GAVIN PUBLISHERS

Advances in Biochemistry and Biotechnology

Mishra S, et al. Adv Biochem Biotechnol 2016: G107

Tools and Applications of Protein Engineering: an Overview

Sanjay Mishra[1*], Amit Kumar Mani Tiwari[1], Abbas Ali Mahdi[2]

[1]Department of Biotechnology, IFTM University, Delhi Road (NH- 24), Moradabad, UP, India

[2]Department of Biochemistry, King George's Medical University, Lucknow, U.P, India

[*]**Corresponding author:** Sanjay Mishra, Department of Biotechnology, IFTM University, Delhi Road (NH- 24), Moradabad, UP, India, Tel: 98 37 09 6059; Email: sanjaymishra@iftmuniversity.ac.in; sanjaymishra66@gmail.com

Citation: Mishra S, Tiwari AKM, Ali Mahdi A (2016) Tools and Applications of Protein Engineering: an Overview. Adv Biochem Biotechnol 2016: G107

Received Date: 7 November, 2016; **Accepted Date:** 28 November, 2016; **Published Date:** 5 December, 2016

Abstract

Protein engineering is an area possessing initiation hot spots in recombinant DNA technology, where managements in gene are expressed as alterations in protein conformation responsible for desired properties. Various techniques for the specific engineering proteins can principally be classified as techniques, which require comprehensive prior acquaintance of protein, founding the concept of rational technique of directed evolution that aids in expression of the progression of natural evolution. Protein engineering so far has been flourishing to produce proteins, which have rewarding applications in industry, health and medicinal sciences, and ultimately in nanobiotechnology in current scenario.

Keywords: Emerging protein engineering; Engineered enzymes; Nanobiotechnology; Natural evolution; Recombinant DNA technology.

Introduction

As a consequence of the insertion of recombinant DNA technology studies associated to changes at protein level gave the imprint to be carried out that led to development of specific area known as Protein Engineering, which includes studies aiming changes in amino acid sequence and their achievable results to yield a diverse protein that has enriched activity and preferred properties [1]. Outcomes facilitated to acquire how to redesign proteins from a set of example proteins, showed the desired behavior. Employing this tactic, authors were able to successfully upsurge extracellular enzyme concentrations of up to 10 by changing the amino acid profile of a protein. The projected methodology has pronounced prospective for improving production rates of other enzymes, probably also in other organisms after assembling an organism-specific classifier. While the approach was applied rapidly to improve enzyme production, the methodology itself is generic: given a set of example proteins and measured properties, sequences can be redesigned to attain certain redesign goals. So far the understanding of protein is restricted to secondary and super secondary structures and protein folding is a complex and molecular chaperons linked process. Although the process of protein folding has been significantly understood by computational techniques, allowing users to create 3-D structures with low energy [2], yet it needs to be comprehended in a way that can lead to manipulate the protein structures at ease for biotechnological applications and protein conformation-functional studies. The process of mutagenesis, based on molecular biology practices, has provided basis for incorporating distinctive mutations at genetic level, which are transcribed to proteins undergoing the process of screening and selection [3]. The amino-acid sequence of a protein describes its native conformation. A protein molecule folds naturally during or after biosynthesis. While these macromolecules may be considered as "folding themselves", the development also rests on the solvent (water or lipid bilayer), the concentration of salts, the pH, the temperature, the possible presence of cofactors and of molecular chaperones. Moreover earlier awareness of proteins, phylogenetic studies and 3-D structures has shaped opportunities to introduce mutations in rational, directed and precise fashion [4]. Besides, the use of in vitro systems and capability to mimic natural progression of evolution has aided to create proteins, which do not exist previously. The proteins having desired properties can find their applications in industries, environmental sciences and agriculture where native or natural proteins might not work competently. In addition to industrial use engineered proteins are also being used as medicine and therapeutics. Protein engineering has potential to contribute to the field of nanobiotechnology [1,5,6]. The beginning section of this review highlights the methodologies used for engineering the proteins followed by subsequent section that focuses on the applications of engineered proteins in different areas.

Protein Engineering Tools

Coherent Protein Designing

Coherent designing of protein is the supreme customary way of protein engineering. Proteins are the most multidimension-

al macromolecules in living systems and have various important functions, including structural, catalytic, sensory, and regulatory functions [4]. Coherent design of enzymes is a great challenge to our understanding of protein structure and physical chemistry and has numerous potential applications. Protein design algorithms have been applied to design or engineer proteins, which fold, fold faster, catalyze, catalyze faster, signal, and adopt preferred conformational states. The field of de novo protein design, however only a few decades old, is a platform to produce sensational results. Developments in this field are already having a noteworthy impression on biotechnology and chemical biology [4]. The coherent designing comprises of the site-directed mutagenesis in which a codon for a specific amino acid is incorporated into the desired gene1. Site-directed mutagenesis is performed in two ways.

Overlap Extension Method

In overlap extension technique, two primer pairs are used. One of the primers from each of the two primer pairs has a incompatible sequence due to a mutation in codon. When these two primer pairs are used in polymerase chain reaction (PCR), two reactions take place in the first cycle with each of the primer pair. These two reactions generate two double-stranded DNA (dsDNA). The denaturation and annealing of these two dsDNA cause in the formation of two hetero-duplex DNA. As one primer from each of the primer pair had a mismatched sequence, the mutated codon will be present in each strand of hetero-duplex DNA. The hetero-duplex DNA strands have overlapping segments, which are filled using DNA polymerase. Then in the second PCR, this mutagenic hetero-duplex DNA is amplified using normal primer pair to generate multiple copies [7,8]. Modern biology research necessitates simple procedures for competent and restriction site-independent modification of genetic material. Classical cloning and mutagenesis strategies are limited by their dependency on restriction sites and the use of complementary primer pairs. Single Oligonucleotide Mutagenesis and Cloning Approach (SOMA) are independent of restriction sites, and thus, require a single mutagenic oligonucleotide to modify a plasmid [8]. In this particular study, the broad application spectrum of SOMA with three examples was effectively demonstrated. Firstly, a novel plasmid was presented, which in a standardized and speedy manner can be used as a template for SOMA to produce GFP-reporters. Such reporters were used to assess the in vivo knock-down quality of morpholinos in Xenopus laevis embryos. Secondly, it was exposed how to use a SOMA-based protocol for restriction-site independent cloning to generate chimeric proteins by domain swapping between the two human hRMD5a and hRMD5b isoforms. Finally, it was shown that SOMA simplifies the generation of randomized single-site mutagenized gene libraries. As an example random-mutagenization of a single codon has been observed to affect the catalytic activity of the yeast Ssy5 endo-protease and recognize a spectrum of tolerated and non-tolerated substitutions. As a result, SOMA represents a highly capable substitute to conventional cloning and mutagenesis strategies [8].

