ISTANBUL TECHNICAL UNIVERSITY ★ FACULTY OF SCIENCE AND LETTERS

DEVELOPMENT OF BIOCATALYSTS BY DIRECTED EVOLUTION FOR SUSTAINABLE PHARMACEUTICAL DRUG PRODUCTION

B.Sc. THESIS

Esra TANRIVERDİ

Department of Molecular Biology and Genetics

Faculty of Science and Letter

JANUARY 2023

ISTANBUL TECHNICAL UNIVERSITY ★ FACULTY OF SCIENCE AND LETTERS

DEVELOPMENT OF BIOCATALYSTS BY DIRECTED EVOLUTION FOR SUSTAINABLE PHARMACEUTICAL DRUG PRODUCTION

B.Sc. THESIS

Esra TANRIVERDİ

(090180006)

Department of Molecular Biology and Genetics

Faculty of Science and Letter

Thesis Advisor: Assoc. Prof. Dr. Halil BAYRAKTAR

JANUARY 2023

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN EDEBİYAT FAKÜLTESİ

SÜRDÜRÜLEBİLİR İLAÇ ÜRETİMİ İÇİN YÖNLENDİRİLMİŞ EVRİM YAKLAŞIMIYLA BİYOKATALİZÖRLERİN GELİŞTİRİLMESİ

LİSANS TEZİ

Esra TANRIVERDİ

Moleküler Biyoloji ve Genetik

Fen Edebiyat Fakültesi

Tez Danışmanı: Doç. Dr. Halil BAYRAKTAR

OCAK 2023

Esra TANRIVERDİ, an undergraduate student of ITU Faculty of Science and Letters student ID 090180006, successfully defended the thesis/dissertation entitled “DEVELOPMENT OF BIOCATALYSTS BY DIRECTED EVOLUTION FOR SUSTAINABLE PHARMACEUTICAL DRUG PRODUCTION”, which she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Assoc. Prof. Dr. Halil BAYRAKTAR**.............................
Istanbul Technical University

Jury Members : **Assoc. Prof. Dr. Halil BAYRAKTAR**...........................
Istanbul Technical University

Assist. Prof. Dr. Abdülhalim KILIÇ………..…………
Istanbul Technical University

Assist. Prof. Dr. Bülent BALTA ………………………
Istanbul Technical University

Date of Submission : 13 January 2023

Date of Defense : 18 January 2023

FOREWORD

This thesis serves as an overview of the four years study leading up to graduation. I am proud of the education that I received from my instructors and being a candidate scientist. I want to express my gratitude to Assoc. Prof. Dr. Halil BAYRAKTAR, my supervisor in my undergraduate thesis project, for taking me on as his student and for the knowledge and experience he has provided me.

I would also like to express my gratitude to my dear family and my friends Umut VOYVODA, Sena Nur UÇAN, Tuğba ÇAVDAR and Zehra ERDEM for always being by my side.

January 2023 Esra TANRIVERDİ

TABLE OF CONTENTS

FOREWORD……………………...……..……..……………………………………...……vii
TABLE OF CONTENTS...……..…..……..……………………………..……..….…ix
ABBREVIATIONS …………………….…………………………….……………….xi
LIST OF FIGURES …....……….………………………………………......……xiii
SUMMARY ……………...………………………………...……….…..……..xv
ÖZET …………………………..……...……………...……………………....xvii
1. INTRODUCTION…………………………………………………………….1
1.1. Use Of Catalysis In Drug Development……………………………………...…1
1.1.1. Transition metal catalysis and limitations………………………………..1
1.2. Biocatalysis……………………………...……………………………….…...3
1.2.1. Use of biocatalysts for sustainable drug production……………….………..3
1.2.2. Limitations of biocatalysis………………………………………………6
1.3. Directed Evolution Of Proteins…………………………………………………7
1.3.1. Library construction………………………………………………………..8
1.3.2. Screening and selection…………………………………………………11
1.3.2.1 Screening methods……………………………………………..12
1.3.2.2 Selection methods…………………………………………...13
1.3.2.2.1 Display-based selection…………………………………..13
1.3.2.2.2 Compartmentalization-based selection…………………..15
1.4. Directed Evoluted Biocatalysts……………………………………………….16
2. REDOX PROTEINS…………………………………………………………..18
2.1. Cytochrome C……………………………………………………………………18
2.2. Cytochrome P450………………………………………………………………..19
2.3. Biomedical Applications Of Cytochromes……………………………………21
2.4 Directed Evolution of Cytochrome Proteins For Pharmaceutical Applications..22
3. CONCLUSION……………………..……………………………...……….…..24
REFERENCES………………………...…………………………………………....26

ABBREVIATIONS

API : Active pharmaceutical ingredient

DAMP : Damage-associated molecular pattern

DMF : Dimethylformamide

EMS : Ethyl methanesulfonate

FACS : Fluorescence-activated cell sorting

FCS : Foward Scatter Channel

GABA : γ-aminobutyric acid

GFP : Green fluorescent protein

GWP : Global-warming potential

IVTC : In vitro compartmentalization

IVVC : In vivo compartmentalization

MS : Mass Spectrometry

NExT : Nucleotide exchange and excision technology

NMR : Nuclear Magnetic Resonance

POI : Protein of interest

PTM : Post translational modifications

RACHITT : Random chimeragenesis on transient templates

REAP : Reconstruction evolutionary adaptive paths

ROM : Reactive oxygenated intermediate

ROS : Reactive oxygen species

SSC : Side Scatter Channel

StEP : Staggered extension process

XMEs : Xenobiotic-metabolizing enzymes

LIST OF FIGURES

Pages

Figure 1.1: A diazepam biocatalysis cascade generated via pathway explorer of RetroBioCat……………………………………………………………………………5

Figure 1.2: The overview of the directed evolution process of a protein………………..9

Figure 1.3: Overview of the DNA shuffling method……………………………………10

Figure 1.4: The display of the variant library on surfaces such as phage, yeast and ribosome and selective pressure for selection of variants…………………………………14

DEVELOPMENT OF BIOCATALYSTS BY DIRECTED EVOLUTION FOR SUSTAINABLE PHARMACEUTICAL DRUG PRODUCTION

SUMMARY

Drug development is an intricate process that requires the utilization of catalysts for the synthesis of pharmaceutical intermediates. The pharmaceutical industry has long been dependent on chemo-catalysts including transition metals for the production of active pharmaceutical ingredients (APIs). Lately, biocatalysts have attracted attention following the advances in the enzyme development through directed evolution and increasing demand for more efficient and faster production. Increasing awareness towards environment due to global warming also puts pressure on companies to aim to reduce their carbon emissions. The advantages of biocatalysts over chemo-catalysts include high selectivity, higher yield at a lower cost, milder reaction conditions and less waste generation, allowing the industry to become much more sustainable and greener. For biocatalysts to completely replace chemical catalysts used in the industry, we must overcome current limitations of these methods. For a certain pharmaceutical process, it is a challenge to screen and identify an enzyme that has all the necessary properties. Also, enzymes are prone to denaturing at harsh environmental conditions that are commonly encountered in industrial processes due to their weak hydrogen bonds in the folded structure. The development of enzymes with precise properties, higher efficiency and improved tolerance to overcome these challenges has been possible by the application of directed evolution. Directed evolution utilizes random or focused mutagenesis to create genetic variants which will then be screened for desired traits. For random mutagenesis, the DNA is damaged with chemical agents and UV irradiation. Replication errors due to DNA polymerases with low fidelity and deactivated proofreading mechanisms can also be used to confer genes with mutations by carrying out error-prone PCR (epPCR). Focused mutagenesis is based on the characterization and targeting of residues of the enzyme involved in catalysis. This method usually involves the incorporation of DNA oligonucleotides with degenerated codons. The genetic variations obtained by these methods are used to generate a library of variants. Each member of the constructed library is expressed in a model organism for screening and identifying a variant with desirable properties. Several classes of proteins including redox proteins and hydrolytic enzymes have been developed by protein engineering through guided evolution in the laboratory in the last decades. For the development of APIs, redox proteins have been the primary focus of directed evolution studies as key biological processes including respiration and photosynthesis involve oxidation-reduction processes which require transfer of electrons. C-type cytochromes are a family of redox proteins containing heme prosthetic group with

iron center. The changes in its redox activity regulates many key cellular processes. One of the members of this family, cytochrome c, is a vital protein for cell processes as it functions in disposal of ROS, respiration and apoptosis depending on where it resides in the cell. Due to its functions in the apoptotic pathway, cytochrome c is widely studied for drug development against cancer. Another distinct member, cytochrome P450, functions in cellular metabolism and detoxification of certain drugs. Due to their electron transfer ability conferred by their heme prosthetic groups and functions in redox processes, cytochrome c and cytochrome P450 have gained considerable interest for drug development. There have been several studies targeting these enzymes for protein engineering through directed evolution for carbene and nitrene transfer reaction catalysis during drug synthesis.

Key Words: Active pharmaceutical ingredient, biocatalysis, catalyst, cytochrome *C*, cytochrome P450, directed evolution, redox proteins

