We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,300

Open access books available

172,000

International authors and editors

190M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Hydrolases: The Most Diverse Class of Enzymes

*Ekta Shukla, Ameya D. Bendre
and Sushama M. Gaikwad*

Abstract

Being the largest and most diverse class of enzymes, hydrolases offer an opportunity to explore the conformational diversity which forms the basis of their differential biological functions. In recent times, there is an urge to re-evaluate and update our existing knowledge on functional and conformational transitions of these enzymes, in the context of emerging scientific trends. In this chapter, we discuss hydrolases in terms of their diversity, classification, and different nomenclature styles that exist. Further, the concepts of protein stability and significance of studying the structure–function relationship of hydrolases are mentioned in detail taking serine protease as an example. The chapter talks about multiple ways by which an enzyme’s structure and function can be explored. The available information and literature survey on hydrolases have been systematically summarized for an easy understanding. Various experimental methods and techniques involving artificial intelligence are introduced in the later sections. The knowledge obtained by these strategies contributes to our current knowledge of the interplay between the stability, structure, and function of these enzymes. This, in turn, can help in designing and engineering these proteins with improved functional and structural features toward the goal of increasing their applicability in biotechnology.

Keywords: hydrolysis, catalysis, nomenclature, structure–function relationship, protein stability

1. Introduction

Hydrolase is a class of hydrolytic enzymes that are commonly used as biochemical catalysts which utilize water as a hydroxyl group donor during the substrate breakdown. In simple words, a hydrolase is an enzyme that catalyzes the hydrolysis of a chemical bond in biomolecules. This, in turn, divides a large molecule into two smaller ones. Hydrolases are hence important for the environment since they digest large molecules into small fragments for the synthesis of biopolymers as well as for the degradation of toxins. In biochemistry,

Hydrolases is the largest and most diverse class of enzymes with more than 200 enzymes that catalyze the hydrolysis of several types of compounds. They catalyze the hydrolytic cleavage of carbon–oxygen (C–O), carbon–nitrogen (C–N), carbon–carbon (C–C), phosphorus–nitrogen (P–N) bonds, etc. Systematic names of hydrolases are formed as “substrate hydrolase.” However, common names are typically in the form ‘*substratease*’, such as nuclease refers to an enzyme that hydrolyses nucleic acids. Examples of some common hydrolases include esterases, proteases, glycosidases, and lipases.

2. Applications/significance of hydrolases

Enzymes of this class carry out important degradative reactions in the body. Hydrolases cleave large molecules into smaller fragments used for synthesis, excretion of waste materials, or as sources of carbon for the production of energy. These are involved in digestion, transport, excretion, regulation and signalling processes, etc.; for example, digestive enzymes like cholinesterase, carboxylesterase, lysosomal hydrolases, etc. To be specific, hydrolase expressed by *Lactobacillus spp.* in the human gut could stimulate the liver to secrete bile salts which facilitate the digestion of food [1].

Hydrolytic enzymes are not only physiologically important, playing role in various cellular processes, but also have myriad commercial applications too. The industrial importance of hydrolases exceeds that of other classes of enzymes holding the highest share of enzymes used for industrial purposes. Almost 75% of all industrial enzymes are hydrolytic enzymes. Carbohydrases, proteases, and lipases dominate the enzyme market, accounting for more than 70% of all enzyme sales. Many industrial sectors, such as the detergent, leather, textiles, pulp and paper, foods and feeds, dairy, biofuels, and waste treatment industries, depend on hydrolases. Proteases remain the dominant enzyme type, because of their extensive use in the detergent and dairy industries. Various carbohydrases (glycosidases), primarily amylases and cellulases, used in industries, such as the starch, textile, detergent, and baking industries, represent the second largest group [2–4].

3. Classification of hydrolases

Apart from the common names given to certain hydrolases, there exist systematic nomenclature systems to name these enzymes.

3.1 Based on enzyme commission (EC) numbers

Hydrolases belong to enzyme class 3 (EC 3) and are further categorized based on the type of bond they cleave [5]. The four-digit code includes the nature of the *bond* hydrolyzed, then the nature of the *substrate*, and lastly the *enzyme* (**Table 1**).

