- Shimada T. 2006. Xenobiotic-metabolizing enzymes involved in the activation and detoxification of carcinogenic polycyclic aromatic hydrocarbons. *Drug Metab Pharmacokinet* 24: 257–276.
- Shimoda-Matsubayashi S et al. 1996. Structural dimorphism in the mitochondrial targeting sequence in the human manganese superoxide dismutase gene. A predictive evidence for conformational change to influence mitochondrial transport and a study of allelic association in Parkinson's disease. *Biochem Biophys Res Commun* 226: 561–565.
- Sibanda BL et al. 1989. Conformation of beta-hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering. *J Mol Biol* 206: 759–777.
- Singh A et al. 1995. Construction and characterization of a chimeric beta-glucosidase. *Biochem J* 305(Pt 3): 715–719.
- Siraj AK et al. 2008. Polymorphisms of selected xenobiotic genes contribute to the development of papillary thyroid cancer susceptibility in Middle Eastern population. *BMC Medical Genetics* 9: 61–69.
- Slade MB et al. 1997. Expression of recombinant glycoproteins in the simple eukaryote *Dictyostelium discoideum*. *Biotechnol Genet Eng Rev* 14: 1–35.
- Soleimani M et al. 2007. Expression of human tissue plasminogen activator in the trypanosomatid protozoan *Leishmania tarentolae*. *Biotechnol Appl Biochem* 48(Pt 1): 55–61.
- Sourrouille C et al. 2009. From Neanderthal to nanobiotech: from plant potions to pharming with plant factories. *Methods Mol Biol* 483: 1–23.
- Stein H et al. 2009. Production of bioactive, post-translationally modified, heterotrimeric, human recombinant type-I collagen in transgenic tobacco. *Biomacromolecules* 10(9): 2640–2645.
- Steiner S, Anderson NL. 2000. Pharmaceutical proteomics. *Ann NY Acad Sci* 919: 48–51.
- Stemmer WP. 1994. Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature* 370: 389–391.
- Stoger E et al. 2002. Plantibodies: applications, advantages and bottlenecks. *Curr Opin Biotechnol* 13: 161–166.
- Su D et al. 2010. Glycosylation-modified erythropoietin with improved half-life and biological activity. *Int J Hematol* 91(2): 238–244.
- Suen KF et al. 2010. Transient expression of an IL-23R extracellular domain Fc fusion protein in CHO vs. HEK cells results in improved plasma exposure. *Protein Expr Purif* 71(1): 96–102.
- Sunley K et al. 2008. CHO cells adapted to hypothermic growth produce high yields of recombinant beta-interferon. *Biotechnol Prog* 24(4): 898–906.
- Surewicz WK, Mantsch HH. 1988. New insight into protein secondary structure from resolution-enhanced infrared spectra. *Biochim Biophys Acta* 952: 115–130.
- Swanson ME et al. 1992. Production of functional human hemoglobin in transgenic swine. *Biotechnology (NY)* 10: 557–559.
- Swartz JR. 2001. Advances in *Escherichia coli* production of therapeutic proteins. *Curr Opin Biotechnol* 12: 195–201.
- Takahashi R, Ueda M. 2010. Generation of transgenic rats using YAC and BAC DNA constructs. *Methods Mol Biol* 597: 93–108.
- Takahashi T et al. 2010. Binding of sulfatide to recombinant hemagglutinin of influenza A virus produced by a baculovirus protein expression system. *J Biochem* 174(4): 459–462.
- Takegawa K et al. 2009. Production of heterologous proteins using the fission-yeast (*Schizosaccharomyces pombe*) expression system. *Biotechnol Appl Biochem* 53(Pt 4): 227–235.
- Tamini RM et al. 2004. Manganese superoxide dismutase polymorphism, plasma antioxidants, cigarette smoking, and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 13: 989–996.
- Tang W et al. 2006. Expression, purification and identification of recombinant mouse interleukin 21 protein in *E. coli*. *Cell Mol Immunol* 3(4): 311–315.
- Tao H, Cornish VW. 2002. Milestones in directed enzyme evolution. *Curr Opin Chem Biol* 6: 858–864.
- Terpe K. 2003. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol* 60: 523–533.
- Terpitz U et al. 2008. Electrofused giant protoplasts of *Saccharomyces cerevisiae* as a novel system for electrophysiological studies on membrane proteins. *Biochim Biophys Acta* 1778(6): 1493–1500.
- The UniProt Consortium. 2008. The universal protein resource (UniProt). *Nucleic Acids Res* 36: D190–D195.
- Tiltscher H, Storr M. 1993. Immobilization of the slime mould *Dictyostelium discoideum* for the continuous production of recombinant human antithrombin III. *Appl Microbiol Biotechnol* 40(2–3): 246–250.
- Tiwari S et al. 2009. Plants as bioreactors for the production of vaccine antigens. *Biotechnol Adv* 27(4): 449–467.