Whole Plasmid Single Round Polymerase Chain Reaction Method

In whole plasmid single round PCR, two oligonucleotide primers are used which are complementary to the dsDNA of plasmid being used as a template [8]. For plasmid manipulations, other site-directed mutagenesis techniques have been supplanted largely by techniques, which are highly efficient but relatively simple, easy to use, and commercially available as a kit. An example of these techniques is the Quick- change method, wherein a pair of complementary mutagenic primers is used to amplify the entire plasmid in a thermo-cycling reaction using a high-fidelity non-strand-displacing DNA polymerase such as pfu polymerase. The reaction generates a nicked, circular DNA. The template DNA must be eliminated by enzymatic digestion with a restriction enzyme such as DpnI, which is specific for methylated DNA. All DNA produced from most Escherichia coli strains would be methylated; the template plasmid, which is biosynthesized in E. coli will, thus, be digested, whereas the mutated plasmid, which is generated in vitro and is hence unmethylated, would be left undigested. It is noticeable that, in these double-strand plasmid mutagenesis methods, while the thermo cycling reaction may be used, the DNA need not be exponentially amplified as in a PCR. Instead, the amplification is linear, and it is therefore inaccurate to describe them as a PCR, since there is no chain reaction. These primers are designed so as to contain the preferred mutation in their sequence. During PCR, DNA polymerase replicates both the strands of the plasmid [8]. As the primers are complementary to dsDNA, they are not displaced from the plasmid and result in the creation of a mutated plasmid. The breaks are present in the mutated plasmid but they do not overlap to each other. To selectively digest the mutated plasmid, a restriction enzyme, DpnI, is used. DpnI produces a nicked, circular plasmid vector. When this nicked plasmid vector is used to transform the competent cells, DNA polymerase repairs the nick in DNA to form circular mutated plasmid. This mutated plasmid is then expressed in the host to produce desired gene product [8].

Directed Evolution

Rational techniques of protein engineering have limitations based on limited understanding of protein folding, hence urging another approach to arise which is known as directed evolution. Directed evolution is based on generating many mutated copies of genes, henceforth their corresponding proteins, using focused or random mutagenesis or computational techniques, consequently generating a library of diverse proteins followed by rigorous screening and selection of favorable ones having desired properties, just mimicking the process of evolution, which has led to existence of a number of diverse proteins families in many years through the process of natural selection. Although this process is

time taking and slow, scientists have created such an analogous system completing it in weeks in laboratories by working on few mutations in a protein as mutation at every codon creates the problem of coverage [9,10].

Random Mutagenesis

Divergence of proteins can be carried out in vitro or in vivo in a random way or in a focused fashion. On the other hand, computer aided techniques are also being employed to analyze available diversity of proteins in order to distinguish the possible useful mutations, can be integrated into gene. Goeddel and co-workers described error-prone polymerase chain reaction (PCR) for the first time that is based on low fidelity of thermo-stable DNA polymerase that lacks proof reading activity and inserts an incorrect nucleotide in newly synthesized strand. The errors can be increased by increasing the concentration of magnesium and manganese ions, or adding unequal concentration of dNTPs, consequently generating the mutant copies of genes, which can be transcribed into diverse proteins, creating libraries applicable to screening [11]. However error-prone is easy to implement yet it does not provide evenly spaced amino acid codon and degeneracy of codons pose problem as only single nucleotide is replaced, therefore number of mutations is not significant. Moreover mutations by polymerase are also biased towards transitions of A and T. To overcome this problem another technique is used called as random approach is sequence saturation mutation (SeSaM) that involves the fragmentation of gene using phosphor-thiolate nucleotides acting as sites for cleavage generating fragments of variable length. Incorporation of deoxyinosine nucleotide at 3' end and their subsequent elongation followed by PCR gives number of mutant copies, majorly having randomly distributed transversions, which cannot be obtained by error-prone PCR [12].

Focused Mutagenesis

Random mutagenesis can generate large number of libraries but they might not be rich in useful proteins. Many of them might have deleterious mutations due to which protein folding may not occur or it might get non-functional. Moreover, for most of the proteins it would be impossible to have their complete coverage. Alternative to random mutagenesis is focused mutagenesis that involves producing mutations at specific sites of proteins probably being a catalytic site or a functional region, therefore, yielding a library of functionally rich proteins [13,14]. One of the famous technique of site directed mutagenesis involves insertion of a cassette compose of oligonucleotides having desired codons into a vector that after transcription forms protein having desired directed amino acids. With series of such eleven cassettes, each having two codons, mutation can be generated at any desired site of gene [15]. Another technique is Site Saturated Mutagenesis, which is carried out at nucleotides in a codon replacing each one to generate all the possible twenty amino acids at that position [16]. This overcomes the problem of codon degeneracy.

Recombination Based Mutagenesis

Recombination in nature is absolutely accountable for variations in genome of organism. Such useful phenomenon that consist of exchange of genetic material directed by complementary DNA strands, finds its applications in a process DNA shuffling where a piece of DNA after fragmentation is reconstructed by overlapping fragments acting as random primers in a PCR reaction [17]. Such technique has been updated with the use of synthetic oligonucleotides as overlapping primers, generating a complete mutated gene product [18]. Another fragment based technique is Nucleotide Exchange and Excision Technology, in which uridine nucleotide is inserted into gene sequence in PCR followed by sequential treatments of uracil glycosylase and a purinic/a pyrimidinic lyases to yield fragments of different lengths, ultimately extending into full-length diverse copies of gene using internal primers [19]. Another method independent of fragments based on premature heat denaturation in PCR is Staggered Extension PCR that yields incomplete extension products that can switch templates, generating variations in an amplicon [20].