3.2 Based on active site

The active site geometry of different hydrolases is different, in spite of the same catalytic method, i.e., hydrolysis. Thus, a Hierarchical classification of hydrolases Catalytic Sites (HCS) has been proposed which is based on the amino acids involved in catalysis [6, 7]. The relation between a class and its subclass in the hierarchy is that the catalytic site of the subclass refines the catalytic site of the base class. Serine hydrolase like esterase, with unusual catalytic dyad Ser-His, belongs to class **S.01** (serine hydrolases with Ser-His dyad) while hydrolases, such as trypsin or subtilisin, are further categorized into subclass **S.01.01** (hydrolases with Ser-His-Asp/Glu triad), i.e. subclass contains all residues of the basic class and some additional ones. Currently, only hydrolases are included in such a classification since they are the most studied and most abundant enzymes (**Table 2**).

Let us understand these nomenclatures by taking an example of a single hydrolase, say a serine protease.

Subclass (hydrolase acting upon)	Sub-subclass example	Enzyme example
3.1 Ester bonds (esterases)	3.1.1 Lipases	3.1.1.3 Triacylglycerol lipase
3.2 Sugars	3.2.1 Glycosidase	3.2.1.1 α-amylase
3.3 Ether bonds	3.3.2 Ether hydrolase	3.3.2.6 Leukotriene-A4 hydrolase
3.4 Peptide bonds (peptidases)	3.4.21 Serine endopeptidase	3.4.21.1 Chymotrypsin
3.5 C-N bonds (other than peptide bonds)	3.5.1 In linear amides	3.5.1.1 Asparaginase
3.6 Acid anhydrides	3.6.1 In P-containing anhydrides	3.6.1.1 Inorganic diphosphatase
3.7 C-C bonds	3.7.1 In ketonic substances	3.7.1.1 Oxaloacetase
3.8 Halide bonds	3.8.1 In C-X compounds	3.8.1.1 Alkylhalidase
3.9 P-N bonds	3.9.1 On P-N bonds	3.9.1.1 Phosphoamidase
3.10 S-N bonds	3.10.1 On S-N bonds	3.10.1.1 N-Sulfoglucosamine sulfohydrolase
3.11 C-P bonds	3.11.1 On C-P bonds	3.11.1.1 Phosphonoacetaldehyde hydrolase
3.12 S-S bonds	3.12.1 On S-S bonds	3.12.1.1 Trithionate hydrolase
3.13 C-S bonds	3.13.1 On C-S bonds	3.13.1.1 UDP-Sulfoquinovose synthase

Adapted from ExplorEnz database: <http://www.enzyme-database.org/downloads/ec3.pdf>

Table 1.
Classification of hydrolase based on EC numbers.

Sr. No.	Base class	Subclass
1	(A) Carboxyl (aspartyl and glutamyl) hydrolases	(A.01) Pepsin-like proteases (A.02) Glycosidases (A.03) Hydrolases with covalent aspartyl-substrate intermediate (A.04) Epoxide hydrolase-like
2	(C) Cysteine hydrolases	(C.01) Cys hydrolases with Cys-His dyad (C.02) N-terminal cysteine hydrolases (C.03) Cys hydrolases with His as a proton donor (C.04) Tyrosine phosphatase-like (C.05) Cys hydrolases with carboxyl group as the proton acceptor
3	(H) Histidine hydrolase	(H.01) Ribonuclease-like (H.02) Hydrolases with covalent His-substrate intermediate (H.03) Hydrolases of carbon–carbon bond
4	(M) Metal-dependent hydrolases	(M.01) Zinc-dependent hydrolases (M.02) Magnesium-dependent hydrolases (M.03) Calcium-dependent hydrolases (M.04) Bimetallic (Zn and Mg-dependent) hydrolases (M.05) Iron-dependent hydrolases (M.06) Manganese-dependent hydrolases (M.X) Hydrolases without specific metal ion requirements
5	(P) N-terminal proline hydrolases	
6	(S) Serine hydrolases	(S.01) Ser hydrolases with Ser-His dyad (S.02) Ser hydrolases with the amino group as the proton acceptor (S.03) Ser hydrolases with carboxyl group as the proton acceptor
7	(T) Threonine hydrolases	(T.01) Asparaginase-like (T.02) N-terminal threonine hydrolases
8	(U) Unclassified hydrolases	(U.01) Proteins without hydrolase activity (U.02) Hydrolases without known catalytic domain structure
9	(Y) Tyrosine hydrolases	(Y.01) Sialidases
10	(Z) Substrate-assisted or cofactor-dependent hydrolases	(Z.01) Phosphatases with substrate’s phosphate as the catalytic base (Z.02) NAD(+)-dependent deacetylases (Z.03) Hydrolases with oxidation/reduction steps

Adapted from <http://www.enzyme.chem.msu.ru/hcs/classes.html>

Table 2.
Classification of hydrolase based on active site residues.

