Biochemical Engineering Journal 174 (2021) 108096

Contents lists available at ScienceDirect

Biochemical Engineering Journal

journal homepage: www.elsevier.com/locate/bej

Review

Protein engineering design from directed evolution to de novo synthesis

Wei Xiong [a], Bo Liu [a], Yujiao Shen [a], Keju Jing [a,*], Thomas R. Savage [b]

[a] Department of Chemical and Biochemical Engineering and the Key Lab for Synthetic Biotechnology of Xiamen City, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, China

[b] Department of Chemical Engineering and Biotechnology, The University of Cambridge, West Cambridge Site, Cambridge CB3 0AS, United Kingdom

ARTICLE INFO

Keywords:
Protein engineering
Directed evolution
Semi-rational design
Artificial synthesis
Computer-Assisted Design

ABSTRACT

With the development of enzyme engineering technology, green enzyme catalysis is widely poised to replace traditional chemical catalysis enabling high enantioselectivity and yielding chiral chemicals. However, some inherent drawbacks still restrict natural enzymes in these applications e.g., enzyme instability, low activity, limited substrate spectra. Herein, protein engineering was exhibited powerful means to reconstruct natural enzymes for promising enzymes with industrial values. Many enzyme engineering design strategies have emerged in recent years (e.g., directed evolution, semi-rational design, artificial synthesis, Computer-Assisted Design (CAD), artificial intelligence) to generate enzymatic diversity and abundance for needs. This review introduces the development of these protein engineering design strategies. This works primarily focus as upon the application of semi-rational design based on sequence information and protein structure, and artificial intelligence through deep learning algorithms altering the function of natural enzymes to obtain superior functionality and specificity industrial enzymes.

1. Introduction

As known, due to functionality and specificity of applications, enzymes own critical values in food, medicine, biofuels, chemical raw materials, and related fields. In particular, enzymes also play a crucial role in biocatalytic production of chiral chemicals [1–3]. Enzymatic reactions have gradually replaced traditional chemical reactions in catalytic systems due to inherent advantages of high selectivity, environmental friendliness, mild reaction conditions, and simplicity of reaction steps. The combination of biotechnology and computer technology has significantly improved the development of related biological disciplines, resulting in a more comprehensive and systematic understanding of the relationship between protein structure and function. Although enzymes have tremendous potential in biocatalytic synthesis, natural enzymes have inherent defects, such as low enzyme activity, instability, limited substrate range, and low selectivity of reactivity. The two most effective ways to solve these problems are locating novel enzymes with specific functionality in diverse environments or modifying existing enzymes to obtain particular desired functionality. Compared to locating new enzymes, it is relatively straightforward to transform natural enzymes at the molecular level, thus protein engineering came into being.

Protein engineering is a powerful tool to modify the structure of enzymes to guarantee a change of enzyme function for practice. Since the concept of protein proposed by Swedish chemist Jacob Berzelius in 1835 to Frances Arnold's "directed evolution of enzymes" awarded the Nobel Prize for the in 2018, protein engineering has been developed for nearly two centuries. Protein engineering has been widely used to improve the specificity, regioselectivity and stereoselectivity of enzyme protein, as well as the thermostability and solvent tolerance of enzyme protein. With industrial perspectives of green sustainable biocatalysis, gradually increasing number of engineered enzymes have been used to synthesized non-natural clinical medicine intermediates and fine chemical products. The application of artificial intelligence to protein engineering has also been established [4] and many protein design strategies have been conceived throughout this long historical process including directed evolution, rational design, semi-rational design, and computer-assisted design (CAD). The schematics of development process is shown in Fig. 1.

2. Directed evolution

The concept of directed evolution emerged as early as the 20th century. More recently, Frances H. Arnold won the 2018 Nobel Prize in

* Corresponding author.
E-mail address: jkj@xmu.edu.cn (K. Jing).

https://doi.org/10.1016/j.bej.2021.108096
Received 1 February 2021; Received in revised form 9 June 2021; Accepted 13 June 2021
Available online 16 June 2021

chemistry due to significant contribution to widely develop the directed evolution of enzymes. Directed evolution aims to simulate the Darwinian evolution process in a test tube. Abundant and diverse mutants are artificially created through random mutations. These mutants are subsequently screened according to specific metrics with the goal of transforming proteins to obtain the specific functionality [5]. Directed evolution is highly efficient as it enables the introduction of random mutations without understanding the underlying structure and function of the enzyme and the catalytic mechanism. As a matter of fact, statistical analysis upon thousands of items in literature regarding the directed evolution of enzymes could be exhibited as indicated in Fig. 2.

For decades, scientists have adopted directed evolution of transforming natural enzymes to study for instance catalytic mechanisms, metabolic pathways, enzymatic properties of enzymes. The statistical analysis could clearly reveal that directed evolution is a powerful and dominant means of transforming enzymes to obtain superior enzymatic functionality. Directed evolution primarily uses random mutation and recombination techniques to create mutants. These methods include error-prone PCR [6], sequence homology-independent protein recombination (SHIPREC) [7], incremental truncation for the creation of hybrid enzymes (ITCHY) [8], recombining the genes by restriction and relegation [9], staggered extension process (StEP) [10], random-priming recombination (RPR) [11], and DNA shuffling [12].

The directed evolution of random mutation combined with high-throughput screening has significantly improved enzyme activity and widened its application in industrial catalysis. A typical example would be the directed evolution developed in 1993 by Arnold's team. A mutant subtilisin E containing six mutation sites was constructed via error-prone PCR and high-throughput screening method. This mutant exhibited 256-fold higher compared to the wild-type (WT) in 60% DMF solvent [13]. This milestone example demonstrated that random mutation of enzyme's gene sequence could significantly improve the performance of enzyme. In subsequent years, A large number of directed evolutionary strategies for combining gene recombination were gradually developed. One such example is the staggered extension process (StEP) for creating mutant via gene recombination. After two rounds of mutant libraries constructed with this method and screening, a mutant subtilisin E was identified via four point mutation that exhibited a 50-fold increase in half-life compared to the wild type at 65 ℃ [10]. The other example is DNA shuffling for in vitro homologous gene recombination from different organisms. Miho Kikuchi et al. [12] reported an in vitro gene recombination novel method called DNA shuffling. It can use single-stranded DNAs (ssDNAs) as template to improve the efficiency of hybridization formation in family shuffling. Of course, directed evolution requires the construction of statistically significant number of gene libraries and the screening of mutant proteins by high-throughput methods, which inevitably leads to a difficult task for certain target protein. Nevertheless, it is undeniable that directed evolution technology is still a powerful tool in modifying protein function.

3. Semi-rational design

In fact, enzyme engineering by directed evolution still depends on an iterative two-step process. Random mutations first generate the molecular diversity and in vitro recombination, the target protein is screened from a large number of libraries by the high-throughput screening method. This approach requires a large amount of work, material, and financial resources to select the desired phenotype. It is still significantly challenging to complete the comprehensive search of sequence space for an average protein. Many researchers have sought methods beyond traditional directed evolution, and thus simple and efficient semi-rational design and computational protein design (CPD) strategies were born [14].

