Introduction

Microbial enzymes catalyze a wide range of chemical reactions essential for life and have gained great interest for their utilization in various industries such as leathers, textile, pharmaceuticals, food, agrochemicals, agriculture, biofuel and other industries and applications. Recently, there is a growing global attention to biocatalysts exploitation as ecofriendly alternative of chemical catalysts in various chemical and industrial sectors (Heux et al., 2015). This is due to the unique advantageous properties of the enzymes over chemical catalysts such as moderate reaction conditions, selectivity, specificity and low cytotoxicity (Jemli et al., 2016). The limitations of natural enzymes application in industrial processes include but not limited to regio- and enantio-selectivity, stability in organic solvents, thermo- and pH stability, specificity and cost-effectiveness. (Porter et al., 2016). Therefore, discovery of new enzymes with better performance and development of new and innovative approaches for improvement of enzyme function and properties are required to increase and expand enzyme applications in various industries (Woodley, 2013; Siddiqui, 2015 and Madhavan et al., 2017).

Recently, there is significant progress in developing new strategies and approaches for novel enzymes discovery and biocatalysts engineering, leading to significant increase of the global industrial enzymes market to about $4.2 billion in 2014, with expectation to reach $6.2 billion by 2020 (Singh et al., 2016). In the past, the industrial bioprocess conditions were modified around the enzymes limitations; however now the biocatalyst is engineered to meet the process requirements through application of various engineering strategies. This review highlights state-of-the-art strategies and approaches for new enzymes discovery and function improvement, as well as the author perspective.

Keywords: Enzyme engineering, Novel enzymes discovery, Directed evolution, Rational design, Metagenomics, Nanobiocatalysis.
Recent approaches for enzyme discovery

Traditionally, microbial biocatalysts discovery was relied on culturable microorganisms including enrichment, isolation and screening for enzymes activities. This strategy has resulted in discovery of numerous enzymes that have been applied in several industrial applications (Winter et al., 2011 and Kaul & Asano, 2012). However, it has been well established that the culturable microbes represent less than 1% of the entire microbial diversity on the planet. This conclusion was made based on the ‘great plate count anomaly’, which refers to huge difference between the bacterial cells number detected in an environmental specimen and that which can be cultured under experimental conditions (Pande & Kost, 2017). For instance, it is estimated that Earth count for bacteria and fungi is about 1 million and 1.5-5.0 million species, respectively. However, <1% of them has been identified to date (Harris, 2012 and Jacque et al., 2017). Hence, only a very tiny fraction of microorganisms is accessed for enrichment and screening for new enzyme activities which in turn would narrow the potential of novel enzymes discovery. Therefore, the not yet culturable microorganisms have a great possibility for new biocatalysts discovery for potential applications in several fields (Porter et al., 2016). The recent advances of cultivation-independent techniques, genes sequencing and enzyme structure databases, have led to the development of: (i) Metagenome analysis and (ii) Proteins database mining (Fig. 1), as effective approaches for novel biocatalysts discovery (Yang & Ding, 2014).

Metagenome-based approach for novel enzymes discovery

The constraint that most microorganisms are not cultivated under the experimental conditions is overcome by the metagenomics approach which became a well-established strategy for discovery of new biocatalysts. Metagenomics is a highly efficient tool to enhance enzymes discovery from the huge diversity of unculturable microorganisms (Mirete et al., 2016). Metagenome analysis is comprised of total genomic DNA extraction directly from environmental samples, construction of metagenomics library of extracted DNA, followed by screening of the generated library for new enzymes activities. Application of metagenome-based approach has resulted in identification of several new enzymes. One of the most remarkable successes was the discovery of polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of Paederus beetles, leading to production of polyketide antibiotics (Piel, 2002 and Deepa et al., 2017).

![Fig. 1. Strategies for novel biocatalysts discovery (Yellow arrow) and engineering (red arrow) (Yang & Ding, 2014).](image-url)
In addition, several novel nitrilases was recently identified and characterized, which catalyze enantioselective hydrolysis of 3-hydroxyglutaronitrile to (R)-cyano-3-hydroxybutyric acid (Fig. 2), an intermediate for the production of Atorvastatin, a lipid-lowering drug (Tarde name: Lipitor; Pfizer, USA) with annual sales of more than $1.0 billion/year (Kau & Asano, 2012). Two main complementary screening strategies are used for novel enzymes discovery from the constructed metagenomics libraries including, (i) Function/activity-based screening (functional metagenomics) (Hosokawa et al., 2015) and (ii) DNA sequencing-based screening (Tomazetto et al., 2015).

**Functional metagenomics:** Functional metagenomics relies on activity-based screening of the metagenomics libraries for identification of new enzymes. The total genomic DNA is extracted directly from environmental sample, digested either mechanically or using suitable endonuclease enzymes, followed by ligation of the DNA fragments into suitable vector and expression in host cells (Fig. 3). Thereafter, the obtained library is screened for new enzymatic activity using suitable screening assay methods (Hosokawa et al., 2015). Application of functional metagenomics has several advantages including: (i) Only the functional genes are detected; (ii) Higher probability of novel genes identification because it does not rely on known genes sequences and (iii) The entire enzyme encoding gene can be obtained (Lopez-Perez & Mirete, 2014). Several novel enzymes with unique properties have been discovered recently by using this strategy that can be utilized in several industrial applications (Table 1). However, the efficiency of this approach depends on the availability of efficient screening assay methods for the desired enzyme activity, in addition to presence of effective expression systems for the enzymes encoding genes. Various vectors such as specific plasmid, phagemid, BAC, and fosmid, as well as heterologous hosts have been recently developed to increase the efficiency of metagenomics genes expression (Liu et al., 2016 and Jose et al., 2017). In order to enhance the detection of functional genes and to expand the functional metagenomics application, reporter genes such as antibiotic resistance marker genes have been recently developed and applied (Ufarté et al., 2015).

