ALLOSTERIC REGULATORY ENZYMES

ALLOSTERIC REGULATORY ENZYMES

by

Thomas Traut

Department of Biochemistry and Biophysics
University of North Carolina
Chapel Hill, NC, USA

Thomas Traut
Department of Biochemistry and Biophysics
University of North Carolina
420 Mary Ellen Jones Bldg.
Chapel Hill, NC 27599-7260
USA

ISBN 978-0-387-72888-9 e-ISBN 978-0-387-72891-9

Library of Congress Control Number: 2007934262

© 2008 Springer Science+Business Media, LLC
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.
The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springer.com

To Karyn Traut

Author's Note

This book has benefited from the careful scrutiny of my colleagues, who offered insightful comments, thoughtful suggestions, and editorial assistance. For their support I am grateful to Richard Wolfenden, Gottfried Schroeder, Randy Stockbridge, Charles Lewis, Jr., and Brian Callahan. The drawing in Figure 4.2 was provided by Jason Traut.

CONTENTS

SECTION 1. OVERVIEW OF ENZYMES

SECTION 3. *V*-TYPE ENZYMES

SECTION 1

OVERVIEW OF ENZYMES

1

INTRODUCTION TO ENZYMES

Summary

All chemical reactions necessary for life are sufficiently slow that one or more unique enzyme catalysts are required to accelerate the reaction and make the needed product almost immediately available. Almost all enzymes are proteins that fold into domains. The majority of enzymes contains one domain (simple enzymes), while many are composed of two or more domains (allosteric enzymes and multifunctional proteins). Most enzymes are designed to function at a constant rate, but allosteric enzymes are sensitive to physiological controls, and thereby adjust their rate and determine the flux through the metabolic pathway that they control. There are two major groups of allosteric enzymes. One group is regulated by changing their affinity for one substrate, while keeping their maximum rate fairly constant (*K*-type enzymes). The second group also demonstrates significant changes in affinity, and in addition has large changes in the maximum rate (*V*-type enzymes). For cells to survive, natural selection has provided that each enzyme is always fast enough, with the slowest enzymes having a rate of $\geq 1\ s^{-1}$.

1.1 Introduction

All enzymes are remarkable for their ability to bind one or two substrates with appropriate specificity, and then facilitate a particular type of chemical reaction, producing one or more new products that are essential for the function of a living cell. Enzymes can be amazingly fast: for normal chemical reactions we have the example of a rate of greater than $10^6\ s^{-1}$ for catalase,[1] and for carbonic anhydrase.[2,3] Enzymes can perform exceedingly difficult reactions: for orotidine monophosphate (OMP) decarboxylase, the rate for the decarboxylation of OMP by the enzyme is 10^{17} faster than the spontaneous rate in the absence of enzyme.[4]

Over 5,000 different enzymes have been characterized, and almost all of these are proteins. If not stated otherwise, it will be assumed that any enzyme is a protein. A

limited number of catalytic reactions have been demonstrated with certain types of RNA molecules, and such catalytic RNAs are now called ribozymes.[5, 6] These first two types of enzymes are normal biological molecules that have evolved to have the features that make them so essential. Based on the properties of these two types of normal catalysts, scientists have explored how to make novel catalysts with DNA and antibodies. The first such DNAzyme was designed to cleave RNA molecules,[7] but no natural DNAzyme has as yet been observed. A limited number of artificial enzymes have also been made by manipulating antibodies to favorably bind a reactive intermediate for some chemical reaction.[8] Such catalytic antibodies are also known as *abzymes*,* and are a demonstration of scientific ingenuity, even though these artificial catalysts are as yet very modest in their catalytic rates.

1.1.1 Why are Enzymes Needed?

Living cells have successfully evolved by adapting to two opposing needs. Their molecules should be stable under most conditions, yet the cell must be able to modify molecules or make new molecules as conditions require this. The organic molecules that have become the basis for cellular metabolism and life must be sufficiently stable to serve as structural units, information storage, catalytic agents and perform various other functions during the lifetime of any cell. These molecules are therefore maintained by bonds that are fairly stable, and such molecules commonly display remarkably long stabilities of many years in an aqueous solution, such as the cytoplasm of a cell. For example, the halftime of hydrolysis ($t_{1/2}$) in aqueous solution is about 400 years for proteins and about 140,000 years for DNA.[9] By comparison RNA has a $t_{1/2}$ of only 4 years.[9] Therefore, except when attacked by some reactive species, most biological molecules are quite stable in their normal cellular environment. At the same time, cells must be dynamic, with the ability to make new proteins and other molecules, and dispose of old ones continuously, in order to be successful in whatever environment they inhabit. The success of living organisms depends on this ability to have a stable cellular environment, as well as catalytic enzymes that can be controlled as to when and how they modify and manipulate all the molecules in the cell.