SÜRDÜRÜLEBİLİR İLAÇ ÜRETİMİ İÇİN YÖNLENDİRİLMİŞ EVRİM YAKLAŞIMIYLA BİYOKATALİZÖRLERİN GELİŞTİRİLMESİ

ÖZET

İlaç geliştirme, farmasötik ara ürünlerin sentezi için katalizörlerin kullanılmasını gerektiren karmaşık bir süreçtir. Farmasötik endüstrisi, aktif farmasötik bileşenlerin (API'ler) üretimi için geçiş metalleri dahil olmak üzere uzun süredir kemo-katalizörlere bağımlıdır. Son zamanlarda, biyokataliz yöntemleri, yönlendirilmiş evrim yoluyla enzim gelişimindeki ilerlemeler ve daha verimli ve daha hızlı üretim için artan talebin ardından dikkatleri üzerine çekmiştir. Küresel ısınma nedeniyle çevreye karşı artan farkındalık, şirketlerin karbon emisyonlarını azaltmayı hedeflemeleri konusunda da baskı oluşturmaktadır. Biyokatalizörlerin kemo-katalizörlere göre avantajları arasında yüksek seçicilik, daha düşük maliyetle daha yüksek verim, daha ılımlı reaksiyon koşulları ve daha az atık üretimi sayılabilir; bu da endüstrinin çok daha sürdürülebilir ve daha çevreci olmasını sağlar. Biyokatalizörlerin endüstride kullanılan kimyasal katalizörlerin tamamen yerini alması için bu yöntemlerin mevcut sınırlamalarının üstesinden gelinmelidir. Belirli bir farmasötik işlem için, gerekli tüm özelliklere sahip bir enzimi taramak ve tanımlamak zorlu bir iştir. Ayrıca enzimler, katlanmış yapıdaki zayıf hidrojen bağları nedeniyle endüstriyel proseslerde yaygın olarak karşılaşılan zorlu çevre koşullarında denatüre olma eğilimindedirler. Bu zorlukların üstesinden gelmek için daha yüksek verimliliğe ve geliştirilmiş toleransa sahip enzimlerin geliştirilmesi, yönlendirilmiş evrim uygulamasıyla mümkün olmuştur. Yönlendirilmiş evrim, sonrasında istenen özellikler için taranacak genetik varyantlar yaratmak için rastgele veya odaklanmış mutagenez kullanır. Rastgele mutagenez yöntemlerinde DNA kimyasal ajanlar ve UV ışıması ile hasar görmektedir. Düşük verime sahip DNA polimerazlarından kaynaklanan replikasyon hataları ve devre dışı bırakılmış düzeltme okuma mekanizmaları, hataya açık PCR (epPCR) gerçekleştirerek mutasyonlu genleri tanıtmak için de kullanılabilir. Odaklanmış mutagenez, katalizde yer alan enzim domainlerinin karakterizasyonuna ve hedeflenmesine dayanır. Bu yöntem genellikle DNA oligonükleotitlerinin dejenere kodonlarla birleştirilmesini içerir. Bu yöntemlerle elde edilen genetik varyasyonlar, bir varyant kütüphanesi oluşturmak için kullanılır. Oluşturulan kitaplığın her bir üyesi, arzu edilen özelliklere sahip bir varyantın taranması ve tanımlanması için bir model organizmada ekspres edilir. Redoks proteinleri ve hidrolitik enzimler de dahil olmak üzere çeşitli protein sınıfları, son on yıllarda laboratuvarda yönlendirilmiş evrim yoluyla protein mühendisliği metodlarıyla geliştirilmiştir. Redoks proteinleri, solunum ve fotosentez dahil olmak üzere temel biyolojik süreçler elektron transferini gerektiren oksidasyon-indirgeme süreçlerini içerdiğinden, yönlendirilmiş evrim çalışmalarının birincil odak noktası

olmuştur. C-tipi sitokromlar, demir iyonu merkezli heme prostetik grubu içeren bir redoks proteini ailesidir. Redoks etkinliğindeki değişiklikler, birçok önemli hücresel işlemi düzenler. Bu ailenin üyelerinden biri olan sitokrom c, hücrede bulunduğu yere bağlı olarak reaktif oksijen türlerinin atılmasında, solunumda ve apoptozda görev yaptığı için hücre süreçleri için hayati bir proteindir. Apoptotik yoldaki işlevleri nedeniyle sitokrom c, kansere karşı ilaç geliştirme için geniş çapta çalışılmaktadır. Diğer bir farklı üye, sitokrom P450, hücresel metabolizmada ve belirli ilaçların detoksifikasyonunda işlev görür. Sitokrom c ve sitokrom P450, heme prostetik grupları tarafından sağlanan elektron transfer yetenekleri ve redoks süreçlerindeki işlevleri nedeniyle, ilaç geliştirme için büyük ilgi görmüştür. İlaç sentezi sırasında karben ve nitren transfer reaksiyonu katalizi için yönlendirilmiş evrim yoluyla protein mühendisliği için bu enzimleri hedefleyen birkaç çalışma yapılmıştır.

Anahtar Kelimeler: API, biyokataliz, katalizör, sitokrom *C*, sitokrom P450, redox proteinleri, yönlendirilmiş evrim

1. INTRODUCTION

1.1 Use of Catalysis In Drug Development

The primary component of any drug product is called the active pharmaceutical ingredient (API) whose development consists of synthesis, extraction and separation steps that rely heavily on catalysis. In basic terms, catalysis increases the rate of a reaction by decreasing its activation energy through utilization of catalysts [1]. Commercial production of pharmaceutical intermediates and APIs requires catalysts that have high reactivity and selectivity which makes the process often costly and time-consuming. New and improved catalysts which are faster, cost-efficient and also, environmentally friendly, are demanded as the structure of APIs become more complex.

1.1.1 Transition metal catalysis and limitations

The pharmaceutical industry mostly relies on transition metal catalysts such as palladium and rhodium, to carry out a variety of chemo- and enantioselective chemical transformations for the production of APIs at industrial scale. The most commonly employed chemical transformations include cross-coupling for C-C and C-X bond formation, oxidation, asymmetric addition and hydrogenation. Buchwald-Hartwig amination and Suzuki-Miyaura coupling comprise a large fraction of the cross-coupling reactions used in the pharmaceutical industry for C-N and C-C bond formations, respectively [2,3]. Their ability of tolerating functional groups and that they have a wide range of substrates allows these reactions to be employed in the production of complex pharmaceutical compounds [4]. These reactions and many others that are frequently utilized in API production require palladium, a precious transition metal, as a catalyst. Since its discovery in 1803 by William Hyde Wollaston, palladium has found many uses for industrial applications including one of its earliest applications as a hydrogenation catalyst due to its high capacity for absorbing hydrogen gas [5].

Although palladium is of great importance in chemical and pharmaceutical industries, its use may have some disadvantages and challenges from a sustainability point of view. The refining process of palladium involves the use of toxic chemicals and also heating cycles at high temperatures which can go up to 1450 ^{o}C [6]. According to research by Nuss and Eckelman (2014), global-warming potential (GWP) per kilogram for palladium is approximately 3,880 kg CO_2-eq/kg which indicates palladium refining can cause the emission of 3,880 kg of CO_2 equivalents for each kilogram of palladium. The GWP value of palladium puts it in the group of industrial metals with highest environmental impact along with many other platinum group elements including ruthenium, rhodium and iridium, which are also commonly used catalysts in the pharmaceutical industry [7]. These platinum group metals also face the danger of becoming depleted in the upcoming decades due to their low earth abundance and high demand by the industry and are therefore, becoming increasingly expensive [8]. All these aforementioned issues suggest that relying on transition metals, especially platinum group metals, for the production of pharmaceutical compounds and drugs can be highly unsustainable, both environmentally and economically.

Additionally, there are also health concerns for using certain transition metals as a catalyst due to the possibility of them being retained in the isolated pharmaceutical compound. The presence of heavy metals in drugs exceeding certain concentrations can be an enormous threat to human health. Palladium has been shown to hinder critical cellular processes including biosynthesis of DNA and RNA. To overcome this problem, palladium is often bound to other materials, or supports, such as charcoal and then removed by filtration, but such applications cannot ensure the complete elimination of palladium from the product [5].

These challenges of transition metal catalysts and increasing awareness towards the environment by public and the governments pressurize pharmaceutical companies, and many others in industrial sectors, to turn to greener, more sustainable options for carrying out such catalytic reactions. Due to recent developments in protein engineering, either by rational or irrational design techniques, biocatalysts have attracted attention for replacing chemo-catalysts. Utilization of enzyme catalysis for production of pharmaceutical compounds have many advantages over chemical catalysis. Arguably, most important

characteristic of biocatalysts is their high-level specificity which allow for highly enantio-, chemo- and regiospecific reactions and mediate synthesis of highly complex compounds [9]. They allow for milder reaction conditions as they are biological compounds that naturally operate under physiological conditions which decreases required energy when compared to transition metal catalysts. As mentioned, one of the disadvantages of transition metal catalysts is toxicity and waste generation which can also be eliminated by use of biocatalysts [10]. In industrial biocatalyzed processes, wild-type or recombinant enzymes, isolated from cultured microorganisms, are mainly utilized but whole-cell catalysis can also be employed for more complex reactions that may require several enzymes [11].

Despite their favorable properties, biocatalysts also have certain limitations that must be overcome for them to completely replace transition metal catalysts. Although biocatalysts allow for high specificity, the synthesis of any pharmaceutical compound is a complex, multi-step process that requires very specific catalytic functions and it can be a challenge to screen and identify enzymes with such properties.

1.2 Biocatalysis

The term "biocatalysis" describes the catalytic transformations performed employing entire cells or isolated enzymes. The sources of biocatalysts can roughly be divided into two which are growing cells (via fermentation) and dead cells [12]. In fermentation catalysis, the aim is to continuously generate new microbial cells and the metabolites necessary for their growth in order to produce the biocatalyst and its substrate. The deactivated biocatalysts and the cofactors are restored by the cell's machinery. Apart from fermentation, biocatalytic transformation can also be carried out by using dead cells or isolated biocatalysts. In contrast with fermentation, the two processes of biocatalyst production and catalytic conversion are independent [12].

1.2.1 Use of biocatalysts for sustainable drug production

Due to increasing awareness towards the environment by public and new environmental policies by governments force industrial companies to take action to reduce their carbon emissions and move towards more sustainable methods. The pharmaceutical industry has

also adopted greener approaches for the production of APIs which causes the generation of high levels of waste, mainly due to utilization of organic solvents, with the conventional manufacturing approaches [13]. The need for greener processes for API production has been addressed by biocatalysts which have many advantages over chemical catalysts. Biocatalysts are biodegradable and non-toxic. They reduce environmental contamination as they work under milder reaction conditions and their utilization reduces the amount of organic solvents that is necessary for the reactions. Since they are from renewable resources, they make for a much more sustainable catalysis both ecologically and economically. Aside from the ecological aspects, biocatalysts are also preferrable due to their high enantio-, stereo- and regioselectivity which is critical as chirality is a major characteristic in APIs [14]. Also, the need of rare precious metals for the conversion and the additional steps and costs related to their elimination to a tolerable level are avoided by biocatalysis [15].

Biocatalysis can be carried out by using isolated wild-type or recombinant enzymes or by whole-cell based systems. Whole-cell catalysis systems can be more cost-efficient compared to cell-free catalysis since enzyme isolation is an intricate and costly process. Also, cells can provide protection from environment and degrading conditions. However, these systems require the transport of substrates across the cell membrane which is an additional step and also, there is the possibility of side-reactions [16]. Therefore, the proper system should be chosen according to the desired product, budget and timescale.

Any industrial biocatalytic process is enabled by reaction design, biocatalyst selection, optimization by directed evolution or rational design and also bioprocess development. Biosynthetic pathway for the target molecule should be designed first either manually or by automatic design using available software programs [17]. One of the widely used programs is called the RetroBioCat which is comprised of tools to automatically design a biocatalytic cascade for any target molecule. While manual design of a cascade requires expert knowledge, RetroBioCat uses the reaction rules that are programmed or readily extracted from available resources to design a cascade starting from the target molecule to the starting material [18]. Two methodologies are provided by RetroBioCat which are network and pathway explorers. The retrosynthesis of such network of potential enzyme-catalyzed reactions is the main emphasis of the network explorer. The pathway explorer

automatically generates a number of alternative routes for the aimed conversion. A weighted assessment is used to evaluate them, and the user is able to adjust it. The process is started by entering the target compound which can easily be drawn in a sketch window or entered as strings. Finally, the program outputs the potential cascades leading from the starting material to the target compound [18]. An example cascade constructed via RetroBioCat for the biocatalysis of an API, diazepam is demonstrated in Figure 1.