![Figure 2](image2.png)

**Fig. 2.** Nitrilase-mediated desymmetrization of prochiral dinitrile to Lipitor intermediate (Kau & Asano, 2012).

![Figure 3](image3.png)

**Fig. 3.** Construction and functional screening of metagenomic libraries (Heux et al., 2015)
Sequence-based metagenomics: In the sequence-based screening approach, designed oligonucleotide probes or primers integrated into PCR protocol are used to target the desired biocatalysts genes, followed by sequencing and read assembly. Bioinformatics analysis of the obtained metagenome sequence is performed for discovery of new enzymes genes through identification of any open reading frames and their functional annotation (Davids et al., 2013 and Ufarté et al., 2015). However, one of the major limitations of sequence-based metagenomics screening strategy is its restriction to discovery of the enzymes with high sequence homology to previously known genes (Tomazetto et al., 2015). Several novel biocatalysts have been discovered using the sequence-based metagenomics approach (Table 1). Jiang et al (2013) reported identification of novel potent prephenate dehydrogenase, playing important role in tyrosine biosynthesis, by sequence-based screening of metagenomics library created from alkaline soil samples. Chen et al (2014) reported the identification of aldehyde dehydrogenase from metagenomic library constructed from hot spring samples, showing high ropionaldehyde oxidative activity. Recently, novel polyhydroxyalkanoate synthase with high activity (64.0U/g proteins) was identified from metagenomics library from limestone soil (Tai et al., 2016).

Database mining approach for novel enzymes discovery

The recent progress on sequencing technology including “next generation sequencing”, has led to uncover the sequences of large number of enzymes which would be identified by "sequence-based" metagenomics and "function-based" metagenomics approaches.
of microbial genomes and metagenomes (Damborsky & Brezovsky, 2014). In addition, the recent advance in the field of bioinformatics and computational biology led to the development of efficient public databases and effective software’s for sequences analysis. Currently, there are more than $10^4$ protein structures available in the Protein Data Bank (http://www.rcsb.org/), more than 6000 microbial genomes and $10^4$ proteins deposited in National Center for Biotechnology Information (NCBI) public database (http://www.ncbi.nlm.nih.gov/). In addition, the recently developed public database, BRENDA (http://www.brenda-enzymes.org/) contain about three million data about enzymes, extracted from publications including enzymes substrates, selectivity, specificity, molecular characterization and biochemical properties (Schomburg et al., 2013). Brookhaven protein structure database (www.pdb.org) is another database for 3D-structure of huge number of proteins which have been not yet characterized regarding their biochemical properties, and it is considered as an unexplored source for novel biocatalysts (Davids et al., 2013).

Mining of these Databases is a promising approach for novel biocatalysts discovery based on the available sequences and enzyme structure data, through sequence-based homology screening and structure similarity analysis (Adesioye et al., 2016). Recently, application of this approach was successful in discovery of several new enzymes including new (R)-selective amine transaminases (Schaetzle et al., 2011), catalyzing conversion of prochiral ketones into chiral (R)-amines (Fig. 4), nitrile hydratase (Pei et al., 2014); vitamin D3-specific hydroxylase (Ban et al., 2014), amidases (Ruan et al., 2016 and Wu et al., 2016), solvent tolerant esterase (Fang et al., 2015); polyketide synthase (Deepa et al., 2017), chitinases (Lin et al., 2017); glutathione S-transferase (Jing et al., 2017); isoprene synthases (Ilmén et al., 2015), glycoside hydrolase family 3 (GH3) aryl β-glucosidases (Kudo et al., 2015) and others (Thevarajoo et al., 2015; Zhang et al., 2015 and Fang et al., 2015).

Recent advances in enhancement of enzyme function

During the evolution time, nature optimized the enzymes functions and properties to meet requirements of the host organism. Thus it is acceptable that natural biocatalysts usually don’t satisfy the industrial process conditions requirement (Yang & Ding, 2014). The limitations of natural enzymes application in industrial processes include but not limited to regio- and enantio-selectivity, stability in organic solvents, thermo- and pH stability, and specificity. (Porter et al., 2016). The main recent strategies for enhancement of enzyme functionality include (i) Enzyme engineering at the molecular level and (ii) Physical stabilization of the enzyme.

Enzyme engineering for enzyme function enhancement

Recent advances and development in the fields of molecular biology, bioinformatics and computational biology allow tuning of the enzyme properties and function through protein engineering, to meet the requirements of the industrial processes conditions and to expand enzymes applications in various industrial dectors (Li & Cirino, 2014). State-of the arte approaches for enzyme engineering includes four main approaches: (i) Directed evolution, (ii)Rational design, (iii) Semi-rational design and (iv) Computational de novo enzyme design (Denard et al., 2015).

![Fig. 4](image)

Fig. 4. (R)-selective amine transaminases, converting prochiral ketones into chiral (R)-amines, discovered through database mining strategy (Schaetzle et al., 2011).
Directed evolution: The concept of directed evolution (DE) is to mimic the natural Darwinian evolution process in the laboratory, i.e. it is in vitro evolution process to improve the enzyme function and properties (Davids et al., 2013 and Arnold, 2015). DE relies on creating a large number of mutants of the enzyme-encoding gene, followed by screening and selection of mutants with the desired improvement using high-throughput screening methods. This process of mutation and selection is repeated until the required biocatalyst function enhancement is achieved (Kumar & Singh, 2013 and Arnold, 2015). In DE strategy, three main methods are applied for creating such enzymes mutants including, (a) Random mutagenesis, (b) Recombination of genes homologs (DNA shuffling) and (c) Cassette mutagenesis (CM). The most common methods for random mutagenesis of the enzyme gene are error-prone PCR and site-saturation mutagenesis (SSM) (Adrio & Demain, 2014). The error-prone PCR produce a huge number of random point mutations in the target gene. It is based on using DNA polymerase with low or no fidelity (i.e. low or no proofreading activity) in PCR, or optimizing the PCR ingredients to enhance nucleotides misincorporation (Wu & Arnold, 2013). Site-saturation mutagenesis (SSM) is an approach for amino acid replacement within the enzyme with all possible amino acids. It can be done for particular residue or randomly within the enzyme. On the other hand, cassette mutagenesis (CM) aimed to create SSM at several adjacent amino acids site. Random mutagenesis approach generates a huge number of enzyme variants library (<10^6) that is subjected to screening for detection of enzyme variants with the desired alteration and improvement (Adrio & Demain, 2014).