4. Classification of serine protease

As per the earlier discussed classification systems for hydrolases, the categoriza-
tion of serine proteases can be viewed in the schematic, as shown in **Figure 1**.

4.1 Based on site of cleavage

Proteases are further subdivided into exopeptidases and endopeptidases depend-
ing on the site of enzyme action. Exopeptidases catalyze the hydrolysis of the peptide
bonds near the N- or C-terminal ends of the substrate and can be classified into
aminopeptidases and carboxypeptidases. Endopeptidases cleave peptide bonds
within and distant from the ends of a polypeptide chain [8, 9]. Serine proteases are
also divided into endo- and exo- serine peptidases.

Figure 1.
Classification scheme of serine protease based on EC numbers and active site residues.

4.2 MEROPS classification system

According to MEROPS database version 9.9 (<https://merops.sanger.ac.uk>) over 183,000 serine proteases are known with >250 structure depositions in PDB (Protein Data Bank).

Clan	Families	Representative members	Fold	Catalytic residues	PDB
PA	12	Trypsin	Greek-key β -barrels	His, Asp, Ser	1DPO
PB	1	Protease from <i>Thermoplasma acidophilum</i>	$\alpha/\beta/\beta/\alpha$	His, Glu, Ser	1PMA
PC	1	Aspartyl dipeptidase	$\alpha/\beta/\alpha$	Ser, His	1FYE
SB	2	Subtilisin, sedolisin	3-layer sandwich	Asp, His, Ser	1SCN
SC	2	Prolyl oligopeptidase	α/β hydrolase	Ser, Asp, His	1QFS
SE	6	D-Ala–D-Ala carboxypeptidase	α -helical bundle	Ser, Lys	3PTE
SF	3	LexA peptidase	all β	Ser, Lys/His	1JHH
SH	2	Cytomegalovirus assemblin	α/β Barrel	His, Ser, His	1LAY
SJ	1	Lon peptidase	$\alpha + \beta$	Ser, Lys	1RR9
SK	2	Clp peptidase	$\alpha\beta$	Ser, His, Asp	1TYF
SP	3	Nucleoporin	all β	His, Ser	1KO6
SQ	1	Aminopeptidase DmpA	4-layer sandwich	Ser	1B65
SR	1	Lactoferrin	3-layer sandwich	Lys, Ser	1LCT
SS	14	L,D-Carboxypeptidase	β -sheet+ β -barrel	Lys, Ser	1ZRS
ST	5	Rhomboid	α -barrel	His, Ser	2IC8

Adapted from reference Rawlings et al. [10].

Table 3.
Known diversity of serine peptidase structure and catalytic mechanism.

This classification system divides peptidases into clans based on catalytic mechanisms and families on the basis of common ancestry. The serine peptidases have been classified into 15 clans comprising numerous families. A summary of catalytic units in all serine peptidase families and their characteristic folds is provided in **Table 3**.

4.2.1 PA clan of serine peptidases

The PA clan (**P**roteases of the mixed nucleophile, superfamily **A**) of endopeptidases is the most abundant, and over two-thirds of this clan are comprised of the S1 family of serine proteases, which bear the archetypal trypsin fold and have a catalytic triad in the order histidine, aspartate, serine. Members have a trypsin/chymotrypsin-like fold

Figure 2. Schematic illustration of the general catalytic mechanism for serine proteases with chymotrypsin as an example.

and similar proteolysis mechanisms but sequence identity of <10%. PA clan proteases share a core motif of two β -barrels arranged perpendicularly with the covalent catalysis occurring at the interface of both barrels. **Figure 2** shows the crystal structure of bovine chymotrypsin, deposited in PDB (PDB Id: 5J4S). The core structure and active site geometry of these proteases are of interest for many applications [10, 11].

