Proteins as Enzymes

Sajad A. Rather, F. A. Masoodi, Jahangir A. Rather, Tariq A. Ganaie, Rehana Akhter, and S. M. Wani

Introduction

Living cells have the cell factories operate as a collection of efficient molecular characteristics. The success of these factories depends on the efficiency of a particular class of biomolecules-protein enzymes (Agarwal 2006). Enzymes are the complex protein molecules that catalyze chemical reactions, i.e. transformations from one or more substrates to one or more products (Bugg 2004). An integrated view of protein structure, dynamics and function is emerging, where proteins are considered as dynamically active machines and internal protein motions are closely linked to function such as enzyme catalysis (Agarwal 2006). Enzymes exhibit the physicochemical properties including solubility, electrophoretic properties, electrolytic behaviors and chemical reactivity of proteins (Lee 2006; Bhatia 2018). The sequence of amino acid of an enzyme also called as primary structure of enzyme plays an important role in enzyme function including substrate/cofactor binding or release (Yadav and Tiwari 2015). Thus the degree of biocatalytic activity chiefly depends on the integrity of the enzymes structure as a protein. The complete biochemically active enzyme is composed of a protein part (apoenzyme) with a co-enzyme or a metal ion and is called a holoenzyme. The co-enzyme in the enzyme structure may bind covalently or non-covalently to the apoenzyme. When the co-enzyme is tightly and permanently bound to protein part (apoenzyme) in this case it is known as a prosthetic group.

$$\underset{(\text{Protein})}{\text{Apoenzyme}} + \underset{(\text{Non-protein})}{\text{Prosthetic group}} = \underset{(\text{Complete Enzyme})}{\text{Holoenzyme}}$$

S. A. Rather · F. A. Masoodi (✉) · J. A. Rather · T. A. Ganaie · R. Akhter · S. M. Wani
Department of Food Science and Technology, University of Kashmir, Srinagar, India

A. Gani, B. A. Ashwar (eds.), *Food biopolymers: Structural, functional and nutraceutical properties*, https://doi.org/10.1007/978-3-030-27061-2_13

An International Commission on enzymes was established by the International Union of Biochemistry [now termed the International Union of Biochemistry and Molecular Biology (IUBMB)] in 1956 to address the problems of enzyme classification and nomenclature based on the overall chemical transformation they catalyze. Enzymes are now named and classified systematically with an EC number to a four level hierarchical description depending on the overall chemical transformation of substrates into products (Cuesta et al. 2015). The EC classification is still made on the basis of the main reaction catalyzed. The EC denotes the six classes of enzymes based on general type of reaction being carried out including (EC-1) oxidoreductases, (EC-2) transferases, (EC-3) hydrolases, (EC-4) lyases, (EC-5) isomerases, and (EC-6) ligases, where EC stands for Enzyme Commission (Kumar et al. 2015).

Enzyme function is intrinsically linked to its structure, determining how it performs substrate binding, catalysis and regulation. The amino acid-based enzymes are globular proteins that range in size from <100 to >2000 amino acid residues. These amino acids can be arranged into polypeptide chains that are folded and bent to form a specific three-dimensional structure (Robinson 2015). Some of the amino acids in enzymes are involved in binding ligands (substrates, intermediates, products, organic cofactors, metal cofactors or allosteric regulators) and some are actively involved in catalysis by interacting with the substrate, intermediate or product of the reaction (Soding et al. 2005). The structures of enzymes can be elucidating by techniques such as spectroscopic methods, X-ray crystallography and more recently, multidimensional NMR methods. The X-ray crystallography has been the most widely used technique for structural characterization of enzymes. The first enzyme to be crystallized and its structure successfully solved was chicken egg lysozyme in 1965. NMR spectroscopy is a powerful tool for elucidating the structure–function relationships of enzymes. It yields detailed information regarding structure of enzyme and the specific ligands which bind to the enzyme. The structure of the ligands at the binding sites of enzymes and the structure of enzyme–ligand complexes can also be obtained, as well as the dynamics of the ligand and the associated structure of the protein binding site (Monasterio 2014). The aim of this chapter is to present and update the existing knowledge about basic principles of enzymes such as proteinaceous nature and substrate binding, detailed description of the enzyme classification and structural characterization.

Proteinaceous Nature of Enzymes and Substrate Binding

All enzymes are proteins made up of amino acids linked together by peptide bonds except small group of RNAase molecules (Bhatia 2018). The structure and reactivity of a protein depends its amino acid sequence, called primary structure, which is genetically determined by the deoxyribonucleotide sequence in the structural gene that codes for it (Illanes 2008). The deoxyribonucleotide sequence is transcribed into a mRNA molecule. The mRNA molecule upon reaching the ribosome of cell is

translated into an amino acid sequence and synthesizes a polypeptide chain. The polypeptide chain is finally transformed into a three dimensional structure, called native structure, which is having the biological functionality (Schumacher et al. 1986; Longo and Combes 1999). The secondary three-dimensional structure is the result of interactions of amino acid residues in the primary structure, mainly by hydrogen bonding of the amide groups. For the globular proteins, like enzymes, these interactions dictate a predominantly ribbon-like coiled configuration termed ɣ-helix. The tertiary three-dimensional structure is the result of interactions of amino acid residues located apart in the primary structure that produce a compact and twisted configuration in which the surface is rich in polar amino acid residues, while the inner part is abundant in hydrophobic amino acid residues. This tertiary structure is essential for the biological functionality of the protein. Some proteins have a quaternary three-dimensional structure, which is common in regulatory proteins, that is the result of the interaction of different polypeptide chains constituting subunits that can display identical or different functions within a protein complex (Dixon and Webb 1979; Creighton 1993).