The stability of a molecule, or its thermodynamic energy, is illustrated in Fig. 1.1. It is the height of this energy barrier $\Delta G^{\ddagger}$ that defines the stability or the reactivity of a molecule. While chemical reactions may be enhanced in the presence of an acid or alkaline solution, or by a metal cation, enzymes have the unique ability to bind molecules with sufficient affinity to transiently stabilize their transition state (denoted by $S^{\ddagger}$ in Fig. 1.1), which greatly reduces the energy barrier, and thereby makes the transition between S and P vastly more favorable. The magnitude of this rate enhancement has been measured for various types of chemical reactions. The catalytic rate of an enzymatic reaction (k_{cat}) is generally at least a billion times greater than the nonenzymatic uncatalyzed reaction (k_{non}), and examples of the remarkable rate enhancement of various enzymes have been defined by Richard Wolfenden and colleagues, and are shown in Fig. 1.2. The most dramatic examples are illustrated by arginine decarboxylase (ADC)

*A contraction from ab (abbreviation for antibody) and enzyme.

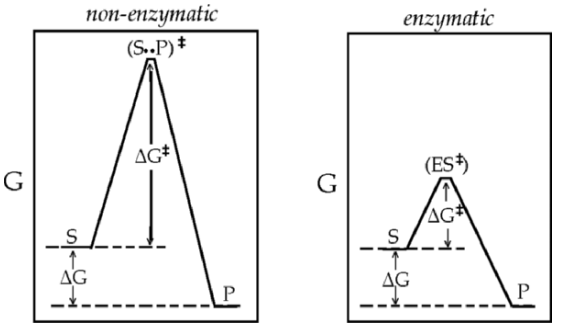

Fig. 1.1. An energy barrier ($\Delta G^{\ddagger}$) prevents the facile interconversion of S and P. An enzyme lowers this energy barrier by stabilizing the transition state between S and P, ES‡

and OMP decarboxylase (ODC). OMP decarboxylase is remarkable in that it has no cofactors to assist in this difficult reaction.[10] Orotidine-5′-monophosphate (OMP) is an intermediate in the biosynthesis of the pyrimidine nucleotide uridine-5′-monophosphate (UMP). OMP has a carboxyl group at carbon 6 of the pyrimidine base, and this must be removed to produce UMP. Without an enzyme to assist the decarboxylation, the elimination of this carboxyl group has a $t_{1/2}$ of 78 million years, demonstrating that this is a very stable bond.[4] The enzyme OMP decarboxylase performs this reaction about 25 times per second, providing a rate enhancement of 17 orders of magnitude.

An additional important point is also demonstrated by Fig. 1.2 with carbonic anhydrase (CAN). The hydration of carbon dioxide to form carbonic acid and bicarbonate is an extremely simple chemical reaction, and occurs with a $t_{1/2}$ of about 5 s in the absence of a catalyst. This spontaneous rate is still not fast enough for living organisms. The function of this enzyme is to hydrate carbon dioxide, a waste product of normal metabolism, and thereby produce carbonic acid, which spontaneously dissociates to bicarbonate, the major buffering agent in most organisms. Carbonic anhydrase performs this reaction in about 1 μs, and is therefore found in all organisms. Humans actually have 11 isozymes of carbonic anhydrase, expressed in our many different tissues.

1.1.2 Allosteric Enzymes

A simplified scheme for three metabolic pathways is illustrated in Fig. 1.3. Depending on various other factors, a specific cell will not need each of the three metabolic end products in equal amounts, at all times. It is therefore desirable to control how much of each of these products is actually made. This control function has evolved in the subset of enzymes known as allosteric regulatory enzymes.

A specific metabolic pathway, as shown in Fig. 1.3, normally includes 3–9 different enzymes in a sequential pathway dedicated to the synthesis of a single necessary molecule. Such a metabolic pathway may be viewed as a linear assembly line, in which each separate enzyme has a unique task in the sequential synthesis of the end product. The figure shows an example of a precursor compound, molecule A, which may be used for the synthesis of three different products, P, Q, and R. The cells' need for each of these