DNA shuffling (genetic recombination) is another powerful approach to create enzyme variants. Basically, DNA shuffling involves fragmentation of related enzymes genes (showing more than 70% homology), followed by random reassembly of the obtained DNA fragments using self-priming PCR reaction (Dalby, 2011). The enzymes mutants created using DNA shuffling tendency to have sequences that retain the native enzyme activity, whereas random mutagenesis approach has higher possibility to create enzyme with novel function. This because while random mutagenesis can generate unnatural mutations, DNA shuffling approaches generate variants with mutations that are often occurs in the parent genes sequences (Illanes et al., 2012). Using this approach, a porcine kidney D-amino acid oxidase was converted to R-amine oxidase, thereby providing a new enzymatic route to S-amines synthesis (Fig. 5 a), which was not previously possible (Yasukawa et al., 2014 and Denard et al., 2015). In another work, Denard group was able to engineer the regioselectivity of hydroxylase activity of cytochrome P450, converting cytochrome P450 from terminal- to a sub-terminal hydroxylase with 99% ee enantioselectivity (Fig. 5 b) (Yang et al., 2014 and Denard et al., 2015).

![Fig. 5. (a) Stereoselective conversion of racemic amines by engineered (R)-amine oxidase variant, (b) hydroxylation of n-octane at subterminal position to produce (S)-2-octanol with 99% ee regioselectivity, using engineered cytochrome P450 variant for highly enantioselective and regioselective activities (Yang et al., 2014; Yasukawa et al., 2014 and Denard et al., 2015).](image-url)
Several properties of enzymes were recently improved using the directed evolution approach, some are summarized in Table 2, including enzyme thermostability (Morimoto et al., 2014 and Acevedo et al., 2017), activity (Plummer et al., 2016 and Yu et al., 2016), substrate specificity (Bosshart et al., 2016 and Sun et al., 2016), regio- and enantio-selectivity (Yang et al., 2014 and Denard et al., 2015). Furthermore, Arnold group recently achieved remarkable success in transforming cytochrome P450 BM3 into an efficient cyclopropanation catalyst, via carbene and nitrene transfer in vivo and in vitro, an activity for which there is no precedent in nature (Coelho et al., 2013a; b and Arnold, 2015).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Target</th>
<th>Results</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellobiohydrolases</td>
<td>Thermostability</td>
<td>Increase of T_{50} by ∼15-20°C; decrease of hydrolysis time by 10-fold</td>
<td>Wu &amp; Arnold (2013)</td>
</tr>
<tr>
<td>Lipase (from Candida antarctica)</td>
<td>Thermostability</td>
<td>14-fold increase of the residual activity</td>
<td>Peng (2013)</td>
</tr>
<tr>
<td>Malic enzyme (from Thermococcus kodakaren)</td>
<td>Thermostability</td>
<td>Switch of cofactor preference from NADP(H) to NAD(H), resulting in drastic improvement of the enzyme thermostability</td>
<td>Morimoto et al. (2014)</td>
</tr>
<tr>
<td>Pyrroloquinoline quinone glucose dehydrogenase</td>
<td>Thermostability</td>
<td>11-fold increase of the half-live time at 50°C</td>
<td>Pu et al. (2017)</td>
</tr>
<tr>
<td>Carbonyl reductase (from Ogataea mimic)</td>
<td>Thermostability</td>
<td>11-fold increase of the half-live time at 50°C</td>
<td>Honda et al. (2017)</td>
</tr>
<tr>
<td>Heparinase</td>
<td>Activity</td>
<td>57.8% higher activity than that from the control clone</td>
<td>Yu et al. (2016)</td>
</tr>
<tr>
<td>Hydrogenases (from Clostridium sp.)</td>
<td>Activity</td>
<td>Increase in hydrogen production</td>
<td>Plummer et al. (2016)</td>
</tr>
<tr>
<td>Cold-adapted xylanase</td>
<td>Thermostability</td>
<td>4.3°C increase in its T_{15-50}</td>
<td>Acevedo et al. (2017)</td>
</tr>
<tr>
<td>Glutathione transferases</td>
<td>Stability</td>
<td>Increased oxidative stability</td>
<td>Axarli et al. (2017)</td>
</tr>
<tr>
<td>Fluorinases from Streptomyces sp.</td>
<td>Conversion of the non-native substrate</td>
<td>3-fold higher activity than wild-type</td>
<td>Sun et al. (2016)</td>
</tr>
<tr>
<td>Tagatose Epimerase</td>
<td>Activity for Two Hexose Substrates</td>
<td>Epimerization of d-fructose and l-sorbose, increased by 9- and 14-fold, respectively</td>
<td>Bosshart et al. (2015)</td>
</tr>
<tr>
<td>Cytochromes P450</td>
<td>Enhance the catalytic promiscuity (novel reaction)</td>
<td>Epoxidation of para-substituted styrenes (Novel reaction)</td>
<td>Li et al. (2013)</td>
</tr>
<tr>
<td>Cytochromes P450</td>
<td>Enhance catalytic promiscuity</td>
<td>Novel activity towards meclofenamic acid (a non-steroidal anti-inflammatory drug).</td>
<td>Venkataraman et al. (2014)</td>
</tr>
<tr>
<td>Cytochromes P450</td>
<td>Enhance the catalytic promiscuity</td>
<td>Novel and improved activity of P450 BM3 towards chrysene and pyrene</td>
<td>Sideri et al. (2013)</td>
</tr>
<tr>
<td>Cytochromes P450</td>
<td>Regio-selectivity</td>
<td>Conversion P450pyr from a terminal to a subterminal hydroxylase</td>
<td>Yang et al. (2014)</td>
</tr>
<tr>
<td>Peroxygenase</td>
<td>Activity</td>
<td>8000-fold increase of the catalytic activity</td>
<td>Molina-Espeja et al. (2014)</td>
</tr>
<tr>
<td>D-amino acid oxidase</td>
<td>Substrate specificity</td>
<td>Convert D-amino acid oxidase into an R-amine oxidase (Novel reaction)</td>
<td>Yatsukawa et al. (2014)</td>
</tr>
<tr>
<td>Cytochromes P450</td>
<td>Novel reaction</td>
<td>P450 BM3-catalyzed cyclopropanation between ethyl diazoacetate and styrene derivatives (Novel reaction)</td>
<td>Coelho et al. (2013 a &amp; b)</td>
</tr>
</tbody>
</table>
Directed evolution is one of the most powerful approaches for improvement of enzymes function as well as creating novel enzymes with no precedent biocatalysis in nature, since it does not require prior knowledge of the target biocatalyst structure. However, one of the main drawbacks of directed evolution approach is its need for highly effective high-throughput screening method for identification of the desired enzyme variants in the constructed huge mutants’ library, which is not always available (Jemli et al., 2016 and Porter et al., 2016).