Screening Methodologies

Followed by mutagenesis, diversified proteins, synthesized in vivo or in vitro undergo the process of screening. Screening gives a set of proteins that are valuable and from which targeted or desired protein with enhanced properties is selected. Proteins which are enzymatic in nature expressed in bacterial cultures or in vitro compartments consisted of water in oil emulsions can be screened by addition of surrogate substrate into media and emulsion respectively that generate a signal related to colorimetric, fluorescent or any other optical property as a result of enzymatic activity [20]. Besides that reporter gene such as GFP can be used to screen proteins in expression mediated way. Alternatively bacterial lysates can also be screened using chromatography techniques or conventional techniques of NMR or X-ray crystallography. For high through put screening yeast surface proteins especially the specific epitopes can be fused with diverse library member that can be identified by fluorescently labeled antibody. Cells exhibiting the epitope antibody complex can be sorted by Fluorescence Activated Cell Sorter (FACS) on the basis of fluorescence.

Selection-A Sophisticated Step

Followed by screening is a process where screened proteins undergo iterative rounds of selection to show their potential for selection, the most favorable one being selected. It is this stage where individual library member is tested and separated in a sophisticated manner. Selection can be based by binding of protein library member with an immobilized target. For simultaneous and

correct selection linkage between gene and its corresponding protein must be maintained [21]. For this purpose cell surface display or phage display methodologies are used, which involves the display of expressed library member fused with cell surface proteins [22] or coat proteins [23], respectively and their subsequent interactions with an immobilized target. Phage display has been used to study protein-protein interactions as well as in discovering new therapeutic antibodies [24]. Selection on binding has been limited mostly to enzymes. In another methodology the replication and activity of diverse protein is associated with the survival of organism as in the case of enzymes inducing antibiotic resistance. Linking the activity of a diverse protein member with expression of antibiotic resistance gene has also been studied. In vivo systems have drawbacks such as host genomic mutation and transformation efficiency. This can be overcome with in vitro methodologies [13]. An in vitro methodology called as Ribosomal display exhibits the stable binding of ribosome with mRNA and synthesized protein in absence of stop codon and controlled conditions thus maintaining linkage between gene and protein [25]. In vitro systems can be used for selection of those enzymes that have DNA or RNA as substrates. Polymerases and nucleases with efficient activity and thermo stability have been generated in such a way [13].

De novo Enzyme Engineering

De novo synthesis of enzymes means that enzymes are being synthesized from the scratch and with respect to their reaction or substrate mechanism; these are not centered on their related parent enzyme. The de novo synthesis can be done by using

i) in silico-rational design; ii) the knowledge of a reaction mechanism; and iii) mRNA display to search large protein libraries [26]. It is far much easier to search de novo proteins from larger libraries using mRNA display method as compared to cell surface and phage display techniques [7], because the mRNA makes covalent bond with the protein encoded by it and makes the direct amplification of desired protein easier [27,28].

Applications of Protein Engineering

Industrial Applications

A broad range of enzymes are being used in different industries like food, paper, leather, cosmetic, pharmaceutical and chemical industry. From early 1990s scientists have been started protein engineering to produce new enzymes for biotechnological industries [29]. Principally, food industry expenditures a diversity of enzymes like proteases, lipases, amylases etc. in food processing. These processes require mostly high temperature, different pH range and also many other compounds are present there, which can inhibit/hinder enzyme activity. So, to overcome these problems and to further enhance their production and activity, properties of enzymes, which include specificity, thermos-stability and catalytic activity, are improved by making the use of new approaches of protein engineering as described above.

Proteases

Are employed in numerous industrial processes for example in paper industry as biofilm removal, in food industry in milk clotting, meat tenderization and to add up flavors and also used in detergents as protein stain removal [30]. Proteins engineers are working to develop engineered proteases, which have ability to act more efficiently at low temperature and alkaline pH. Mesophilic subtilis in proteases from B. sphaericusis modified using direct evolution to work at low temperature. These subtilisin-like proteases show 9.6 times more catalytic efficiency at 10°C [31]. Mutations in more than 275 amino acids of subtilisin have been reported. Subtilisin BPN, subtilisin E and savinase are most mutagenized proteases used industrial processes [32]. Purafect, Maxapem and Durazym are new bacterial alkaline proteases having the enhanced catalytic activity and better stability at higher temperature, against varying washing conditions and oxidizing agents. These are developed by creating mutations using site direct and random mutagenesis [33]. Protein engineering and cloning techniques have made possible to produce commercial proteases with required characters of pH and temperature activity and stability. It has also modified the bacterial species to produce large quantities of enzymes under different stress conditions [34].

Amylases

Are used in many industries to multiply functions for example it is used in food industry to soften bread, adjust flour, for liquefaction and scarification of starch and juice treatment. In detergent and paper industry, these enzymes are used to remove starch stains and de-inking [30]. For the production of different food and industrial products starch is converted into bioethanol or into food ingredients like fructose, glucose and organic acids in microbial fermenters which require biocatalysts such amylase for the liquefaction and scarification. So to improve the activity and stability of amylases at harsh conditions, protein engineering and DNA recombinant technology are have been used. Scientists have been developed engineered Bacillus α-amylase by creating hybrids [35], introducing proline residues in loop regions [36] and random mutagenesis. Akoh et al., reported rice as an instance for the production of industrial useful biocatalysts from raw material of agriculture [37]. Yeast Pichia pastorisis a value able host for enhanced expression of recombinant α-amylase gene [38].