In enzymes, proteins (apoenzyme) can be conjugated or associated with other molecules like, co-enzyme or co-factor or a prosthetic group (Fig. 1). However catalysis always occurs in the protein portion of an enzyme. The co-enzyme in the enzyme structure may bind covalently or noncovalently to the apoenzyme. When the co-enzyme is tightly and permanently bound to protein part (apoenzyme) in this case it is known as a prosthetic group (Yadav and Tiwari 2015). Prosthetic groups may be organic macromolecules, like carbohydrates (glycoproteins), lipids (lipoproteins) and nucleic acids (nucleoproteins), or simple in organic entities, like metalions. Prosthetic groups are tightly bound (usually covalently) to the apoenzyme and do not dissociate during catalysis (Union of Pure and Applied Chemistry

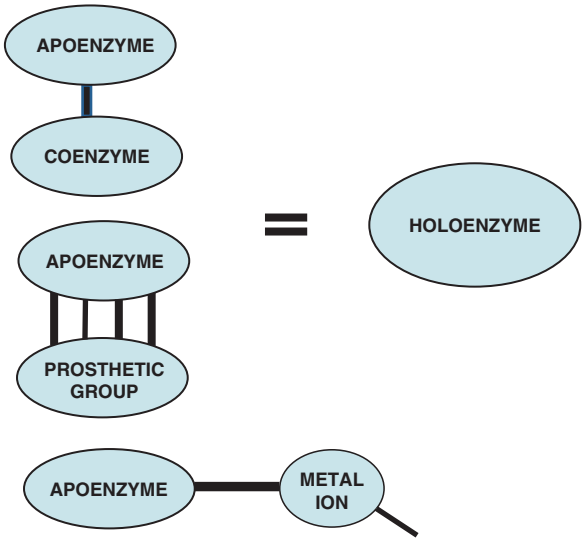

Fig. 1 The components of a holoenzyme

2005–2009; Illanes 2008). Although there are also prosthetic groups that are not cofactors (e.g. retinal in light receptors), only those prosthetic groups that are located in the active site of an enzyme are denoted cofactors. Therefore a prosthetic group is distinguished from a coenzyme in that it stays with the enzyme over many catalytic cycles, possibly until the enzyme is degraded. The coenzyme, on the other hand, binds to the enzyme at the beginning of each catalytic cycle and leaves at the end of it (Union of Pure and Applied Chemistry 2005–2009).

Small portion of the enzyme (active site) is involved in catalysis which is usually formed by very few amino acid residues. In enzymatic reaction substrate binds to the enzyme at the active site and produces changes in the distribution of electrons in its chemical bonds which lead to the reactions that result to the formation of products. The products formed are then released from the enzyme and is ready for the next catalytic cycle. It is the shape and charge properties of the active site of enzyme which enable it to bind to a specific substrate molecule, and demonstrate it specificity in catalytic activity (Whitehurst and van Oort 2009). According to the early lock and key hypothesis proposed by the German chemist Emil Fischer in 1894, the active site has a unique geometric shape that is complementary to the geometric shape of the substrate molecule that fits into it. However this rigid hypothesis hardly explains many experimental evidences of enzyme biocatalysis (Sonkaria et al. 2004). Later on through some techniques such as X-ray crystallography, it became clear that enzymes are quite flexible but not rigid structures. In the light of this finding, induced-fit theory was proposed by Daniel Koshland in 1958 according to which the substrate induces a change in the enzyme conformation after binding that may orient the catalytic groups in a way prone for the subsequent reaction. This theory has been extensively used to explain enzyme catalysis (Yousef et al. 2003). Since, it is the active site alone that binds to the substrate. The rest of protein acts to stabilize the active site and provide an appropriate environment for interaction of the site with the substrate molecule (Robinson 2015). According to the transition-state theory, enzyme catalysis is the transition state complementariness, which considers the preferential binding of the transition state rather than the substrate or product (Benkovic and Hammes-Schiffer 2003).

Classification of Enzymes

Classifying enzymes in different groups based on the type of reaction they catalyze is a possible way to gain an understanding of the bonds they create or break. Classification of enzymes is developing constantly and one current issue is that the recommendations for enzyme classification and nomenclature are inappropriate for several enzyme groups (e.g. carbohydrate-active enzymes), especially in case of enzymes with multiple substrate specificity and for isoenzymes. The enzyme classification system is being constantly updated with new enzymes or corrections to existing entries and the details of recommendations for enzyme classification are provided. Because of the growing complexity in the naming of enzymes, the

International Union of Biochemistry [now termed the International Union of Biochemistry and Molecular Biology (IUBMB)] set up the Enzyme Commission (EC) for providing a systematic approach to the naming of enzymes and published first report in 1961. The sixth edition, published in 1992, contained details of nearly 3200 different enzymes, and supplements published annually have now extended this number to over 5000 (Robinson 2015). The E.C. number classification is a four level hierarchical system of an enzyme's overall reaction or function. The E.C. first level corresponds to six classes according to the type of reaction being carried out includes oxidoreductases catalyze oxidation/reduction reactions (EC 1), transferases transfer a chemical group (EC 2), hydrolases perform hydrolysis of chemical bonds (EC 3), lyases also cleave chemical bonds by other means than by oxidation or hydrolysis (EC 4), isomerases catalyze geometric and structural changes between isomers (EC 5), and ligases joins two compounds with associated hydrolysis of a nucleoside triphosphate molecule (EC6). The next two classification levels are subclass and sub-sub-class (level 2 and level 3) depends on a various criteria such as chemical bond cleaved or formed, the reaction center, the transferred chemical group or the cofactor used for catalysis. The final level (fourth) gives a serial number for each enzyme reaction, substrate specificity. One E.C. number denotes an overall chemical reaction of an enzyme. Thus, several enzymes, which may be non-homologous, may be identified by the same E.C. number if they catalyze the same overall reaction. For example, the enzyme with the trivial name lactate dehydrogenase has the EC number 1.1.1.27, is an oxidoreductase (indicated by the first digit) with the alcohol group of the lactate molecule as the hydrogen donor (second digit) and NAD+ as the hydrogen acceptor (third digit), and is the 27th enzyme to be categorized within this group (fourth digit). The basic E.C. number classification layout of enzymes is described in Table 1.

The EC classification is still made on the basis of main reaction being catalyzed (Cuesta et al. 2015). Nowadays the assignment of EC numbers to enzyme is a common routine in the functional annotation of proteins and protein-coding genes in databases such as UniprotKB (UniProt Consortium 2013) and Ensembl (Kersey et al. 2014) and has been adopted by the widely uses Gene Ontology (GO) (Ashburner et al. 2000). However possible changes between EC classes are observed. There are some preferences such as transferases (EC 2) becoming oxidoreductases (EC 1), hydrolases (EC 3) and lyases (EC 4) (Martınez Cuesta et al. 2014). Exchanges between different EC classes suggest that the chemistry of enzymes is more complex than previously classified with close relationships between enzymes with radically different EC numbers. The substrate specificity of enzyme is represented by the last digit of the EC number, while the first three digits describe the type of the reaction. In case the sequence identity is below 70%, all the four digits of the EC number start to diverge quickly (Rost 2002). This creates an urgent need to choose alternative methods to sub-group enzymes that reflects their function or substrate specificity. The chemistry of related enzyme functions can now be explored using robust computational approaches like EC-BLAST (Rahman et al. 2014). This tool searches and compares reactions on the basis of bond charges, reaction centers, and structures of substrates and products (Cuesta et al. 2015; Rausch et al. 2005).