5. Catalytic mechanism of hydrolases

To understand the mechanism in a simple way, let us again take the example of a serine protease. Serine proteases are widely distributed in nature and found in all kingdoms of cellular life as well as many viral genomes. Over one-third of all known proteolytic enzymes are serine peptidases [9]. All of the serine proteases contain three residues at their active site—a serine, a histidine, and an aspartate, comprising the characteristic ‘catalytic triad’. Some serine proteases are synthesized as larger, inactive, precursors. As an example, chymotrypsinogen is converted to chymotrypsin by the excision of two dipeptides, 14–15 and 147–148 [12]. Interestingly, the structures of chymotrypsinogen and chymotrypsin are almost superimposable, i.e., the conformational change involved in the conversion process appears to be fairly small. The implication is that even relatively small structural changes can result in dramatic changes in activity. The serine proteases also differ in their sequence and substrate specificity. For instance, the bacterial protease subtilisin will cleave essentially any substrate, while other enzymes, Factor Xa (involved in blood clotting) requires a specific residue recognition sequence, Ile-Glu-Gly-Arg, to uniquely hydrolyze its polypeptide substrate after the arginine. Similarly, trypsin is specific for cleavage after Lys and Arg residues.

Almost all clan PA peptidases utilize the canonical catalytic triad of Ser195, Asp-102, and His-57 (chymotrypsin numbering). Catalysis proceeds through the formation of an H-bond between Asp-102 and His-57, which facilitates the abstraction of the proton from Ser195 and generates a potent nucleophile [10]. The catalytic triad is stabilized through a network of additional H-bonds formed by conserved amino acid residues surrounding the triad, which are Thr54, Ala56, and Ser214. The reaction pathway involves two tetrahedral intermediates. Initially, the hydroxyl O atom of Ser195 attacks the carbonyl of the peptide substrate as a result of His57 in the catalytic triad acting as a base. The backbone N atoms of Gly193 and Ser195 stabilize the tetrahedral intermediate and generate a positively charged pocket within the active site known as the oxyanion hole. The tetrahedral intermediate collapse results in the formation of an acyl-enzyme intermediate. In the second half of the mechanism, a water molecule displaces the free polypeptide fragment and attacks the acyl-enzyme intermediate. Again, the oxyanion hole stabilizes the second tetrahedral intermediate of the pathway and the collapse of this intermediate liberates a new C terminus in the substrate.

In the next section, we discuss how an enzyme’s or protein’s structure is maintained. There are various kinds of inter- and intra-molecular forces that are involved in preserving the native functional structure of a protein.

6. Forces involved in enzyme’s (protein’s) stability

Protein stability is predominantly dictated by forces that help in maintaining the native structure of a protein which include covalent interactions, such as disulfide bonds, and weak (non-covalent) interactions, such as hydrogen bonds, hydrophobic and ionic

Figure 3. (a) Molecular interactions which stabilize the protein structure. (b) Effect of various physicochemical conditions on protein structure. (c) Sequence-structure-function triad important for proteins.

interactions (**Figure 3a**). For instance, the primary structure is associated with the covalent bonds (peptide bonds) between the amino acid residues, making up the protein backbone. The secondary structures involve primarily hydrogen bonding between the atoms, thereby creating stable local conformations and structures. Sometimes, it also involves disulfide linkages between two cysteine residues of the same or different chains in a protein [13].

The ultimate three-dimensionally folded tertiary structure of a whole globular protein is formed and maintained by various weak ionic and hydrophobic interactions. Hydrophobic interactions play an important role in stabilizing a protein conformation where the interior of a protein generally consists of a densely packed core of hydrophobic amino acid side chains. Though covalent bonds (such as disulfide bonds) are much stronger than individual weak interactions, i.e., approximately 200–460 kJ/mol, are required to break a single covalent bond, whereas weak interactions can be disrupted by a mere 4–30 kJ/mol. Yet, due to their sheer number, the weak interactions predominate as the stabilizing force in protein structure [14]. The protein folding code is, thus, written in the side chains and not in the backbone hydrogen bonding, because it is through the side chains that one protein differs from another. The number of protein conformational diseases that are now recognized is an indication of the importance of proteins achieving and maintaining their correct fold.

7. Structure–function relationship of proteins

For a protein to be functional, it needs to fold into a specific three-dimensional native structure. The sequence of amino acids determines the structure of a protein which ultimately governs its function. Protein folding is an extremely active field of research, which requires converging expertise from biology, chemistry, computer science, and physics. Understanding the relationship between protein structure and function is a complicated puzzle and remains a primary focus in structural biology. Protein molecules display a remarkable relationship between their amino acid sequence, their

three-dimensional structure, and their function at the molecular level (**Figure 3c**) [15]. A polypeptide can also adopt a less rigid or more flexible conformation, different from its functional native form, responding to changes in the environment. To understand this structure–function paradigm, one of the widely used approaches is to subject the native and active conformation of a protein to various physicochemical stress conditions and monitor the changes occurring in its conformation and function at each step.