Lipases

Are also used intensively by food and detergent industries such as for lipid stain removal, chees flavor, dough stability and as contaminants controller in paper & pulp industry. For food processes toxicologically safe lipases are required which are obtained from Candida rugose. Different commercial isoforms of lipases are

produced by DNA shuffling, computer modeling and protein engineering [37]. Later on a comprehensive study was accomplished on mutagenesis and protein engineering to enhance the catalysis of microbial lipases [39].

Environmental Applications

Oxygenases, laccases and peroxidases are three major classes of enzymes, which have significant role in environmental applications for biodegradation of organic and toxic pollutants. But mostly, these enzymes face problems like enzyme denaturation by toxic compounds, inhibition of ES (enzyme- substrate) complex and low catalytic activity. Scientists have done intensive work to overcome these problems by developing engineered enzymes by recombinant technology and rational enzyme design [40].

Medical and Clinical Applications

Protein engineering has enormous applications in the area of therapeutics. Previously protein engineering is done to obtain second generation recombinant protein having substantial properties in clinical applications [41]. Mutation, DNA shuffling and recombinant DNA approach were used in protein engineering to get improved results of therapeutic protein [42]. Later advancements in protein engineering resulted in production of secreted therapeutic proteins such as interferon, insulin, etc. [43], use of combinatorial proteins for therapeutics [44], and also development in gene therapy by inducing recombination using mega nucleases and DNA double-strand breaks [45,46]. Development of therapeutics against cancer is the major field of interest in protein engineering. One of potential treatment recommended for cancer is pre-targeted immunotherapy in which radiation toxicity is thought to be minimized. By using protein engineering, the use of this pre-targeted immunotherapy was expected to be an efficient treatment for cancer [47]. Advancements in recombinant DNA technology and protein engineering enable the synthesis of novel antibodies, which can be used as anti-cancer drugs. These unique antibodies are engineered such a way that they precisely identify and bind with higher affinity with their cancerous antigenic markers, and aid in eliminating the cancerous cell with greater accuracy [48]. Improvement in protein engineering gives rise to some of its other significant medical applications. One of them is protein cationization technique that helps in development of future therapeutics [49]. Tissue regeneration and polymer based drug delivery system was another milestone of protein engineering [50]. Targeted drug delivery remains the prominent feature of a novel biopharmaceutical to achieve successful therapies. Functional proteins and peptides are engineered offering an efficient vehicle for adequate and targeted delivery of drug in this concern. Emerging cancer therapies are the gift of this "modular protein engineering" involving the application of highly specific, smart protein based targeted drug delivery [51]. Indisputably, health care can be more functioning if the diagnosis is rapid, precise and insightful [52]. About twelve hundred genetic disorders have been reported. Most of human contain a few genes with no signs of disease and many of them are responsible for susceptibility, though molecular basis of majority of these diseases is still obscure [52]. Effective efforts have been made in last more than three decades in view of diagnosing genetic disorders prior to embryo implantation in humans, and gathered a lot of worth [53]. As a consequence of the discovery of complete genome synchronizing with relevant protein sequences of Mycoplasma genetalium, approximately 10% error rate in the explanation for more than 300 genes was monitored [54]. The outcome and certain results can easily be monitored if such genetic disorder frequencies are extrapolated to the human genome. In view of overcoming such errors, authentication of the gene products with the help of protein-based tools is mandatory [52]. Beadle and Tatum proposed one gene-one hypothesis, was condemned later; has shown that certain genes result in dozens of proteins [55]. These may be produced either in traces with a very short half-life, fragmented, chemically altered or the fragments of various genes are restructured. Such modulations are in process to be the vital elements to understand functional features of various diagnostically noteworthy proteins [53].

In the beginning of 20th century, it has been tried to fractionate certain enzyme proteins (PP1γ2) and protein-protein complexes viz. PP1γ2-sds22, PP1γ2-14-3-3 and PP1γ2-hsp90 [56-58], which seem to be the crucial markers for regulation of sperm maturation, motility, capacitation and fertilization phenomena. As far as confirmatory depictions of protein-protein interactions, protein cross-linking and post translational modulations of these and/or certain diagnostically important proteins, are concerned, cannot be obtained through genetics. Therefore, in such conditions gene analysis is not suitable in clinical diagnosis of the proteins, and ultimately proteomics requires the characterization of certain proteins, which are active agents of a cell and gene products. These agents directly contribute to the drug development as all drugs are directed against proteins, except a few, interfere in DNA replication in cancer cells and RNA in AIDS virus multiplicity. The estimation of proteins are not yet sensitive enough to detect minute quantities present in the tissues and/or biological fluids, although protein based diagnosis are in craze. Thus, improvements in protein detection and characterization techniques would help in diagnosing diseases with precision and sensitivity. Henceforth, developments in protein nanotechnologies, which have been carried out in recent years, are reviewed here.

It is quite significant to register the protein concentration in a biological sample before studying for its functional activity. The precise quantification of low abundance protein is the biggest challenge, which has been overcome by nanotechnology. Nano-orange reagent technique [59], binding of silver particles to glutaralde-

hyde proteins [60], fluorometric assay [61], ELISA [62], radioimmunoassay [63] and immunofluorosence detection techniques are the tools to quantify the proteins in nano quantity and even less, though, except spectro fluorometric [61] technique, those are multi-step, complicated and time-consuming methods.