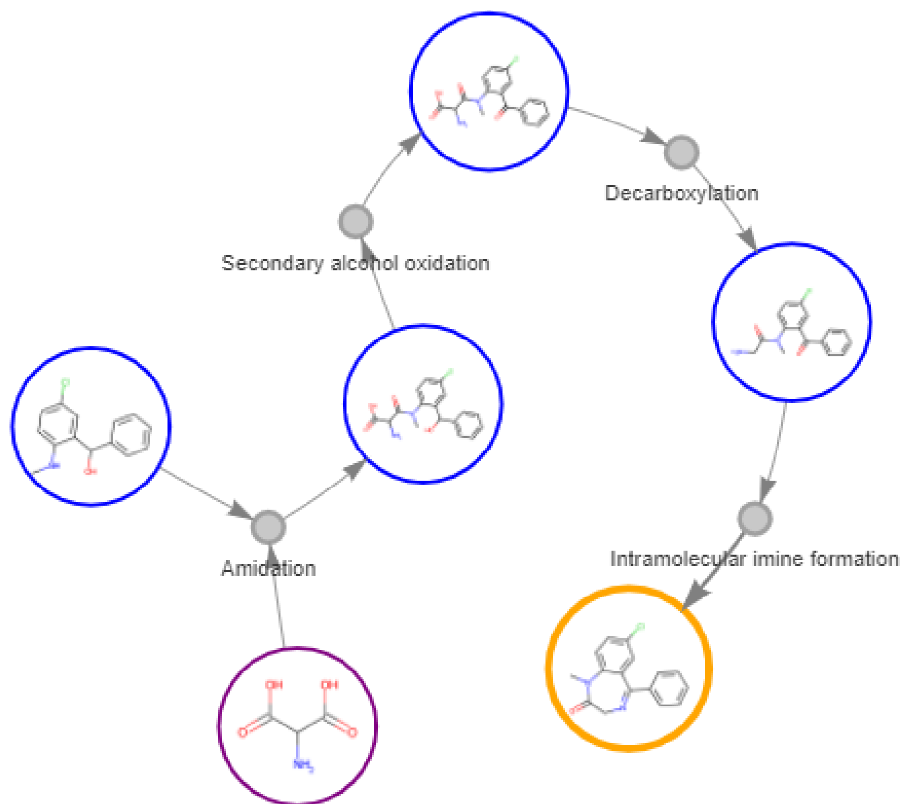

Figure 1.1: A diazepam biocatalysis cascade generated via pathway explorer of RetroBioCat [18].

Diazepam, a benzodiazepine class drug, is often used to overcome seizures, anxiety and muscular spasms resulted by the irregularities in the γ-aminobutyric acid (GABA) levels. GABA, a unique substance with ability to transmit impulses throughout nervous system, is made more active by benzodiazepines such diazepam [19]. The catalytic pathway for the synthesis of diazepam via enzyme catalysis is generated by RetroBioCat after the entry of the diazepam molecular structure which is encircled in yellow. The cascade is comprised of four distinct reactions of amidation, secondary alcohol oxidation, decarboxylation and intramolecular imine formation beginning from the starting material

aminomalonic acid, a buyable compound, encircled in purple [18]. As proposed by the software, amidation, secondary alcohol oxidation and decarboxylation reactions can be catalyzed enzymatically by ANL-family amide synthetases, a keto reductase and a decarboxylase, respectively [18]. Softwares such as RetroBioCat can speed up the cascade design step and render the process much easier by providing buyable starting materials and enzymes required for the biocatalytic conversions.

After a biosynthetic pathway is designed, a suitable biocatalyst should be chosen and optimized for each step. Many enzymes are commercially available for direct purchase as purified enzymes or crude cell lysates. Also, gene synthesis can be employed to obtain any enzyme of choice if the amino acid sequence is known. Once the enzymes are determined, their properties should be optimized since wild-type enzymes are often limited in their operating conditions and not favorable for large-scale industrial processes due to temperature levels and solvents. Optimization of any enzyme aims to improve its stability, selectivity and substrate specificity [17]. Rational design and directed evolution are the main approaches that can be utilized to improve and optimize an enzyme. Finally, after design and experimentations, bioprocess development can be achieved with the suitable physiological and environmental parameters.

1.2.2 Limitations of biocatalysis

Until the major breakthroughs in biotechnology, the quantity of commercially accessible biocatalysts posed a significant barrier to their broad adoption in the industrial scale. Some animal-sourced glycosidases, lipases and proteases were primarily those already in use in detergent and food sectors. Now, recombinant DNA technology has advanced to the point where the gene of a target enzyme can be synthesized and cloned into a host in a short period of time at a minimal expense which enables the widespread use of enzymes for catalytic conversions industrially [15]. Despite all the advantages mentioned, biocatalysts also have some inadequacies that limit their use at the industrial level. In order to be able to transform their biological substrates efficiently in vivo, wild-type enzymes have undergone evolution millions of years. They are evolved to operate on natural substrates in aqueous media and at physiological conditions. Therefore, it is expected that

they would struggle maintaining their biological function at harsher industrial conditions and operate on non-natural substrates. In order to employ enzymatic conversions in the pharmaceutical industry for the production of APIs, these limitations must be overcome. The manufacturing of APIs and pharmaceutical intermediates from non-natural substrates require better, more efficient biocatalysts with higher stability under non-physiological conditions. The adoption of biocatalytic conversions in the pharmaceutical industry also requires an enzyme recovery process and should be cost-efficient [12]. All of these requirements have been made possible by directed evolution which is based on proteins' ability to adjust and evolve depending on the selective pressure. The use of directed evolution has made it feasible to create biocatalysts with greater efficiency, refined characteristics and enhanced resistance to severe external conditions [20].

1.3 Directed Evolution of Proteins

Proteins have been the solution to many challenges that cellular life faces since their existence for millions of years. All cellular processes including DNA replication and repair, photosynthesis, metabolism, respiration and cell signaling involve complex molecular mechanisms mediated by protein function. The ability of proteins to mediate such complex processes involving biochemical and biophysical interactions is also utilized by humans in the medical field, agriculture, waste-management and many industries including pharmaceutical, food, chemical and textile. Since industrial applications require very specific enzymatic functions and environmental conditions, it is a challenge to screen and identify an enzyme that has all the necessary properties.

The development of enzymes with precise properties, higher efficiency and improved tolerance to harsh environmental conditions has been possible by the applications of directed evolution since proteins are able to evolve and adapt under selective pressure. Directed evolution is a protein engineering technique that starts with identifying a starting protein for the targeted application and creating genetic variants, by random or focused mutagenesis, to be expressed in a model organism for screening. For the identification of a variant with desired traits, a natural selection process is generated in the laboratory by selective methodologies [20]. Directed evolution allows for protein functions, that nature

has given over millions of years, to be enhanced and optimized in the laboratory for certain academic or industrial applications in much shorter periods of time.

The idea of mimicking the evolutionary process in the laboratory for production of optimized enzymes was first proposed by Manfred Eigen and William Gardiner in 1984. In their paper, Eigen and Gardiner indicated a procedure work flow of guided evolution following the Darwinian logic based on RNA self-replication but also noted that finding a variant of a single protein molecule with desired traits would be a challenge since phenotypic advantages of translation products would not be expressed as replication efficiency as in nature. Therefore, they proposed a controlled optimization procedure based on iterative rounds of small gene library construction by mutagenesis, amplification, expression and screening for optimal genotypes [21].

The first successful application of directed evolution of an enzyme was reported by Frances H Arnold in 1993. In his paper, Arnold described a method for the directed evolution of protease Subtilisin E to obtain a variant that can function in a highly-denaturing environment containing high concentrations of dimethylformamide (DMF), a polar organic solvent. After several rounds of mutagenesis, the PC3 subtilisin E variant was obtained with a k_{cat}/K_m that was 256 times of the wild type in a DMF concentration of 60% [22]. In directed evolution, advantageous mutants are discovered by high-throughput screening and/or selection in the constructed library of variants [20].

1.3.1 Library construction

The first step of directed evolution is the selection of a parent protein which exerts protein functions related with the desired function. The chosen starting sequence is subjected to mutagenesis in order to generate a library of enzyme variants. The genetic variations can be random or focused and be created by computer-guided mutagenesis and homologous or non-homologous recombination [20].

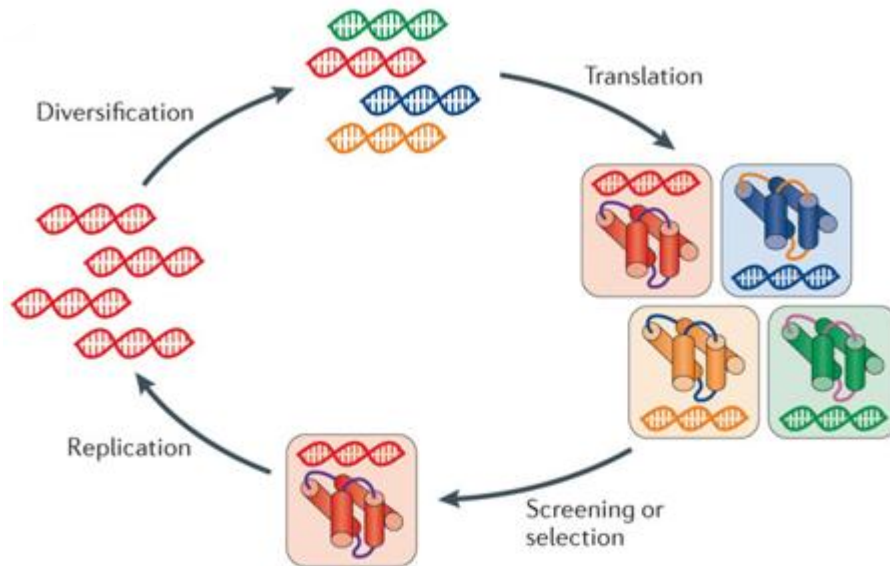

Figure 1.2: The overview of the directed evolution process of a protein [23]