7.1 Studying folding/unfolding transitions of proteins

So, we understood that any disturbance in the delicate balance of these interactions could lead to loss of the native structure of the protein. Therefore, understanding the protein folding and the unfolding mechanism is equally essential as learning its function. A protein exists in equilibrium with unfolded conformational states in solution with the folded ensemble being favored at ambient conditions. This equilibrium between the folded and the unfolded states can be perturbed by changing the thermodynamic state of the system (temperature, pressure, and pH) or by changing the composition by the addition of co-solvents to the solution [16]. Interestingly, the effect of co-solvents on the protein can alter this equilibrium in any direction. For example, urea and guanidium hydrochloride (GdnHCl) induce disorder and favor the unfolded state of proteins, and are, therefore, known as denaturants/chaotropes. On the other hand, protective osmolytes/kosmotropes, such as trimethylamine N-oxide (TMAO), dimethyl sulphoxide (DMSO), glycine, betaine, glycerol, and sugars, induce stabilization of the folded proteins (**Figure 3b**). Studies on solvent-mediated structural and conformational transitions of proteins can provide insight into their stability, folding pathways, and intermolecular aggregation behavior [16, 17].

Native proteins are generally marginally stable, i.e. free energy gap separating the folded and unfolded states in typical proteins under physiological conditions is quite small (20 to 65 kJ/mol). Therefore, when the delicate balance between the interactions involved in stabilizing or destabilizing a particular structure is disturbed by harsh environments, such as extreme temperature, pH, and chaotropes; it may lead to structural and functional alterations in protein [14]. A loss of the three-dimensional structure of a protein, sufficient to cause loss of function is called denaturation. The denatured state does not always equate with the complete unfolding of the protein (**Figure 3b**). Denaturation can be either partial or complete and it can also be reversible or irreversible. Under most conditions, denatured proteins exist in a set of partially folded states that are poorly understood [18]. In some cases, the structure of an enzyme remains stable, but the labile active site tends to lose its geometry and hence the activity. Contrary to this, the active site may get unusually stabilized and highly active. A polypeptide can also adopt a less rigid or more flexible conformation different from its functional native form, responding to changes in the environment [17, 19, 20].

Exploring structure–function relationships of proteins/enzymes can help in establishing the factors responsible for their stability. Furthermore, knowledge of the overall stability of protein molecules is important, especially when the protein in question is useful in industrial-scale biotechnology, where they may be subject to conditions, such as high temperature, low pH, and presence of co-solvents [21]. The optimization of biological stability is also an important criterion while considering the application of biomolecules (such as proteins/enzymes) as therapeutic agents [22]. Interestingly, novel proteins are now designed as variants of existing proteins or from non-natural amino acids or *de novo*. Moreover, new polymeric materials called **foldamers** are finding applications in biomedicine as antimicrobials, lung surfactant replacements, etc. (reviewed in [23]).

7.2 Tools for probing protein structure and conformational transitions/dynamics

The protein folding problem fascinates the scientist, the educated layman, and the entrepreneur. The full understanding of a molecular system comes from careful examination of the sequence–structure–function triad. Over the last 30 years, detailed experimental and theoretical studies of a number of proteins have advanced our understanding of protein folding and dynamics.

Technique	Structural parameter probed
Fluorescence	
Intrinsic	Environment of Trp and Tyr
ANS binding	Exposure of hydrophobic surface area
Substrate binding	Formation of the active site
FRET	Inter-residue distances
Anisotropy	Depolarization of the fluorescence emission
Fluorescence Correlation Spectroscopy	Autocorrelation analysis of fluctuations in fluorescence emission due to internal dynamics
2-D fluorescence lifetime correlation spectroscopy	Correlation of the fluorescence photon pairs with respect to the excitation–emission delay times
Single-molecule spectroscopy (Sm-FRET and sm-PET)	Distance between fluorophores dynamics
Red Edge Excitation Shift	Rate of solvent relaxation around an excited state fluorophore in a protein
Circular dichroism	
Far UV	Secondary-structure information
Near UV	Tertiary-structure information
Protein engineering	Role of individual residues in stabilizing intermediates and transition states
Small-angle X-ray scattering (SAXS)	Dimension and shape of a polypeptide
Absorbance (near UV)	Environment of aromatic residues or co-factors
FTIR	Secondary-structure information
NMR	
Real time	Environment of individual residues
Dynamic NMR	Lineshape analysis provides folding–unfolding rates close to equilibrium
Native state HX	Global stability and metastable states
Pulsed HX ESI MS	Folding populations
Force spectroscopy (AFM/optical tweezers)	Unfolding forces and unfolding-rate constants of single molecules
Differential scanning calorimetry (DSC)	Energetics
Differential scanning fluorimetry (DSF)	Environment of fluorescent dye and intrinsic fluorescence
Differential static light scattering (DSLS)	Temperature of aggregation of a protein

Table modified from references [24, 25].