Semi-rational design strategy is based on the structural information of enzyme protein, combined with the analysis of conserved amino acid sequence. It focuses on specific amino acid positions to create a small and precise mutation library. A significant advantage of this method is convenient to screen target protein from small and accurate mutation samples, thus avoiding the time-consuming and labor-consuming high-throughput assay process. Of course, the limitation of semi-rational design lies in the need for accurate structural information, focusing only on conserved amino acid sites and missing non-conserved sequences related to catalytic function. Currently, the semi-rational design strategy is generally based on two distinct methodologies: sequence-based enzyme redesign and structure-based enzyme redesign, with the specific methods shown in Fig. 3.

The former aims to locate the critical conserved residues binding to substrate through multiple sequence alignments (MSA) of homologous protein sequences. Structural analysis infers the essential interactions with substrate or cofactors for predicting functional hot spots [15–18]. Its potential benefits of enzyme engineering could be elucidated in three following examples. Kataoka and co-workers performed multiple sequence alignment of phenylalanine (PheDH), glutamate (GluDH), and leucine dehydrogenase (LeuDH) to determine key substrate binding sites of LeuDH. The variant A113G/V291S successfully expanded the substrate spectrum, and increased enzyme activity 130-fold compared with the wild type for methionine as the substrate [19]. Nakano et al. identified two critical residues through D-lactate dehydrogenase sequence alignment. As a result, variant T75L/ A234S was constructed and showed 6.8-fold enhanced catalytic efficiency compared with wild type

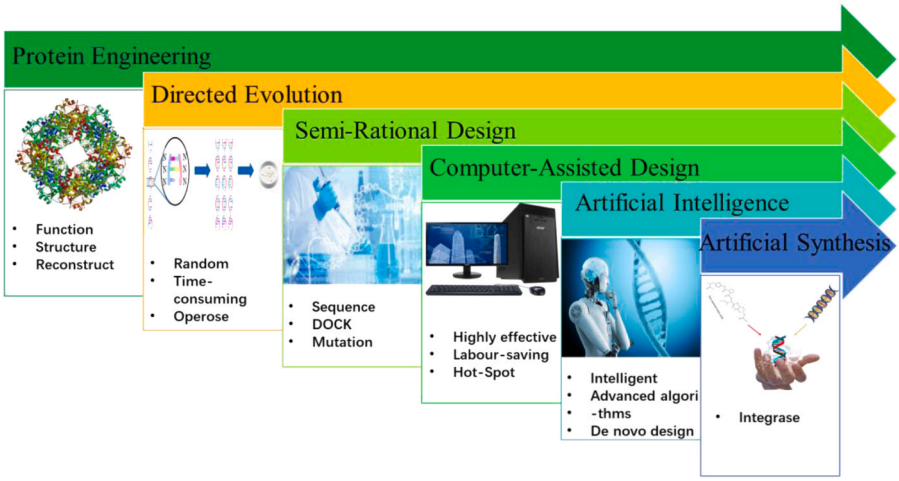

Fig. 1. Design strategy development of protein engineering.

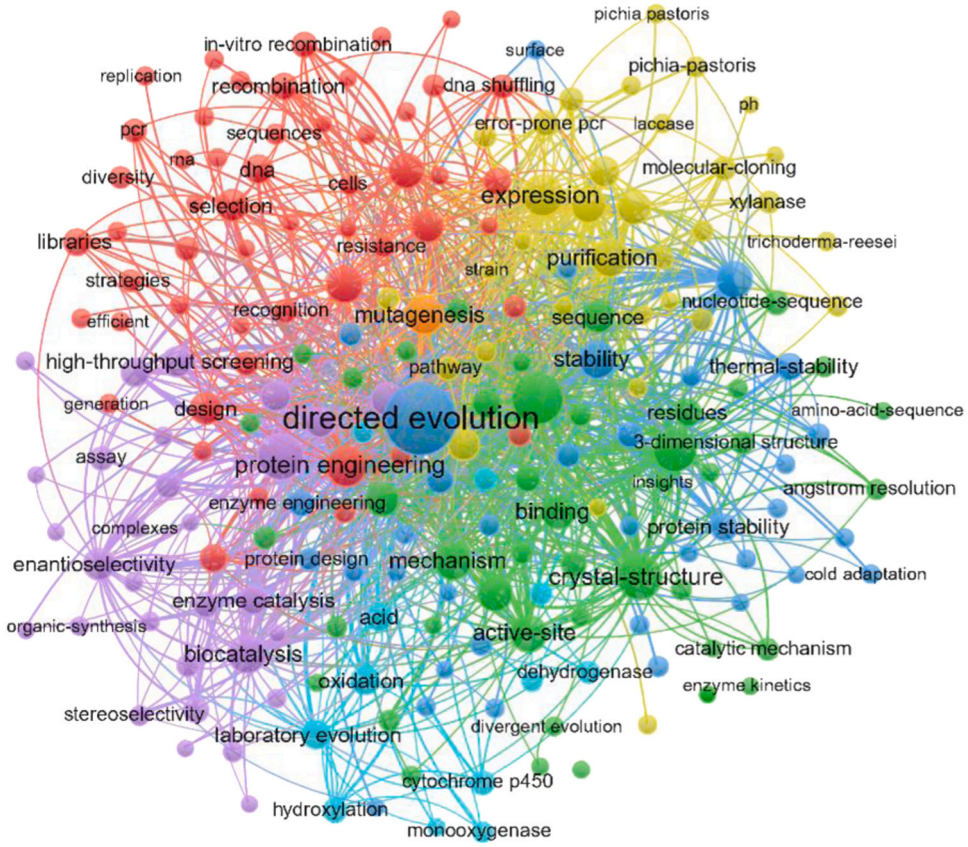

Fig. 2. Knowledge mapping related to directed evolution in the last decade.

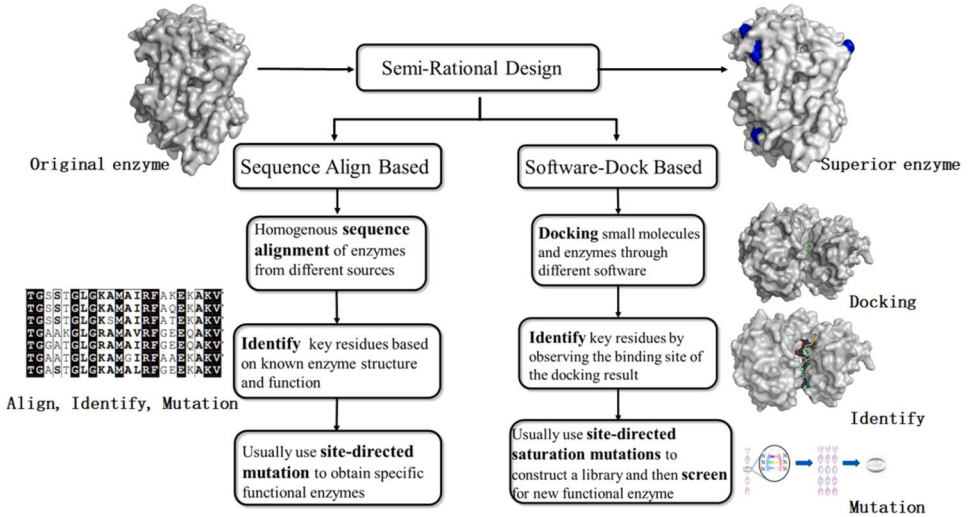

Fig. 3. Schematic diagram of semi-rational design for sequence alignment and structure analysis method.

