catalysts

MDPI

Review

Protein Engineering for Industrial Biocatalysis: Principles, Approaches, and Lessons from Engineered PETases

Konstantinos Grigorakis, Christina Ferousi and Evangelos Topakas *

IndBioCat Group, Biotechnology Laboratory, School of Chemical Engineering, National Technical University of Athens, Zografou Campus, 9 Iroon Polytech Str., 15772 Athens, Greece; k_grigorakis@chemeng.ntua.gr (K.G.); cferousi@chemeng.ntua.gr (C.F.)

* Correspondence: vtopakas@chemeng.ntua.gr

Abstract: Protein engineering has emerged as a transformative field in industrial biotechnology, enabling the optimization of enzymes to meet stringent industrial demands for stability, specificity, and efficiency. This review explores the principles and methodologies of protein engineering, emphasizing rational design, directed evolution, semi-rational approaches, and the recent integration of machine learning. These strategies have significantly enhanced enzyme performance, even rendering engineered PETase industrially relevant. Insights from engineered PETases underscore the potential of protein engineering to tackle environmental challenges, such as advancing sustainable plastic recycling, paving the way for innovative solutions in industrial biocatalysis. Future directions point to interdisciplinary collaborations and the integration of emerging machine learning technologies to revolutionize enzyme design.

Keywords: protein engineering; biocatalysis; rational design; directed evolution; semi-rational design; machine learning; plastic-degrading enzymes

Academic Editor: Jose M. Palomo

Received: 13 January 2025
Revised: 30 January 2025
Accepted: 1 February 2025
Published: 4 February 2025

Citation: Grigorakis, K.; Ferousi, C.; Topakas, E. Protein Engineering for Industrial Biocatalysis: Principles, Approaches, and Lessons from Engineered PETases. *Catalysts* **2025**, *15*, 147. https://doi.org/10.3390/catal15020147

1. Introduction

For billions of years, enzymes have served as Nature's catalysts, driving countless biochemical reactions essential for life. Archeological findings indicate that humans have exploited enzymatic processes for daily applications such as bread fermentation and beer brewing since prehistorical times [1,2].

Today, these biological catalysts are being harnessed and engineered to solve some of the most pressing challenges in industrial biotechnology. For industrial use, enzymes must not only be cost-effective but also exhibit high performance, high specificity or promiscuity, and stability under the specific conditions required for their application. These stringent requirements often expose the limitations of wild-type (WT) enzymes that frequently fail to meet industrial demands due to low catalytic rates [3], poor thermal [4] or pH stability [5], inadequate organic solvent tolerance [6], restricted substrate range [7], susceptibility to inhibition by their substrates or products [8–10], and incompatible optimal reaction pH [11]. Subsequently, advancements in biocatalytic performance across a wide range of operational conditions are crucial in meeting the demands of large-scale industrial applications.

To overcome these challenges, protein engineering seeks to ameliorate these issues by introducing novel enzymatic activities, enhancing catalytic efficiencies, broadening or changing substrate specificities, and optimizing enzymatic stability under harsh operational conditions, such as high temperatures or diverse pH environments [8]. Protein engineering strategies are predominantly divided into three main categories: (i) rational design, which relies on detailed structural and mechanistic knowledge of enzymes to either introduce

targeted modifications or to design entirely novel catalysts; (ii) directed evolution, which mimics natural selection by iteratively mutating and screening enzymes for improved properties; and (iii) semi-rational design, which prioritizes the design of smart libraries through evolutionary insights from homologous proteins. In addition, (iv) machine learning (ML) and deep learning (DL) methods have recently emerged as promising alternative strategies, leveraging vast amounts of genomic, structural, and functional data to predict mutations that enhance enzymatic properties [12,13].

The transformative potential of protein engineering is highlighted by the example of engineered industrial polyethylene terephthalate (PET) hydrolases (PETases). PET accounts for around 5% of total global plastic production [14] and is the most recycled plastic worldwide [15], although traditional thermomechanical recycling methods downgrade PET and produce inferior recycled products [16]. The discovery of novel plastic-degrading enzymes has significantly advanced enzymatic recycling technologies, offering advantages such as selective recycling from plastic mixtures and producing final products with virgin PET quality [3,17]. Despite this progress, naturally occurring PETases exhibit limitations in their efficiency and stability, restricting their use in industrial settings; therefore, recent advances in enzymatic recycling of PET have been driven not only by the discovery of new PETases but also by significant protein engineering efforts. These advancements culminated in engineered enzymes, such as the leaf and branch compost cutinase (LCC) variant LCC^{ICCG}, which is the first PETase to be industrialized for PET bio-recycling [3,4,18], an example that highlights the capability of protein engineering to improve enzyme efficiency and also expand industrial enzymatic applications beyond what Nature alone provides.