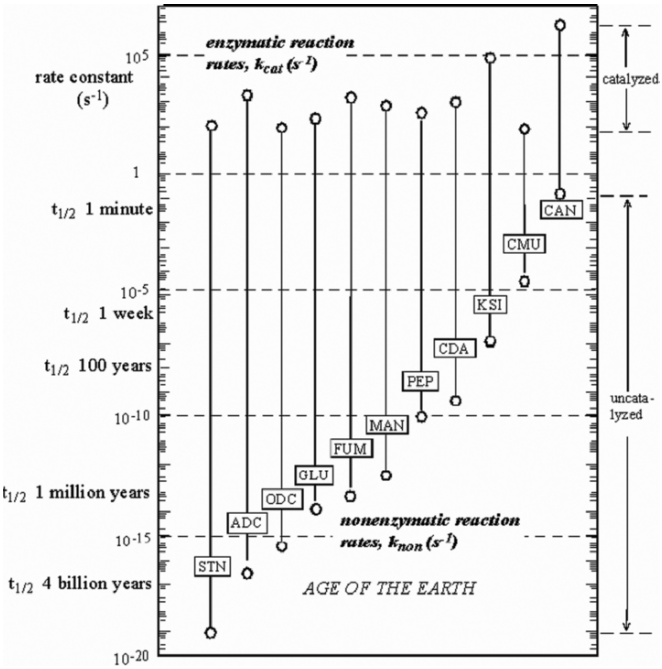

Fig. 1.2. Rate enhancements measured for various enzymes: ADC, arginine decarboxylase; CAN, carbonic anhydrase; CDA, cytidine deminase; CMU, chorismate mutase; FUM, fumarase; GLU, α-glucosidase; KSI, ketosteroid isomerase; MAN, mandelate racemase; ODC, OMP decarboxylase; PEP, carboxypeptidse B; STN, staphylococcal nuclease (figure courtesy of Richard Wolfenden)

final products may vary at different times, so that the three pathways have evolved to be independently regulated. The first enzyme that distinctly leads to that end product is normally the enzyme that commits the use of the substrate (B, in this figure) for the specific final product. Therefore, enzyme E_2 is the committed enzyme for the pathway leading to P. Enzyme E_2 is usually regulated by the end product, P. In this example, binding of compound P by E_2 would lead to this enzyme being inhibited, since this would occur only when P is at a high concentration, and its continued synthesis is no longer necessary. As the concentration of P becomes lower, since P is itself consumed over time, this inhibition diminishes, and the synthesis of P resumes. Such feedback inhibition by end products of the committed enzyme in a pathway is a standard feature in metabolism.

Enzymes that are able to be regulated by binding specific ligands are defined as allosteric (from the Greek: *allos* = other, and *stereos* = shape). This describes the key feature of such enzymes, their ability to change between two or more structural shapes that vary in their ability to bind a substrate, or in their ability to position a critical catalytic side chain, and therefore in their rate of catalysis. In Fig. 1.3, enzymes E_3 and E_4 would normally also be allosteric, but regulated by the end products Q and R.

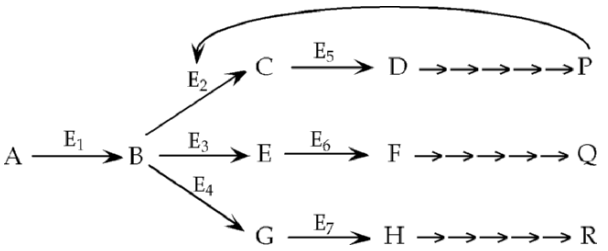

Fig. 1.3. A branched metabolic system. Enzyme E_2 is at the committed step for the synthesis of compound P. This end product normally acts as an allosteric inhibitor of the specific enzyme initiating the pathway for its specific synthesis

The typical examples presented show regulation by inhibition. It is common for these regulated allosteric enzymes to also respond in a positive fashion, with higher activity, to increased binding of a normal substrate, or an activator. It is by changing their rate, faster or slower, in response to changing concentrations of the specific cellular metabolites that these enzymes recognize and bind that enables such enzymes to be sensitive to some metabolic aspect of the cell. Since they respond by appropriately altering their activity, allosteric regulatory enzymes act as pacemakers for their pathway. We think of them as *regulatory*, since they regulate the pathway in which they function, and also because they are themselves regulated by the binding of physiological effectors.

Clearly, the needed feature for these pacemaker regulatory enzymes is that their activity or rate can be altered, and Nature has evolved two major strategies for regulating enzyme activity. Many enzymes are able to alter the affinity for their substrate, their K_m, with a conformational change. While such *K-type* enzymes have a fairly constant V_{max}, if at a fixed cellular concentration of substrate their affinity is made poorer, their rate must be slower, and as their affinity is improved, their rate will be faster. For the *V-type* enzymes the conformational change leads to a change in their affinity as well as in their maximum velocity. This may be accomplished by different means, such as the displacement, or the appropriate positioning of an important catalytic residue. Change in V_{max} may also occur by any factor that binds and sterically hinders access of the normal substrate to the catalytic site. Examples of these will be discussed in Chap. 1.3.