[20]. Recently, Jianhe Xu developed a cofactor specificity reversal strategy (e.g., small smart library design (CSR-SaSLid)). It aims to change the coenzyme preference of 7β-Hydroxysteroid dehydrogenase from *Ruminococcus torques* (Rt7β-HSDH) by constructing a small and precise mutation library based on homologous sequence alignment. In this strategy, the binding domain of the 7β-HSDH enzyme and coenzyme 2′- phosphate (β2α3) was analyzed through MSA of homologous 7β-HSDH sequence with different nicotinamide coenzyme preferences. The double mutant G39D/T17A was precisely constructed, which reversed NAD(P)H's preference and increased the enzyme activity 223-fold for the biocatalytic synthesis of ursodeoxycholic acid [21]. These examples all suggested the successful application of a semi-rational design strategy based on sequence-based enzyme redesign to pinpoint functional hot spots for improving activity, selectivity, substrate spectral specificities, and changing coenzyme preference.

In general, enzyme function is linked to the sequence of amino acids and closely related to the three-dimensional structure of the enzyme. The second method of semi-rational design is based on three-dimensional structure design using various available docking software with analysis of the substrate, transition state or product binding in the

active site to more effectively pinpoint the functional hot spots that may control the reaction [16,22,23]. For example, Gong and co-workers obtained the multi-docking sites by docking the cyclohexylamine oxidase (CHAO) with D-valine, D-phenylalanine, and L-valine ethyl ester using Autodock software. Through site-directed mutagenesis and screening Y32I/ M226T, the resultant catalytic efficiency of D-valine was 30 times higher than that of the wild-type. The yield and ee value of D-valine reached 95% and 99%, respectively [24].

In addition, Tournier and co-workers used molecular docking and surface analysis methods to study the interaction of leaf-branch compost cutinase (LCC) (protein database (PDB ID: 4EB0)) with its substrate mono (2-hydroxyethyl) terephthalic acid (2-HE $(MHET)_3$). The saturated mutations of four sites were screened from candidates of 15 critical amino acids, and disulfide bonds were introduced at 238 and 283 sites of mutants. Finally, the best variants F243I/D238C/S283C/Y127G (ICCG) and F243W/D238C/S283C/Y127G (WCCG) could be stably screened. After process optimization, ICCG and WCCG mutants were found to reach 90% degradation rates within 9.3 h and 10.5 h, respectively. The mutant enzymes successfully catalyzed the degradation of 1000 kg of waste PET to obtain 863 kg of terephthalic acid. This structure-based design strategy has become an effective method to tailor the biocatalyst of LCC to help solve the issue of plastic pollution, a significant issue that humanity has faced for a long time [25].

4. Computer-Assisted Design (CAD)

With the exponential increase of computing power and the advance of protein design algorithms, computer-assisted design has become an indispensable tool to model protein structures to tailor to specific biocatalysts. Computer-Assisted Design (CAD) generally focuses on the movement of the overall protein structure in different situations. Through the use of fundamental theories of chemistry, physics, and statistics, the changes of energy and interaction between each amino acid in the movement is revealed, as well as the “lock and key” [26] mode formed by the process of substrate entering the binding pocket. With CAD approach, researchers design and predict the changes of free energy of engineering proteins and bind the substrate and coenzyme using methods of simulated annealing, three-dimensional structure searching methods, and quantum mechanical calculation methods represented by Monte Carlo approaches [27]. The engineered protein is thus expressed and analyzed for verification of the design results and finally the target enzyme is tailored to meet the unique needs. A classic example is the development of antibodies against novel coronavirus (COVID-19). In 2020, the novel coronavirus has spread rapidly and seriously threatened human health all over the world. To effectively combat the novel coronavirus, it is of great importance to analyze the protein structure of novel coronavirus and to develop effective inhibitors for health and medical research sectors. Rao and Jiang firstly identified the high-resolution crystal structure of key hydrolase Mpro of COVID-19. Furthermore, they used computer-assisted technology to discover inhibitors and virtual screening of antibodies based on the structure. The screening results found 30 active candidate ingredients, including Indinavir, Lopinavir, and Ritonavir, which have potential therapeutic effects on COVID-19. That is, CAD has inevitably become an essential tool in the development of vaccines. [28–31].

Both CAD and semi-rational design are based on the in-depth understanding of the relationship between protein structure and functionality, enabling proteins to be tailored to obtain specific functions by using a variety of computing tools. Table 2 introduces the most popular protein design engineering software in recent years.

Aside from the synthesis of unnatural compounds, existing natural and engineered enzymes have gradually failed to meet our needs, therefore alterations of the de novo design of proteins have been paid for study. The following table lists some computer-assisted de novo protein design cases over decades (Table 3).

Recently utilized protein computing designing software predominantly includes Rosetta design [32], Transcent [33], Tinker [34], IPRO [35], and Osprey [36]. According to the protein structure and function analysis, multiple peptide Motifs were explored to determine protein function. Therefore, the difficulty and challenges of computer design of proteins with unknown functions would come from the comprehensive search of protein sequence space. For example, attempts to change coenzyme or substrate preferences have been successful in many cases [37–40]. However, efforts to improve the catalytic affinity of natural enzymes to their preferred substrates have been found more difficult, with only a few successful examples [41,42].

Compared to directed evolution, computed protein design can provide a clear transformation plan and significantly reduce the workload of building and screening mutant libraries. At present, many achievements have been made in de novo protein design, enzyme-substrate selectivity, and thermal stability design with some engineered enzymes reaching the industrial application level [43].

5. Machine learning design protein

Data-driven protein design methods based on computational machine learning (ML), in particular deep learning, have been applied to protein structure design with the accumulation of interaction information relating to protein structure and functionality. Recently, Harvard University’s G. Church established the UniRep neural network. UniRep can extract the basic characteristics of protein structures directly from amino acid sequences through deep learning, and accurately predict the impact of amino acid replacement on protein function, saving a significant amount of work, material, and financial resources. Based on the model of Victoria jellyfish green fluorescent protein (GFP), 64,800 variants of the protein were analyzed by UniRep, each with 1–12 mutations. The results show that UniRep can accurately predict how the mutation will change the protein’s brightness [44].

ML also showed powerful application in the field of protein engineering, predicting a beneficial mutant that can improve enzyme

Table 1
Summary of random mutation and recombination techniques.