Protein Engineering in Nanobiotechnology

The applications of protein engineering in nanobiotechnology are progressing with the time. Nanotechnology was not getting appraisal due to their problematic synthesis and assembly in functional systems. Then came the studies related to biomolecular structural organizations, which revealed their hierarchical arrangements from nano to macro levels. Proteins, lipids and carbohydrates are the biological macromolecules, which are used for biosynthetic formation of tissues under regulated gene expressions. Proteins are the most significant of them as they are the structural components during tissue formation and assist in the transport and arrangement of building materials. Thus proteins are the major focus for nanotechnological systems in their regulated synthesis and assembly. The combinatorial methods of biology employed in protein engineering such as the technologies of bacterial cell surface display and phage display also find their applications in nanobiotechnology to screen selectively binding polypeptide sequences to inorganic surfaces. Individual clones, which are specific in their binding to an inorganic material surface, are revealed through stepwise washings of phages or cells in the biological technique named as Bio-panning. Sequencing of these clones is performed in order to get the amino acid sequences of these polypeptides, specifically bind to semi-metal oxides and other nanotechnology surfaces1 Nanobiotechnology excelled further through another technique using Genetically Engineered Proteins for inorganics i.e. GEPIs suggesting self-arrangement of molecular systems [52]. After then, a number of specific peptides, which bind to certain surfaces like quartz and gold, have been selected characterized [64,65]. Computational methods were combined with experimental approaches in order to better engineer the binding of peptides and assembly of nanotechnology systems giving higher function specific peptides, which can be applied in therapeutics, tissue engineering and nanotechnologies utilizing biological, organic and inorganic materials [66]. Protein engineered peptides are used in biosensors, used as molecular motors and transducers, in the generation of biocompatible nano materials. Bioinformatics analyses have also great impact in this emerging field of protein engineering [67]. Amyloid fibrils are also an important and fascinating application of protein engineering in the construction of nano wires as they serve as the templates. This is a property of many of the proteins that they form an organized aggregate of fibrils named as amyloid fibrils. This quality of well-organized non-covalent aggregate formation ability of amyloid fibrils allows their utilization in nanotechnology where self-assembly and organization of small molecules is critical [68].

Other Emerging Applications

Innovative proteins known as affibody binding proteins, which are of non-immunoglobulin (Ig) origin, have been developed using protein engineering techniques. They have high affinity and are used in diagnostics, viral targeting, bioseparation and tumor imaging [69,70]. For development of new biosensors for analytical diagnosis, insertional protein engineering has been observed to immerge during last 10 years [71,72]. The amino acid sequence of a protein affects both its conformation and function. Thus, the ability to modify the sequence, and hence the structure and activity, of individual proteins in a systematic way, opens up many opportunities, both scientifically and for exploitation in biocatalysis. Modern techniques of synthetic biology, whereby progressively large sequences of DNA can be synthesized de novo, permit an unparalleled ability to engineer proteins with novel functions. Nevertheless, the number of possible proteins is far too large to test individually, so certain means are needed for navigating the 'search space' of probable protein sequences competently and consistently in order to find anticipated activities and other properties. Enzymologists differentiate binding (Kd) and catalytic (kcat) stages. In a similar way, judicious approaches have blended design (for binding, specificity and active site modeling) with the more empirical techniques of classical directed evolution (DE) for improving kcat (where natural evolution rarely pursues the highest values), particularly with regard to residues distant from the active site and where the functional linkages underpinning enzyme dynamics are both unknown and hard to predict. Epistasis (where the 'best' amino acid at one site depends on that or those at others) is a noteworthy feature of directed evolution. The objective of this overview is to highlight some of the approaches, which are being developed to allow using directed evolution to improve enzyme properties, often dramatically. It has been noticed that directed evolution varies in a number of ways from natural evolution, including in particular the existing mechanisms and the prospective selection pressures. Thus, it is hereby stressed on opportunities afforded by techniques that enable protein engineer or enzymologist to map sequence to (structure and) activity in silico, as an effective means of modeling and exploring protein landscapes. Because known landscapes may be assessed and reasoned about as a whole, simultaneously, this offers opportunities for protein improvement not readily available to natural evolution on speedy timescales. Intelligent landscape triangulation, well-versed by sequence-activity interactions and coupled to the emerging techniques of synthetic biology, offers scope for the development of novel biocatalysts, which are both highly active and robust. Besides, for gene expression analysis, zinc finger protein engineering is becoming attractive for molecular biologists. Later on a three-finger protein was successfully engineered to study the expression of an oncogene in mouse cell line [73,74]. The understanding of gene regulation and

the structure and function of the human genome amplified dramatically at the end of the 20th century. However, the technologies for manipulating the genome have been slower to develop. For example, the arena of gene therapy has been focused on correcting genetic diseases and increasing tissue repair for more than 40 years. However, with the exception of a few very low efficacy methodologies, conventional genetic engineering approaches have only been capable to supplement auxiliary genes to cells. This has been a considerable complication to the clinical success of gene therapies and has also directed to severe unintended concerns in several cases. Consequently, technologies that facilitate the defined modification of cellular genomes have diverse and remarkable implications in many facets of research and are significant for translating the products of the Genomic Revolution into noticeable benefits for medicine and biotechnology. To address this requirement, in the 1990s, a mission was embarked to develop technologies for engineering protein-DNA interactions with the purpose of generating custom tools capable of targeting any DNA sequence. The goal has been to allow researchers to reach into genomes to specifically regulate, knock out, or replace any gene. To realize these goals, it has primarily been focused on understanding and manipulating zinc finger proteins. In particular, it is sought to create a simple and straightforward method that enables unspecialized laboratories to engineer custom DNA-modifying proteins employing only defined modular components, a web-based usefulness and standard recombinant DNA technology. Two substantial challenges faced so far were

- The development of zinc finger domains that target sequences not recognized by naturally occurring zinc finger proteins and
- Determining how individual zinc finger domains could be chained together as polydactyl proteins to recognize unique locations within complex genomes.

Various researchers have since used this modular assembly method to engineer artificial proteins and enzymes that activate, repress, or create definite changes to user-specified genes in human cells, plants, and other organisms. Besides, engineered certain novel techniques for externally controlling protein activity and delivery have been successfully developed [74], as well as developed new approaches for the directed evolution of protein and enzyme function. This compilation highlights independent studies, which have successfully used the modular assembly approach to create proteins with novel function, and focuses on emerging alternate procedures for genomic targeting, including transcription activator-like effectors, and how they complement the synthetic zinc finger protein technology. However, in biofuel industry, to obtain bio fuels from lingo cellulosic materials, such cellulose enzymes are produced by protein engineering, which have improved catalytic activity and reduced the production costs of biofuels [75]. Protein cysteine modification, an approach of protein engineering, produces proteins with diverse functions [76,77]. The usage of proteins as therapeutics has a long history and is becoming ever more common in modern medicine. Despite the fact that number of protein-based drugs is growing every year, major problems still remain with their application. Among these complications are quick degradation and excretion from patients, consequently requiring recurrent dosing that in turn increases the chances for an immunological response as well as increasing the cost of therapy. One of the main strategies to improve these problems is to link a polyethylene glycol (PEG) group to the protein of interest. This procedure, called PEGylation, has grown strongly in recent years occasioning in several approved drugs. Installing a single PEG chain at a definite site in a protein is quite challenging. There has been substantial research into several approaches for the site-specific PEGylation of proteins. After introducing the site-specific PEGylation, recent developments using chemical methods have been comprehended. That is followed by a more extensive discussion of bio-orthogonal reactions and enzymatic labeling. More specifically, such novel proteins are frequently used to develop new therapeutic proteins, which show improved half-life and reduced toxicity [76,77].