1.2 The Structures and Conformations of Proteins

1.2.1 Protein Conformations

A brief review of protein structure will help to explain enzyme binding sites, and the possibilities for allosteric effects. Most of the proteins in the cell, especially enzymes, normally fold so as to have an overall globular form. This comes naturally from the sequence of the protein, in which about one half of the amino acids are hydrophobic, and only when the protein folds so as to have these hydrophobic amino acids buried in the interior, away from the aqueous medium, will the form or structure of the protein be stable.

Each protein is a compact ensemble of secondary structures: the helices, beta strands, and loops that together comprise the total protein. While there is always an arrangement of these structural elements that is thermodynamically most favorable, variations from this most favored structure may not differ much in stability, so that most proteins are actually somewhat flexible, transiently converting into two or more somewhat similar structural shapes. Loops are especially mobile, and the opening and closing of loops at a catalytic binding site is frequently the rate-limiting feature for positioning of the substrate at the catalytic site. Binding sites are normally clefts or pockets in the surface of the protein, and may be formed by the proper positioning of adjacent secondary structure elements. Since proteins are flexible, the shape of the binding site may be transiently altered, as the overall shape of the protein varies. This feature provides the basis for regulation, by varying the fraction of the total enzyme population that has the correct shape or conformation to bind the desired substrate, and therefore the fraction of the total enzyme population that is competent to perform catalysis.

In all discussions about enzyme activity, and its regulation, it is important to think of each enzyme as a large population of molecules. Since enzymes generally have a cellular concentration above nanomolar, this denotes at least 10^9 enzyme molecules per microliter of cell volume for each specific enzyme. Never, under physiological conditions, will all of these molecules of the same enzyme have the same shape or conformation. The population will always include a mixture of several conformations or structural shapes, that is altered only by factors that may stabilize one of these conformational states, and thereby make it more abundant.[11,12] The illustration in Fig. 1.4, panel A shows the classical model of an allosteric enzyme that may have two conformations in the absence of a ligand, R and T. T is at a lower energy state and therefore the more stable and the more abundant form. For allosteric enzymes R represents the active form, while T is the less active or inactive form. In the absence of any ligands, T is normally the dominant species for *K*-type enzymes, while for *V*-type enzymes the dominant species may be either form, depending on the individual enzyme. The presence of a substrate, S, or an activator, A, will stabilize the R conformation, while an inhibitor, I, will stabilize the less active T conformation. Detailed examples of such allosteric features will be presented in later chapters.

Also, in Chap. 4 we will explore in greater detail the fact that all enzymes, whether allosteric or not, have multiple conformations. For the understanding of Fig. 1.4, the important point is that for normal enzymes there is only one active state under physiological conditions, and its abundance is not altered by any feature of the enzyme assay. Allosteric enzymes may often be represented by two conformations, since the key feature is the availability of regulatory effectors to bind to and stabilize the active or the inactive conformation. As the availability of the effectors changes, the distribution of the enzyme between the two principle conformations is changed, and this provides the basis for allosteric regulation.

Figure 1.4b illustrates the various equilibria[†] between these forms. In the absence of any ligands, the thermodynamic equilibrium favors conformation T, and therefore only a

[†]A true chemical equilibrium does not occur within cells, and a steady-state ratio of the two conformations is a more accurate description. The term *equilibrium* will be used since that is generally more convenient, in that it covers all simple chemical systems.

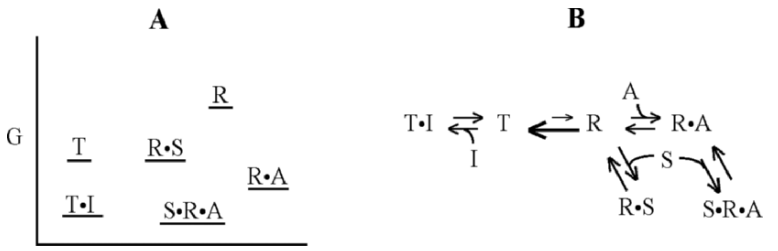

Fig. 1.4. Thermodynamic stability of enzyme conformations. G represents the free energy associated with any molecule. When proteins fold, they reach a stable tertiary structure that reflects their lowest free energy. T and R represent the inactive, and active forms of the enzyme. S, substrate; A, activator; I, inhibitor. The *dark arrow* in (**B**) emphasizes that allosteric enzymes will be proportionately more in the T form, since that is the more stable form. Note that ligands always stabilize (lower G) that conformation of the enzyme that binds the ligand

small fraction of the total enzyme population will be in the R conformation, which has better activity. Should the substrate become more abundant, it would bind to and stabilize the R conformation, making this species more abundant, and thereby increasing enzyme activity. This also demonstrates that the energy difference between these two conformations is very modest, since it cannot be greater than the binding energy of the substrate, which is normally in the range of 3–6 kcal/mol.