Methods	Advantages	Disadvantages	Ref.
Error-prone PCR	Rapid construction of any gene sequence libraries by introducing random mutation, accumulating positive mutation by iteration.	The mutation frequency cannot be controlled, negative mutation more than positive mutation.	[6]
DNA shuffling	Quick obtaining positive mutants t through the combination of mutations in the gene groups.	Homology of gene sequences must be more than 70%.	[12]
RPR	Using a small number of single chain DNA or mRNA as the template, reducing the parent components and facilitating to screen.	High homology of gene sequences is necessary.	[11]
StEP	Mixing different-point mutation templates to create full-length recombined gene fragments without obtaining and purifying short fragments.	High homology of gene sequences is necessary.	[10]
ITCHY	Using ITCHY to create multiple crossover mutants, without the homology of DNA sequences.	Creating recombination only between two different parents, and generating low proportion of functional chimeric progenies.	[8]
SHIPREC	An enhanced version of the ITCHY, improving the proportion of functional chimeric progenies and creating chimeric protein combination mutations with low sequence identity or even no sequence identity.	Creating recombination only between two different parents.	[7]

Table 2
The most commonly used software in protein design engineering.

Software	Description	Website	Ref.
Discovery Studio	Used in protein structure and function research, also used in drug discovery, with docking, simulation and other modules	http://softwaretopic.informer.com/discovery-studio-4-0-server/	[62]
AutoDock Vina	Used a simpler scoring function to speed up the search	https://www.cgl.ucsf.edu/chimera/data/downloads/1.7/docs/ContributedSoftware/vina/vina.html	[63]
AutoDock	Free, mainly used for docking of large molecule and small molecule ligands	http://autodock.scripps.edu/	[64]
MOE-Dock	Mass molecular attribute calculation and simulation, widely used in drug molecular design	https://www.chemcomp.com/Products.htm	[65]
Rosetta	Widely used in de novo protein design	https://www.rosettacommons.org/	[66]
NAMD	Free, parallel computing efficiency is very high	http://www.ks.uiuc.edu/Research/namd/	[67]
Amber	Commercial, commonly used in molecular dynamics simulations to calculate energy changes	http://ambermd.org/	[68]
GROMACS	Free, commonly used in molecular dynamics simulations to calculate energy changes	http://www.gromacs.org/	[69]
Sybyl	Commercial, commonly used in molecular dynamics simulations to calculate energy changes	https://en.freedownloadmanager.org/Windows-PC/SYBYL-X.html	[70]
YASARA	Commercial, commonly used in molecular dynamics simulations to calculate energy changes	http://www.yasara.org/	[71]
Consensus Finder	Constructing consensus sequence by consensus analysis	http://kazlab.umn.edu/	[72]
I-mutate 2.0	Predicting mutation sites that can affect thermal stability	http://folding.biofold.org/i-mutant/i-mutant2.0.html	[73]
FoldX	Evaluating the effect of amino acid residues on protein stability by analyzing changes in free energy	http://foldx.crg.es	[74]
POVME	Calculating binding pocket volume	https://github.com/POVME/POVME	[75]
MSDsite	Prediction of substrate-protein binding sites	https://www.ebi.ac.uk/pdbe-site/pdbemotif/	[76]
POCASA	Predicting protein binding pockets	http://g6altair.sci.hokudai.ac.jp/g6/Research/POCASA_e.html	[77]
CaverDock	Analysis of small molecular transport processes in proteins	https://loschmidt.chemi.muni.cz/caverdock/	[78]

activity and persistent stability, which was successfully verified [45]. Arnold used GB1 protein binding data to build a model and obtained an optimal mutant through the use of machine learning. Subsequently, the method was applied to the directed evolution of enzyme enantioselectivity. The phenyl-dimethyl silane reaction with 2-diazopropanoic acid ethyl ester (Me EDA) was catalyzed by Rma NOD enzyme. It successfully predicted and verified that a mutant enzyme capable of producing only S or R enantiomers was obtained. The ee values of the two products were 93% and 79%, respectively [46]. With the development of advanced computer technologies, low-cost data-driven protein design will gradually replace protein engineering and resolve the current defects.

Table 3
Case study of computer-aided protein de novo design.

Software	Description	Ref.
Rosetta	A new protein called Neo-2/15 was created based on simulating the role of key immunomodulatory interleukins IL-2 and IL-5, which can be used for anti-cancer treatment without toxicity.	[79]
Rosetta	For the first time created a transmembrane protein not found in nature.	[80]
Rosetta	De novo design of a β-barrel round protein that can bind to the fluorescent compound of DFHBI, which can bind and interact with small molecule targets with high precision.	[81]
Rosetta	Designed and created self-assembled spiral protein filaments from scratch for the first time.	[82]
Rosetta	Inventively designed a protein that responds to acids, using pH-dependent destruction of lipid membranes, is expected to provide new strategies for delivering drug molecules into cells.	[83]
Rosetta	Designed Retro-Aldol reaction enzyme, which catalyzed the C-C bond cleavage reaction rate is 4 orders of magnitude and multiple turnovers higher than that of the enzyme-free reaction system.	[84]
Rosetta	The Kemp elimination enzyme was designed, and its catalytic elimination rate was 5 orders of magnitude higher than that of the enzyme-free reaction system , then combined with directed evolution, its K_{cat}/k_m value increased by 200 times.	[85]
Rosetta	The Diels-Alder reaction enzyme is designed to catalyze the bimolecular Diels-Alder reaction, and the enantioselectivity of the product is up to 97%.	[86]
SCADS	Designed for the first time a single-chain protein containing four helical fragments that can be folded to form a four-helix bundle.	[87]
Rosetta	Design small molecule binding sites and use it to produce protein binder steroid digoxigenin.	[88]
Rosetta	Design protein function switches controlled by inducible conformational changes.	[89]
TopoBuilder/ Rosetta	Develop strategies for assembling protein topologies without templates to adapt to functional motifs, thereby designing proteins with complex structural motifs from scratch.	[90]
Rosetta	The heterodimer module designed from scratch is used as a component, and the logic gate circuit is constructed according to the needs through different combinations.	[91]
Rosetta	Designed a novel protein PS1 that can bind to a highly electron-deficient unnatural porphyrin at 100 ℃.	[92]
Rosetta	Generic design method of protein homo-oligomers based on hydrogen-bond networks module.	[93]

6. Introduction of unnatural amino acids in protein engineering

With a deeper understanding of active unnatural amino acids, a new enzyme design strategy introducing unnatural amino acids as the catalytic amino acid residue has begun to be applied in site-directed mutagenesis for enzyme redesign [47–49]. The enzyme redesigned by this new strategy imitates the natural enzyme and promotes the target reaction [50,51]. Clemens Mayer et al. introduced the unnatural amino acid PAF (p-Aminophenylalanine) at the 15th position of the resistance regulatory factor of Lactococcus lactis (LmrR) which is called LmrR-PAF. The variant LmrR_PAF_RMH (containing the A92R/N19M/F93H mutation) was then obtained through multiple rounds of mutation of the residue binding to the substrate pocket. The K_{cat} of LmrR_ PAF_RMH was found to be 91-fold over the wild type [52].