Random mutagenesis techniques for library diversification include chemical mutagenesis, use of mutator strains and error-prone PCR (epPCR). Chemical mutagenesis is aided by chemical and physical agents including UV irradiation, nitrous acid, ethyl methanesulfonate (EMS), bisulfite and base analogues. In chemical mutagenesis methods, the mutation rates are dose-dependent and thus, can be easily controlled. However, the mutation rates can be low and these techniques may require the use of hazardous chemicals [23]. Another approach to random mutagenesis is the use of mutator strains such as the commonly used XL1-Red. XL1-Red is a strain of *E. coli* with high mutation rates due to its defective DNA repair function which makes it a tool for introducing mutations to parent sequence. In this technique, the gene of the interested protein is inserted into a plasmid which is then propagated in the XL1-Red strain for a certain period. For screening of a desired phenotype, the library of mutated plasmids is isolated and transformed into a model organism. Although mutator strains are commonly preferred for creating libraries of variants, their genetic instability and the risk of inducing undesired mutations on host genome limit their use [24]. Another commonly used technique for random mutagenesis is error-prone PCR (epPCR) which can help achieve high mutation rates and an even mutational spectrum. epPCR utilizes DNA polymerases with low fidelity and deactivated proofreading mechanisms to confer genes with mutations due

replication errors. In epPCR, unequal concentrations of dNTPs, mutagenic base analogues, high concentrations of Mg^{2+} and Mn^{2+} are introduced to the PCR conditions to reduce the fidelity of base-pairing [25, 26]. Cycle number can also be increased to achieve high mutation rates since mutations will accumulate at each cycle [23]. Focused mutagenesis techniques for library diversification include site-directed saturation mutagenesis and computational design strategies. Site-directed saturation mutagenesis can be used to generate high-quality focused libraries by randomization of one or more codons in the sequence using mutagenic primers with degenerate codons [27]. Site-directed mutagenesis allows for targeting only the residues relevant with the desired function but requires the structural data and generates large libraries which may complicate the selection process. Computational design techniques utilize tools such as RosettaDesign and FoldX for molecular modelling to determine beneficial codon substitutions [28, 29]. Reconstructing evolutionary adaptive paths (REAP) analysis and SCHEMA algorithm are also commonly used for designing libraries with variants that have high possibility of exhibiting desired properties [30, 31].

Apart from random and focused mutagenesis techniques, library diversification can also be achieved by homologous or non-homologous recombination methods. Homologous recombination is a process contributing to genetic diversity and is evolutionarily conserved. Utilization of homologous recombination for enzyme improvement was first put into practice by William P Stemmer in 1994. Stemmer and colleagues developed a method called DNA shuffling for homologous recombination of previously-mutated DNA segments by fragmentation via DNase and PCR re-assembly [32].

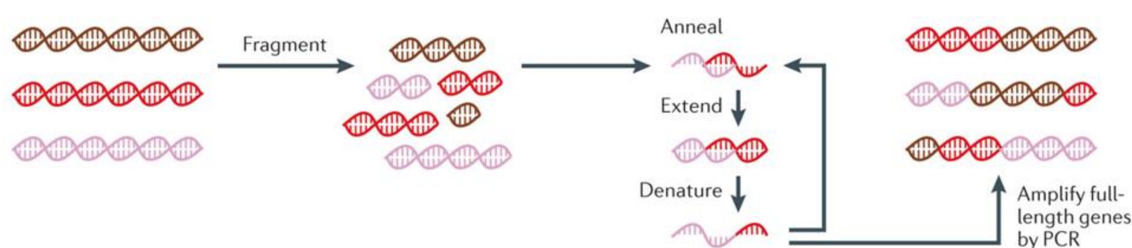

Figure 1.3: Overview of the DNA shuffling method [23]

Other methods for homologous recombination include staggered extension process (StEP), random chimeragenesis on transient templates (RACHITT) and nucleotide exchange and excision technology (NExT). Contrary to homologous recombination, non-homologous recombination methods including ITCHY and NRR can be used for shuffling of distantly related proteins [33, 34].

1.3.2 Screening and selection

Thanks to rapid developments in the directed evolution methodologies and sophisticated computational techniques, large libraries consisting of more than a billion gene variants can be constructed in a short period of time. Building a structured methodology for thoroughly evaluating these gene libraries and efficiently identifying desired phenotypes is a major issue in directed evolution. After the library of variants is constructed by random, focused or recombination-based mutagenesis, the created library is then analyzed by either screening or selection. Analytical methods, either biophysical or biochemical, are used to individually test each variant in the gene library for the desired function in the screening methods. Since experimental conditions such as the reaction environment and substrates can easily be modified at an industrial scale, the screening methods are flexible and adaptable for a variety of catalytic conversions. One of the drawbacks of this approach, however, is the low throughput which is in the range of 10^4-10^6 variants [35]. Another approach for analyzing a gene library is selection, which is desirable when the function of the engineered enzyme is necessary for the host organism survival or growth. An example of a protein function useful as a selection criteria is the detoxification of a certain substance whose presence may prevent growth. During the selection process, the nonfunctional proteins will be immediately eliminated only to select the functional variants. Spectroscopic or colorimetric enzyme assays for the desired property or improved activity can also be used in conjunction with the selection method. Proteins exhibiting the desired features can be identified from ~10^9 variants in a given gene library via selection [36]. Therefore, this approach can offer a high-throughput analysis of the created variants in a short time.

1.3.2.1 Screening methods

Screening-based analysis assesses every variant individually as opposed to the selection-based approaches which significantly lowers the likelihood of missing a desirable variant since the scope of candidates is extensive. In screening methods, 10^4 to 10^6 variants can be individually analyzed and screened, therefore, the throughput and scale of the library analysis are among the drawbacks of this approach [34]. The screening of a given variant library can be by microtiter plate-based, fluorescence-activated cell sorting (FACS)-based, mass spectrometry (MS)-based, microfluidics-based or by digital imaging. In microplate-based methods, by using a plate reader to measure fluorescence or UV absorbance, the substrates and products can be distinguished easily. A variety of enzyme properties can be screened by coupling fluorometric and colorimetric assays using 96- to 9600-well containing microtiter plates [37].

FACS can also be used to assess each variant for the desired enzyme characteristic. Up to 10^5 cells/sec can be analyzed using FACS in combination with a of screening methods to quantitatively investigate various enzyme properties [38]. A mix of cells can be separated via FACS into different groups based on their unique fluorescence and light scattering behaviors. In this method, a straight line of cells is passed across a laser beam to collect data on the fluorescence and cell size via measurements of side scatter channel (SSC) and forward scatter channel (FSC), respectively, by the detectors [37]. After cells pass through the laser beam, positively- or negatively-charged cell droplets are produced as the cells travel through a vibrating nozzle. An electromagnetic field directs the charged droplet carrying the target cell into the appropriate container [38].

Screening of gene variants can also be carried out via MS-based methods. In order to characterize complicated mixtures of engineered enzymes, MS can be used to carry out high-throughput measurements using the suitable ligands and substrates. The MS-based screening should be designed according to the nature and characteristics of the desired enzyme product. The assays can be designed to be either surface-based or solution-based [39, 40]. Microtiter plates are a popular tool for solution-based assays while microarrays are frequently employed for assays designed to be surface-based. After designing the assay, capillary electrophoresis or various chromatographic techniques can be utilized to

separate the desired enzyme from a combination of variants. Then, the sample is ionized by the ion source of the spectrometer and generated gas-phase ions are carried through a magnet to be separated by their *m/z* ratio and then through the detector. Therefore, a mass spectrum is generated by plotting the relative ion abundance against *m/z* ratios [41].

Droplet microfluidics can also be benefitted for high-throughput screening of gene libraries. Operations including dilution, sorting and addition can be carried out by microfluidic techniques after encapsulation of each variant within droplets where enzymatic reactions occur. Various analytical methods such as fluorescence, absorbance, FACS, MS and NMR can be coupled to microfluidic systems for obtaining readouts of the samples [37].

1.3.2.2 Selection methods

As opposed to the aforementioned screening methods, selection methods are able to rapidly eliminate the variants of no interest via application of a selective pressure which enables the assessment of larger gene libraries up to ~10^{11} in a shorter period of time. Since, each variant is not assessed individually, the possibility of missing a desired variant is more likely than the screening methods. Selection methods for the assessment of a library can be carried out via display-based technologies or compartmentalization [37].

1.3.2.2.1 Display-based selection

The connection between the protein and the DNA molecule that it is synthesized from determines how proteins are displayed on an organism's surface. Given that the displayed variant library is exposed to the external environment, the enzymes with the target property and function can be subjected to selection. The selected variants can easily be traced back to the complementary genes in the library which can then be amplified. Display-based selection includes plasmid display, SNAP display, phage display, mRNA display, ribosome display, cell surface display and retrovirus display [42].

A DNA-binding protein is linked to the encoded enzyme non-covalently in the plasmid display method. Following cell expression, the fusion protein attaches to the plasmid that

encodes it. Then, the formed complex of protein/plasmid is selected after cell lysis via the binding of recognition sequences on the DNA-binding domain. GAL4 DBD is commonly used for selection via plasmid display method which was constructed by Choi and colleagues (2005) [43]. High affinity DNA-binding can be achieved for obtaining target fusion proteins since the dissociation constant of the DBD is in nanomolar range and it has been widely employed in the yeast two-hybrid systems as DNA-binding partner [43]. In another technique called SNAP display, each protein and its encoding sequence is linked covalently via a SNAP-tag. While enzyme of interest is fused to the SNAP-tag, the template DNA is linked to the its substrate which is benzylguanine [44]. After that, the complex is enclosed in droplets of a water-in-oil emulsion, in which the protein is expressed. Then, the SNAP-tag and its substrate is bound which results in the coupling of the expressed protein and its encoding sequence [42].

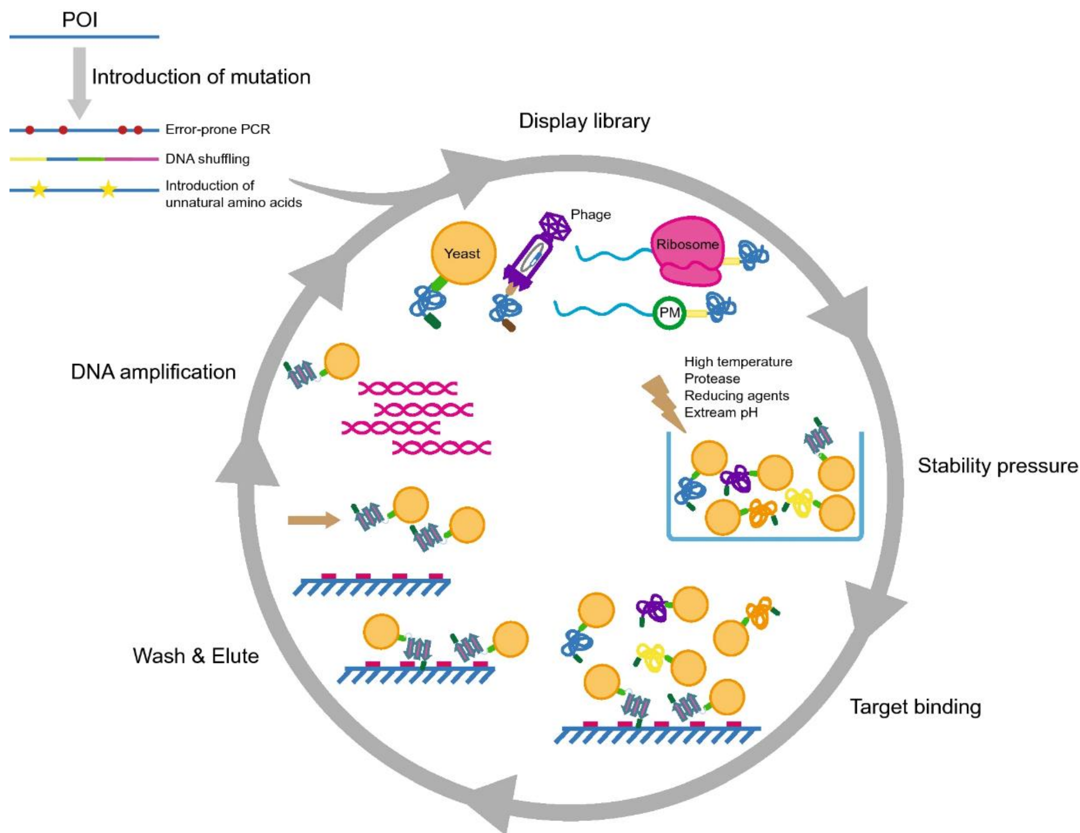

Figure 1.4: The display of the variant library on surfaces such as phage, yeast and ribosome and selective pressure for selection of variants [45].