An activator that binds at a separate regulatory site would also increase the concentration of the active conformation. Overall, some of the enzyme molecules will always sample the different conformations, since energetically they are not that different. Depending on the enzyme, additional minor conformations may occur. Figure 1.4b is intended to illustrate the simplest system with only two conformations, though most proteins have more conformational states. However, if additional conformational states are not normally at a significant frequency, then the system may be simplified by considering only the conformations that are important for the observed enzyme activity.

It is important to note that the two conformations for active and inactive enzymes normally exist in the absence of regulatory effectors. The importance of such effectors is that they alter the equilibrium between the two conformations, and therefore alter the overall number of enzymes in the active conformation.

1.2.2 Protein Structures

The structure of a protein defines its function. A limited number of proteins form linear molecules, which serve as structural elements on a macromolecular scale. Silk and collagen are examples of such structural molecules that function in an extracellular environment, while myosin and fibroin are intracellular. Enzymes are almost always globular proteins, and they display a remarkable range of sizes, both for their subunits, and for the complete enzyme complex that many form. Protein structure is defined at four levels. The *primary* structure is the linear amino acid sequence of a protein; *secondary* structure defines the normal small structural elements such as alpha helices, beta strands, and loops; *tertiary* structure defines a single folded protein chain (equals a protein subunit); *quaternary* structure refers to the complex of two or more protein subunits.

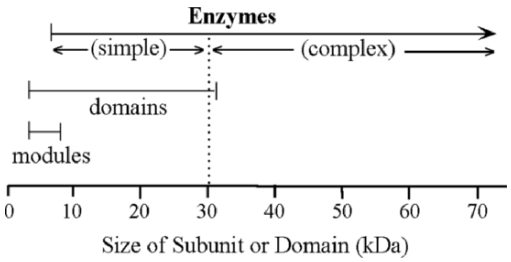

Fig. 1.5. Variation in the size of proteins, and in the size of structural components. The demarcation at 30 kDa is an approximation for single domain enzymes

Because of the large size range of proteins, illustrated in Fig. 1.5, additional terms have evolved to provide more specific descriptions about structure/function units within a protein. These are summarized in Table 1.1. It must be emphasized that currently there is no established consensus for the use of these terms. Different authors use these terms with somewhat distinct meanings, depending on what they wish to emphasize. In the following discussion, a protein's size or mass will always be for the single protein chain, or subunit, to avoid confusion with the size of large protein complexes. Enzymes that are proteins[‡] have sizes from as small as about 9 kDa for the HIV protease, and up to 565 kDa for the calcium channel in muscle cells.

In crystal structures of larger proteins (usually greater than 30 kDa), two or more distinct globular portions are frequently evident, and these are domains. Larger proteins always contain two or more domains. However, the term domain is also used to define a subcomponent of the protein by other criteria: the region that contains the catalytic site, or a portion of the protein that is easily cleaved by a protease, or the section of the protein that is involved in subunit contacts to form a dimer, and so forth. With an awareness of the different meanings associated with these terms, a reader can usually interpret the specific meaning by the context in which the term is used. At least half of all enzymes have a subunit mass of ≤30 kDa, and generally do not give evidence for containing more than

Table 1.1. Definitions for protein structural units

3° structure size	Term	Definition	M_r (kDa)
Large	Domain	Some subcomponent of total protein; "obviously" distinct	3–30
Small	Subdomain	Smaller local unit of 3° structure	3–20
	Module	Ligand binding unit Exon-coded unit	3–7
	Motif	An identified sequence associated with a specific structure/function	1.5–6

[‡]Proteins are extended chains of amino acids, and commonly when such chains are about 50 amino acids or less, they are defined as simple polypeptides, and begin to be called proteins as they become larger. There is no absolute size limit for the term protein.

one globular structural region. These are the simple enzymes shown in Fig. 1.5. Because the term domain have multiple definitions, domains overlap in size with simple enzymes, though they may occasionally be smaller. To help us with the discussion of protein evolution to follow, I will state that simple enzymes are one-domain proteins, and complex enzymes always contain two or more domains.