This review aims to provide a guide for exploring protein engineering concepts and commonly employed strategies. It explores the structure-function interplay behind enzymatic activity, substrate specificity, and stability, along with the latest and most relevant approaches employed to optimize or create industrially relevant biocatalysts, categorized into rational design, directed evolution, semi-rational design, and ML approaches. Highlighting recent advancements and proposing future research directions, this review seeks to contribute to the development of innovative, efficient, and industrially relevant enzymes and protein engineering solutions while inspiring and equipping future scientists with the knowledge and tools needed to innovate in this emerging field during its transformative prime.

2. Fundamental Principles for Engineering Protein Activity, Specificity, and Stability

Enzymes function as biological catalysts, exhibiting remarkable specificity and catalytic efficiency [19], expediting chemical reactions by lowering the activation energy barrier required for the conversion of reactants into products, primarily through the stabilization of the reaction's transition state(s) [20]. A comprehensive understanding of the primary principles governing enzyme activity, specificity, and stability is essential for functional enzyme engineering. Key factors influencing enzymatic performance include the reaction conditions, predominantly temperature and pH, their affinity to the substrate, the employed catalytic mechanism, and the thermodynamic stability of the enzyme under reaction conditions. An industrially relevant enzyme may need to exhibit high activity that ensures efficient, rapid, and economical catalysis, high substrate specificity or promiscuity, depending on the application, and high stability, which allows for higher temperatures, thus enhancing reaction rates, reactant solubility, and reducing microbial contamination risks [21]. This section discusses critical physicochemical parameters underlying enzymatic performance.

2.1. Dependence on Temperature

Temperature is one of the most critical factors affecting enzyme activity, stability, and overall performance. As biological catalysts, enzymes exhibit a delicate balance between enhanced activity at higher temperatures and the risk of denaturation. Understanding how temperature influences enzymatic function is essential for designing robust enzymes that balance optimum activity and thermal stability under industrial conditions.

2.1.1. Optimum Temperature (T_{opt})

At low temperatures, enzymatic reactions proceed slowly due to reduced molecular motion and limited substrate-enzyme collisions. Increasing the temperature accelerates molecular movement, facilitating faster reactions up to the enzyme's optimal temperature, T_{opt}, where the rate of reaction reaches its peak [22,23]. The traditional approach describes this temperature dependence as a factor that accelerates reaction rates until the enzyme denatures, according to the Arrhenius equation (Equation (1)):

$$k_{cat}(T) = A_{cat}e^{-\frac{E_a}{RT}} \quad (1)$$

where k_{cat} is the rate constant, the units of which depend on the reaction order, A_{cat} is the pre-exponential factor, E_a is the activation energy, R is the universal gas constant, and T is the temperature in K.

As the temperature increases, enzyme activity rises exponentially until irreversible thermal denaturation dominates, converting the active enzyme (E_{act}) into an irreversibly inactivated form (X) (Equation (2)):

$$E_{act} \overset{k_{inact}}{\rightarrow} E_{inact} \quad (2)$$

The reaction rate constant for this inactivation, k_{inact}, also follows the Arrhenius equation (Equation (1)), with A_{inact} being the pre-exponential factor and E_{inact} being the activation energy for denaturation. In this case, as this is a first-order reaction, the half-life of the enzyme can also be calculated as such (Equation (3)):

$$t_{1/2} = \frac{ln2}{k_{inact}} \quad (3)$$

However, many experimental observations have shown that the relationship between temperature, catalytic rates, and enzyme stability is more nuanced than a simple gain in rate that is offset by irreversible thermal denaturation [24,25].

In the early 20th century, the formulation of Transition State Theory by Eyring, Polanyi, and others culminated in the derivation of the Eyring equation (Equation (4)) for rate constants. For a first-order rate constant, the equation can be expressed as:

$$k_{cat}(T) = \kappa\frac{k_B T}{h}e^{\frac{-\Delta G^{\ddagger}}{RT}} \quad (4)$$

where $\Delta G^{\ddagger}$ represents the Gibbs free energy difference between the reactants and the transition state, k_B and h are the Boltzmann's and Planck's constants, respectively, and κ is the transmission coefficient, which accounts for the probability of successful passage through the transition state, often assumed to be 1, reflecting a system where every trajectory crossing the transition state proceeds to product formation [26,27].

The "Equilibrium Model" introduced the idea that even before irreversible denaturation at higher temperatures, enzymes can adopt a reversible equilibrium between a fully

active form, E_{act}, and an inactive form, E_{inact}, that can undergo thermal inactivation to the denatured state, X [28,29]:

$$E_{act} \rightleftarrows E_{inact} \rightarrow \mathrm{X} \tag{5}$$

By accounting for a reversible conformational change, the equilibrium model provided markedly improved the fit to enzymatic activity data across a range of temperatures compared to older approaches, thereby reaffirming the validity of an optimal temperature that is not merely a point before catastrophic thermal denaturation but, rather, a meaningful characteristic intrinsic to the enzyme's dynamic structural landscape [30].