Table 4.
Experimental techniques used to measure folding/unfolding and dynamics.

7.2.1 Experimental approach

The experimental techniques for studying protein structural transitions monitor the gradual folding/unfolding of proteins and observe conformational changes under various conditions. **Table 4** summarizes a few of the standard biophysical techniques based on fluorescence, absorbance and circular dichroism, etc. which are often used to probe such transitions in protein structure.

Figure 4 shows conformational and functional transitions in a serine protease isolated from *Conidiobolus brefeldianus* upon thermal denaturation. The changes in the protein structure are quite clear upon increasing temperature which is in corroboration with the loss in activity of the enzyme. These conformational changes were probed with fluorescence spectroscopy and Far-UV circular dichroism spectroscopy while functional activity assays were carried out using casein proteolysis. The enzyme apparently begins to lose its structure and function above 55°C. Thus, it becomes clear that a change in the native conformation of a protein affects its function.

Figure 4. Thermal denaturation of serine protease from *Conidiobolus brefeldianus*: (a) fluorescence spectra at different temperatures showing the redshift in λ_{max} of intrinsic fluorescence at increasing temperatures. (b) Far-UV CD spectra at different temperatures clearly mark the change in protein conformation above 55°C. (c) Activity profile over various time points at different temperatures. This is in accordance with the loss of native structure above 55°C which hampers the catalytic activity of the enzyme correspondingly. (d) Cartoon representation of the changes occurring in structure and function of the enzyme at higher temperatures. (Figure credit: Shukla et al. [26]).

Figure 5.

Cartoon representation of homology model of a serine protease from rat (Prss30): It selectively cleaves synthetic peptide substrates of trypsin and activates the epithelial sodium channel. The model is derived from the AlphaFold database (<https://alphafold.ebi.ac.uk>). The protein structure for this physiologically and commercially significant serine protease has not been solved. Thus, the predicted model provides us with valuable information about the protein folds, secondary structure arrangements, and catalytic triad of this important enzyme. β -sheets are shown in red color, cyan color marks the α -helices, and magenta color highlights the loops.

7.2.2 Theoretical approach

New theoretical and computational approaches have emerged, including various bioinformatics tools, artificial intelligence (AI) based methods, deep evolutionary analysis, structure-prediction web servers, physics-based force fields, etc. These techniques are employed to complement the experiments in providing an overall picture of the protein structure. The computer-based protein-structure prediction has been advanced by Molt and colleagues, in an event initiated in 1994 called CASP: Critical Assessment of protein Structure Prediction [27]. Currently, all successful structure-prediction algorithms are based on the assumption that similar sequences lead to similar structures. These methods depend heavily on the PDB for template sequences. There are several computational methods for protein structure determination, including homology modeling, fold recognition via threading and *ab initio* methods [15, 28]. In recent times, computational methods that can predict protein structures with atomic accuracy, even in cases where no similar structure is available, have been developed. Recently developed programs and databases, such as AlphaFold and RosettaFold which are neural-based networks, are a great success in this field (Figure 5) [29, 30].

The tremendous increase in the amount of sequence and structural data of proteins, together with the advances in the experimental and bioinformatics methods are improving our knowledge about the relationship between the protein sequence, structure, dynamics, and function [31]. This knowledge, in turn, helps us to understand how proteins interact with their substrate and other molecules, such as small molecules or ligands, which can become a drug candidate [32]. Predicting the binding modes and affinities of different compounds upon interaction with the protein binding sites is the main goal of ‘structure-based drug design’ and is achieved by the

‘docking’ approach. There are a number of programs written to carry out such analysis. In general, a large number of conformations are generated for the small molecule (substrate or ligand), either prior to docking or during the docking routine. Each conformation is positioned in the active site in a variety of orientations, the combination of conformation and orientation being known as a ‘pose’. Further, many such poses are selected and ranked by a scoring function to determine the overall best pose [15, 33] and the binding energy and affinity being calculated. The new frontiers now lie in physics-based modeling and AI to predict conformational changes, understand protein dynamics, design synthetic proteins, and improve protein modeling based on the laws of physics.