1.2.3 Multidomain Proteins

Ligand binding is one of the special functions of all enzyme domains, and when an enzyme has more than one domain, each domain commonly has a different ligand to bind. The term ligand (from the Latin *ligare* = to bind) includes all cellular metabolites that are substrates or effectors for enzymes, as well as macromolecules such as proteins, chromosomes, or membrane surfaces, to which enzymes may bind. It is a general feature that a protein's size is determined by how many ligand-binding sites it needs for its normal biological function. In other words, while enzymes may vary in size from 9 kDa to about 565 kDa, each enzyme is about the right size for its normal functions.

In the distribution of protein enzymes in simple bacteria we see that most of the enzymes are small. There is normally only one gene coding for each type of enzyme, but genes for enzymes that function together in a metabolic pathway are frequently clustered into an operon, a region of DNA that has the advantage that its genes are controlled by a single inducer region. When the gene for a catalytic subunit is adjacent to a second gene for a regulatory subunit that has the ability to bind to and alter the conformation of the catalytic subunit, then gene fusion can lead to these separate protein subunits becoming joined into a single protein subunit.

Gene fusion results when a termination signal at the end of the first gene is deleted or altered. Now, during transcription of this extended DNA segment the polymerase continues after the end of the reading frame for the first gene (no termination signal) and extends this RNA until it reaches the end of the second gene, producing a single mRNA that now codes for two domains, equivalent to the original two separate proteins. When the mRNA is translated, the two original proteins will no longer be separate proteins, but two domains that are joined by a short polypeptide chain encoded by the sequence of RNA between the two genes that was previously not transcribed or at least not translated. Naturally, to keep the second gene in the correct reading frame, the linker RNA region must contain 3*n* nucleotides. Also, the stop codon normally at the end of the first gene must be mutated to code for an amino acid, to assure continuity of the total polypeptide chain.

This simplest example would result in the formation of an allosteric regulatory enzyme, since it has combined the domain for a catalytic activity with the domain for binding regulatory effectors. By the same process, two or more genes for enzyme catalytic centers can become fused, if those genes are already sequential along a section of DNA.

1.2.3.1 Evolution of Multidomain Proteins

Gene duplication is a common event in most eukaryotes, and for living organisms in general it has been estimated that at least 50% of all genes were duplicated.[13] For humans,

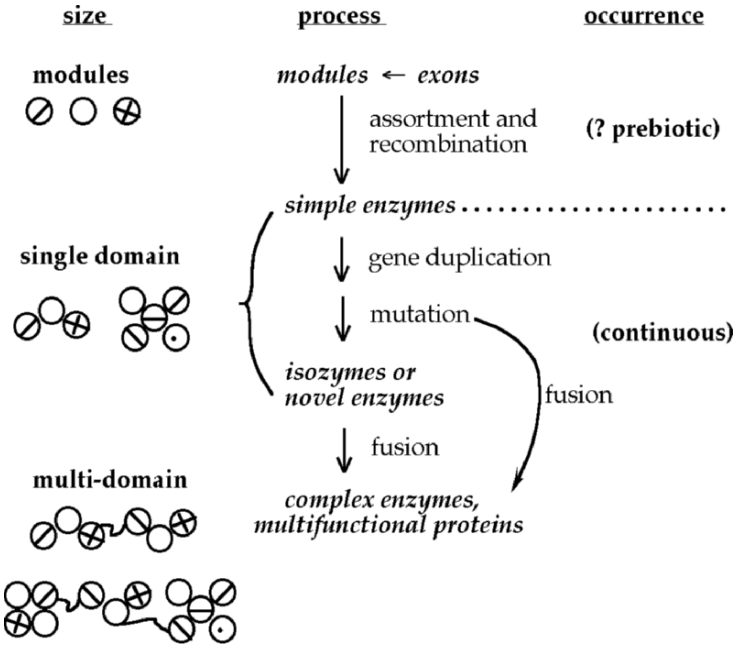

Fig. 1.6. Recombination of exon-coded modules, or larger DNA segments leads to various larger proteins

over 80% of our genes contain protein coding regions that are found in at least one other gene.[14] These extra copies of a gene (isogenes) may continue to code for essentially the same enzyme activity. However, since they initially are extra copies, then chance mutations may occur which modify the binding or catalytic rate of one of the duplicated enzymes, especially when this altered form of the same enzyme may also become preferentially expressed in a tissue or cell where the newer properties provide a benefit. Though the majority of such duplications are made nonfunctional by mutation, there are many examples of useful isozymes in mammals. It is this ability to benefit from mutations in extra genes that has led to new variants of the same catalytic function, or to important new enzyme activities.