Table 1 The E.C. classification layout of enzymes according to the IUBMB enzyme nomenclature

Class	Sub-class	sub-sub-class	Reaction type
EC 1: Oxidoreductases			
	EC 1.1		Acting on the CH–OH group of donors
		EC 1.1.1	NAD or NADP as acceptor
		EC 1.1.2	Cytochrome as acceptor
		EC 1.1.3	Oxygen as acceptor
		EC 1.1.4	Disulfide as acceptor
		EC 1.1.5	Quinine or similar compound as acceptor
		EC 1.1.99	Other acceptors
	EC 1.2		Acting on the aldehyde or oxo group of donors
		EC 1.2.1	NAD or NADP as acceptor
		EC 1.2.2	Cytochrome as acceptor
		EC 1.2.3	Oxygen as acceptor
		EC 1.2.4	Disulfide as acceptor
		EC 1.2.7	Iron–sulfur protein as acceptor
		EC 1.2.99	Other acceptors
	EC 1.3		Acting on the CH–CH group of donors
		EC 1.3.1	NAD or NADP as acceptor
		EC 1.3.2	Cytochrome as acceptor
		EC 1.3.3	Oxygen as acceptor
		EC 1.3.5	Quinine or similar compound as acceptor
		EC 1.3.7	Iron–sulfur protein as acceptor
		EC 1.3.99	Other acceptors
	EC 1.4		Acting on the CH–NH_2 group of donor
		EC 1.4.1	NAD or NADP as acceptor
		EC 1.4.2	Cytochrome as acceptor
		EC 1.4.3	Oxygen as acceptor
		EC 1.4.4	Disulfide as acceptor
		EC 1.4.7	Iron–sulfur protein as acceptor
		EC 1.4.99	Other acceptors
	EC 1.5		Acting on the CH–NH group of donors
		EC 1.5.1	NAD or NADP as acceptor
		EC 1.5.3	Oxygen as acceptor
		EC 1.5.4	Disulfide as acceptor
		EC 1.5.5	Quinine or similar compound as acceptor
		EC 1.5.99	Other acceptors
	EC 1.6		Acting on NADH or NADPH
		EC 1.6.1	NAD or NADP as acceptor
		EC 1.6.2	Cytochrome as acceptor
		EC 1.6.4	Disulfide as acceptor
		EC 1.6.5	Quinine or similar compound as acceptor
		EC 1.6.6	Nitrogenous group as acceptor

(continued)

Table 1 (continued)

Class	Sub-class	sub-sub-class	Reaction type
		EC 1.6.8	Flavin as acceptor
		EC 1.6.99	Other acceptors
	EC 1.7		Acting on other nitrogenous compounds as donors
		EC 1.7.2	Cytochrome as acceptor
		EC 1.7.3	Oxygen as acceptor
		EC 1.7.7	Iron–sulfur protein as acceptor
		EC 1.7.99	Other acceptors
	EC 1.8		Acting on a sulfur group of donors
		EC 1.8.1	NAD or NADP as acceptor
		EC 1.8.2	Cytochrome as acceptor
		EC 1.8.3	Oxygen as acceptor
		EC 1.8.4	Disulfide as acceptor
		EC 1.8.5	Quinine or similar compound as acceptor
		EC 1.8.7	Iron–sulfur protein as acceptor
		EC 1.8.99	Other acceptors
	EC 1.9		Acting on a heme group of donors
		EC 1.9.3	Oxygen as acceptor
		EC 1.9.6	Nitrogenous group as acceptor
		EC 1.9.99	Other acceptors
	EC 1.10		Diphenols and related substances as donors
		EC 1.10.1	NAD or NADP as acceptor
		EC 1.10.2	Cytochrome as acceptor
		EC 1.10.3	Oxygen as acceptor
		EC 1.10.99	Other acceptors
	EC 1.11		Acting on a peroxide as acceptor
		EC 1.11.1	NAD or NADP as acceptor
	EC 1.12		Acting on hydrogen as donor
		EC 1.12.1	NAD or NADP as acceptor
		EC 1.12.2	Cytochrome as acceptor
		EC 1.12.99	Other acceptors
	EC 1.13		Acting on single donors with incorporation of molecular oxygen (oxygenases)
		EC 1.13.11	Incorporation of two atoms of oxygen
		EC 1.13.12	Incorporation of one atom of oxygen

(continued)

Table 1 (continued)

Class	Sub-class	sub-sub-class	Reaction type
		EC 1.13.99	Miscellaneous
	EC 1.14		Acting on paired donors, with incorporation or reduction of molecular oxygen
		EC 1.14.11	2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors
		EC 1.14.12	NADH2 or NADPH2 as one donor, and incorporation of two atoms of oxygen into one donor
		EC 1.14.13	NADH2 or NADPH2 as one donor, and incorporation of one atom of oxygen
		EC 1.14.14	Reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen
		EC 1.14.15	Reduced iron–sulfur protein as one donor, and incorporation of one atom of oxygen
		EC 1.14.16	Reduced pteridine as one donor, and incorporation of one atom of oxygen
		EC 1.14.17	Ascorbate as one donor, and incorporation of one atom of oxygen
		EC 1.14.18	Another compound as one donor, and incorporation of one atom of oxygen
		EC 1.14.99	Miscellaneous
	EC 1.15		Acting on superoxide radicals as acceptor
	EC 1.16		Oxidising metal ions
		EC 1.16.1	NAD or NADP as acceptor
		EC 1.16.3	Oxygen as acceptor
	EC 1.17		Acting on CH or CH_2 groups
		EC 1.17.1	NAD or NADP as acceptor
		EC 1.17.3	Oxygen as acceptor
		EC 1.17.4	Disulfide as acceptor
		EC 1.17.99	Other acceptors
	EC 1.18		Acting on iron-sulfur proteins as donors
		EC 1.18.1	NAD or NADP as acceptor
		EC 1.18.6	Dinitrogen as acceptor
		EC 1.18.99	H+ as acceptor
	EC 1.19		Acting on reduced flavodoxin as donor
		EC 1.19.6	Dinitrogen as acceptor

(continued)

Table 1 (continued)