Most commonly used technique for the selection of protein of interest is the phage display technique. The capability of filamentous phages such as M13 to infect bacterial hosts without destroying them makes them useful for the phage display technique [42]. This method involves the creation of a gene library based on phagemid DNA and its transformation into the host cells. Expressed proteins are displayed on the surface of the phage and available for selection. Similar to the phage display, in the cell display technique, the encoded proteins are displayed on the surface of cells which are then distinguished via a sorting instrument. Fluorescent molecules are useful for enzymes to bind for the selection step. The aforementioned display methods are restricted by transformation efficiency as they require an organism for the display of the target protein [46]. The additional step of transformation also requires more time than the mRNA/ribosome display methods. In the ribosome display method, in vitro transcription and translation of the gene occurs and a complex forms with the mRNA, expressed protein and the ribosome. Real-time PCR is used to recover the targeted gene [47]. In this method, selection step is rather quick since there is no additional step for transformation.

1.3.2.2.2 Compartmentalization-based selection

In this method, the gene and enzyme are spatially restricted to exist in a certain single compartment through compartmentalization. Existing cells and phage particles can be used for this method which is called *in vivo* compartmentalization (IVVC), or synthetic compartments can be produced such as emulsion droplets which is called *in vitro* compartmentalization (IVTC). While IVVC-based selection is restricted by the transformation efficiency, IVTC-based selection is not. A major drawback of the IVTC applications is the lack of post-translational modifications (PTM), which most enzymes require to be functional, therefore, IVVC is commonly preferred over IVTC [42]. One of the selection methods based on IVVC is called the growth complementation. In this method, the selection of enzymes engineered via directed evolution can only be selected by using living cells. Only the cells possessing the target enzyme variants survive when a selective pressure is applied since the assessed enzyme characteristic is associated with the survival of the host cell. In vivo compartmentalization has certain drawbacks one of

which is extreme temperatures or pH levels limit the use for detecting the desired enzyme property [48].

A reporter can also be utilized for selection in *in vivo* compartmentalization (IVVC). The target enzyme activity is intended to regulate the reporter activity which enables the selection of variants since cells harboring the desired enzymes will show reporter activity [49]. Commonly used reporters include LacZ which shows colorimetric activity and green fluorescent protein (GFP) which shows fluorescence activity [50, 51]. Lux operons such as luxCDABE can also be used as reporters to monitor bioluminescent activity of the cells [52]. Several strategies can be employed for reporter-based selection of the desired variants. These strategies can utilize transcriptional or translational activities of the reporter via indirect activity of the desired enzyme. For example, the desired variant can activate reporter transcription by binding to the transcriptional regulator or desired variant can directly modify the transcriptional regulator which can also result in the reporter transcription. The activity can also be translational as mentioned above and in this case, the desired variant can bind to a riboswitch/ribozyme and cause reporter gene to be translated [49].

1.4 Directed Evolution of Biocatalysts

The applications of directed evolution to produce biocatalysts for pharmaceutical industry aims high catalytic efficiency, stability, substrate specificity and necessarily enantioselectivity. Since understanding the pharmacological significance of stereoisomerism, the testing of chiral drugs for pure enantiomers has been made mandatory by the relevant authorities [13]. Protein engineering methods have been widely employed for creating biocatalysts suitable for pharmaceutical applications aiming to manufacture pure enantiomers. In an earlier study by Reetz and colleagues (1997), epPCR was conducted by altering reaction conditions to subject a lipase enzyme to random mutagenesis. The aim was to obtain a mutant clone of the 933 base-paired lipase gene to enhance the enantioselectivity of the enzyme. The reaction conditions were adjusted to introduce up to two amino acid changes per enzyme. The mutant enzymes were tested for their enantioselectivity and 12 out of 1000 mutants showed improvement in their

enantioselectivity and a mutant named PlBOl-E4 demonstrated an enantioselectivity of 31% for S-configuration which is 2% for the wild-type lipase [53]. By the developments in the directed evolution methodologies over the years, the enhancement of enantioselectivity of enzymes become more successful and engineered novel enzymes emerged with enantioselectivity up to >98% *ee*.

In a more recent study, a phenylalanine dehydrogenase enzyme was engineered via saturation mutagenesis to improve its substrate specificity and enantioselectivity for the asymmetric reductive amination of phenyl-2-butanone and phenylacetone. The reductive amination of these ketones are important reactions for producing precursors and intermediates used in the manufacturing of active pharmaceutical ingredients [54, 55]. Target products were (R)-1-methyl-3-phenylpropylamine and (R)-amphetamine with high enantiopurity via reductive aminations of phenyl-2-butanone and phenylacetone, respectively, by the engineered enzyme. *Rhodococcus* sp. M4 was chosen as the source organism for the phenylalanine dehydrogenase and double-site saturation mutagenesis was performed via PCR at the residues Lys66 and Asn262. Two mutants with improved specific activities were selected as starting proteins for directed evolution via single-site saturation mutagenesis to further improve the enantioselectivity and amination activity of the enzyme. A mutant named K66Q/ S149G/N262C showed even more enhanced amination activity and all three mutants showed >98% *ee* for both (R)-1-methyl-3-phenylpropylamine and (R)-amphetamine [56]. More and more enzymes with higher enantioselectivity have been developed via directed evolution as suitable biocatalysts for the production of pharmaceutical ingredients over the years. The diverse class of proteins being engineered through directed evolution and their enhanced catalytic functions can revolutionize the pharmaceutical industry and render it much more sustainable.

Another notable enzyme that has been engineered to enhance enantioselectivity is a hemoprotein derived from wild-type *Rma* cytochrome *c* for the catalyzation of alkenes into amino alcohols in a single step [57]. Numerous bioactive substances and pharmaceutical drugs including antibiotics, anti-HIV and anti-malarial drugs and also neurotransmitters contain amino alcohol motifs that are optically active [58].

2. REDOX PROTEINS

Key biological processes including cell signaling, photosynthesis, metabolism and respiration involve oxidation-reduction processes which require transfer of electrons. These processes are catalyzed by a class of proteins called redox enzymes whose catalytic sites function as electron donors or acceptors [59]. Redox activity of redox enzymes is largely conferred by the metal ions that are redox-active. Proteins that accommodate a metal ion are called metalloproteins which are the predominant ET centers for cell processes [60].

2.1 Cytochrome *C*

Cytochrome c is a type of metalloprotein as it contains a heme prosthetic group with heme iron at the center. The F atom of the heme group confers the cytochrome c with its ability of electron transfer as it switches between its Fe^{2+} and Fe^{3+} states [61]. Cytochrome c is a vital redox center for cell processes as it functions in disposal of ROS, respiration and cell apoptosis depending on where it resides in the cell [62]. The structural similarity of cyt c among many species of organisms is one of its distinguishing features. Since there haven't been many significant changes to this protein's sequence during evolution, the domains and residues that are essential to its stability and function have been maintained [63]. It has different roles according to its localization in the cell including being a redox agent in the ETC, triggering apoptosis and maintaining the levels of ROS [64]. The relatively high redox potential required for cyt c to act in the respiratory chain is provided by the axial bond between the heme-iron and one of its axial ligands Met80. Due to its extremely poor stability, the Met80-containing region is one among the first to unfold which makes the linkage between the Met80 and heme-iron rather weak [63]. The weakness of this axial bond allows the substitution of Met80 with other residues. This property of cyt c confers its different functions according to the residue which Met80 is switched to [62]. Apart from its role as a redox protein, another important function of cyt c is its role in the caspase cascade activation due to the interactions with Apaf-1 in the cytoplasm. The activation of these apoptosis precursors by the activities of cyt c makes it essential for the apoptotic

intrinsic pathway. For the apoptotic pathway to begin, the cyt c has to be released from mitochondria to the cytosol which is enabled by the permeabilization of the membrane. The permeabilization of the mitochondrial membrane occurs via the peroxidase activity of cyt c which is exerted after cyt c is bound to a phospholipid called cardiolipin. The interactions between cyt c and cardiolipin disrupts the bond between Met80 and heme iron and confers cyt c its peroxidase activity. In the cardiolipin-bound cyt c, the Met80 residue is replaced with a histidine residue which alters the redox potential of the protein. The newly gained function of cyt c enables the peroxidation of cardiolipin and causes the permeability of the mitochondrial membrane [65, 66]. After cyt c diffuses into the cytosol, Apaf-1 goes through oligomerization in the presence of dATP and forms an apoptosome of a heptameric nature which can activate caspase-9. When caspase-9 is activated, a protease cascade that includes caspases 3 and 7 is also initiated, causing target proteins to be disintegrated and the cell to die in a controlled manner [67]. Although many residues in the cyt c is known to be interacting with Apaf-1, the knock-out studies show that the lysine residue at the 72 position is mainly responsible for the relation between cyt c and the WD40 region of Apaf-1 [68].

Another important function of cyt c in the eukaryotic cells is the scavenging of reactive oxygen species (ROS). When present in low quantities, ROS like H_2O_2 and superoxide anion are known to have significant signaling functions. But they can significantly harm the cell and its constituents such as DNA and lipids when present above certain concentration and lead to death, particularly when hydroxyl radicals are produced. As the primary sources of intracellular ROS, mitochondria are known to produce superoxide anions in both the intermembrane space and the matrix [69]. Experimental evidence demonstrates the cyt c's detoxifying role in the clearance of H_2O_2 and superoxide anion in mitochondria and its function in the regulation of ROS levels [70].