An additional benefit of such duplications and recombinations is that these events may also be used in a new context, if they lead to the fusion of genes to produce proteins with two or more catalytic domains. These are known as multifunctional proteins, to emphasize that they contain more than one enzyme function. A simple scheme is illustrated in Fig. 1.6 to suggest how small protein modules, which normally are coded by a single exon, or larger domain-sized units may become fused to produce larger enzymes. In the majority of multifunctional proteins, the fused catalytic centers represent distinct different enzymes. There is frequently a ready comparison possible when the catalytic centers remain as individual enzymes in microbes, while having become fused into a single protein in higher eukaryotes.

One of the more dramatic examples of what is possible is given by the enzyme fatty acid synthetase. In bacteria eight genes code for eight different proteins,[15] of which seven have different catalytic activities that sequentially function to attach a two carbon acetyl group to a growing fatty acid chain, which is transiently anchored to the eighth member of this complex, the acyl carrier protein. Yeast give evidence that gene fusion has occurred, as yeast have only two genes, coding for five and three of the eight proteins for this metabolic sequence.[16] The fusion has become complete in mammals, as now a single gene codes for a single appropriately large protein containing all eight protein domains.[17] Since the seven catalytic activities of fatty acid synthetase work in concert on the acyl chain attached to the acyl carrier protein, but function sequentially, this organization of the eight distinct proteins in microbes into a single coherent protein in mammals represents greater efficiency.

The simplest system, and presumably the earliest version, is to produce eight proteins that function separately. Bacteria have partially improved on the simplest version, by having the separate proteins evolve sites to recognize and bind other members of this metabolic sequence, and then form a complex. Since all the enzyme subunits are required in comparable quantities for this entire assembly to be formed in a steady manner, fusion guarantees that each domain will be made in the same quantity as the others. By being linked into a single protein, each catalytic center is always present to optimize the steady and continued production of fatty acids.

Fatty acid synthetase is an example of the fusion of enzymes sequential in a pathway, and also arranged into a continuous group within an operon. There are also many examples of gene fusion of the same gene, when the second copy has occurred by gene duplication. Initially, this should only produce a protein with identical catalytic domains, for which there would be no clear benefit. But as with other isozymes, one of these two domains is now free to experience mutations, and if any beneficial mutation occurs this will then be selected. A spectrum of what is possible with the fusion of duplicated genes is in Table 1.2. With hexokinase II the duplicated domain retains the same activity, but one of the domains has an altered affinity for the same substrate, so that the duplicated enzyme now has a wider response range for the substrate than is normal for a single binding site.[18]

With carbamoyl-phosphate synthetase, the duplication of the carboxy kinase domain has led to modest changes so as to make it bind carbamate, which is formed from a product of the first domain, carboxy-phosphate.[19] With Hexokinase I the duplicated domain has lost catalytic activity, but now binds the product more tightly, and also still communicates with the active domain, so as to provide better inhibition by the product, glucose-6-phosphate.[20] Phosphofructokinase also uses the duplicated domain for several new regulatory sites to increase the number of effectors that control the rate of the active domain.[21] Finally, with the OPET decarboxylase/HHDD isomerase (Table 1.2) we have one example where sufficient changes were made in the duplicated domain to have it now perform a slightly altered type of chemical reaction. This is done without changing the normal ligand, since it isomerizes the product made by the first domain.[22]

For convenience, in the preceding discussion I have sometimes referred to changes occurring in the duplicated enzyme, or enzyme domain. When comparing such isozymes, we can only ascertain that they are highly related, and thereby derived from a common

Table 1.2. Possible new functions derived from gene duplication plus fusion

Added property	Enzymes	Function of new domain	Refs.
Same catalytic activity; extra substrate sensitivity	Hexokinase II	Same activity, but altered K_m for glucose	18
Same catalytic mechanism; different substrate	Carbamoyl-phosphate synthetase	A carboxy-kinase becomes a carbamate-kinase	19
New regulatory features	Hexokinase I	Inhibition by glucose-6-P	20
New catalytic activity	OPET decarboxylase/ HHDD isomerase[a]	New catalytic activity	22

[a]HHDD, 2-hydroxyhepta-2,4-diene-1,7-dioate; OPET, 5-oxopent-3-ene-1,2,5-tricarboxylate

ancestor. All the isogenes are equally susceptible to mutations. Once one of the isogenes develops beneficial mutations, it may replace the ancestor, or both may be altered and continue to function if their different attributes are beneficial and thereby maintained by natural selection.