- Tobin JF et al. 1993. Transfected *Leishmania* expressing biologically active IFN-gamma. *J Immunol* 150(11): 5059–5069.
- Tracewell CA, Arnold FH. 2009. Directed enzyme evolution: climbing fitness peaks one amino acid at a time. *Curr Opin Chem Biol* 13(1): 3–9.
- Trometer C, Falson P. 2010. Mammalian membrane protein expression in baculovirus-infected insect cells. *Methods Mol Biol* 601: 105–117.
- Trueman LJ. 1995. Heterologous expression in yeast. *Methods Mol Biol* 49: 341–354.
- Turner NJ. 2009. Directed evolution drives the next generation of biocatalysts. *Nat Chem Biol* 5(8): 567–573.
- Twyman RM et al. 2003. Molecular farming in plants: host systems and expression technology. *Trends Biotechnol* 21: 570–578.
- Twyman RM et al. 2005. Transgenic plants in the biopharmaceutical market. *Expert Opin Emerg Drugs* 10(1): 185–218.
- Umeno D, Arnold FH. 2004. Evolution of a pathway to novel long-chain carotenoids. *J Bacteriol* 186: 1531–1536.
- Vahedi F et al. 2009. Characterization of *Bacillus anthracis* spores isolates from soil by biochemical and multiplex PCR analysis. *East Mediterr Health J* 15(1): 149–156.
- van Beilen JB, Li Z. 2002. Enzyme technology: an overview. *Curr Opin Biotechnol* 13(4): 338–344.
- van Berkel PH et al. 2002. Large scale production of recombinant human lactoferrin in the milk of transgenic cows. *Nat Biotechnol* 20: 484–487.
- van den Hombergh JP et al. 1997. Aspergillus as a host for heterologous protein production: the problem of proteases. *Trends Biotechnol* 15: 256–263.

- van Ooyen AJ et al. 2006. Heterologous protein production in the yeast *Kluyveromyces lactis*. *FEMS Yeast Res* 6(3): 381–392.
- Velander WH et al. 1992. Production of biologically active human protein C in the milk of transgenic mice. *Ann NY Acad Sci* 665: 391–403.
- Verdoes JC et al. 1993. Glucoamylase overexpression in *Aspergillus niger*: molecular genetic analysis of strains containing multiple copies of the *glaA* gene. *Transgenic Res* 2: 84–92.
- Vitale A, Pedrazzini E. 2005. Recombinant pharmaceuticals from plants: the plant endomembrane system as bioreactor. *Mol Interv* 5(4): 216–225.
- Wada M et al. 2003. Directed evolution of *N*-acetylneuraminic acid aldolase to catalyze enantiomeric aldol reactions. *Bioorg Med Chem* 11: 2091–2098.
- Wagner CR, Benkovic SJ. 1990. Site directed mutagenesis: a tool for enzyme mechanism dissection. *Trends Biotechnol* 8: 263–270.
- Waldman JM et al. 1991. Analysis of human exposure to benzo(a)pyrene via inhalation and food ingestion in the Total Human Environmental Exposure Study (THEES). *J Expo Anal Environ Epidemiol* 1: 193–225.
- Walsh G. 2003. Pharmaceutical biotechnology products approved within the European Union. *Eur J Pharm Biopharm* 55: 3–10.
- Wang CY et al. 1999. *N*-Acetyltransferase expression and DNA binding of *N*-hydroxyheterocyclic amines in human prostate epithelium. *Carcinogenesis* 20: 1591–1595.
- Wang L, Schultz PG. 2002. Expanding the genetic code. *Chem Commun (Camb.)* 1: 1–11.
- Wang L et al. 2005. Bioprocessing strategies to improve heterologous protein production in filamentous fungal fermentations. *Biotechnol Adv* 23(2): 115–129.
- Ward M et al. 1990. Improved production of chymosin in *Aspergillus* by expression as a glucoamylase-chymosin fusion. *Biotechnology (NY)* 8: 435–440.
- Ward PP et al. 1995. A system for production of commercial quantities of human lactoferrin: a broad spectrum natural antibiotic. *Biotechnology (NY)* 13: 498–503.
- Wardell AD et al. 1999. Characterization and mutational analysis of the helicase and NTPase activities of hepatitis C virus full-length NS3 protein. *J Gen Virol* 80(Pt 3): 701–709.
- Watson JV, Dive C. 1994. Enzyme kinetics. *Methods Cell Biol* 41: 469–507.
- Wells JA et al. 1987. Designing substrate specificity by protein engineering of electrostatic interactions. *Proc Natl Acad Sci USA* 84: 1219–1223.
- Werner RG et al. 2007. Glycosylation of therapeutic proteins in different production systems. *Acta Paediatr* 96(455): 17–22.