Class	Sub-class	sub-sub-class	Reaction type
	EC 1.20		Acting on phosphorus or arsenic in donors
	EC 1.21		Acting on X–H and Y–H to form an X–Y bond
	EC 1.22		Acting on halogen in donors
	EC 1.97		Other oxidoreductases
EC 2: Transferases			
	EC 2.1		Transferring one-carbon groups
		EC 2.1.1	Methyltransferases
		EC 2.1.2	Hydroxymethyl-, formyl- and related transferases
		EC 2.1.3	Carboxyl- and carbamoyltransferases
		EC 2.1.4	Amidinotransferases
	EC 2.2		Transferring aldehyde or ketonic groups
		EC 2.2.1	a single subclass containing the transaldolases
	EC 2.3		Acyltransferases
		EC 2.3.1	Acyltransferases
		EC 2.3.2	Aminoacyltransferases
	EC 2.4		Glycosyltransferases
		EC 2.4.1	Hexosyltransferases
		EC 2.4.2	Pentosyltransferases
		EC 2.4.99	Transferring other glycosyl groups
	EC 2.5		Transferring alkyl or aryl groups, other than methyl groups
		EC 2.5.1	A single subclass that includes a rather mixed group of such enzymes
	EC 2.6		Transferring nitrogenous groups
		EC 2.6.1	Transaminases (aminotransferases)
		EC 2.6.3	Oximinotransferases
		EC 2.6.99	Transferring other nitrogenous groups
	EC 2.7		Transferring phosphorus-containing groups
		EC 2.7.1	Phosphotransferases with an alcohol group as acceptor
		EC 2.7.2;	Phosphotransferases with a carboxyl group as acceptor
		EC 2.7.3	Phosphotransferases with a nitrogenous group as acceptor
		EC 2.7.4	Phosphotransferases with a phosphate group as acceptor
		EC 2.7.6	Diphosphotransferases
		EC 2.7.7	Nucleotidyltransferases
		EC 2.7.8	Transferases for other substituted phosphate groups
		EC 2.7.9	Phosphotransferases with paired acceptors
	EC 2.8		Transferring sulfur-containing groups
		EC 2.8.1	Sulfurtransferases
		EC 2.8.2	Sulfotransferases

(continued)

Table 1 (continued)

Class	Sub-class	sub-sub-class	Reaction type
		EC 2.8.3	CoA-transferases
	EC 2.9		Transferring selenium-containing groups
	EC 2.10		Transferring molybdenum- or tungsten-containing groups
EC 3: Hydrolases			
	EC 3.1		Acting on ester bonds
		EC 3.1.1	Carboxylic ester hydrolases
		EC 3.1.2	Thiolester hydrolases
		EC 3.1.3	Phosphoric monoester hydrolases
		EC 3.1.4	Phosphoric diester hydrolases
		EC 3.1.5	Triphosphoric monoester hydrolases
		EC 3.1.6	Sulfuric ester hydrolases
		EC 3.1.7	Diphosphoric monoester hydrolases
		EC 3.1.8	Phosphoric triester hydrolases
		EC 3.1.11	Exodeoxyribonucleases producing 5′-phosphomonoesters
		EC 3.1.13	Exoribonucleases producing 5′-phosphomonoesters
		EC 3.1.14	Exoribonucleases producing other than 5′-phosphomonoesters
		EC 3.1.15	Exonucleases active with either ribo- or deoxyribonucleic acids and producing 5′-phosphomonoesters
		EC 3.1.16	Exonucleases active with either ribo- or deoxyribonucleic acids and producing other than 5′-phosphomonoesters
		EC 3.1.21	Endodeoxyribonucleases producing 5′-phosphomonoesters
		EC 3.1.22	Endodeoxyribonucleases producing other than 5′-phosphomonoesters
		EC 3.1.25	Site-specific endodeoxyribonucleases specific for altered bases
		EC 3.1.26	Endoribonucleases producing 5′-phosphomonoesters
		EC 3.1.27	Endoribonucleases producing other than 5′-phosphomonoesters
		EC 3.1.30	Endonucleases active with either ribo- or deoxyribonucleic acid and producing 5′-phosphomonoesters
		EC 3.1.31	Endonucleases active with either ribo- or deoxyribonucleic acid and producing other than 5′-phosphomonoesters
	EC 3.2		Glycosidases
		EC 3.2.1	Hydrolysing O-glycosyl compounds
		EC 3.2.2	Hydrolysing N-glycosyl compounds
		EC 3.2.3	Hydrolysing S-glycosyl compounds
	EC 3.3		Acting on ether bonds
		EC 3.3.1	Thioether hydrolases
		EC 3.3.2	Ether hydrolases
	EC 3.4		Acting on peptide bonds (peptidases)
		EC 3.4.11	Aminopeptidases
		EC 3.4.13	Dipeptidases

(continued)

Table 1 (continued)

Class	Sub-class	sub-sub-class	Reaction type
		EC 3.4.14	Dipeptidyl-peptidases and tripeptidyl-peptidases
		EC 3.4.15	Peptidyl-dipeptidases
		EC 3.4.16	Serine-type carboxypeptidases
		EC 3.4.17	Metallocarboxypeptidases
		EC 3.4.18	Cysteine-type carboxypeptidases
		EC 3.4.19	Omega peptidases
		EC 3.4.21	Serine endopeptidases
		EC 3.4.22	Cysteine endopeptidases
		EC 3.4.23	Aspartic endopeptidases
		EC 3.4.24	Metalloendopeptidases
		EC 3.4.99	Endopeptidases of unknown catalytic mechanism
	EC 3.5		Acting on carbon-nitrogen bonds, other than peptide bonds
		EC 3.5.1	In linear amides
		EC 3.5.2	In cyclic amides
		EC 3.5.3	In linear amidines
		EC 3.5.4	In cyclic amidines
		EC 3.5.5	In nitriles
		EC 3.5.99	In other compounds
	EC 3.6		Acting on acid anhydrides
		EC 3.6.1	In phosphorus-containing anhydrides
		EC 3.6.2	In sulfonyl-containing anhydrides
	EC 3.7		Acting on carbon-carbon bonds
		EC 3.7.1	In ketonic substances
	EC 3.8		Acting on halide bonds
		EC 3.8.1	In C-halide compounds
	EC 3.9		Acting on phosphorus-nitrogen bonds
	EC 3.10		Acting on sulfur-nitrogen bonds
	EC 3.11		Acting on carbon-phosphorus bonds
	EC 3.12		Acting on sulfur-sulfur bonds
EC 4: Lyases			
	EC 4.1		Carbon-carbon lyases
		EC 4.1.1	Carboxy-lyases
		EC 4.1.2	Aldehyde-lyases
		EC 4.1.3	Oxo-acid-lyases
		EC 4.1.99	Other carbon–carbon lyases
	EC 4.2		Carbon-oxygen lyases
		EC 4.2.1	Hydro-lyases
		EC 4.2.2	Acting on polysaccharides

(continued)

Table 1 (continued)