**Rational design:** In contrast to directed evolution (DE), rational design approach (RD) is based on knowledge and understanding of the enzyme structure-function relationship in order to achieve the desired enzyme changes (Denard et al., 2015). The accurate data about protein/DNA sequences, 3-dimensional enzyme structure; and the amino acids residues responsible for the catalytic activity, stability, specificity and selectivity, in addition to other features, is essential for successful application of the rational design approach (Adrio & Demain; 2014). Based on this information, various molecular biology techniques are applied to achieve the desired enzyme modifications. Therefore, the recent remarkable development and success of RD approach is due to the progress on enzyme structure determination, protein modeling and structure–function relation understanding (Wang et al., 2016). Furthermore, the recent advances in enzyme modeling techniques, including homolog structure prediction, molecular dynamics and free energy perturbation determination, have led to possible predication of the suitable point or multiple mutations to achieve desired enzyme modifications (Acevedo et al., 2017). Moreover, using modeling techniques, it is possible to evaluate and predict, *in silico*, the influence of site-directed mutagenesis on the enzyme properties and hence design for less number of mutations (Illanes et al., 2012). However, in some cases, it is difficult to have enough information about enzyme structure particularly in case of multiple subunits enzymes, non-ribosomal peptide synthetases, and membrane bound enzymes. Hence, it would be difficult to predict and determine the proper mutations for the desired alteration (Yang & Ding, 2014). In this case, another approach called “consensus sequence design” developed recently can be applied for prediction of the correct enzyme point mutations to achieve the desired biocatalyst changes (Sullivan et al., 2012). In “consensus sequence design” approach, conserved amino acid residues in sets of homologous enzymes are determined and targeted for site-directed mutagenesis (Sullivan et al., 2012). The hypothesis behind this approach is that the frequency of certain residues distributed among the homologous proteins is positively correlated with its participation in enzyme stability during the biocatalyst natural evolution according to the "Natural Selection Theory". In another word, the protein amino acids residues that tolerated the natural mutation over billions of years are valid targets for enhancing enzyme stabilization (Sullivan et al., 2012 and Yang & Ding et al., 2014).

One of most remarkable achievement using RD approach was reported by Savile et al. (2010) for antidiabetic drug Sitagliptin (Trade name: Januvia®) production. Sitagliptin production by chemical methods is unfavorable due to several problems such as high cost, product contamination and stereoselectivity. There is no natural transaminase that can use the Sitagliptin precursor, prositagliptin ketone, as a substrate. However, by using RD approach, substrate specificity of *Arthrobacter* sp. transaminase was modified and enzyme variants were obtained that can use such bulky ketone structure as a substrate (Fig. 6 A). Furthermore, the obtained transaminase variant was highly stable in high organic solvent concentrations and high temperature, required to enhance the solubility of prositagliptin ketone, leading to the commercial production of the antidiabetic Sitagliptin drug (Januvia®) (Savile et al., 2010 and Kau & Asano, 2012). RD of xylose dehydrogenase for improving enzyme thermostability resulted in obtaining variants with superior thermostability, showing 5-fold higher enzyme half-life (Feng et al., 2016). Recently, Hu et al (2017) reported improvement of several properties of β-Mannanase using RD strategy for enzyme N-glycosylation, where the proper sites for N-glycosylation were selected in the protein loop area and carbohydrate chain was successfully linked to the enzyme. The obtained modified β-mannanase variants showed significant enhancement of the enzyme thermostability, protease-resistance and pH stability (Hu et al., 2017). Substrate specificity of succinic semialdehyde reductase was altered using RD approach for conversion of levulinc acid (cellulosic biomass derivative), which
is not the natural substrate for the enzyme, to 4-hydroxyvaleric acid, a precursor to bio-polymers synthesis (Fig. 6 B). Based on RD, an amino acid residue (Met13) located in the substrate binding site was substituted with six hydrophobic amino acids. One of the obtained enzyme variant showed 7.0 fold increase in catalytic efficiency for conversion of levulinic acid to 4-hydroxyvaleric acid with $K_m$ value of 9.0 and 28.1 mM for mutant and wild type enzyme, respectively (Yeon et al., 2015).

Semi-rational design: Semi-rational design (SRD) is a combination of rational design and directed evolution, which needs only preliminary enzyme structure-function information to determine restricted number of amino acids residues to be targeted for mutation (such as residue in active sites and legend binding sites). Thus, SRD is an approach to create “small but smart library”, i.e. with small number of enzyme variants but highly concentrated and enriched with most likely effective mutations (Zhuang et al., 2017). Semi-rational design is the approach of choice particularly if highly effective high throughput screening method is not available for the library screening to identify the desired enzyme mutants. Combinatorial active-site saturation testing (CASTing) and iterative saturation mutagenesis (ISM) are the most effective protocols used for semi-rational design (Parra et al., 2013). CASTing method involves saturation mutagenesis of multiple sites within enzyme active site. Therefore, it can investigate possible synergistic effects of multiple mutations within the enzyme. However, the obtained library number using this approach could be higher than the capacity of the available screening system (Shen et al., 2015).