Of course, enzyme's stability is critical to influence its industrial application [53], as a highly stable enzyme can accept more extensive mutations while retaining its natural structure thus showing great potential for directed evolution [54]. High enzyme stability is the synergistic result of non-covalent interactions, covalent interactions, salt bridges, aromatic interactions, and hydrophobic interactions [55]. Homoserine o-succinyl-transferase (MetA) is an important protein related to methionine synthesis. The instability of MetA causes the growth of *Escherichia coli* to be hindered above 40 ℃. Jack C.L

introduced a pisothiocyanate phenylalanine (pNCSF) containing isothiocyanate group at the interface of MetA dimer, which can establish a thiourea connection between residues at high temperatures. Through temperature and pressure screening, the author screened F264pNCSF mutants, which kept aspects of the catalytic activity of wild-type enzymes whilst resulting in a significant increase in melting point compared with the wild-type enzyme. Meanwhile, the connection between the unnatural amino acids and proline at the second position in the dimer was established by mass spectrometry [56].

7. Artificial synthesis

With the development of biotechnology, a technical term coined 'artificial synthetase' began to emerge. As the "energy factory" of cells, mitochondria are indispensable and important organs. The mitochondria dysfunction caused by electron accumulation due to the interruption of the electron transfer chain (ETC), can lead to more than 300 rare genetic diseases, such as Leigh syndrome (a serious nervous system disease that may occur in early infancy) and MELAS (which can lead to muscle weakness, diabetes and strokes, usually before the age of 40). Excess electrons eventually spill into the blood circulatory system in the form of lactate, which can be used as a marker of intracellular diseases. Patgiri A et al. [57] created an artificial enzyme (known as LOXCAT) to alleviate intracellular redox imbalance by directly cycling redox metabolites. The new synthetase was constructed base on the crystal structure of two kinds of lactate oxidase (LOX) and catalase (CAT) [58, 59], which derived from bacterial proteins and the biochemical data [60,61]. They added LOXCAT to a human cell culture medium with mitochondrial defects and found that the artificial enzyme LOXCAT converts lactate into pyruvate, which enters the cell and absorbs electrons, thereby reducing the build-up. Pyruvate is then converted into lactic acid, which is released from cells. LOXCAT converts this lactic acid into pyruvate again, thus forming a cycle. This emergency of artificial enzyme will provide a new dawn for the clinical treatment of some mitochondrial dysfunction diseases.

8. Summary and prospects

Apparently, green and environmentally sustainable renewable energy technologies were developed to replace traditional industrial catalysis. Of course, green enzyme catalysis technology is a primary choice of green technology. However, due to limitations of the existing scientific and technological conditions and environment, there are still many unexplored enzymes that have not been implemented. That is, the functions of existing enzymes cannot currently meet the needs of modern industrial catalysis. Therefore, engineering enzyme transformations has become an essential means for us to obtain specific enzyme functions. This review introduced the successful transformation of natural enzymes within protein engineering and the application of data-driven artificial intelligence to practice in life science. With AI technology development, our vision has gradually changed from the known protein sequence-structure information. It is a significant leap in life sciences to reconstruct the unknown protein sequence-structure information through engineering. At present, the application of artificial intelligence in life science is at an early development stage. It is expected that proteins based on AI technology will be able to meet future industrial needs and resolve current technical bottlenecks in industrial circle. With the advancement of science and technology, the intersection of these disciplines will inevitably bring us more protein engineering technologies and provide us a clearer understanding of the structure and function of enzymes. The application of artificial intelligence in the process of enzyme transformation is the best example. The enzyme function could be significantly manipulated according to our needs by using continuous development of novel protein engineering technologies, conforming to the trend of green catalysis.

CRediT authorship contribution statement

BL and YJS collected and collated the literature, TRS revised the manuscript, WX and KJJ wrote the manuscript. All the authors read and approved the final manuscript.

Declaration of Competing Interest

We have no conflicts of interest in this work.

Acknowledgments

The authors are grateful for the financial support received for the National Natural Science Foundation of China (No. 21776232 and No. 21978244). In addition, we would like to thank Professor Bor-Yann Chen for revising the grammar of the article.

References

[1] O. Kirk, T.V. Borchert, C.C. Fuglsang, Industrial enzyme applications, Curr. Opin. Biotechnol. 13 (2002) 345–351.
[2] R.O. de Souza, L.S. Miranda, M.L. Gutarra, Recent advances in applied biocatalysis and biotechnology, Biotechnol. Adv. 33 (2015) 371.
[3] Y.P. Xue, C.H. Cao, Y.G. Zheng, Enzymatic asymmetric synthesis of chiral amino acids, Chem. Soc. Rev. 47 (2018) 1516–1561.
[4] R. Chowdhury, C.D. Maranas, From directed evolution to computational enzyme engineering—A review, AIChE J. 66 (2019) 1–17.
[5] F.H. Arnold, Directed evolution: bringing new chemistry to life, Angew. Chem. Int. Ed. Engl. 57 (2018) 4143–4148.
[6] R.C. Caldwell, G.F. Joyce, Randomization of genes by PCR mutagenesis, PCR Methods Appl. 2 (1992) 28–33.
[7] V. Sieber, C.A. Martinez, F.H. Arnold, Libraries of hybrid proteins from distantly related sequence, Nat. Biotechnol. 19 (2001) 456–460.
[8] W.S. Hu, E.H. Bowman, K.A. Delviks, Homologous recombination occurs in a distinct retroviral subpopulation and exhibits high negative interference, J. Virol. 71 (1997) 6028–6036.
[9] J.C. Moore, F.H. Arnold, Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents, Nat. Biotechnol. 14 (1996) 458–467.
[10] Z. Shao, H. Zhao, L. Giver, A. Joseph, F.H. Arnold, Molecular evolution by staggered extension process (StEP)in vitro recombination, Nat. Biotechnol. 16 (1998) 258–261.
[11] Z. Shao, H. Zhao, L. Giver, F.H. Arnold, Random-priming in vitro recombinationan effective tool for directed evolution, Nucleic Acids Res. 26 (1998) 681–685.
[12] M. Kikuchi, K. Ohnishi, S. Harayama, An effective family shuffling method using single-stranded DNA, Gene 243 (2000) 133–137.
[13] K. Chen, F.H. Arnold, Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide, Proc. Natl. Acad. Sci. U.S.A. 90 (1993) 5618–5622.
[14] R.A. Sheldon, P.C. Pereira, Biocatalysis engineering: the big picture, Chem. Soc. Rev. 46 (2017) 2678–2691.
[15] H. Jochens, U.T. Bornscheuer, Natural diversity to guide focused directed evolution, Chembiochem 11 (2010) 1861–1866.
[16] M.C. Ebert, J.N. Pelletier, Computational tools for enzyme improvement: why everyone can - and should - use them, Curr. Opin. Chem. Biol. 37 (2017) 89–96.
[17] A. Beier, S. Bordewick, M. Genz, S. Schmidt, C. Peters, H.J. Joosten, U. T. Bornscheuer, Switch in cofactor specificity of a Baeyer-Villiger monooxygenase, Chembiochem 17 (2016) 2312–2315.
[18] J. Bendl, J. Stourac, E. Sebestova, O. Vavra, M. Musil, J. Brezovsky, J. Damborsky, HotSpot wizard 2.0: automated design of site-specific mutations and smart libraries in protein engineering, Nucleic Acids Res. 44 (2016) W479–W487.
[19] K. Kataoka, K. Tanizawa, Alteration of substrate specificity of leucine dehydrogenase by site-directed mutagenesis, J. Mol. Catal. B: Enzym. 23 (2003) 299–309.
[20] K. Nakano, S. Sawada, R. Yamada, T. Mimitsuka, H. Ogino, Enhancement of the catalytic activity of d-lactate dehydrogenase from Sporolactobacillus laevolacticus by site-directed mutagenesis, Biochem. Eng. J. 133 (2018) 214–218.
[21] Z. You, Q. Chen, S.C. Shi, J. Pan, X.L. Qian, C.X. Li, J.H. Xu, Switching cofactor dependence of 7β-hydroxysteroid dehydrogenase for cost-effective production of ursodeoxycholic acid, ACS Catal. 9 (2018) 466–473.
[22] S. Lutz, Beyond directed evolution-semi-rational protein engineering and design, Curr. Opin. Biotechnol. 21 (2010) 734–743.
[23] H.J. Wijma, R.J. Floor, S. Bjelic, S.J. Marrink, D. Baker, D.B. Janssen, Enantioselective enzymes by computational design and in silico screening, Angew. Chem. Int. Ed. Engl. 54 (2015) 3726–3730.
[24] R. Gong, P. Yao, X. Chen, J. Feng, Q. Wu, P.C.K. Lau, D. Zhu, Novel access to D-valine synthesis by improved variants of bacterial cyclohexylamine oxidase, ChemCatChem 10 (2018) 387–390.
[25] V. Tournier, C.M. Topham, A. Gilles, B. David, C. Folgoas, E. Kamionka, M. L. Desrousseaux, H. Texier, S. Gavalda, M. Cot, E. Guemard, M. Dalibey, J. Nomme, G. Cioci, S. Barbe, M. Chateau, I. Andre, S. Duquesne, A. Marty, An engineered PET