- Wharton CW. 1983. Some recent advances in enzyme kinetics. *Biochem Soc Trans* 11: 817–825.
- Wilkins MR et al. 1999. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* 112: 531–552.
- Williams GJ et al. 2004. Directed evolution of enzymes for biocatalysis and the life sciences. *Cell Mol Life Sci* 61: 3034–3046.
- Wolfenden R. 2003. Thermodynamic and extrathermodynamic requirements of enzyme catalysis. *Biophys Chem* 105: 559–572.
- Wong TS et al. 2006. The diversity challenge in directed protein evolution. *Comb Chem High Throughput Screen* 9(4): 271–288.
- Wu ZL et al. 2009. Enhanced bacterial expression of several mammalian cytochrome P450s by codon optimization and chaperone coexpression. *Biotechnol Lett* 31(10): 1589–1593.
- Xie S et al. 2009. Emerging affinity-based techniques in proteomics. *Expert Rev Proteomics* 6(5): 573–583.
- Xiong H et al. 2009. *E. coli* expression of a soluble, active single-chain antibody variable fragment containing a nuclear localization signal. *Protein Expr Purif* 66(2): 172–180.
- Xu C et al. 2002. CYP2A6 genetic variation and potential consequences. *Adv Drug Deliv Rev* 54: 1245–1256.
- Yang H et al. 2009. Recombinant human antithrombin expressed in the milk of non-transgenic goats exhibits high efficiency on rat DIC model. *J Thromb Thrombolysis* 28(4): 449–457.
- Yang P et al. 2008. Cattle mammary bioreactor generated by a novel procedure of transgenic cloning for large-scale production of functional human lactoferrin. *PLoS One* 3(10): e3453.
- Yen MR et al. 2009. Bioinformatic analyses of transmembrane transport: novel software for deducing protein phylogeny, topology, and evolution. *J Mol Microbiol Biotechnol* 17(4): 163–176.
- Yeo DS et al. 2004. Strategies for immobilization of biomolecules in a microarray. *Comb Chem High Throughput Screen* 7: 213–221.
- Yeung PK. 2000. Technology evaluation: transgenic antithrombin III (rhAT-III), Genzyme Transgenics. *Curr Opin Mol Ther* 2: 336–339.
- Yuan L et al. 2005. Laboratory-directed protein evolution. *Microbiol Mol Biol Rev* 69: 373–392.
- Yurimoto H, Sakai Y. 2009. Methanol-inducible gene expression and heterologous protein production in the methylotrophic yeast *Candida boidinii*. *Biotechnol Appl Biochem* 53(Pt 2): 85–92.
- Zaks A. 2001. Industrial biocatalysis. *Curr Opin Chem Biol* 5: 130–136.
- Zeng QK et al. 2009. Reversal of coenzyme specificity and improvement of catalytic efficiency of *Pichia stipitis* xylose reductase by rational site-directed mutagenesis. *Biotechnol Lett* 31(7): 1025–1029.
- Zerbs S et al. 2009. Bacterial systems for production of heterologous proteins. *Methods Enzymol* 463: 149–168.
- Zhang R et al. 2009. Ser67Asp and His68Asp substitutions in *Candida parapsilosis* carbonyl reductase alter the coenzyme specificity and enantioselectivity of ketone reduction. *Appl Environ Microbiol* 75(7): 2176–2183.
- Zhang Y et al. 2009a. Genetic variations in xenobiotic metabolic pathway genes, personal hair dye use, and risk of non-Hodgkin lymphoma. *Am J Epidemiol* 170: 1222–1230.
- Zhang Y et al. 2009b. Bioinformatics analysis of microarray data. *Methods Mol Biol* 573: 259–284.
- Zhang Z et al. 2004. A new strategy for the synthesis of glycoproteins. *Science* 303(5656): 371–373.
- Zhang ZJ et al. 2003. Interaction between polymorphisms of the dopamine D3 receptor and manganese superoxide dismutase genes in susceptibility to tardive dyskinesia. *Psychiatr Genet* 13: 187–192.
- Zhao P et al. 2010. Heterologous expression, purification, and biochemical characterization of a greenbug (*Schizaphis graminum*) acetylcholinesterase encoded by a paralogous gene (*ace-1*). *J Biochem Mol Toxicol* 24(1): 51–59.

- Zhao H et al. 1998. Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat Biotechnol* 16: 258–261.
- Zhuge J et al. 2004. Stable expression of human cytochrome P450 2D6*10 in HepG2 cells. *World J Gastroenterol* 10(2): 234–237.
- Zoonens M, Miroux B. 2010. Expression of membrane proteins at the *Escherichia coli* membrane for structural studies. *Methods Mol Biol* 601: 49–66.
- Zweiger G, Scott RW. 1997. From expressed sequence tags to ‘epigenomics’: an understanding of disease processes. *Curr Opin Biotechnol* 8: 684–687.