In the iterative saturation mutagenesis (ISM) developed by Reetz et al. (2010 a &b), the important sites of the enzyme are determined (based on the structural data) and divided into 2-4 groups, each group consists of 1-3 amino acids (Fig. 7). These groups are then subjected individually to saturation mutagenesis, creating four different libraries of enzyme variants (in case of four sites). Thereafter, the obtained libraries are screened for enzyme mutants exhibiting the desired modification. The improved enzyme mutants are then served as template for a second cycle of saturation mutagenesis at another group and screened for mutants showing further improvement. The process is repeated until achieving the desired enzyme function modification. Using ISM approach, the thermostability of lipase was drastically enhanced (Fig. 7), obtaining two hyperthermostable mutants with $T_{m}$ values of 89 and 93°C (Reetz et al., 2010 a, b and Li & Cirino, 2014). Thermal stability of several enzymes was also improved by employing the semi-rational approach including α-amylase, about 3.4 fold enhancement, (Yang et al., 2016); cis-epoxysuccinate, 33.5 fold, (Qiao et al., 2015); cellobiose 2-epimerase (Shen et al., 2015) and cold-active xylanase (Acevedo et al., 2017). Recently, several properties of β-glucosidase were enhanced using SRD strategy including tolerance to ethanol (2.4 fold); half-life (4.3 fold), pH- and thermal stability (Fang et al., 2016). Moreover, using SRD strategy many properties of several biocatalysts were successfully improved including catalytic activity (Wang et al., 2016), substrate specificity, regio-and enantio-selectivity (Li et al., 2014) and ionic tolerance (Dabirmanesh et al., 2015).

Fig. 6. (A) Synthesis of the antidiabetic drug Sitagliptin (Januvia®) by the developed novel transaminase variant (Savile et al., 2010). The author and Merck and Codexis were awarded the Presidential Green Chemistry Award from US Environmental Protection Agency for such remarkable work. (B) The main reaction of wild-type succinic semialdehyde reductase (a) and conversion levulinic acid to of 4-hydroxyvaleric acid by engineered enzyme variant (b) (Yeon et al., 2015).
Computational de novo enzyme design: Enzyme engineering, using directed evolution, rational design and semi rational design approaches, has achieved significant success in improving functionality and properties of many existing biocatalysts such as thermostability, selectivity and specificity. However, it does not show similar success in creating novel enzyme function, i.e. de novo enzyme. Recently, computational enzyme design has been developed as a promising strategy to create “made-to-order” enzyme with new catalytic function which is not present in nature (Kries et al., 2013 and Zanghellini, 2014). Generally, computational enzyme design has two major goals including, (i) Automated sequence design, i.e. re-design of existing biocatalysts, and (ii) de novo enzyme design creating novel biocatalyst with new function (Jemli et al., 2016).

In automated sequence design (ASD), a new catalytic site is generated in a natural biocatalyst through mutation of a set of amino acids residues, whereas de novo enzyme approach aims to design the entire protein backbone structure and sequences (Frushichev et al., 2014). In contrast to rational design which includes limited number of mutation, ASD involves creating of novel sequences libraries with more than 50% mutation of the wild enzyme sequence while retaining a stable protein that can be expressed (Jemli et al., 2016). On the other hand, in de novo enzyme design both substrate binding and catalytic sites are introduced simultaneously to protein scaffolds that have no catalytic activity (Frushichev et al., 2014 and Madhavan et al., 2017).

Computational design approach for de novo enzyme include the following steps (Fig. 8 A): (i) The target reaction is selected and a suitable catalytic mechanism is identified, including all residues side chains and functional groups, (ii) Employing quantum mechanics model to optimize the active site in order to maximize enzyme-substrate complex transition state stabilization, (iii) The obtained structure, so-called “theozymes”, that has optimum structure and residues for catalytic activity, but in need of a protein backbone to support them are docked into variety of protein scaffolds from the Protein Data Bank, (iv) in silico optimization of the obtained complexes by searching for mutation of the amino acids surrounding the active site that can lead to better packing and finally (v) in vitro production and biochemical characterization of most promising enzyme designs, based on the in...
levels using directed evolution, rational design, procedures of enzyme engineering at the molecular level. Phenylbutyrate was converted to 2-oxo-4-phenylbutyrate with ethyl (R)-2-hydroxy-4-phenylbutyrate. The obtained enzyme mutant was able to convert 700/S, which is even higher than several enzymatic activities. The catalytic activity was increased by about 10³ fold by further application of rational design, showing a turnover number of 700/S, which is even higher than several natural biocatalysts. The initial activity of computationally designed enzyme was measured using computational enzyme design and RD approaches, leading to remarkable progress in this field such as SCHEMA, ProSAR, and ROSETTA, the latter is the most successful computational algorithm (Mak & Siegel, 2014).

Computational enzyme design has succeeded in creating several de novo enzymes for some non-biological reactions (Fig. 9) such as the Diels–Alder reaction (Siegel et al., 2010), Kemp elimination reaction (Blomberg et al., 2013) retro-aldol reaction (Eiben et al., 2012 and Giger et al., 2013), Morita–Baylis–Hillman catalyzing protein (Bjelic et al., 2013) and de novo esterase for unnatural substrate (Richter et al., 2012). Recently, combination of computationally designed enzyme, directed evolution and rational design approaches resulted in improvement de novo enzymes performance to the level of natural biocatalysts. The initial activity of computationally designed de novo Kemp eliminases (catalyze conversion of 5-nitrobenzisoxazole to cyanophenol) was extremely low. However, the catalytic activity was increased by about 10³ fold by further application of rational design, showing a turnover number of 700/S, which is even higher than several natural biocatalysts (Khersonsky et al., 2012 and Blomberg et al., 2013). Deng et al. (2016) reported inversion of the carbonyl reductase enantioselectivity by using computational enzyme design approach through substrate-enzyme docking and molecular dynamics simulation to identify the molecular basis of carbonyl reductase enantio-preference. Accordingly, three binding amino acid residues were determined that play a crucial role in the enzyme enantioselectivity; then those were targeted for point mutation. One of the obtained enzyme mutant was able to convert 2-oxo-4-phenylbutyrate to ethyl (R)-2-hydroxy-4-phenylbutyrate with ee>99% (Deng et al., 2016).