At last but not the least, it is mandatory to discuss about the thioredoxins, which are recognized as ubiquitous proteins containing a conserved -Trp-Cys-Gly-Pro-Cys-Lys- redox catalytic site. Mammalian thioredoxin family members include thioredoxin-1 (Trx1), mitochondrial thioredoxin-2 (Trx2), and a larger thioredoxin-like protein, p32TrxL [78-81]. Thioredoxin is reduced by NADPH and thioredoxin reductase and, in turn reduces oxidized cysteine groups on certain specific proteins. When thioredoxin levels are elevated there is increased cell growth and resistance to the normal mechanism of programmed cell death [78]. An increase in thioredoxin levels seen in many human primary cancers compared to normal tissue appears to contribute to increased cancer cell growth and resistance to chemotherapy [79]. Mechanisms by which thioredoxin increases cell growth include an increased supply of reducing equivalents for DNA synthesis, activation of transcription factors that regulate cell growth and an increase in the sensitivity of cells to other cytokines and growth factors [80]. The mechanisms for the inhibition of apoptosis by thioredoxin are just now being elucidated [81]. Because of its role in stimulating cancer cell growth and as an inhibitor of apoptosis, thioredoxin offers a target for the development of certain specific drugs to treat and prevent cancer [78-81].

Conclusion

Protein engineering is one of the applications of recombinant DNA technology. Rational design that requires the prior knowledge, has gained importance because of computational algorithms and techniques generating useful output from protein sequence. Directed evolution on the other hand is a lengthy process involv-

ing screening and selection but provides a fair chance to have protein that might not be present in nature. Though conventional techniques have always been proven useful, protein engineering has contributed to study functional properties in more diverse way. Classes of engineered enzymes such as proteases and amylases have substantial applications in food, detergent, paper and several other industries. Other classes such as peroxidases and oxygenases are being used in environmental studies. Pharmaceutical products such as engineered antibodies have also been in market. Novel engineered proteins are being used in diagnostics and biosensors. Besides that nanobiotechnology is also getting benefit through this field. Protein engineering will keep on as a source for creating diversity in proteins to be used as experimental tools in metabolic engineering and protein studies. Further improvements in protein engineering are anticipated through the applications of advanced 'omics' technologies covering from genes to metabolites of biotechnological significance. Lastly, because of the role in stimulating cancer cell growth and as an inhibitor of apoptosis, thioredoxin offers a target for the development of novel drugs to treat and prevent cancer.

Acknowledgements

Author is grateful to Prof. R.M. Dubey (Vice Chancellor, IFTM University, Moradabad, India) for providing Institutional Promotion Grant to the Department of Biotechnology, IFTM University, Moradabad. Besides, Prof. M.I.H. Ansari (Director, School of Engineering & Technology, IFTM University, Moradabad, India) is duly acknowledged for time-to-time inspiration so that enabling to accomplish the compilation embodied in the present paper.

References

1. Van den Berg BA, Reinders MJT, Van der Laan J-M, Roubos JA, de Ridder D (2014) Protein redesign by learning from data. Protein Engineering, Design & Selection 27: 281-288.
2. Dill KA and Maccallum JL (2012) The protein-folding problem, 50 years on. Science 338: 1042-1046.
3. Zawaira A, Pooran A, Barichievy S, Chopra D (2012) A discussion of molecular biology methods for protein engineering. Molecular Biotechnology 51: 67-102.
4. Tiwari MK, Singh R, Singh RK, Kim I-W, Lee J-K (2012) Computational approaches for rational design of proteins with novel functionalities. Computational and Structural Biotechnology Journal 2: 1-13.
5. Ali MA, Muzaffar A, Awan MF, Din S, Nasir IA, et al. (2014) Genetically modified foods: Engineered tomato with extra advantages. Adv. Life Sci 1: 139-152.
6. Ali MA, Rehman I, Iqbal A, Din S, Rao AQ, et al. (2014) Nanotechnology, a new frontier in Agriculture. Adv Life Sci 1: 129-138.
7. Arnold FH (1993) Engineering proteins for no natural environments. The FASEB Journal 7: 744-749.
8. Antikainen NM and Martin SF (2005) Altering protein specificity: techniques and applications. Bioorganic & Medicinal Chemistry 13: 2701-2716.
9. Affholter J and Arnold FH (1999) Engineering a revolution. Chemtech 29: 34-39.
10. Israr M, Ilyas M, Rahman Z (2014) Copy number variation in forensic science. Advancements in Life Sciences 1: 71-72.
11. Leung DW, Chen E, Goeddel DV (1989) A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. Technique 1: 11-15.
12. Wong TS, Tee KL, Hauer B, Schwaneberg U (2004) Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution. Nucleic Acids Research 32: e26-e26.
13. Packer MS and Liu DR (2015) Methods for the directed evolution of proteins. Nature Reviews Genetics 16: 379-394.
14. Saif R, Khan E, Azhar A, Choudhary S, Hussain T, et al. (2016) Insight of Tp53 mutations and their effect on protein in different feline and canine neoplasms. Advancements in Life Sciences 3: 42-50.
15. Kegler-Ebo DM, Docktor CM, Dimaio D (1994) Codon cassette mutagenesis: a general method to insert or replace individual codons by using universal mutagenic cassettes. Nucleic Acids Research 22: 1593-1599.
16. Siloto RM and Weselake RJ (2012) Site saturation mutagenesis: Methods and applications in protein engineering. Biocatalysis and Agricultural Biotechnology 1: 181-189.
17. Stemmer WP (1994) Rapid evolution of a protein in vitro by DNA shuffling. Nature 370: 389-391.
18. Ness JE, Kim S, Gottman A, Pak R, Krebber A, et al. (2002) Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently. Nature Biotechnology 20: 1251-1255.
19. Muller KM, Stebbel SC, Knall S, Zipf G, Bernauer HS, et al. (2005) Nucleotide exchange and excision technology (NExT) DNA shuffling: a robust method for DNA fragmentation and directed evolution. Nucleic Acids Research 33: e117.
20. Zhao H, Giver L, Shao Z, Affholter JA, Arnold FH (1998) Molecular evolution by staggered extension process (StEP) in vitro recombination. Nature Biotechnology 16: 258-261.
21. Silwoski G, Kothiwale SK, Meiler J, Lowe Jr EW (2014) Computational methods in drug discovery. Pharmacol Rev 66: 334-395.
22. Bessette PH, Rice JJ, Daugherty PS (2004) Rapid isolation of high-affinity protein binding peptides using bacterial display. Protein Engineering Design and Selection: 17: 731-739.
23. Mccafferty J, Griffiths AD, Winter G, Chiswell DJ (1990) Phage antibodies: filamentous phage displaying antibody variable domains. Nature 348: 552-554.
24. Liu JKH (2014) The history of monoclonal antibody development–Progress, remaining challenges and future innovations. Annals of Medicine and Surgery 3: 113-116.
25. Hanes J and Pluckthun A (1997) In vitro selection and evolution of functional proteins by using ribosome display. Proceedings of the National Academy of Sciences 94: 4937-4942.
26. Seelig B and Szostak JW (2007) Selection and evolution of enzymes form a partially randomized noncatalytic scaffold. Nature 448: 828-831.
27. Golynskiy MV and Seelig B (2010) De novo enzymes: From computational design to mRNA display. Trends in Biotechnology 28: 340-345.