depolymerase to break down and recycle plastic bottles, Nature 580 (2020) 216–219.
[26] U.d.G.d. Alkohole, Mittheilungen, (1893) 2400–2412.
[27] A. Romero-Rivera, M. Garcia-Borras, S. Osuna, Computational tools for the evaluation of laboratory-engineered biocatalysts, Chem. Commun. 53 (2016) 284–297.
[28] T. Park, S.Y. Lee, S. Kim, M.J. Kim, H.G. Kim, S. Jun, S.I. Kim, B.T. Kim, E.C. Park, D. Park, Spike protein binding prediction with neutralizing antibodies of SARS-CoV-2, bioRxiv, (2020).
[29] S.F. Ahmed, A.A. Quadeer, M.R. McKay, Preliminary identification of potential vaccine targets for the COVID-19 Coronavirus (SARS-CoV-2) based on SARS-CoV immunological studies, Viruses 12 (2020) 254–268.
[30] W. Dai, B. Zhang, X.M. Jiang, H. Su, J. Li, Y. Zhao, Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease, Science 368 (2020) 1331–1335.
[31] Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L.W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao, H. Yang, Structure of Mpro from COVID-19 Virus and Discovery of its Inhibitors, bioRxiv, (2020).
[32] Y. Liu, B. Kuhlman, RosettaDesign server for protein design, Nucleic Acids Res. 34 (2006) W235–W238.
[33] A. Fischer, N. Enkler, G. Neudert, M. Bocola, R. Sterner, R. Merkl, TransCent: computational enzyme design by transferring active sites and considering constraints relevant for catalysis, BMC Bioinform. 10 (2009) 54–69.
[34] J.A. Rackers, Z. Wang, C. Lu, M.L. Laury, L. Lagardere, M.J. Schnieders, J. P. Piquemal, P. Ren, J.W. Ponder, Tinker 8: software tools for molecular design, J. Chem. Theory Comput. 14 (2018) 5273–5289.
[35] R.J. Pantazes, M.J. Grisewood, T. Li, N.P. Gifford, C.D. Maranas, The iterative protein redesign and optimization (IPRO) suite of programs, J. Comput. Chem. 36 (2015) 251–263.
[36] P. Gainza, K.E. Roberts, I. Georgiev, R.H. Lilien, D.A. Keedy, C.Y. Chen, F. Reza, A. C. Anderson, D.C. Richardson, J.S. Richardson, B.R. Donald, OSPREY: protein design with ensembles, flexibility, and provable algorithms, Methods Enzym. 523 (2013) 87–107.
[37] A. Lehmann, J.G. Saven, Computational design of four-helix bund leproteins that bind nonbiological cofactors, Biotechnol. Prog. 24 (2008) 74–79.
[38] G.A. Khoury, H. Fazelinia, J.W. Chin, R.J. Pantazes, P.C. Cirino, C.D. Maranas, Computational design of Candida boidinii xylose reductase for altered cofactor specificity, Protein Sci. 18 (2009) 2125–2138.
[39] D. Cui, L. Zhang, S. Jiang, Z. Yao, B. Gao, J. Lin, Y.A. Yuan, D. Wei, A computational strategy for altering an enzyme in its cofactor preference to NAD (H) and/or NADP(H), FEBS J. 282 (2015) 2339–2351.
[40] F. Garcia-Guevara, I. Bravo, C. Martinez-Anaya, L. Segovia, Cofactor specificity switch in Shikimate dehydrogenase by rational design and consensus engineering, Protein Eng. Des. Sel. 30 (2017) 533–541.
[41] C.Y. Chen, I. Georgiev, A.C. Anderson, Computational structure-based redesign of enzyme activity, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 3764–3769.
[42] M.J. Grisewood, N.J. Hernandez Lozada, J.B. Thoden, N.P. Gifford, D. Mendez-Perez, H.A. Schoenberger, M.F. Allan, M.E. Floy, R.Y. Lai, H.M. Holden, C. D. Maranas, Computational redesign of Acyl-ACP thioesterase with improved selectivity toward medium-chain-length fatty acids, ACS Catal. 7 (2017) 3837–3849.
[43] R. Li, H.J. Wijma, L. Song, Y. Cui, M. Otzen, Y. Tian, J. Du, T. Li, D. Niu, Y. Chen, J. Feng, J. Han, H. Chen, Y. Tao, D.B. Janssen, B. Wu, Computational redesign of enzymes for regio- and enantioselective hydroamination, Nat. Chem. Biol. 14 (2018) 664–670.
[44] E.C. Alley, G. Khimulya, M. AlQuraishi, G.M. Church, Unified rational protein engineering with sequence-based deep representation learning, Nat. Methods 16 (2019) 1315–1322.
[45] R.J. Fox, S.C. Davis, E.C. Mundorff, L.M. Newman, V. Gavrilovic, S.K. Ma, L. M. Chung, C. Ching, S. Tam, S. Muley, J. Grate, J. Gruber, J.C. Whitman, R. A. Sheldon, G.W. Huisman, Improving catalytic function by ProSAR-driven enzyme evolution, Nat. Biotechnol. 25 (2007) 338–344.
[46] Z. Wu, S.B.J. Kan, R.D. Lewis, B.J. Wittmann, F.H. Arnold, Machine learning-assisted directed protein evolution with combinatorial libraries, Proc. Natl. Acad. Sci. U.S.A. 116 (2019) 8852–8858.
[47] M. Pott, T. Hayashi, T. Mori, P.R.E. Mittl, A.P. Green, D. Hilvert, A. Noncanonical, Proximal heme ligand affords an efficient peroxidase in a globin fold, J. Am. Chem. Soc. 140 (2018) 1535–1543.
[48] T. Hayashi, D. Hilvert, A.P. Green, Engineered metalloenzymes with non-canonical coordination environments, Chemistry 24 (2018) 11821–11830.
[49] H. Yang, A.M. Swartz, H.J. Park, P. Srivastava, K. Ellis-Guardiola, D.M. Upp, G. Lee, K. Belsare, Y. Gu, C. Zhang, R.E. Moellering, J.C. Lewis, Evolving artificial metalloenzymes via random mutagenesis, Nat. Chem. 10 (2018) 318–324.
[50] N.M. Okeley, W.A. van der Donk, Novel cofactors via post-translational modifications of enzyme active sites, Chem. Biol. 7 (2000) 159–171.
[51] M.J. Appel, C.R. Bertozzi, Formylglycine, a post-translationally generated residue with unique catalytic capabilities and biotechnology applications, ACS Chem. Biol. 10 (2015) 72–84.
[52] C. Mayer, C. Dulson, E. Reddem, A.W.H. Thunnissen, G. Roelfes, Directed evolution of a designer enzyme featuring an unnatural catalytic amino acid, Angew. Chem. Int. Ed. Engl. 58 (2019) 2083–2087.
[53] A.S. Bommarius, J.K. Blum, M.J. Abrahamson, Status of protein engineering for biocatalysts: how to design an industrially useful biocatalyst, Curr. Opin. Chem. Biol. 15 (2011) 194–200.