**High-throughput screening tools:** The procedures of enzyme engineering at the molecular level using directed evolution, rational design, or computational enzyme design, usually include generation of mutants' libraries consisting of a huge number of enzyme variants. Therefore, efficient high throughput screening (HTS) methods are required in order to identify the desired enzyme variants in short time, easy procedure and low cost (Fig. 10 A). Hence, availability and development of high-throughput screening technology is considered as a limiting step in the success of enzyme engineering process (Pitzler et al., 2014 and Wohlgemuth et al., 2015). High-throughput enzyme screening assays is usually carried out in 96-well microtiter plates, where enzymes or substrates are bound to the wells surface (Fig. 10 A). Recently, different efficient HTS methods were developed such as microfluidic and fluorescence activated cell sorting (FACS) devices that have the ability to screen up to 10⁷ events per hour (Jacques et al., 2017). In microfluidic devices, enzymatic reaction is carried out in nano- to picoliter reaction volumes, and the compounds to be screened are injected consecutively into the microfluidic capillary that is connected with a proper detector (Wohlgemuth et al., 2015). On the other hand, the flow cytometry-based screening system is based on detection of fluorescent product in a compartment such as cells-aqueous/oil/aqueous emulsions (Pitzler et al., 2014 and Jacques et al., 2017).

FACS based screening system with an *in vitro* compartmentalization in double emulsion was developed to screen proteases variants libraries generated by directed evolution approach for improvement of the enzyme resistance to various proteases inhibitors (Tu et al., 2011). The created enzyme mutants’ library was expressed and screened in double emulsions using a fluorogenic peptide substrate. After several cycles of FACS sorting, an enzyme variant containing six mutations was found and separated, showing improvement of the enzyme resistance to protease inhibitors by about 2.6-fold (Tu et al., 2011). An interesting FACS based approach for cell sorting according to cell viability was developed by Davids et al. (2013) for screening of highly enantioselective esterase variants within esterase mutants library expressed in E. coli. In this approach, one of the substrate enantiomer was linked to a toxic compound while the other substrate enantiomer is linked to the bacterial carbon source. Hence, the viability of E. coli cells indicates expression of the desired enantioselective esterase mutants which will be sorted out by FACS (Davids et al., 2013).
Fig. 8. (A) Computational enzyme design methodology (Kries et al., 2013). (B) Complete flowchart of the computational design for de novo enzyme, to create "made-to-order" enzyme biocatalytic activity (Zanghellini, 2014).
Fig. 9. Examples of computationally designed enzymes. (a) Novel esterase converted from protein thioredoxin (from \textit{E. coli}); (b) Enzyme for carbon–carbon bond cleavage via retroaldolization; (c) Enzyme for proton transfer from carbon and (d) Diels-Alder cycloaddition (Kries et al., 2013).

Fig. 10. (A) General strategy used to library screening for enzymes with the desired modification (Jacques et al., 2017), (B) Microfluidic droplets system for screening of enzyme libraries [(1) The enzyme mutants library is in a host; (2) Each drop enclose one cell that express one enzyme mutant together with the proper substrate and cell lysis agents, cell lysis, enzyme release and substrate conversion to product; (3) Detection of the droplet fluorescence or absorbance and sorting out by FACS and (5) DNA recovery and analyzed (Mair et al., 2017)].

Recently, combination of microfluidic and FACS systems resulted in ultra-high throughput screening system (UHTS) (Wojcik et al., 2015). In this UHTS system, drop-based microfluidics system formed by aqueous drops, acting as reaction vessels of picoliter volume, dispersed in oil (Fig. 10 B). Each droplet enclosed one cell that expressed one enzyme mutant together with the proper substrate and cell lysis agents. Hence, the vessels “drop” which contained enzyme mutant exhibiting the desired modified enzymatic activity, was sorted out by FACS (Wojcik et al., 2015 and Jacques et al., 2017). Using Microfluidic-FACS UHTS system, new mutants of horseradish peroxidase were identified from \(10^8\) enzymes mutants’ library, created by the directed evolution approach, within short time and low cost (Agresti et al., 2010 and Jacques et al., 2017).
Recent advance in physical stabilization of enzyme

Enzymes have several advantages over industrial chemical catalysts including mild reaction conditions, high specificity, selectivity. However, the catalytic activity, enzyme stability in high temperature or organic solvents, storage, and reusability of the soluble enzymes are of major concerns which usually limit their industrial applications due to its impact on the bioprocess cost (Siddiqui, 2015 and Ibrahim et al., 2016). Therefore, introduction and development of new and innovative biocatalyst stabilization approaches will increase the enzymes utilization in various industrial applications (Woodley, 2013). Immobilization of soluble enzymes on water-insoluble supports is one of the most effective approaches for biocatalyst stabilization. Application of immobilized enzymes in industrial bioprocesses can offer several unique advantages over the soluble enzyme such as enzyme reusability, enhancement of enzyme stability, simplifying product purification and better bioprocess control, resulting in reduction of the bioprocess cost (Ibrahim et al., 2014 and Madhavan et al., 2017). Previously several solid supports have been used as carriers for enzyme immobilization; however as a result of recent progress in nanotechnology diverse nanostructured materials have been provided that potentially support the enzymes immobilization technology and opened a new field, the so-called nanobiocatalysis (Ansari & Husain, 2011; Raival et al., 2014 and Mukhopadhyay et al., 2015).