28. Ali MA and Abbas A (2016) Analysis of reporter proteins GUS and DsRed driven under the control of CaMV35S promoter in syncytia induced by beet cyst nematode Hetero deraschachtii in Arabidopsis roots. Advancements in Life Sciences 3: 89-96.
29. Wiseman A (1993) Designer enzyme and cell applications in industry and in environmental monitoring. Journal of Chemical Technology and Biotechnology 56: 3-13.
30. Kirk O, Borchert TV, Fuglsang CC (2002) Industrial enzymes' applications. Current Opinion in Biotechnology 13: 345-351.
31. Wintrode PL, Miyazaki K, Arnold FH (2000) Cold adaptation of a mesophilic subtilisin-like protease by laboratory evolution. Journal of Biological Chemistry 275: 31635-31640.
32. Bryan PN (2000) Protein engineering of subtilisin. Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology 1543: 203-222.
33. Gupta R, Beg Q, Lorenz P (2002) Bacterial alkaline proteases: molecular approaches and industrial applications. Applied Microbiology and Biotechnology 59: 15-32.
34. Schallmey M, Singh A, Ward OP (2004) Developments in the use of Bacillus species for industrial production. Canadian Journal of Microbiology 50: 1-17.
35. Suzuki Y, Ito N, Yuuki T, Yamagata H, Udaka S (1989) Amino acid residues stabilizing a Bacillus alpha-amylase against irreversible thermoinactivation. Journal of Biological Chemistry 264: 18933-18938.
36. Matthews BW, Nicholson H, Becktel W (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. Proceedings of the National Academy of Sciences: 84: 6663-6667.
37. Akoh CC, Chang S-W, Lee G-C, Shaw J-F (2008) Biocatalysis for the production of industrial products and functional foods from rice and other agricultural produce. Journal of Agricultural and Food Chemistry 56: 10445-10451.
38. Tull D, Gottschalk TE, Svendsen I, Kramfoft B, Phillipson BA, et al. (2001) Extensive N-glycosylation reduces the thermal stability of a recombinant alkalophilic Bacillus α-amylase produced in Pichia pastoris. Protein Expression and Purification 21: 13-23.
39. Verma M, Azmi W, Kanwar S (2008) Microbial lipases: at the interface of aqueous and nonaqueous media: a review. Acta Microbiologica et Immunologica Hungarica 55: 265-294.
40. Karigar CS and Rao SS (2011) Role of microbial enzymes in the bioremediation of pollutants: a review. Enzyme Research 2011: 1-11.
41. Buckel P (1996) Recombinant proteins for therapy. Trends in Pharmacological Sciences: 17: 450-456.
42. Kurtzman AL, Govindarajan S, Vahle K, Jones JT, Heinrichs, et al. (2001) Advances in directed protein evolution by recursive genetic recombination: applications to therapeutic proteins. Curr Opin Biotechnol 12: 361-370.
43. Bonnin-Debs AL, Boche I, Gille H, Brinkmaan U (2004) Development of secreted proteins as biotherapeutic agents. Expert opinion on Biological Therapy 4: 551-558.
44. Lowe D and Jermutus L (2004) Combinatorial protein biochemistry for therapeutics and proteomics. Current Pharmaceutical Biotechnology 5: 17-27.
45. Paques F and Duchateau P (2007) Mega nucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. Current Gene Therapy 7: 49-66.
46. Shahzad M, Shafique M, Hussain M, Shan MA, Perveen R, et al. (2016) Successful DNA profiling for identification of burnt families from their bones using AmpFℓSTR identifiler RPlus Kit. Advancements in Life Sciences 3: 59-62.
47. Lam L, Liu X, Cao Y (2003) Pre targeted radio immunotherapy, a potential cancer treatment. Drugs of the Future 28: 167-173.
48. Zafir-Lavie I, Michaelil Y, Reiter Y (2007) Novel antibodies as anticancer agents. Oncogene 26: 3714-3733.
49. Futami J, Kitazoe M, Murata H and Yamada H (2007) Exploiting protein cationization techniques in future drug development. Expert Opinion on Drug Discovery 2: 261-269.
50. Liu S, Maheshwari R, Kick KL (2009) Polymer-based therapeutics. Macromolecules 42: 3-13.
51. Vazquez E, Ferrer-Mirrales N, Mangues R, Corchero JL, Schwartz J, et al. (2009) Modular protein engineering in emerging cancer therapies. Current Pharmaceutical Design 15: 893-916.
52. Mishra Sand Maheshwari SK (2008) Development of protein nanotechnology: Applications in disease diagnosis at proteomics level. Nanobiotechnology; Pointer Publishers 70-76.
53. Verlinsky Y, Rechitsky S, Schoolcraft W, Strom C, Kuliev A (2001) Preimplantation diagnosis for Fanconi anemia combined with HLA matching. JAMA 285: 3130-3133.
54. Bocs S, Danchin A, Médigue C (2002) Re-annotation of genome microbial CoDing-Sequences: finding new genes and inaccurately annotated genes. BMC Bioinformatics 3: 5-8.
55. Davis RH (2007) Beadle's progeny: Innocence rewarded, innocence lost. Journal of Biosciences 32: 197-205.
56. Mishra S, Somanath PR, Huang Z, Vijayaraghavan S (2003) Binding and inactivation of the germ cell-specific protein phosphatase PP1γ2 by sds22 during epididymal sperm maturation. Biology of Reproduction 69: 1572-1579.
57. Mishra S, Huang Z, Vijayaraghavan S (2003) Binding to sds22 and inactivation of catalytic activity of the protein phosphatase PP1 gamma 2 occurs during sperm motility initiation in the epididymis. Biology of Reproduction 68: 170-170.
58. Myers K, Somanath PR, Berryman M, Vijayaraghavan S (2004) Identification of chloride intracellular channel proteins in spermatozoa. FEBS letters 566: 136-140.
59. Movileanu L, Howorka S, BrahaO, Bayley H (2000) Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore. Nature Biotechnology 18: 1091-1095.
60. Smith GS, Schloesser R, Brodie JD, Dewey SL, Logan J, et al. (1998) Glutamate modulation of dopamine measured in vivo with positron emission tomography (PET) and C-raclopride in normal human subjects. Neuropsychopharmacology 18: 18-25.
61. Sangwan RS, Mishra S, Kumar S (1998) Direct fluorometry of phase-extracted tryptamine-based fast quantitative assay of L-tryptophan decarboxylase from Catharanthus roseus leaf. Analytical Biochemistry 255: 39-46.
62. Lequin RM (2005) Enzyme immunoassay (EIA)/enzyme-linked immunosorbent assay (ELISA). Clinical Chemistry 51: 2415-2418.
63. Goldsmith SJ (1975) Radioimmunoassay: review of basic principles. Semi Nucl Med 5: 125-152.
64. Sekker E, Reed ML, Begley MR (2009) Nanoporous gold: fabrication, characterization, and applications. Materials 2: 2188-2215.