1.2.3.2 Interaction Between Domains

The first successful model to describe cooperativity was presented in 1965 in a seminal paper by Monod, Wyman, and Changeux.[23] These authors introduced the simplest model, with two conformations for an enzyme, designated *R*, for the active conformation, and *T*, for the less active conformation. This model was quickly followed by a more extended treatment by Koshland, Némethy, and Filmer (KNF model)[24] who introduced the two visual depictions to represent the enzyme's conformation illustrated in Fig. 1.7, using a circle for *T* and a square for *R*. These visual shapes have now become icons for symbolizing the conformations of an enzyme, but they may be misleading in how readers visualize a protein structure. That is, the change from one conformation to the other suggests a complete alteration in the shape of the protein's subunit. When this model for allosteric regulation and cooperative conformational changes was presented in 1966, our knowledge of protein structures was still in its earliest phase, so that domains were not perceived as a general feature of protein structure.

In the ensuing years we have obtained thousands of protein structures, and a more advanced understanding of the details and variations that are possible. An important general finding is that conformational change in a protein is largely in the movement of

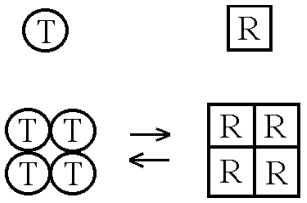

Fig. 1.7. The two classic conformational states introduced in the MWC model of Monod et al.[23] The visual icons for these conformations were introduced by Koshland et al.[24] Since allosteric enzymes are almost always oligomeric, an example of a tetramer is shown

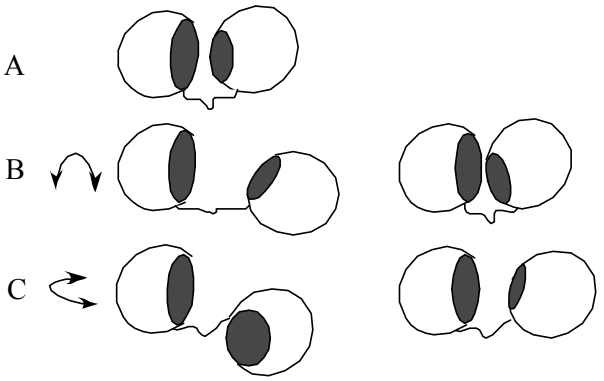

Fig. 1.8. Domain movement produces conformational change. In (**A**) the proper position of two domains forms the active binding site. Since the domains are connected by a short, but flexible polypeptide segment, then one domain may rotate within the plane of the figure, (**B**), to make the site too open or too closed. In (**C**) rotation perpendicular to the plane of (**B**) will again deform the proper shape of the binding site. Either domain may rotate without significantly changing its own conformation

one domain in a subunit relative to a connected domain (Fig. 1.8). This figure shows a normal binding pocket, such as a catalytic site, formed at the junction of two domains. To change the enzyme's ability to bind a ligand at this site, it needs only to be modified a little. This is most commonly done by rotation of one domain relative to its partner, since a short and sufficiently flexible polypeptide segment links them.

Since domains are comparable in size to entire subunits for smaller enzymes, a similar architecture in enzymes that are oligomeric is possible for the formation or disruption of a catalytic site. An example of this is illustrated in Fig. 1.9, where it is evident that the subunit by itself is not designed to have the binding pockets for its two substrates in an appropriate juxtaposition to facilitate catalysis. If these subunits normally join to form a dimer that is symmetric at the interface between the subunits, then now a complete catalytic pocket is formed across the dimer interface. Naturally, there will be two fully competent catalytic sites for the dimer, consistent with the almost universal stoichiometry of one catalytic site per subunit.

It is often possible to detect if enzymes have such binding sites that form between subunits, by exploring the activity of the enzyme as it dissociates from the oligomer to its subunits, as reviewed by Traut.[25] Important variables that influence enzyme dissociation include: enzyme concentration, ligand concentration, other cellular proteins, pH, and temperature. All these variables can be readily manipulated in vitro, but normally only the first two are physiological variables. The only constraints on how far the enzyme may be diluted come from the inherent activity of the enzyme, and the sensitivity of the enzyme assay. If very dilute enzyme can still form a small amount of product during the assay time, this must be enough product to be detectable, or an incorrect conclusion will be made as to whether the dissociated subunit is still active. Despite these constraints, more than 40 enzymes have now shown a change in activity as a function of their oligomeric state.[25] No single database has compiled all the enzymes known to show the feature illustrated in Fig. 1.9, but this feature is likely to be more widely used.