Nanostructured materials and Nanobiocatalysis: The field of nanobiocatalysis is an integration of biotechnology, biocatalysis, material science and nanotechnology. Nanoscale materials provide several unique and physicochemical capabilities for enzyme immobilization over the traditional macro supports. Nanostructured materials can provide the upper limits in balancing the key factors that determine the efficiency of biocatalysts due in part to high surface area/volume ratio that enhance catalysis; high loading capacity and effective mass transfer of the nanomaterials (Verma et al., 2013 and Ding et al., 2015). Furthermore, the confining effect of the nanopores, Brownian motion and self-assembling of the nanoparticles provide nanobiocatalyst systems with unique behaviors and properties in comparison to the classical immobilization using macro carriers (Madhavan et al., 2017). A variety of nanomaterials have been explored for biocatalysts immobilization such as nanoparticles, nanofibers, nanotubes and nanoporous matrices, using different immobilization approaches including covalent attachment, adsorption and confinement within the nanopores (Ibrahim et al., 2015).

Silica based nanoparticles is one of the most successful nanomaterials for enzyme immobilization. It is characterized by uniform pore system, tunable and easy to be functionalized, biocompatibility and low cytotoxicity (Zhou & Hartmann, 2013). In addition, magnetic silica can be prepared by silica coating on magnetic nanoparticle cores and hence the immobilized enzyme can be easily separated from the reaction solution by application of an external magnetic field (Ibrahim et al., 2014 and Kumari & Singh, 2016). My group and I have explored utilization of several silica based nanoparticles for immobilization of different enzymes such as immobilization of cyclodextrin glucanotransferase (CGTase) onto amino propyl functionalized silica coated superparamagnetic nanoparticles (Ibrahim et al., 2013) and functionalized magnetic double mesoporous core-shell silica nanospheres (Ibrahim et al., 2014). The immobilized CGTase exhibited significant improvement of enzyme kinetics and stability parameters, high operation stability and can be recovered easily from the bulk solution using magnetic field (Fig. 11 A). Recently, my group and I reported development of novel robust nanobiocatalyst for detergents formulations and the other applications of alkaline protease through covalent enzyme immobilization onto hollow core-mesoporous shell silica nanospheres (Ibrahim et al., 2016 a), and amino-functionalized rattle-type magnetic core@mesoporous shell silica nanoparticles (Ibrahim et al., 2016 b) (Fig. 11 B). The developed alkaline protease nanobiocatalyst exhibited significant enhancement in thermostability, catalytic activity, substrate affinity and stability in a variety of organic solvents, surfactants and commercial laundry detergents. Immobilization of several enzyme into nanomaterial have been reported recently, showing significant improvement of enzymes properties including lipase/mesoporous silica (Zheng et al., 2017), laccase/magnetic nanoparticles (Fortes et al., 2017), pectinase/chitosan magnetic nanoparticles (Sojitra et al., 2017), α-amylase/metallic/bimetallic magnetic nanoparticle (Singh et al., 2016), transaminase/PVA-Fe₃O₄ nanoparticle (Jia et al., 2016), trypsin/magnetic nanofibers and others (Lee et al., 2016).
RECENT TRENDS FOR DISCOVERY AND ENHANCEMENT OF...

**Single enzyme nanobiocatalyst:** One of the most recent innovative approaches of application of nanotechnology for biocatalysts stabilization is the development of single enzyme nanoparticles (SENs), where a porous matrix (organic or inorganic) of few nanometers thickness is built up on the enzyme molecule (Hegedus & Nagy, 2015). The constructed nanoscale shield surrounding the enzyme molecules can provide the enzyme with unique properties including biocatalyst stabilization, with no negative effect on the substrate/product mass transfer and enzyme kinetics (Gupta et al., 2012 and Zhou & Hartmann, 2013). Recently, the thermostability and half-life time of cellulase and hemicellulase (from *Thermobifida fusca*) were significantly improved by coating every enzyme molecule with 2-3nm layer of polyacrylamide polymer (Hegedus & Nagy, 2015). Hong et al. (2017) reported development of SENs of lipase and α-chymotrypsin by deposition and assembly of uniform nanolayer of silicate on the lipase molecules surface. The applied approach for SENs construction yielded robust lipase system, showing drastic increase of the enzyme stability, while retaining similar catalytic activity relative to the native enzyme (Fig. 12).

**Future outlook for discovery and enhancement of enzyme function (The author perspective)**

Due to the increase of global awareness of the climate changes, environment protection and health, there is a growing attention to biocatalysts exploitation as ecofriendly alternative of chemical catalysts in various industrial processes. This is due to the several advantageous properties of enzymes including high selectivity, specificity, moderate reaction conditions and low cytotoxicity. Therefore, in the last decade the enzymes market exhibited significant growth of about 8.0% per year. However, the limitations of natural enzymes application in industrial processes include not limited to regio- and enantio-selectivity, stability in organic solvents, thermo- and pH stability, and specificity and cost-effectiveness. In addition, there are several chemical reactions that can’t be catalyzed by natural biocatalysts. In this regards, two main approaches are used to expand the enzymes utilization in various industrial applications including (i) Searching nature for discovery of better performing novel biocatalysts, and (ii) Improvement of the enzyme function and properties.
(i) Searching nature for discovery of better performing new biocatalyst

As the microbial origin of the biocatalysts is the key factor in determining its activity, stability and other properties, scientists continue to search nature for enzymes with superior performance for the commercial exploitations and application of enzymes. However, I think to increase the probability of novel enzymes discovery, microbial isolation and screening process should be carried out using new or unexplored environments. One of such environments is the extreme environments in which several groups of extremophilic microorganisms can survive and grow optimally. Extremophiles are microorganisms that have evolved to live in a variety of extreme environments and they fall into a number of different groups including hyper thermophiles, alkaliphiles, acidophiles, halophiles, psychrophiles, xerophiles and piezophile. During the evolution time the extremophiles have adapted to survive and grow in ecological niches with harsh conditions such as high and very low temperature, high salt concentrations, highly alkaline and acidic conditions, and high pressure. As a result of microbial adaptation to such extreme environments, extremophiles produce unique biocatalysts, so-called extremozymes, which function and catalyze reactions under harsh conditions in which their counterparts, thriving in normal conditions cannot work. Therefore, ‘extremozymes’ can be utilized in various industrial processes without the need for alteration of the process parameters, to meet the biocatalysts requirements, resulting in more efficient industrial processes.