Class	Sub-class	sub-sub-class	Reaction type
		EC 4.2.99	Other carbon–oxygen lyase
	EC 4.3		Carbon-nitrogen lyases
		EC 4.3.1	Ammonia-lyases
		EC 4.3.2	Amidine-lyases
		EC 4.3.3	Amine-lyases
		EC 4.3.99	Other carbon–nitrogen-lyases
	EC 4.4		Carbon-sulfur lyases
	EC 4.5		Carbon-halide lyases
	EC 4.6		Phosphorus-oxygen lyases
	EC 4.99		Other lyases
EC 5: Isomerases			
	EC 5.1		Racemases and epimerases
		EC 5.1.1	Acting on amino acids and derivatives
		EC 5.1.2	Acting on hydroxy acids and derivatives
		EC 5.1.3	Acting on carbohydrates and derivatives
		EC 5.1.99	Acting on other compounds
	EC 5.2		cis-trans-Isomerases
	EC 5.3		Intramolecular isomerases
		EC 5.3.1	Interconverting aldoses and ketoses
		EC 5.3.2	Interconverting keto- and enol-groups
		EC 5.3.3	Transposing C=C bonds
		EC 5.3.4	Transposing S–S bonds
		EC 5.3.99	Other intramolecular oxidoreductases
	EC 5.4		Intramolecular transferases (mutases)
		EC 5.4.1	Transferring acyl groups
		EC 5.4.2	Phosphotransferases (phosphomutases)
		EC 5.4.3	Transferring amino groups
		EC 5.4.99	Transferring other groups
	EC 5.5		Intramolecular lyases
	EC 5.99		Other isomerases
EC 6: Ligases			
	EC 6.1		Forming carbon—oxygen bonds
		EC 6.1.1	Ligases forming aminoacyl-tRNA and related compounds
	EC 6.2		Forming carbon—sulfur bonds
		EC 6.2.1	Acid–thiol ligases
	EC 6.3		Forming carbon—nitrogen bonds
		EC 6.3.1	Acid–ammonia (or amine) ligases (amide synthases)
		EC 6.3.2	Acid–amino-acid ligases (peptide synthases)
		EC 6.3.3	Cyclo-ligases

(continued)

Table 1 (continued)

Class	Sub-class	sub-sub-class	Reaction type
		EC 6.3.4	Other carbon–nitrogen ligases
		EC 6.3.5	Carbon–nitrogen ligases with glutamine as amido-N-donor
	EC 6.4		Forming carbon—carbon bonds
	EC 6.5		Forming phosphoric ester bonds
	EC 6.6		Forming nitrogen—metal bonds

Table 2 Enzyme classification attempts based on sequence similarity, structural similarity and protein descriptors

Method	Feature used	Classification accuracy/result	References
BLAST, FASTA	Sequence information	40% of enzyme classes predicted correctly	Shah and Hunter (1997)
BLAST	Sequence information	Found putative analogy of 40.5% for all EC classes	Audit et al. (2007)
Bayesian	Structural information	45% of enzyme classes predicted correctly	Borro et al. (2006)
Support vector machine	Structural properties	60% accuracy in functional annotation of enzymes	Dobson and Doig (2005)
Structure template matching	Structural information	87% accuracy in functional annotation of enzymes	Kristensen et al. (2008)
Nearest neighbor algorithm	Sequence Descriptor: Amino acid composition	95% accuracy to the level of enzyme class	Nasibov and Kandemir-Cavas (2009)
Nearest neighbor algorithm	Domain composition and pseudo amino acid composition	98% accuracy to the level of enzyme class	Cai et al. (2005)
Self-organizing maps	Reaction descriptors	Accuracies up to 92%, 80% and 70% for class, subclass and sub-subclass levels, respectively	Latino et al. (2008)
Support vector machine	Amino Acid Composition and Conjoint triad feature	81–98% accuracy in predicting the first three EC digits	Wang et al. (2011)
Recursive feature elimination technique (RFE)	sequence information	Accuracies up to 97.8%, 87.3%, and 85.6%, for the first, second and third level	Kumar et al. (2015)

For a dataset of functionally known protein sequences belonging to different enzyme groups, group-specific features can be extracted to build models using machine learning algorithms or computational approaches to predict the function of an unknown protein sequence or to assign a group label to it (Juncker et al. 2009; Ong et al. 2007). Table 2 shows the enzyme classification attempts based on sequence similarity, structural similarity and protein descriptors.

Structural Characterization of Enzymes

The proteins in enzyme molecules fold into three-dimensional structures determining how it performs substrate binding, catalysis and regulation. Some of the amino acids are involved in binding ligands (substrates, intermediates, products, organic cofactors, metal cofactors or allosteric regulators) and some are actively involved in catalysis by interacting with the substrate, intermediate or product of the reaction (Soding et al. 2005). Thus the catalytic activity of enzymes depends on the integrity of their native protein conformation. The structures of enzymes can be elucidating by techniques such as spectroscopic methods, x-ray crystallography and more recently, multidimensional NMR methods.

X-ray Crystallography

X-ray crystallography has been the most explored technique for obtaining three-dimensional structures of proteins and in particular enzymes. Knowledge of three-dimensional structures is essential to understand reaction mechanisms at the atomic level (Feiten et al. 2017). One of the pioneers of enzyme crystallography was David Blow (1931–2004); he shared the Wolf Prize in Chemistry in 1987 for this research along with David Phillips (1924–1999), who first successfully solved the structure of chicken egg lysozyme in 1965 (Helliwell 2017). The Wolf Prize 1987 citation stated "*for their contributions to protein X-ray crystallography and to the elucidation of structures of enzymes and their mechanisms of action"*. Its structure was solved to a resolution of 2°A. The diffraction of X-rays caused by a single protein molecule is too weak to be measured (Rhodes 2006). Therefore, protein crystals are used for X-ray structure determination to amplify the signal. A protein crystal contains many copies of the molecule neatly arranged in a highly ordered regular three dimensional array or crystal lattice (Rhodes 2006). The suitability of enzyme crystals for structure determination is based on their ability to interact with X-rays. In the experimental setup (Fig. 2) a narrow beam of monochromatic X-rays of suitable wavelength is directed to the crystal which either traverses straight through the crystal, in between the enzyme molecules, or hit the electron clouds of the atoms in the enzyme molecules. The molecules arranged side-by-side in a periodic way form a lattice from which the waves diffracted to the same directions accumulate and strengthen each other to produce diffraction maxima that can be recorded by sensitive detectors (Petsko and Ringe 2004). Enzyme crystals are almost invariably frozen during the X-ray crystallography achieved by directing a cold stream of nitrogen gas onto the crystal or soaking in a solution called "cryoprotectant" so that, when frozen, vitrified water, rather than crystalline ice, is formed. Freezing makes the crystal tolerant to damage by the radiation and usually allows a higher quality and higher resolution diffraction data, while providing more accurate structural information (Ilari and Savino 2008). Additionally, freezing may sometimes help in