8. Conclusions

Hydrolases could participate in a variety of biological processes due to their diversification. Being the largest and most diverse class of enzymes, hydrolase offers an opportunity to explore the conformational/topological diversity which forms the basis of their differential biological activity. Thus, there is an urge to re-evaluate our existing knowledge on the functional and conformational transitions of these enzymes, in the context of emerging scientific trends. In this chapter, we discuss hydrolases in terms of their diversity, classification, the importance of the structure–function relationship of hydrolases taking serine protease as an example. The ongoing pandemic (SARS CoV-2 infections) further illustrates the importance to study hydrolases from the therapeutic point of view. To let the virus enter the host cell, viral spike protein plays a very important role and it is further activated by the serine protease 2 (TMPRSS2). The host proteases thus are involved in an intricate play in SARS-CoV-2 infection along with other viral infections and in designing antiviral therapeutic strategies [34].

Figure 6.
The sequence–structure–function triad necessary for understanding enzymes. In this chapter, it has been discussed wrt hydrolases.

Furthermore, the available information and literature survey on selected hydrolases have been systematically summarized for easy understanding. Knowledge of the relationships between protein structure and function at the molecular level remains a primary focus in structural biology. So, various experimental and *in silico* methods and techniques have been mentioned in the chapter which contributes to our knowledge of the interplay among the stability, structure, and function of these enzymes (**Figure 6**). It can serve as a structural toolbox to improve their efficiency in the future by helping in engineering these proteins with improved functional and structural features.

Acknowledgements

Authors thank Dr. Radha Chauhan, Scientist, NCCS, Pune, for her kind support.

Author details

Ekta Shukla^{1*}, Ameya D. Bendre² and Sushama M. Gaikwad³

1 National Centre for Cell Science, S.P. Pune University Campus, Pune, India

2 Department of Chemistry, S.P. Pune University, Pune, India

3 CSIR-National Chemical Laboratory, Pune, India

*Address all correspondence to: ektaashukla19@gmail.com

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

References

- [1] Chand D, Avinash VC, Yadav Y, Pundle AV, Suresh CG, Ramasamy S. Molecular features of bile salt hydrolases and relevance in human health. *Biochimica et Biophysica Acta - General Subjects*. 2017;**1861**:2981-2991
- [2] Li S, Yang X, Yang S, Zhu M, Wang X. Technology prospecting on enzymes: Application, marketing and engineering. *Computational and Structural Biotechnology Journal*. 2012;**2**:3
- [3] Shimizu K, Morse DE. Silicatein: A unique silica-synthesizing catalytic triad hydrolase from marine sponge skeletons and its multiple applications. In: Moore BS, editor. *Methods in Enzymology*. Vol. 605. Massachusetts, US: Academic Press; 2018. pp. 429-455
- [4] Mustafa M, Ali L, et al. Heterologous expression and characterization of glycoside hydrolase with its potential applications in hyperthermic environment. *Saudi Journal of Biological Sciences*. 2021. DOI: 10.1016/j.sjbs.2021.09.076. (in press)
- [5] Bairoch A. The ENZYME database in 2000. *Nucleic Acids Research*. 2000;**28**(1):304-305
- [6] Gariev A, Varfolomeev SD. Hierarchical classification of hydrolases catalytic sites. *Bioinformatics*. 2006;**22**(20):2574-2576
- [7] McDonald AG, Boyce S, Tipton KF. ExplorEnz: The primary source of the IUBMB enzyme list. *Nucleic Acids Research*. 2009;**37**:D593-D597
- [8] Berg JM, Tymoczko JL, Stryer L. *Biochemistry*. In: Section 10.5, Many Enzymes Are Activated by Specific Proteolytic Cleavage. 5th ed. New York: W H Freeman; 2002 Available from: <https://www.ncbi.nlm.nih.gov/books/NBK22589/>
- [9] Rao MB, Tanksale AM, Ghatge MS, Deshpande VV. Molecular and biotechnological aspects of microbial proteases. *Microbiology and Molecular Biology Reviews*. 1998;**62**(3):597-635
- [10] Rawlings ND, Barrett AJ, Bateman A. MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research*. 2012;**40**(D1):D343-D350
- [11] Page MJ, Di Cera E. Serine peptidases: Classification, structure and function. *Cellular and Molecular Life Sciences*. 2008;**65**(7-8):1220-1236
- [12] Gurung N, Ray S, Bose S, Rai V. A Broader View: Microbial Enzymes and their Relevance in Industries, Medicine, and beyond. London, UK: Hindawi Publishing Corporation, BioMed Research International; 2013 Article ID 329121
- [13] Berg JM, Tymoczko JL, Stryer L. Chapter 2: Protein composition and structure. In: *Biochemistry*. 7th ed. New York: W.H. Freeman; 2007
- [14] Nelson D, Cox M. Chapter 4: Three dimensional structure of proteins. In: *Lehninger Principles of Biochemistry*. 4th ed. New York: W.H. Freeman and Company; 2005
- [15] Breda A, Valadares NF, Norberto de Souza O, et al. Protein structure, modelling and applications [updated 2007 Sep 14]. In: Gruber A, Durham AM, Huynh C, et al., editors. *Bioinformatics in Tropical Disease Research: A Practical and Case-Study Approach*. Bethesda (MD): National Center for Biotechnology Information (US); 2006 2008. Chapter A06
- [16] Varshney N, Rabbani G, Badr G, Khan RH. Cosolvents induced unfolding