The natural habitats of extremophiles include extreme hot spring, soda lakes, salt brines, ice, deep sea, niches with toxic waste and organic solvent and others which were previously considered as unsuitable environments for life. Therefore, I think more attention should be paid for such extreme environments as promising sources for discovery of novel biocatalysts with superior performance. My group and I reported isolation and characterization of several novel industrial enzymes with unique properties from
various extremophiles, mainly thermophiles and alkaliophiles from hypersaline soda lakes and hot springs (located in Egypt), including alkaline protease, chromate reductase, alkaline cyclodextrin glucanotransferase (CGTase) and thermophilic cellulases.

However, due to the fact that only a very tiny fraction of microorganisms (less than 1%) on the Earth can be cultured under the laboratory conditions, metagenomics approach has been developed as strategy to overcome microbial culturability problem. Therefore, it is possible to isolate and screen environmental DNA for discovery of new biocatalysts without the need for growing the microbes. Once again, extreme environments represent promising sources, as mentioned above, for new enzymes discovery using metagenomics approach either function-based or sequencing-based metagenomics. However, for successful application of the metagenomics, more research and work are required for development of efficient high- and ultrahigh throughput screening methods to identify the enzymes within the generated metagenomics libraries.

(ii) Improvement of enzyme function and properties

In order to enhance the enzyme function and properties to meet the industrial processes conditions requirements, several enzyme engineering approaches can be used to modify enzymes at the molecular level such as directed evolution, rational design and computational enzyme design. However, the main limitation of these methods is perhaps the availability of high throughput screening methods for identification of the desired enzyme variants within the enzyme mutants’ libraries that usually contain a huge number of variants.

Therefore, I think that using semi-rational approach for enzyme engineering could be more effective in achieving the desired biocatalyst modification. Semi-rational approach is based on only preliminary data about enzyme structure-function relationship and computer algorithms, to create ‘small but smart enzyme library’, i.e. library with much less mutants number but highly enriched with mutation at most likely effective positions, predicted using computer modeling. Hence, screening assay methods with reasonable throughput can be sufficient for identification of enzyme mutants with the desired alteration.

Computational de novo enzyme design is a recently developed tool, as a promising approach for generating made-to-order enzyme, catalyzing novel reaction for which there is no natural biocatalyst. Despite this approach is still in its beginning of development, I think it is very promising and can play important role in expanding the enzymes applications in new industries in the near future. However, for more effective exploitation of computational design for de novo enzyme, in addition to directed evolution and rational design approaches; I think more multidisciplinary research and effort using innovative ideas should be performed for the development of efficient high- and ultrahigh throughput screening methods, allowing effective screening of large enzyme mutant’s libraries in short time and cost-effectiveness. Screening assays methods using microfluidic and FACS technologies are in early beginning and considered as promising approaches for high-throughput screening technologies. However, it should be emphasized that that development of enzyme assays and HTS methods needs multidisciplinary collaborative workflow from several fields including microbiology, biochemistry, chemistry, cell biology, material science and biophysics, for fruitful outcomes.

In the area of novel enzyme function discovery and development, I think more attention should be made to the recently emerged term, ‘enzyme promiscuity’. According to evolutionary biologists, enzyme promiscuity is defined as the ability of enzyme to catalyze secondary reactions, which are physiologically irrelevant either due to too low efficiency to affect the cell biology and fitness or owing to the fact that the biocatalysts never hit substrates. Promiscuous activities are of great significance and can be a starting point for evolution and developing new enzymatic functions and hence expanding the biotechnological application of enzymes.

Despite enzyme immobilization approach has been previously applied for long time, the recent advances of nanotechnology has provided us with varieties of nanostructured materials suitable for enzymes immobilization which can reshape our vision of enzyme immobilization technology; and led to the development of nanobiocatalysis field. Nanotructred materials provide several unique and physicochemical capabilities for enzyme immobilization over the traditional
macro supports. Therefore, more efforts and collaborative research between material scientists, microbiologists, enzymologist, are required to achieve more success in the area of nanobiocatalysis, as a new field for enhancement of enzymes function and properties.

**References**


Wojcik, M., Telzerow, A., Quax, W.J. and Boersma, Y.L. (2015) High throughput screening in


(Received 5/2/2018; accepted 13/9/2018)
RECENT TRENDS FOR DISCOVERY AND ENHANCEMENT OF...

Abdel-Nasser Salah El-Ebary El-Adham
Department of Chemical Engineering, Faculty of Engineering, Cairo University, Giza, Egypt.

In the recent years, there has been a growing interest in using enzymes as a environmentally friendly alternative to chemical catalysts in various industrial applications. Enzymes are used to improve and expand their application in different areas, such as (i) the discovery of new enzymes with improved properties for industrial applications, and (ii) the development and optimization of industrial enzymes. The recent developments in enzyme discovery and design have led to the increase in the efficiency and diversity of enzymes. Therefore, the application of enzymes has expanded to include industrial processes, which were not possible in the past due to the use of classical chemical catalysts. Additionally, there has been a significant success in developing "custom-made" enzymes that meet the specific requirements of industrial processes, leading to more efficient and cost-effective processes. In this paper, the authors focus on the strategies and tools for the discovery of new enzymes and the improvement of their properties, including the design of enzymes at the molecular level through different methods such as directed evolution, relative design, and computer-aided design, as well as the field of nanomaterials (small molecules).

Egypt J. Microbiol. 53 (2018)