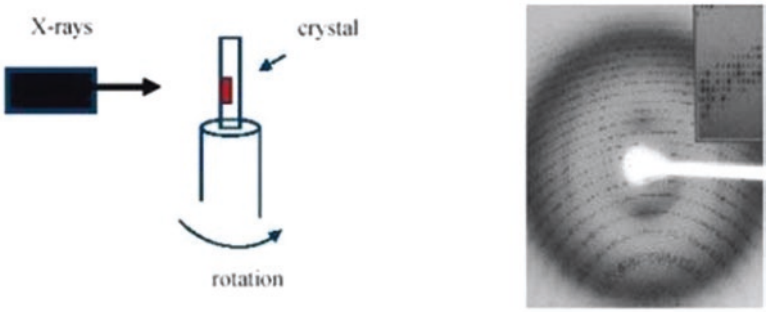

Fig. 2 Structural characterization of enzymes by X-ray crystallography

trapping substrates or other molecules that bind to the enzyme to become part of the structure, which is fundamental for structure-function studies (Rhodes 2000).

Atomic' resolution at ≥1.2°A resolution allows the placement of atoms with fewer geometrical restraints and gives a better picture of the protein structure. Advances in X-ray sources and cryo-crystallography have led to increasing numbers of structures solved at these high resolutions (Kleywegt et al. 1996). The three-dimensional representation of the protein may be displayed in a molecular structure viewer as a model that was created by the crystallographer to be chemically realistic and to match the observed electron density as precisely as possible. The resolution of a crystal structure is measured in angstrom and refers to the minimum distance between two points that can be distinguished. Although there is a large number of quality assessment methods available, resolution is a straightforward and robust parameter to assess the quality of a protein structure model (Kleywegt et al. 2004).

Nuclear Magnetic Resonance (NMR) Spectroscopy

NMR spectroscopy is a powerful tool for elucidating the structure–function relationships of substrates, peptides, proteins and in particular enzymes. It yields detailed information regarding structure of enzyme and the specific ligands which bind to the enzyme. The structure of ligands at the binding sites of enzymes and the structure of enzyme–ligand complexes can also be obtained, as well as the dynamics of the ligand and the associated structure of the protein binding site. The tertiary structures of proteins can now be obtained independently of diffraction data in solution by homo nuclear and hetero nuclear multi-dimensional NMR. In principle one can investigate the magnetic nuclei of each of the atoms within the molecule of the enzyme (^{1}H, ^{13}C, ^{15}N, …) or ligands which bind to the enzyme (^{1}H, ^{19}F, ^{31}P, ^{13}C, …), or of the environment of the active-site (solvent $^{1}H^{2}O$, $^{2}D^{2}O$, ^{23}Na, ^{39}K, ^{35}Cl, …) (Monasterio 2014). Until recently, NMR spectroscopy has yielded structures of protein complexes with small and medium size (~30 to 40 kDa). Major breakthroughs during recent past especially in isotope-labeling techniques, have enabled NMR

characterization of large protein systems with molecular weights of hundreds of kDa. This has provided unique insights into the binding, dynamic, and allosteric properties of enzymes (Huang and Kalodimos 2017).

The useful approach to study enzyme structure by protein NMR is the observation of the resonances from histidine. The C-2 and C-5 proton resonances are downfield from the aromatic protons (Markley 1975). The classical use of these properties was with the small enzyme (Mr = 23,500) RNAase (Meadows and Jardetzky 1986) and the large enzyme (Mr = 237,000) pyruvate kinase (Meshitsuka et al. 1981). The C-2 proton resonance is especially sensitive to the ionization state of the imidazole nitrogens, thus the pKa for each individual histidine within the native enzyme can be obtained from titration studies. The binding of a ligand or metal ion to a specific histidine or histidines could result in a change in the magnetic environment (chemical shift) of the resonance and an alteration in the pKa. This application of NMR has been useful in some limited number of enzymes. Enzymes enriched with ^{13}C and ^{15}N have been used to increase the range of chemical shifts of these nuclei in order to enhance spectral dispersion and increases the possibility of resolving more resonances. The detailed structural and dynamic studies of larger proteins have been done with ^{13}C and ^{15}N isotope labels through NMR and nuclear Overhauser effect (Redfield et al. 1989). This type of studies is routine for determining the structure of enzymes and their dynamics using multidimensional NMR (Kevin et al. 1998; Bachovchin 2001). An alternative approach is use of a reporter group such as ^{19}F on the enzyme or on the substrate to obtain information regarding enzyme structure and the effects of ligand binding on the enzyme (Geric 1981; Danielson and Falke 1996). ^{19}F nucleus is 83% as sensitive as ^{1}H, and has a large range of chemical shifts in addition there are no back ground resonances of ^{19}F to cause interference. The ^{19}F reporter groups can be incorporated by different methods. A fluorinated amino acid i.e. fluorotyrosine, fluor-oalanine can be added to growth medium and incorporated into the protein (Sykes and Weiner 1980). The amino acids i.e. tyrosines, alanines containing the ^{19}F are labeled and will exhibit a resonance. The hetero dimer of tubulin, the principal protein of microtubules, fluoro tyrosine can be incorporated to α-subunit on the C-terminal amino acid through the reaction catalyzed by tubulin–tyrosine-ligase (Monasterio et al. 1995). An alternative approach is to covalently label the enzyme at a specific residue with a fluorine-containing reagent like trifluoroacetic anhydride, trifluoroacetyliodide, or 3-bromo-1, 1, 1-trifluoro-propanone. The chemical shift and/or the line width (1/T2) of the ^{19}F label, a "reporter" for a change in the enzyme structure, must reflect ligand binding and/or catalysis. In case ^{19}F resonance is sensitive to conformational changes in the enzyme then site-specific modification of groups at the active site will be reflected by changes in the ^{19}F resonance. The method of using reporter groups can be also be elucidated by using other labels like ^{2}H or ^{13}C labels. However, most other labels are less sensitive than fluorine. A potential strength of using these labels is the incorporation of ^{2}H for ^{1}H or ^{13}C for ^{12}C into the protein will have a very minor, if any, effect on the protein itself. Use of reporter groups yield information regarding the environment of the group. But not the specific structural features of the enzyme, comparative structural changes can be studied by photo-chemically induced nuclear polarization (photo

CIDNP) originating from free radical reactions. This has been developed as a sensitive method to measure structural changes on the surface of proteins (Kaptein 1982; Berliner 1989). Photo-chemically induced nuclear polarization (photo CIDNP) requires a modified spectrometer and a proper light source (laser) to begin to probe surface changes. This technique has the advantage of high sensitivity, and it yields general conformation information (Monasterio 2014).