2.2 Cytochrome P450

Nearly all genomes include the CYP gene family which code for one of the major superfamilies of enzymes called cytochrome P450s, or formally hemethiolate monooxygenases. There are 57 functional P450 genes in the human genome whose

upregulation is promoted via various compounds such as environmental pollutants, steroid hormones, their byproducts and xenobiotics [71]. Their primary function is the metabolism of xenobiotics including environmental toxins and drugs which is executed by xenobiotic-metabolizing enzymes (XMEs) via functionalization and conjugation, also called phases I and II, respectively. P450 enzymes make up the majority of phase I XMEs [72]. Metabolic activation and detoxification of foreign compounds are carried out upon increase in the gene expressions of phase I and II XMEs due to the interactions of xenobiotics with XME receptors. The metabolic activation is performed by phase I XMEs via addition of a functional group to the substrate. The substrates turn into electrophilic intermediates after being acted on by phase I XMEs. These electrophilic intermediates are then conjugated with certain moieties by the phase II XMEs in order to create easily excretable hydrophilic compounds [73].

The fifth coordination site of P450 is where the heme iron is bound, and there exists a conserved cysteine residue which is involved in oxygen insertion to the substrate. The oxygenated product is then released which can become stable or stay as an unstable reactive oxygenated intermediate (ROM). These may bind covalently to DNA due to being extremely reactive and cause oxidative stress, mutations and toxicity [72]. A diverse range of substrates, both endo- and exogenous, are catalyzed by P450s into toxic metabolites in the liver. The accumulation of these metabolites due to the catalytic activities of P450 enzymes increase oxidative stress and cause toxicity to the cell which may lead to tumorigenesis, even though P450 enzymes' main objective is to facilitate the clearance of xenobiotics from the system by converting them into hydrophilic, easily excretable molecules [74]. The generation of toxic load in the cell via catalytic activities of P450s affects cell signaling pathways. Some other eicosanoid-mediated mechanisms and inflammation are additional ways whereby P450s contribute to the development and spread of tumors [72, 74].

Apart from drug metabolism in the liver, CYP enzymes have other endogenous functions worth mentioning. Cellular processes including the biosynthesis of vitamin D3, steroids, cholesterol, bile acids and their metabolism all require CYP enzymes [72]. The reactions

that can be catalyzed via P450s include sulfoxidation, oxidative deamination, aromatic epoxidation and hydroxylation, both aromatic and aliphatic [71].

2.3 Biomedical Applications of Cytochromes

Due to their involvement in vital biological processes in the body, therapeutic proteins have attracted a lot of interest for their potential clinical applications. There are currently hundreds of therapeutic proteins approved for use and also many are in clinical trials for the treatment of diseases including infections, immune disorders, diabetes, anemia and cancer [75]. Since proteins are extremely selective and specific, they can target particular stages in the pathophysiology of a disease. Therefore, the beneficial immune reactions can be maintained since protein therapeutics target only the key cellular and molecular mechanisms associated with disease as opposed to small-molecule drugs that suppress both harmful and beneficial immune reactions [76]. Depending on the disease, recombinantly produced interleukins, interferons, hormones, enzymes, thrombolytics, blood factors and antibodies are among most used protein therapeutics [75]. Using protein therapeutics against cancer is especially prominent since current cancer drugs have detrimental side-effects due to their off-target activities. Cytochrome c has the potential to be exploited as an anticancer therapeutic protein because of its proapoptotic activities. A cyt c-transferrin conjugate was developed by Saxena and colleagues (2018) to be used against lung cancer [77]. The conjugation of the enzyme to transferrin was shown to enable the cancer cells to be actively targeted due to the overexpression of transferrin receptor on their surface. It was demonstrated that the cyt c-transferrin conjugate was taken up by the cancer cells that show transferrin receptor overexpression which then underwent apoptosis due to caspase-3 activation [77]. The results indicate that cytochrome c has a promising potential to be used as an anticancer drug via utilizing active drug delivery systems.

Another possible biomedical application of cyt c comes from its presence in the extracellular space upon release from cells due to structural damage which enables cyt c to be useful as a damage-associated molecular pattern (DAMP). Since released cyt c eventually enters bloodstream, its levels in the blood serum can be monitored to serve as a biomarker for significant mitochondrial and cellular damage. The high levels of cyt c

may be the indicator of many diseases including multiple organ failure, autoimmune diseases and cancer [78].

CYP family of cytochromes also has many possible applications in the biomedical field due their biocatalytic properties and significance in drug metabolism. The biomedical implications of human CYPs often focus on their involvement in drug clearance, despite the fact that they are also crucial for the production of steroidal compounds and cholesterol. More and more clinical studies are utilizing P450-based systems to conduct drug testing and investigate drug–drug interactions [79]. Screening the P450 substrates of a drug enables one to identify potentially hazardous drug-food or drug-drug interactions, which are frequently brought on by the altering of P450 metabolism via inhibition or induction [80]. Therefore, P450-based biosensors might be helpful for screening pharmaceutical compounds. Their ability to carry out a wide range of biocatalytic transformations also makes them potential tools for drug discovery and synthesis. Dehalogenation, nitration, sulfoxidation and hydroxylation are among the biocatalytic processes that can be conducted by wild-type and evolutionarily engineered CYPs [79]. Due to their significant regio- and chemoselectivity and capacity of adding oxygen to a wide range of molecules, P450s are often subjected to protein engineering in order to be repurposed in industrial synthesis of chemicals and pharmaceutical compounds in a sustainable fashion. The engineering of P450s evolutionarily mainly aims improving their stability, efficiency and also substrate specificity [81].

2.4 Directed Evolution Of Cytochrome Proteins For Pharmaceutical Applications

Cytochrome proteins have been a focus in the directed evolution studies for the last two decades because of their ability of carrying complex catalytic conversions and operating on a wide range of substrates. They are able to work under mild conditions and act with high regio- and chemo-selectivity which are critical factors in the pharmaceutical applications. The majority function as monooxygenases, inserting an oxygen atom to a substrate while reducing the second to water by presence of O_2 and a cofactor [82]. When O2 attaches to their heme iron site, the O-O linkage is enzymatically disrupted,

incorporating one O to product while reducing other to H_2O. The critical intermediate which allows this transformation to occur is the Fe(IV)- oxo-porphyrin cation radical. The P450s are able to oxidize C-H bonds that are typically inert [83]. They have similar limitations in solvent compatibility and stability as other mesophilic catalysts. In an industrial setting, the process calls for enzymes to be active for prolonged periods of time at a wide range of temperature to accomplish maximum productivity. Therefore, enhancement of thermostability has been a goal of protein engineering studies involving P450s as well as improvement of solvent compatibility, turnover rates and enabling operation with preferred cofactors. In an earlier study by Kumar and colleagues (2005), for the purpose of creating more potent P450s for the processing of cancer prodrugs ifosfamide and cyclophosphamide in the body, P450 2B1 was exposed to engineering via random mutagenesis. With cyclophosphamide, the mutants L209A/S334P exhibited a 2.8-fold rise in k(cat)/K(m), while V183L/L209A showed a 3.5-fold improvement with ifosfamide in their metabolism via the clinically favorable 4-hydroxylation mechanism [84]. Random mutagenesis was used to further enhance these mutants which were later screened for improved tolerance dimethyl sulfoxide (DMSO) and temperature. Industrial applications will be made significantly more efficient by improving wild-type P450s' resistance to industrial solvents and temperature in addition to successfully substituting expensive redox partners with alternative oxidant such H_2O_2. Eight mutants with >2-fold improvements in reaction rates were identified after screening more than 150 distinct mutants at the T_{50} (temperature at which 50% of parental activity is retained) and $[DMSO]_{50}$ (DMSO concentration at which 50% of parental activity is retained). These mutants were also able operate with the H_2O_2 as the redox partner [85]. In another study, P450 BM3 was engineered at a single residue to catalyze synthetic conversions which often require transition metals. A P450 BM3 mutant was obtained that, through a putative carbene transfer reaction, can catalyze the remarkably enantio- and diastereoselective olefin cyclopropanation of double bond. The carbene transfer via P450 BM3 mutant with the co-substrate diazoester was achieved with the mutation of Thr168Ala [86]. The research focused on engineering cytochrome proteins for enantioselective synthetic catalyzations that normally require transition metals can

revolutionize the pharmaceutical industry and render is more sustainable both environmentally and economically.

Cytochrome c has also been engineered as it has many critical functions at cellular scale and is able to carry out a wide range of catalytic conversions. Thanks to protein' ability to adapt, they can be rendered competent of transformations unrelated to their native roles such as generating carbon-silicon linkages. A revolutionary study presented a cyt c mutant that is able insert carbene to silicon-hydrogen bonds to catalyze the synthesis of organosilicon compounds. Using chiral transition metal compounds based on iridium, rhodium and some other metals, synthetic conversions like carbene integration to silanes can be made enantioselective. Although enantioselectivity is achieved by these methodologies, their turnover rate is limited and require specific solvents which makes engineered enzymes desired for these conversions [87]. Biocatalysis via enzymes is increasingly desired as substitute for conventional synthesis techniques due to their specificity and selectivity in chemical conversions.

3. CONCLUSIONS

The manufacturing of pharmaceutical intermediates is a delicate multistep process that demands the use of catalysts. The pharma industry has relied on chemo-catalysts such as transition metals for the production of APIs for a long time. Following the advancements in the directed evolution, biocatalysis have gained interest for a more efficient process of API manufacturing. The benefits of biocatalysts over chemo-catalysts include excellent selectivity, low-cost process, sustainability and reduced waste output. Growing environmental consciousness as a result of global warming also puts pressure on the industry to seek to cut their carbon footprints. Biocatalysis enables the industry to become significantly more sustainable and eco-friendlier. For the widespread adoption of biocatalysis in the pharmaceutical sector, its limitations such as low stability and narrow substrate range should be improved which can be achieved via protein engineering. A family of redox proteins known as C-type cytochromes has heme prosthetic groups with iron centers. Their multifunctional nature and abundance in the genomes of almost all

species make them suitable candidates for improvement via directed evolution. Numerous research has focused on using cyt c and P450 proteins to create functional variants through mutagenesis and there have been excellent advancements in their functionality and efficiency of catalytic transformations. The functionalization of these novel cytochrome variants in the catalytic conversions for the synthesis of APIs can revolutionize the pharmaceutical industry and help render it more sustainable and greener.

REFERENCES

1. Busacca, C., Fandrick, D., Song, J., & Senanayake, C. (2012). Transition Metal Catalysis in the Pharmaceutical Industry. *Applications Of Transition Metal Catalysis In Drug Discovery And Development*, 1-24. doi: 10.1002/9781118309872.ch1

2. Hartwig, J. F. (2008). Evolution of a fourth-generation catalyst for the amination and thioetherification of aryl halides. *Accounts of chemical research*, *41*(11), 1534-1544.