[54] J.D. Bloom, S.T. bLabthavikul, C.R. Otey, Protein stability promotes evolvability, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 5869–5874.
[55] Z. Xu, Y.K. Cen, S.P. Zou, Y.P. Xue, Y.G. Zheng, Recent advances in the improvement of enzyme thermostability by structure modification, Crit. Rev. Biotechnol. 40 (2020) 83–98.
[56] J.C. Li, F. Nastertorabi, W. Xuan, G.W. Han, R.C. Stevens, P.G. Schultz, A. Single, Reactive noncanonical amino acid is able to dramatically stabilize protein structure, ACS Chem. Biol. 14 (2019) 1150–1153.
[57] A. Patgiri, O.S. Skinner, Y. Miyazaki, G. Schleifer, E. Marutani, H. Shah, R. Sharma, R.P. Goodman, T.L. To, X. Robert Bao, F. Ichinose, W.M. Zapol, V.K. Mootha, An engineered enzyme that targets circulating lactate to alleviate intracellular NADH: NAD(+) imbalance, Nat. Biotechnol. 38 (2020) 309–313.
[58] S.J. Li, Y. Umena, K. Yorita, T. Matsuoka, A. Kita, K. Fukui, Y. Morimoto, Crystallographic study on the interaction of L-lactate oxidase with pyruvate at 1.9 Angstrom resolution, Biochem. Biophys. Res. Commun. 358 (2007) 1002–1007.
[59] W. Melik-Adamyan, J. Bravo, X. Carpena, J. Switala, I. Fita, P.C. Loewen, Substrate flow in catalases deduced from the crystal structures of active site variants of HPII from Escherichia coli, Proteins 44 (2001) 270–281.
[60] I. Taurino, R. Reiss, M. Richter, M. Fairhead, L. Thöny-Meyer, G. De Micheli, S. Carrara, Comparative study of three lactate oxidases from Aerococcus viridans for biosensing applications, Electrochim. Acta 93 (2013) 72–79.
[61] C. Andre, S.W. Kim, X.H. Yu, J. Shanklin, Fusing catalase to an alkane-producing enzyme maintains enzymatic activity by converting the inhibitory byproduct H_2O_2 to the cosubstrate O_2, Proc. Natl. Acad. Sci. U.S.A. 110 (2013) 3191–3196.
[62] V. Temml, V. Kaserer, Z. Kutil, P. Landa, D. Schuster, Pharmacophore modeling for COX-1 and -2 inhibitors with ligandscout in comparison to discovery studio, Future Med. Chem. 6 (2014) 1869–1881.
[63] O. Trott, A.J. Olson, AutoDock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455–461.
[64] F. Osterberg, G.M. Morris, M.F. Sanner, A.J. Olson, D.S. Goodsell, Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock, Proteins 46 (2002) 34–40.
[65] C.R. Corbeil, C.I. Williams, P. Labute, Variability in docking success rates due to dataset preparation, J. Comput. Aided Mol. Des. 26 (2012) 775–786.
[66] A. Leaver-Fay, M. Tyka, S.M. Lewis, O.F. Lange, J. Thompson, R. Jacak, K. Kaufman, P.D. Renfrew, C.A. Smith, W. Sheffler, I.W. Davis, S. Cooper, A. Treuille, D.J. Mandell, F. Richter, Y.E. Ban, S.J. Fleishman, J.E. Corn, D.E. Kim, S. Lyskov, M. Berrondo, S. Mentzer, Z. Popovic, J.J. Havranek, J. Karanicolas, R. Das, J. Meiler, T. Kortemme, J.J. Gray, B. Kuhlman, D. Baker, P. Bradley, ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules, Methods Enzymol. 487 (2011) 545–574.
[67] I. Kass, C.F. Reboul, A.M. Buckle, Computational methods for studying serpin conformational change and structural plasticity, Methods Enzymol. 501 (2011) 295–323.
[68] R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the Amber biomolecular simulation package, Wiley Interdiscip. Rev. Comput. Mol. Sci. 3 (2013) 198–210.
[69] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1–2 (2015) 19–25.
[70] K.W. Hahn, W.A. Klis, J.M. Stewart, SYBYL® software:driving biopolymer breakthroughs!, Science 248 (1990) 1544–1547.
[71] J.M. Walker, YASARA: A Tool to Obtain Structural Guidance in Biocatalytic Investigations, (2018).
[72] M. Lehmann, L. Pasamontes, S.F. Lassen, M. Wyss, The consensus concept for thermostability engineering of proteins, Biochim. Biophys. Acta 1543 (2000) 408–415.
[73] E. Capriotti, P. Fariselli, R. Calabrese, R. Casadio, Predicting protein stability changes from sequences using support vector machines, Bioinformatics 2 (2005) ii54-58.
[74] R. Guerois, J.E. Nielsen, L. Serrano, Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations, J. Mol. Biol. 320 (2002) 369–387.
[75] J.D. Durrant, L. Votapka, J. Sorensen, R.E. Amaro, POVME 2.0: an enhanced tool for determining pocket shape and volume characteristics, J. Chem. Theory Comput. 10 (2014) 5047–5056.
[76] A. Golovin, D. Dimitropoulos, T. Oldfield, A. Rachedi, K. Henrick, MSDsite: a database search and retrieval system for the analysis and viewing of bound ligands and active sites, Proteins 58 (2005) 190–199.
[77] D.G. Levitt, L.J. Banaszak, POCKET:a computer graphics method for identifying and displaying protein cavities and their surrounding amino acids, J. Mol. Graph. 10 (1992) 229–234.
[78] O. Vavra, J. Filipovic, J. Plhak, D. Bednar, S.M. Marques, J. Brezovsky, J. Stourac, L. Matyska, J. Damborsky, CaverDock: a molecular docking-based tool to analyse ligand transport through protein tunnels and channels, Bioinformatics 35 (2019) 4986–4993.
[79] D.A. Silva, S. Yu, U.Y. Ulge, J.B. Spangler, K.M. Jude, C. Labao-Almeida, L.R. Ali, A. Quijano-Rubio, M. Ruterbusch, I. Leung, T. Biary, S.J. Crowley, E. Marcos, C. D. Walkey, B.D. Weitzner, F. Pardo-Avila, J. Castellanos, L. Carter, L. Stewart, S. R. Riddell, M. Pepper, G.J.L. Bernardes, M. Dougan, K.C. Garcia, D. Baker, De novo design of potent and selective mimics of IL-2 and IL-15, Nature 565 (2019) 186–191.
[80] L. Peilong, M. Duyoung, F. DiMaio, K.Y. Wei, M.D. Vahey, S.E. Boyken, Z. Chen, J. A. Fallas, G. Ueda, W. Sheffler, V.K. Mulligan, J.U. Bowie, D. Baker, Accurate