65. Oren M and Rotter V (2009) Mutant p53 gain-of-function in cancer. Cold Spring Harbor Perspectives in Biology.

66. Tamerler C, Khatayevich D, Gungormus M, Kacar T, Oren EE, et al. (2010) Molecular biomimetics: GEPI-based biological routes to technology. Peptide Science 94: 78-94.

67. Banta S, Megeed Z, Casali M, Rege K, Yarmush ML (2007) Engineering protein and peptide building blocks for nanotechnology. Journal of Nanoscience and Nanotechnology 7: 387-401.

68. Hamada D, Yanagihara I, Tsumoto K (2004) Engineering amylo idogenicity towards the development of nano fibrillar materials. Trends in Biotechnology 22: 93-97.

69. Nygren PA (2008) Alternative binding proteins: Affibody binding proteins developed from a small three-helix bundle scaffold. Febs Journal 275: 2668-2676.

70. Mouratou B, Behar G, Pecorari F (2015) Artificial affinity proteins as ligands of immunoglobulins. Biomolecules 5: 60-75.

71. Ferraz R, Vera A, Aris A, Villaverde A (2006) Insertional protein engineering for analytical molecular sensing. Microbial Cell Foactries 5: 15.

72. Currin A, Swainston N, Day PJ, Kell DB (2015) Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. Chem Soc Rev 44: 1172-1239.

73. Klug A (2010) The discovery of zinc fingers and their applications in gene regulation and genome manipulation. Annual Review of Biochemistry 79: 213-231.

74. Gersbach CA, Gaj T, Barbas III CF (2014) Synthetic Zinc Finger Proteins: The Advent of Targeted Gene Regulation and Genome Modification Technologies. Acc Chem Res 47: 2309-2318.

75. Wen F, Nair NU, Zhao H (2009) Protein engineering in designing tailored enzymes and microorganisms for biofuels production. Current Opinion in Biotechnology 20: 412-419.

76. Nagahara N, Matsumura T, Okamoto R, Kajihara Y (2009) Protein cysteine modifications: (2) reactivity specificity and topics of medicinal chemistry and protein engineering. Current Medicinal Chemistry 16: 4490-4501.

77. Dozier JK and Distefano MD (2015) Site-specific PEGylation of therapeutic proteins. Int J Mol Sci 16: 25831-25864.

78. Powis G, Mustacich D, Coon A (2000) The role of the redox protein thioredoxin in cell growth and cancer. Free Radic Biol Med 29: 312-322.

79. Biaglow JE and Miller RA (2005) The thioredoxin reductase/thioredoxin system: novel redox targets for cancer therapy. Cancer Biol Ther 4: 6-13.

80. Powis G and Kirkpatrick DL (2007) Thioredoxin signaling as a target for cancer therapy. Curr Opin Pharmacol 7: 392-397.

81. Dal Piaz F, Braca A, Belisario MA, De Tommasi N (2010) Thioredoxin system modulation by plant and fungal secondary metabolites. Curr Med Chem 17: 479-494.