Conclusions

Enzymes are proteins responsible for catalysis of biochemical reactions. The classification information-rich EC number given by the Enzyme Commission as a simple identifier still persists. However robust approaches to quantitatively compare catalytic reactions or to accurately predict enzyme mechanisms are just beginning to appear. Further combining bond changes and reaction centers with structural information about the substrates, products and mechanisms are needed to capture the essence of enzyme chemistry in a functional classification.

X-ray Crystallography and NMR are most explored technique for structural characterization of proteins and in particular enzymes. Recent technical advances in crystallography, as well as better computational programs have made it much more rapid in solving enzyme crystal structures. Modern NMR spectroscopy techniques make extensive use of isotopically enriched proteins and should prove a powerful approach for structural characterization of proteins in particular enzymes in the future. Further technological advances are needed to establish NMR as the primary tool for obtaining atomic structures of challenging systems with even higher complexity. The accumulating data on enzyme structures—and novel approaches, particularly genome projects and bioinformatics—are expected to increase our understanding of enzyme function and mechanisms in the future.

References

Agarwal, P. K. (2006). Enzymes: An integrated view of structure, dynamics and function. *Microbial Cell Factories, 5*, 2. https://doi.org/10.1186/1475-2859-5-2

Ashburner, M., Ball, C. A., & Sherlock, G. (2000). The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nature Genetics, 25*, 25–29.

Audit, B., Levy, E. D., Gilks, W. R., Goldovsky, L., & Ouzounis, C. A. (2007). CORRIE: Enzyme sequence annotation with confidence estimates. *BMC Bioinformatics, 8*(Suppl 4), S3.

Bachovchin, W. W. (2001). Contributions of NMR spectroscopy to the study of hydrogen bonds in serine protease active sites. *Magnetic Resonance in Chemistry, 39*, 199–213.

Benkovic, S. J., & Hammes-Schiffer, S. (2003). A perspective on enzyme catalysis. *Sci, 301*(5637), 1196–1202.

Berliner, L. J. (1989). Laser chemically induced dynamic nuclear polarization studies in proteins: α-lactalbumin. *Archivos de Biología y Medicina Experimentales, 22*, 123–128.

Bhatia, S. (2018). Enzymes, proteins and bioinformatics. In *Introduction to pharmaceutical biotechnology* (2nd ed.). Bristol, UK: IOP Publishing. https://doi.org/10.1088/978-0-7503-1302-5ch1

Borro, L. C., Oliveira, S. R. M., Yamagishi, M. E. B., Mancini, A. L., Jardine, J. G., Mazoni, I., … Neshich, G. (2006). Predicting enzyme class from protein structure using Bayesian classification. *Genetics and Molecular Research, 5*(1), 193–202.

Bugg, T. D. H. (2004). *Introduction to enzyme and coenzyme chemistry* (2nd ed.). Oxford, UK: Blackwell Publishing Ltd.

Cai, Y. D., Zhou, G. P., & Chou, K. C. (2005). Predicting enzyme family classes by hybridizing gene product composition and pseudo-amino acid composition. *Journal of Theoretical Biology, 234*(1), 145–149.

Creighton, T. E. (1993). *Proteins* (2nd ed.). New York, NY: W. H. Freeman and Company, 507pp.

Cuesta, S. M., Rahman, S. A., Furnham, N., & Thornton, J. M. (2015). The classification and evolution of enzyme function. *Biophysical Journal, 109*, 1082–1086.

Danielson, M. A., & Falke, J. J. (1996). Use of 19F NMR to probe protein structure and conformational changes. *Annual Review of Biophysics and Biomolecular Structure, 25*, 163–195.

Dixon M, Webb EC (1979) Enzymes 3rd ed. Academic Press, New York, NY, 1116pp. Drauz K, Waldmann H (2002) Enzyme catalysis in organic synthesis, vol II. Weinhein, Germany: Wiley VCH, 998pp.

Dobson, P. D., & Doig, A. J. (2005). Predicting enzyme class from protein structure without alignments. *Journal of Molecular Biology, 345*(1), 187–199.

Feiten, M. C., Luccio, M. D., Santos, K. F., de Oliveira, D., & Oliveira, J. V. (2017). X-ray crystallography as a tool to determine three-dimensional structures of commercial enzymes subjected to treatment in pressurized fluids. *Applied Biochemistry and Biotechnology, 182*(2), 429–451.

Geric, J. T. (1981). Fluorine magnetic resonance in biochemistry. In L. J. Berliner & J. Reuben (Eds.), *Biological magnetic resonance* (Vol. 1, pp. 139–203). New York, NY: Plenum Press.

Helliwell, J. R. (2017). New developments in crystallography: exploring its technology, methods and scope in the molecular biosciences. *Bioscience Reports, 37*, BSR20170204. https://doi.org/10.1042/BSR20170204

Huang, C., & Kalodimos, C. G. (2017). Structures of large protein complexes determined by nuclear magnetic resonance spectroscopy. *Annual Review of Biophysics, 46*, 317–336.

Ilari, A., & Savino, C. (2008). Protein structure determination by X-ray crystallography. In J. M. Keith (Ed.), *Bioinformatics, volume I: Data, sequence analysis, and evolution*. Totowa, NJ: Humana Press. https://doi.org/10.1007/978-1-60327-159-2

Illanes, A. (2008). *Enzyme biocatalysis: Principles and applications*. Dordrecht, Netherlands: Springer.

Juncker, A. S., Jensen, L. J., Pierleoni, A., Bernsel, A., Tress, M. L., Bork, P., … Brunak, S. (2009). Sequence-based feature prediction and annotation of proteins. *Genome Biology, 10*(2), 206.

Kaptein, R. (1982). Photo-CIDNP studies. In L. J. Berliner & J. Reuben (Eds.), *Proteins in biological magnetic resonance* (Vol. 4, pp. 145–191). New York, NY: Plenum Press.

Kersey, P. J., Allen, J. E., & Staines, D. M. (2014). Ensembl Genomes 2013: scaling up access to genome-wide data. *Nucleic Acids Research, 42*, D546–D552.

Kevin, H., Lewis, G., & Kay, E. (1998). The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. *Annual Review of Biophysics and Biomolecular Structure, 27*, 357–406.

Kleywegt, G., Hoier, H., & Jones, T. (1996). A re-evaluation of the crystal structure of chloromuconate cycloisomerase. *Acta Crystallographica. Section D, Biological Crystallography, 52*, 858–863.