computational design of multipass transmembrane proteins, Science 359 (2018) 1042–1046.
[81] J. Dou, A.A. Vorobieva, W. Sheffler, L.A. Doyle, H. Park, M.J. Bick, B. Mao, G. W. Foight, M.Y. Lee, L.A. Gagnon, L. Carter, B. Sankaran, S. Ovchinnikov, E. Marcos, P.S. Huang, J.C. Vaughan, B.L. Stoddard, D. Baker, De novo design of a fluorescence-activating beta-barrel, Nature 561 (2018) 485–491.
[82] H. Shen, J.A. Fallas, E. Lynch, W. Sheffler, B. Parry, N. Jannetty, J. Decarreau, M. Wagenbach, De novo design of self-assembling helical protein filaments, Science 362 (2018) 705–709.
[83] S.E. Boyken, M.A. Benhaim, F. Busch, M. Jia, M.J. Bick, H. Choi, J.C. Klima, Z. Chen, C. Walkey, A. Mileant, A. Sahasrabuddhe, K.Y. Wei, E.A. Hodge, S. Byron, B. Sankaran, N.P. King, D. Baker, De novo design of tunable, pH-driven conformational changes, Science 364 (2019) 658–664.
[84] L. Jiang, E.A. Althoff, F.R. Clemente, L. Doyle, D. Röthlisberger, J.L. Gallaher, J. L. Betker, F. Tanaka, C.F. Barbas III, D. Hilvert, K.N. Houk, D. Baker, De Novo computational design of retro-aldol enzymes, Science 319 (2008) 1387–1391.
[85] D. Rothlisberger, O. Khersonsky, A.M. Wollacott, L. Jiang, J. DeChancie, J. Betker, J.L. Gallaher, E.A. Althoff, A. Zanghellini, O. Dym, S. Albeck, K.N. Houk, D. S. Tawfik, D. Baker, Kemp elimination catalysts by computational enzyme design, Nature 453 (2008) 190–195.
[86] J.B. Siegel, A. Zanghellini, H.M. Lovick, G. Kiss, A.R. Lambert, J.L. Gallaher, D. Hilvert, B.L. Stoddard, K.N. Houk, F.E. Michael, D. Baker, Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction, Science 329 (2010) 309–313.
[87] H.C. Fry, A. Lehmann, L.E. Sinks, I. Asselberghs, A. Tronin, V. Krishnan, J.K. Blasie, K. Clays, W.F. DeGrado, J.G. Saven, M.J. Therien, Computational de novo design and characterization of a protein that selectively binds a highly hyperpolarizable abiological chromophore, J. Am. Chem. Soc. 135 (2013) 13914–13926.
[88] C.E. Tinberg, S.D. Khare, J. Dou, L. Doyle, J.W. Nelson, A. Schena, W. Jankowski, C.G. Kalodimos, K. Johnsson, B.L. Stoddard, D. Baker, Computational design of ligand-binding proteins with high affinity and selectivity, Nature 501 (2013) 212–216.
[89] R.A. Langan, S.E. Boyken, A.H. Ng, J.A. Samson, G. Dods, A.M. Westbrook, T. H. Nguyen, M.J. Lajoie, Z. Chen, S. Berger, V.K. Mulligan, J.E. Dueber, W.R. P. Novak, H. El-Samad, D. Baker, De novo design of bioactive protein switches, Nature 572 (2019) 205–210.
[90] F. Sesterhenn, C. Yang, J. Bonet, J.T. Cramer, X. Wen, Y. Wang, C.I. Chiang, L. A. Abriata, I. Kucharska, G. Castoro, S.S. Vollers, M. Galloux, E. Dheilly, S. Rosset, P. Corthesy, S. Georgeon, M. Villard, C.A. Richard, D. Descamps, T. Delgado, E. Oricchio, M.A. Rameix-Welti, V. Mas, S. Ervin, J.F. Eleouet, S. Riffault, J. T. Bates, J.P. Julien, Y. Li, T. Jardetzky, T. Krey, B.E. Correia, De novo protein design enables the precise induction of RSV-neutralizing antibodies, Science 368 (2020) 730–739.
[91] Z. Chen, R.D. Kibler, A. Hunt, F. Busch, J. Pearl, M. Jia, Za.L. VanAernum, B.I. M. Wicky, G. Dods, H. Liao, M.S. Wilken, C. Ciarlo, S. Green, J. Stamatoyannopoulos, V.H. Wysocki, M.C. Jewett, S.E. Boyken, D. Baker, De novo design of protein logic gates, Science 368 (2020) 78–84.
[92] N.F. Polizzi, Y. Wu, T. Lemmin, A.M. Maxwell, S.Q. Zhang, J. Rawson, D. N. Beratan, M.J. Therien, W.F. DeGrado, De novo design of a hyperstable non-natural protein-ligand complex with sub-A accuracy, Nat. Chem. 9 (2017) 1157–1164.
[93] S.E. Boyken, Z. Chen, B. Groves, R.A. Langan, G. Oberdorfer, A. Ford, J.M. Gilmore, F. DiMaio, G. Seelig, P.H. Zwart, D. Baker, De novo design of protein homo-oligomers with modular hydrogen-bond network–mediated specificity, Science 352 (2016) 680–687.