3. Miyaura, N., & Suzuki, A. (1995). Palladium-catalyzed cross-coupling reactions of organoboron compounds. *Chemical reviews*, *95*(7), 2457-2483.

4. Blakemore, D. C., Doyle, P. M., & Fobian, Y. M. (2016). Suzuki–Miyaura Coupling. In *Synthetic methods in drug discovery - volume 1* (pp. 1-69). Cambridge: Royal Society of Chemistry.

5. Garrett, C. E., & Prasad, K. (2004). The art of meeting palladium specifications in active pharmaceutical ingredients produced by Pd-catalyzed reactions. *Advanced Synthesis & Catalysis*, *346*(8), 889-900.

6. McCarthy, S., Braddock, D., & Wilton-Ely, J. (2021). Strategies for sustainable palladium catalysis. *Coordination Chemistry Reviews*, *442*, 213925. doi: 10.1016/j.ccr.2021.213925
7. Nuss, P., & Eckelman, M. (2014). Life Cycle Assessment of Metals: A Scientific Synthesis. *Plos ONE*, *9*(7), e101298. doi: 10.1371/journal.pone.0101298
8. Hayler, J. D., Leahy, D. K., & Simmons, E. M. (2018). A pharmaceutical industry perspective on sustainable metal catalysis. *Organometallics*, *38*(1), 36-46.

9. Martínez-Montero, L., Schrittwieser, J., & Kroutil, W. (2018). Regioselective Biocatalytic Transformations Employing Transaminases and Tyrosine Phenol Lyases. *Topics In Catalysis*, *62*(17-20), 1208-1217. doi: 10.1007/s11244-018-1054-7

10. Bornscheuer, U., Huisman, G., Kazlauskas, R., Lutz, S., Moore, J., & Robins, K. (2012). Engineering the third wave of biocatalysis. *Nature*, *485*(7397), 185-194. doi: 10.1038/nature11117

11. Arroyo, M., de la Mata, I., García, J., & Barredo, J. (2017). Biocatalysis for Industrial Production of Active Pharmaceutical Ingredients (APIs). *Biotechnology Of Microbial Enzymes*, 451-473. doi: 10.1016/b978-0-12-803725-6.00017-0

12. Sheldon, R., & Brady, D. (2018). The limits to biocatalysis: pushing the envelope. *Chemical Communications*, *54*(48), 6088-6104. doi: 10.1039/c8cc02463d

13. Erythropel, H., Zimmerman, J., de Winter, T., Petitjean, L., Melnikov, F., & Lam, C. et al. (2018). The Green ChemisTREE: 20 years after taking root with the 12 principles. *Green Chemistry*, *20*(9), 1929-1961. doi: 10.1039/c8gc00482j

14. Siedentop, R., & Rosenthal, K. (2022). Industrially Relevant Enzyme Cascades for Drug Synthesis and Their Ecological Assessment. *International Journal Of Molecular Sciences*, *23*(7), 3605. doi: 10.3390/ijms23073605

15. Sheldon, R., & Pereira, P. (2017). Biocatalysis engineering: the big picture. *Chemical Society Reviews*, *46*(10), 2678-2691. doi: 10.1039/c6cs00854b

16. De Carvalho, C. (2011). Enzymatic and whole cell catalysis: Finding new strategies for old processes. *Biotechnology Advances*, *29*(1), 75-83. doi: 10.1016/j.biotechadv.2010.09.001

17. Bell, E., Finnigan, W., France, S., Green, A., Hayes, M., & Hepworth, L. et al. (2021). Biocatalysis. *Nature Reviews Methods Primers*, *1*(1). doi: 10.1038/s43586-021-00044-z

18. Finnigan, W. (2022). RetroBioCat. Retrieved 26 October 2022, from https://retrobiocat.com/automated_cascade_design

19. Enna, S. The GABA Receptors. *The GABA Receptors*, 1-21. doi: 10.1007/978-1-59745-465-0_1

20. Romero, P., & Arnold, F. (2009). Exploring protein fitness landscapes by directed evolution. *Nature Reviews Molecular Cell Biology*, *10*(12), 866-876. doi: 10.1038/nrm2805

21. Eigen, M., & Gardiner, W. (1984). Evolutionary molecular engineering based on RNA replication. *Pure And Applied Chemistry*, *56*(8), 967-978. doi: 10.1351/pac198456080967

22. Chen, K., & Arnold, F. (1993). Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proceedings Of The National Academy Of Sciences*, *90*(12), 5618-5622. doi: 10.1073/pnas.90.12.5618

23. Packer, M., & Liu, D. (2015). Methods for the directed evolution of proteins. *Nature Reviews Genetics*, *16*(7), 379-394. doi: 10.1038/nrg3927

24. Selifonova, O., Valle, F., & Schellenberger, V. (2001). Rapid Evolution of Novel Traits in Microorganisms. *Applied And Environmental Microbiology*, *67*(8), 3645-3649. doi: 10.1128/aem.67.8.3645-3649.2001

25. Cadwell, R., & Joyce, G. (1992). Randomization of genes by PCR mutagenesis. *Genome Research*, *2*(1), 28-33. doi: 10.1101/gr.2.1.28

26. Zaccolo, M., Williams, D., Brown, D., & Gherardi, E. (1996). An Approach to Random Mutagenesis of DNA Using Mixtures of Triphosphate Derivatives of Nucleoside Analogues. *Journal Of Molecular Biology*, *255*(4), 589-603. doi: 10.1006/jmbi.1996.0049

27. Reetz, M., & Carballeira, J. (2007). Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nature Protocols*, *2*(4), 891-903. doi: 10.1038/nprot.2007.72

28. Das, R., & Baker, D. (2008). Macromolecular Modeling with Rosetta. *Annual Review Of Biochemistry*, *77*(1), 363-382. doi: 10.1146/annurev.biochem.77.062906.171838

29. Buß, O., Rudat, J., & Ochsenreither, K. (2018). FoldX as Protein Engineering Tool: Better Than Random Based Approaches?. *Computational And Structural Biotechnology Journal*, *16*, 25-33. doi: 10.1016/j.csbj.2018.01.002

30. Chen, F., Gaucher, E., Leal, N., Hutter, D., Havemann, S., & Govindarajan, S. et al. (2010). Reconstructed evolutionary adaptive paths give polymerases accepting reversible terminators for sequencing and SNP detection. *Proceedings Of The National Academy Of Sciences*, *107*(5), 1948-1953. doi: 10.1073/pnas.0908463107

31. Voigt, C., Martinez, C., Wang, Z., Mayo, S., & Arnold, F. (2002). Protein building blocks preserved by recombination. *Nature Structural Biology*. doi: 10.1038/nsb805

32. Stemmer, W. (1994). Rapid evolution of a protein in vitro by DNA shuffling. *Nature*, *370*(6488), 389-391. doi: 10.1038/370389a0

33. Ostermeier, M., Shim, J., & Benkovic, S. (1999). A combinatorial approach to hybrid enzymes independent of DNA homology. *Nature Biotechnology*, *17*(12), 1205-1209. doi: 10.1038/70754

34. Bittker, J., Le, B., Liu, J., & Liu, D. (2004). Directed evolution of protein enzymes using nonhomologous random recombination. *Proceedings Of The National Academy Of Sciences*, *101*(18), 7011-7016. doi: 10.1073/pnas.0402202101

35. Arnold, F., & Moore, J. (1997). Optimizing industrial enzymes by directed evolution. *New Enzymes For Organic Synthesis*, 1-14. doi: 10.1007/bfb0103300

36. Clackson, T., & Wells, J. (1994). In vitro selection from protein and peptide libraries. *Trends In Biotechnology*, *12*(5), 173-184. doi: 10.1016/0167-7799(94)90079-5

37. Wang, Y., Xue, P., Cao, M., Yu, T., Lane, S. T., & Zhao, H. (2021). Directed evolution: methodologies and applications. *Chemical reviews*, *121*(20), 12384-12444.

38. Chiu, F. W., & Stavrakis, S. (2019). High-throughput droplet-based microfluidics for directed evolution of enzymes. *Electrophoresis*, *40*(21), 2860-2872.

39. Liu, M., Van Voorhis, W., & Quinn, R. (2021). Development of a target identification approach using native mass spectrometry. *Scientific Reports*, *11*(1). doi: 10.1038/s41598-021-81859-4

40. O'Connell, T. N., Ramsay, J., Rieth, S. F., Shapiro, M. J., & Stroh, J. G. (2014). Solution-Based Indirect Affinity Selection Mass Spectrometry-A General Tool For High-Throughput Screening Of Pharmaceutical Compound Libraries. *Analytical Chemistry*, *86*(15), 7413-7420.

41. Si, T., Xue, P., Choe, K., Zhao, H., & Sweedler, J. V. (2021). High-Throughput Mass Spectrometry Complements Protein Engineering. *Protein Engineering: Tools and Applications*, 57-79.

42. Xiao, H., Bao, Z., & Zhao, H. (2015). High throughput screening and selection methods for directed enzyme evolution. *Industrial & engineering chemistry research*, *54*(16), 4011-4020.

43. Choi, Y., Pack, S., & Yoo, Y. (2005). Development of a Plasmid Display System using GAL4 DNA Binding Domain for the in Vitro Screening of Functional Proteins. *Biotechnology Letters*, *27*(21), 1707-1711. doi: 10.1007/s10529-005-2735-4

44. Cole, N. B. (2013). Site-Specific Protein Labeling with SNAP-Tags. *Current protocols in protein science*, *73*(1), 30-1.

45. Ren, C., Wen, X., Mencius, J., & Quan, S. (2019). Selection and screening strategies in directed evolution to improve protein stability. *Bioresources And Bioprocessing*, *6*(1). doi: 10.1186/s40643-019-0288-y

46. Marintcheva, B. (2018). Phage Display. *Harnessing The Power Of Viruses*, 133-160. doi: 10.1016/b978-0-12-810514-6.00005-2

47. Thom, G., & Groves, M. (2012). Ribosome Display. *Methods In Molecular Biology*, 101-116. doi: 10.1007/978-1-61779-931-0_6

48. Aharoni, A., Griffiths, A., & Tawfik, D. (2005). High-throughput screens and selections of enzyme-encoding genes. *Current Opinion In Chemical Biology*, *9*(2), 210-216. doi: 10.1016/j.cbpa.2005.02.002

49. van Rossum, T., Kengen, S. W., & van der Oost, J. (2013). Reporter-based screening and selection of enzymes. *The FEBS journal*, *280*(13), 2979-2996.

50. Mohn, W. W., Garmendia, J., Galvao, T. C., & De Lorenzo, V. (2006). Surveying biotransformations with a la carte genetic traps: translating dehydrochlorination of lindane (gamma-hexachlorocyclohexane) into lacZ-based phenotypes. *Environmental microbiology*, *8*(3), 546-555.