Kleywegt, G. J., Harris, M. R., Zou, J. Y., Taylor, T. C., Wahlby, A., & Jones, T. A. (2004). The Uppsala Electron-Density server. *Acta Crystallographica. Section D, Biological Crystallography, 60*, 2240–2249.

Kristensen, D. M., Ward, R. M., Lisewski, A. M., Erdin, S., Chen, B. Y., Fofanov, V. Y., … Lichtarge, O. (2008). Prediction of enzyme function based on 3D templates of evolutionarily important amino acids. *BMC Bioinformatics, 9*, 17.

Kumar, S., Bhola, A., & Tiwari, A. K. (2015). Classification of enzyme functional classes and subclasses using support vector machine. In: *2015 1st International Conference on Futuristic trend in Computational Analysis and Knowledge Management (ABLAZE-2015)*.

Latino, D. A. R. S., Zhang, Q. Y., & Aires-De-Sousa, J. (2008). Genome-scale classification of metabolic reactions and assignment of EC numbers with self-organizing maps. *Bioinformatics, 24*(19), 2236–2244.

Lee, H. C. (2006). Structure and enzymatic functions of human CD38. *Molecular Medicine, 12*, 317–323.

Longo, M., & Combes, D. (1999). Thermostability of modified enzymes: A detailed study. *Journal of Chemical Technology and Biotechnology, 74*, 25–32.

Markley, J. L. (1975). Observation of histidine residues in proteins by means of nuclear magnetic resonance spectroscopy. *Accounts of Chemical Research, 8*, 70–80.

Martınez Cuesta, S., Furnham, N., & Thornton, J. M. (2014). The evolution of enzyme function in the isomerases. *Current Opinion in Structural Biology, 26*, 121–130.

Meadows, D. H., & Jardetzky, O. (1986). Nuclear magnetic resonance studies of the structure and binding sites of enzymes IV. Cytidine 30-monophosphate binding to ribonuclease. *Proceedings of the National Academy of Sciences, 61*, 406–413.

Meshitsuka, S., Smith, G. M., & Mildvan, A. S. (1981). Proton NMR studies of the histidine residues of rabbit muscle pyruvate kinase and of its phosphoenol pyruvate complex. *The Journal of Biological Chemistry, 256*, 4460–4465.

Monasterio, O. (2014). Nomenclature for the applications of nuclear magnetic resonance to the study of enzymes. *Perspectives on Science, 1*, 88–97.

Monasterio, O., Nova, E., Lopez-Brauet, A., & Lagos, R. (1995). Tubulin–tyrosine ligase catalyzes covalent binding of mfluorotyrosine to tubulin. Kinetic and 19F-NMR Studies. *FEBS Letters, 374*, 165–168.

Nasibov, E., & Kandemir-Cavas, C. (2009). Efficiency analysis of KNN and minimum distance-based classifiers in enzyme family prediction. *Computational Biology and Chemistry, 33*(6), 461–464.

Ong, S. A., Lin, H. H., Chen, Y. Z., Li, Z. R., & Cao, Z. (2007). Efficacy of different protein descriptors in predicting protein functional families. *BMC Bioinformatics, 8*, 300.

Petsko, G. A., & Ringe, D. (2004). *Protein structure and function*. London, UK: New Science Press. 195pp.

Rahman, S. A., Cuesta, S. M., & Thornton, J. M. (2014). EC-BLAST: A tool to automatically search and compare enzyme reactions. *Nature Methods, 11*, 171–174.

Rausch, C., Weber, T., Kohlbacher, O., Wohlleben, W., & Huson, D. H. (2005). Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Research, 33*(18), 5799–5808.

Redfield, A. G., Mclntosh, L. P., & Dahlquist, F. W. (1989). Use of 13C and 15N isotope labels for proton nuclear magnetic resonance and nuclear Overhauser effect. Structural and dynamic studies of larger proteins and nucleic acids. *Archivos de Biología y Medicina Experimentales, 22*, 129–138.

Rhodes, G. (2000). *Crystallography made crystal clear*. San Diego, CA: Academic Press, 269pp.

Rhodes, G. (2006). *Crystallography made crystal clear–A guide for users of macromolecular models* (3rd ed.). London, UK: Academic Press Publications.

Robinson, P. K. (2015). Enzymes: Principles and biotechnological applications. *Essays in Biochemistry, 59*, 1–41.

Rost, B. (2002). Enzyme function less conserved than anticipated. *Journal of Molecular Biology, 318*(2), 595–608.

Schumacher, G., Sizmann, D., & Haug, H. (1986). Penicillin acylase from *E. coli*: Unique gene–protein relation. *Nucleic Acids Research, 14*(14), 5713–5727.

Shah, I., & Hunter, L. (1997). Predicting enzyme function from sequence: A systematic appraisal. *Proceedings of the International Conference on Intelligent Systems for Molecular Biology, 5*, 276–283.

Soding, J., Biegert, A., & Lupas, A. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research, 33*, 244–248.

Sonkaria, S., Boucher, G., & Fl'orez-Alvarez, J. (2004). Evidence for 'lock and key character in an anti-phosphonate hydrolytic antibody catalytic site augmented by non-reaction centre recognition: Variation in substrate selectivity between an anti-phosphonate antibody, an anti-phosphate antibody and two hydrolytic enzymes. *The Biochemical Journal, 381*, 125–130.

Sykes, E. D., & Weiner, J. H. (1980). Biosynthesis and 19F NMR characterization of fluoro amino acid containing proteins. In J. S. Cohen (Ed.), *Magnetic resonance in biology* (Vol. 1, pp. 1–196). New York, NY: Wiley.

Union of Pure, I. & Applied Chemistry. (2005–2009). IUPAC compendium of chemical terminology - The gold book. http://goldbook.iupac.org/.

UniProt Consortium. (2013). Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Research, 41*, D43–D47.

Wang, Y. C., Wang, Y., Yang, Z. X., & Deng, N. Y. (2011). Support vector machine prediction of enzyme function with conjoint triad feature and hierarchical context. *BMC Systems Biology, 5*, S6.

Whitehurst, R. J., & van Oort, M. (2009). *Enzymes in Food Technology* (2nd ed.). Chichester, UK: Wiley-Blackwell.

Yadav, S. K., & Tiwari, A. K. (2015). Classification of enzymes using machine learning based approaches: A review. *Machine Learning and Applications: An International Journal, 2*, 30–49.

Yousef, M. S., Clark, S. A., Pruett, P. K., Somasundaram, T., Ellington, W. R., & Chapman, M. S. (2003). Induced fit in guanidino kinases – Comparison of substrate-free and transition state analog structures of arginine kinase. *Protein Science, 12*, 103–111.