and aggregation of keyhole limpet Hemocyanin. *Cell Biochemistry and Biophysics*. 2014;**69**(1):103-113

[17] Suresh Kumar R, Suresh CG, Brannigan JA, Dodson GG, Gaikwad SM. Bile salt hydrolase, the member of ntn-hydrolase family: Differential modes of structural and functional transitions during denaturation. *IUBMB Life*. 2007;**59**(2):118-125

[18] Bischof JC, He X. Thermal stability of proteins. *Annals. New York Academy of Sciences*. 2005;**1066**:12-33

[19] Rohamare SB, Dixit VS, Nareddy PK, Sivaramakrishna D, Swamy MJ, Gaikwad SM. Polyproline fold—in imparting kinetic stability to an alkaline serine endopeptidase. *Biochimica et Biophysica Acta*. 2013;**1834**:708-716

[20] Dalal S, More SV, Shankar S, Laxman RS, Gaikwad SM. Subtilase from *Beauveria* sp.: Conformational and functional investigation of unusual stability. *European Biophysics Journal*. 2014;**43**:393-403

[21] Eijsink VG, Bjørk A, Gåseidnes S, Sirevåg R, Synstad B, van den Burg B, et al. Rational engineering of enzyme stability. *Journal of biotechnology*. 2004;**113**(1):105-120

[22] Gecse KB, Khanna R, van den Brink G, Ponsioen CY, Löwenberg M, Jairath V, et al. Biosimilars in IBD: Hope or expectation? *Gut*. 2013;**62**(6):803-807

[23] Dill K, MacCallum J. The protein-folding problem, 50 years on. *Science*. 2012;**338**:1042-1046

[24] Radford SE. Protein folding: Progress made and promises ahead. *Trends in Biochemical Sciences*. 2000;**25**(12):611-618

[25] Senisterra GA, Finerty PJ Jr. High throughput methods of assessing protein

stability and aggregation. *Molecular BioSystems*. 2009;**5**(3):217-223

[26] Shukla E, Agrawal SB, Gaikwad SM. Conformational and functional transitions and *in silico* analysis of a serine protease from *Conidiobolus brefeldianus* (MTCC 5185). *International Journal of Biological Macromolecules*. 2017;**98**:387-397

[27] Brockwell DJ, Smith DA, Radford SE. Protein folding mechanisms: New methods and emerging ideas. *Current Opinion in Structural Biology*. 2000;**10**:16-25

[28] Laskar A, Rodger EJ, Chatterjee A, Mandal C. Modeling and structural analysis of PA clan serine proteases. *BMC Research Notes*. 2012;**5**:256

[29] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;**596**:583-592

[30] Baek M et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*. 2021;**373**(6557):871-876

[31] John M, Krzysztow F, Andriy K, Torsten S, Anna T. Critical assessment of methods of protein structure prediction (CASP)-round x. *Proteins*. 2014;**82**(0 2):1-6

[32] Dodson EJ. Computational biology: Protein predictions. *Nature*. 2007;**450**:176-177

[33] Alonso H, Bliznyuk AA, Gready JE. Combining docking and molecular dynamic simulations in drug design. *Medicinal Research Reviews*. 2006;**26**:531-568

[34] Saadat YR, Khatibi SMH, Vahed SZ, Ardalan M. Host serine proteases: A potential targeted therapy for COVID-19 and influenza. *Frontiers in Molecular Biosciences*. 2021;**8**:725528