51. Uchiyama, T., & Miyazaki, K. (2010). Product-induced gene expression, a product-responsive reporter assay used to screen metagenomic libraries for enzyme-encoding genes. *Applied and environmental microbiology*, *76*(21), 7029-7035.

52. Craney, A., Hohenauer, T., Xu, Y., Navani, N. K., Li, Y., & Nodwell, J. (2007). A synthetic luxCDABE gene cluster optimized for expression in high-GC bacteria. *Nucleic acids research*, *35*(6), e46. https://doi.org/10.1093/nar/gkm086

53. Reetz, M. T., Zonta, A., Schimossek, K., Jaeger, K. E., & Liebeton, K. (1997). Creation of enantioselective biocatalysts for organic chemistry by in vitro evolution. *Angewandte chemie international edition in english*, *36*(24), 2830-2832.

54. Abrams, P., Speakman, M., Stott, M., Arkell, D., & Pocock, R. (1997). A dose-ranging study of the efficacy and safety of tamsulosin, the first prostate-selective; 1A-adrenoceptor antagonist*, in patients with benign prostatic obstruction. *British journal of urology*, *80*(4), 587-596.

55. Clifton, J. E., Collins, I., Hallett, P., Hartley, D., Lunts, L. H., & Wicks, P. D. (1982). Arylethanolamines derived from salicylamide with. alpha- and beta-adrenoceptor blocking activities. Preparation of labetalol, its enantiomers and related salicylamides. *Journal of Medicinal Chemistry*, *25*(6), 670-679.

56. Ye, L. J., Toh, H. H., Yang, Y., Adams, J. P., Snajdrova, R., & Li, Z. (2015). Engineering of amine dehydrogenase for asymmetric reductive amination of ketone by evolving Rhodococcus phenylalanine dehydrogenase. *Acs Catalysis*, *5*(2), 1119-1122.

57. Cho, I., Prier, C. K., Jia, Z. J., Zhang, R. K., Görbe, T., & Arnold, F. H. (2019). Enantioselective aminohydroxylation of styrenyl olefins catalyzed by an engineered hemoprotein. *Angewandte Chemie International Edition*, *58*(10), 3138-3142.

58. Bergmeier, S. C. (2000). The synthesis of vicinal amino alcohols. *Tetrahedron*, *17*(56), 2561-2576.

59. Prabhulkar, S., Tian, H., Wang, X., Zhu, J., & Li, C. (2012). Engineered Proteins: Redox Properties and Their Applications. *Antioxidants & Redox Signaling*, *17*(12), 1796-1822. doi: 10.1089/ars.2011.4001

60. Liu, J., Chakraborty, S., Hosseinzadeh, P., Yu, Y., Tian, S., Petrik, I., ... & Lu, Y. (2014). Metalloproteins containing cytochrome, iron–sulfur, or copper redox centers. *Chemical reviews*, *114*(8), 4366-4469.

61. Tuppy, H., & Kreil, G. (2013). Cytochrome c. *Encyclopedia Of Biological Chemistry*, 599-601. doi: 10.1016/b978-0-12-378630-2.00374-1

62. Santucci, R., Sinibaldi, F., Cozza, P., Polticelli, F., & Fiorucci, L. (2019). Cytochrome c: An extreme multifunctional protein with a key role in cell fate. *International Journal Of Biological Macromolecules*, *136*, 1237-1246. doi: 10.1016/j.ijbiomac.2019.06.180

63. Zaidi, S., Hassan, M., Islam, A., & Ahmad, F. (2013). The role of key residues in structure, function, and stability of cytochrome-c. *Cellular And Molecular Life Sciences*, *71*(2), 229-255. doi: 10.1007/s00018-013-1341-1

64. Delinois, L. J., De León-Vélez, O., Vázquez-Medina, A., Vélez-Cabrera, A., Marrero-Sánchez, A., Nieves-Escobar, C., Alfonso-Cano, D., Caraballo-Rodríguez, D., Rodriguez-Ortiz, J., Acosta-Mercado, J., Benjamín-Rivera, J. A.,

González-González, K., Fernández-Adorno, K., Santiago-Pagán, L., Delgado-Vergara, R., Torres-Ávila, X., Maser-Figueroa, A., Grajales-Avilés, G., Méndez, G. I., Tinoco, A. D. (2021). Cytochrome c: Using biological insight toward engineering an optimized anticancer Biodrug. *Inorganics*, *9*(11), 83. https://doi.org/10.3390/inorganics9110083

65. Patriarca, A., Polticelli, F., Piro, M., Sinibaldi, F., Mei, G., & Bari, M. et al. (2012). Conversion of cytochrome c into a peroxidase: Inhibitory mechanisms and implication for neurodegenerative diseases. *Archives Of Biochemistry And Biophysics*, *522*(1), 62-69. doi: 10.1016/j.abb.2012.03.028

66. Milazzo, L., Tognaccini, L., Howes, B. D., Sinibaldi, F., Piro, M. C., Fittipaldi, M., Baratto, M. C., Pogni, R., Santucci, R., & Smulevich, G. (2017). Unravelling the non-native low-spin state of the cytochrome *c*–cardiolipin complex: Evidence of the formation of a his-ligated species only. *Biochemistry*, *56*(13), 1887–1898. https://doi.org/10.1021/acs.biochem.6b01281

67. Hu, Y., Benedict, M. A., Ding, L., & Núñez, G. (1999). Role of cytochrome c and dATP/ATP hydrolysis in Apaf-1-mediated caspase-9 activation and apoptosis. *The EMBO journal*, *18*(13), 3586-3595.

68. Yu, T., Wang, X., Purring-Koch, C., Wei, Y., & McLendon, G. (2001). A Mutational Epitope for Cytochrome c Binding to the Apoptosis Protease Activation Factor-1. *Journal Of Biological Chemistry*, *276*(16), 13034-13038. doi: 10.1074/jbc.m009773200

69. Starkov, A. A. (2008). The role of mitochondria in reactive oxygen species metabolism and signaling. *Annals of the New York Academy of Sciences*, *1147*(1), 37-52.

70. Min, L., & Jian-xing, X. (2007). Detoxifying function of cytochrome c against oxygen toxicity. *Mitochondrion*, *7*(1-2), 13-16. doi: 10.1016/j.mito.2006.11.011

71. Shankar, K., & Mehendale, H. (2014). Cytochrome P450. *Encyclopedia Of Toxicology*, 1125-1127. doi: 10.1016/b978-0-12-386454-3.00299-2

72. Williams, R. T. (1971). The metabolism of certain drugs and food chemicals in man. *Annals of the New York Academy of Sciences*, *179*(1), 141-154.

73. Nebert, D. W., & Dalton, T. P. (2006). The role of cytochrome P450 enzymes in endogenous signalling pathways and environmental carcinogenesis. *Nature reviews. Cancer*, *6*(12), 947–960. https://doi.org/10.1038/nrc2015

74. Raunio, H., Kuusisto, M., Juvonen, R., & Pentikainen, O. (2015). Modeling of interactions between xenobiotics and cytochrome P450 (CYP) enzymes. *Frontiers In Pharmacology*, *6*. doi: 10.3389/fphar.2015.00123

75. Dimitrov, D. (2012). Therapeutic Proteins. *Methods In Molecular Biology*, 1-26. doi: 10.1007/978-1-61779-921-1_1

76. Patil, S., Narvekar, A., Puranik, A., Jain, R., & Dandekar, P. (2019). Formulation of therapeutic proteins: strategies for developing oral protein formulations. *Innovative dosage forms: design and development at early stage*, 391-432.

77. Saxena, M., Delgado, Y., Sharma, R., Sharma, S., Guzmán, S., Tinoco, A., & Griebenow, K. (2018). Inducing cell death in vitro in cancer cells by targeted delivery of cytochrome c via a transferrin conjugate. *PLOS ONE*, *13*(4), e0195542. doi: 10.1371/journal.pone.0195542

78. Eleftheriadis, T., Pissas, G., Liakopoulos, V., & Stefanidis, I. (2016). Cytochrome c as a potentially clinical useful marker of mitochondrial and cellular damage. *Frontiers in immunology*, *7*, 279.

79. Valikhani, D., Bolivar, J. M., & Pelletier, J. N. (2021). An Overview of Cytochrome P450 Immobilization Strategies for Drug Metabolism Studies, Biosensing, and Biocatalytic Applications: Challenges and Opportunities. *ACS Catalysis*, *11*(15), 9418-9434.

80. Schneider, E., & Clark, D. S. (2013). Cytochrome P450 (CYP) enzymes and the development of CYP biosensors. *Biosensors and Bioelectronics*, *39*(1), 1-13.

81. Behrendorff, J. B., Huang, W., & Gillam, E. M. (2015). Directed evolution of cytochrome P450 enzymes for biocatalysis: exploiting the catalytic versatility of enzymes with relaxed substrate specificity. *Biochemical Journal*, *467*(1), 1-15.

82. Behrendorff, J., Huang, W., & Gillam, E. (2015). Directed evolution of cytochrome P450 enzymes for biocatalysis: exploiting the catalytic versatility of enzymes with relaxed substrate specificity. *Biochemical Journal*, *467*(1), 1-15. doi: 10.1042/bj20141493

83. Rittle, J., & Green, M. T. (2010). Cytochrome P450 compound I: capture, characterization, and C-H bond activation kinetics. *Science (New York, N.Y.)*, *330*(6006), 933–937. https://doi.org/10.1126/science.1193478

84. Kumar, S., Chen, C. S., Waxman, D. J., & Halpert, J. R. (2005). Directed evolution of mammalian cytochrome P450 2B1: mutations outside of the active site enhance the metabolism of several substrates, including the anticancer prodrugs cyclophosphamide and ifosfamide. *The Journal of biological chemistry*, *280*(20), 19569–19575. https://doi.org/10.1074/jbc.M500158200

85. Kumar, S., Sun, L., Liu, H., Muralidhara, B. K., & Halpert, J. R. (2006). Engineering mammalian cytochrome P450 2B1 by directed evolution for enhanced catalytic tolerance to temperature and dimethyl sulfoxide. *Protein engineering, design & selection : PEDS*, *19*(12), 547–554. https://doi.org/10.1093/protein/gzl042

86. Coelho, P. S., Brustad, E. M., Kannan, A., & Arnold, F. H. (2013). Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. *Science (New York, N.Y.)*, *339*(6117), 307–310. https://doi.org/10.1126/science.1231434

87. Kan, S., Lewis, R., Chen, K., & Arnold, F. (2016). Directed evolution of cytochrome c for carbon–silicon bond formation: Bringing silicon to life. *Science*, *354*(6315), 1048-1051. doi: 10.1126/science.aah6219