More recently, Macromolecular Rate Theory (MMRT) described the temperature dependence of enzyme-catalyzed reactions independent of stability or regulatory processes, purely based on thermodynamics and the role of changing heat capacity ($\Delta C_p^{\ddagger}$) between the enzyme-substrate complex (ES) and the enzyme-transition state complex ($ETS^{\ddagger}$), where the heat capacity (C_p) for ES is generally larger. This negative $\Delta C_p^{\ddagger}$ reflects a reduction in low-frequency vibrational modes in the transition state. Based on experimental observations, $\Delta C_p^{\ddagger}$ is assumed to be constant, independent of temperature and $\kappa = 1$, for simplicity [26]. Thus, given Equation (4), if [31–33]:

$$\Delta G^{\ddagger}(T) = \Delta H^{\ddagger}(T) - T\Delta S^{\ddagger}(T) \tag{6}$$

$$\Delta H(T) = \Delta H(T_o) + \int_{T_o}^{T} \Delta C_p dT' \tag{7}$$

$$S(T) = \Delta S(T_o) + \int_{T_o}^{T} \frac{\Delta C_p}{T'} dT' \tag{8}$$

then the reaction rate constant, k_{cat}, is given by (Equation (9)):

$$k_{cat}(T) = \frac{k_B T}{h} e^{\frac{-\Delta H^{\ddagger}(T_o) - \Delta C_p^{\ddagger}(T - T_o)}{RT} + \frac{\Delta S^{\ddagger}(T_o) + \Delta C_p^{\ddagger}(\ln(T) - \ln(T_o))}{R}} \tag{9}$$

Due to the increasing dominance of the entropic contribution ($\Delta S^{\ddagger}(T)/R$) over the enthalpic term ($-\Delta H^{\ddagger}(T)/RT$), the reaction rate declines above the optimum temperature (T_{opt}), even in the absence of enzyme denaturation. In thermophilic enzymes, as T_{opt} approaches 100 °C, $\Delta C_p^{\ddagger}$ diminishes to 0, causing the temperature dependence to approximate Arrhenius behavior as the enthalpic and entropic terms become less temperature-sensitive. Conversely, for mesophilic and psychrophilic enzymes that maintain relatively high unfolding temperatures, a negative $\Delta C_p^{\ddagger}$ leads to a curved temperature dependence in their catalytic rates [26].

By incorporating a simple correction term accounting for enzymatic denaturation (k_{inact}), MMRT can accurately deconvolute the intrinsic thermodynamic effects arising from $\Delta C_p^{\ddagger}$ on the reaction rate constant (k_{cat}) from those due to unfolding, thereby providing a comprehensive theoretical framework for understanding the temperature dependence of enzymatic reactions [26].

2.1.2. Melting Temperature (T_m)

Enzyme inactivation theory proposes an initial equilibrium phase where the E_{act} undergoes unfolding to the reversibly denatured and inactive state E_{inact}, which retains the potential for either refolding to its native conformation or progressing to irreversible inactivation state X (Equation (5)) [29].

Melting Temperature (T_m) is the temperature at which 50% of the enzyme population transitions from its active to its inactivated, but still reversibly denatured, state [34]. Close to or above the T_m, reversibility of inactivation rapidly decreases [29]. T_m is typically measured

with differential scanning calorimetry [35], optical methods such as circular dichroism [36], dynamic light scattering [37], nano-differential scanning fluorimetry (nanoDSF) based on either intrinsic fluorescence or utilizing dyes, such as SYPRO Orange [38], or directly from enzymatic activity assays at different temperatures [29]. Typically, T_m falls between 5 °C and 15 °C above the T_{opt} [39]. T_m serves as a crucial benchmark for assessing an enzyme's thermal resilience, as enzymes with higher T_m values are generally more stable under elevated temperatures, making them more suitable for industrial processes that operate under harsh conditions [40].

2.2. Dependence on pH

Similarly to temperature, pH also influences enzymatic activity and stability by affecting the ionization states of titrating residues. As pH affects the overall protein conformation, the enzyme–substrate interactions and the catalytic residues, enzymes exhibit specific pH optima and working ranges that must be carefully considered to maintain optimal performance.

2.2.1. Optimum pH (pH_{opt})

The activity of enzymes is intrinsically linked to the pH of their environment, as the protonation states of catalytic residues and substrates are critical for enzymatic function. The intrinsic pK_a values of ionizable groups such as carboxyls, amines, hydroxyls, thiols, and imidazoles determine their protonation state, as shown in Figure 1 [41].

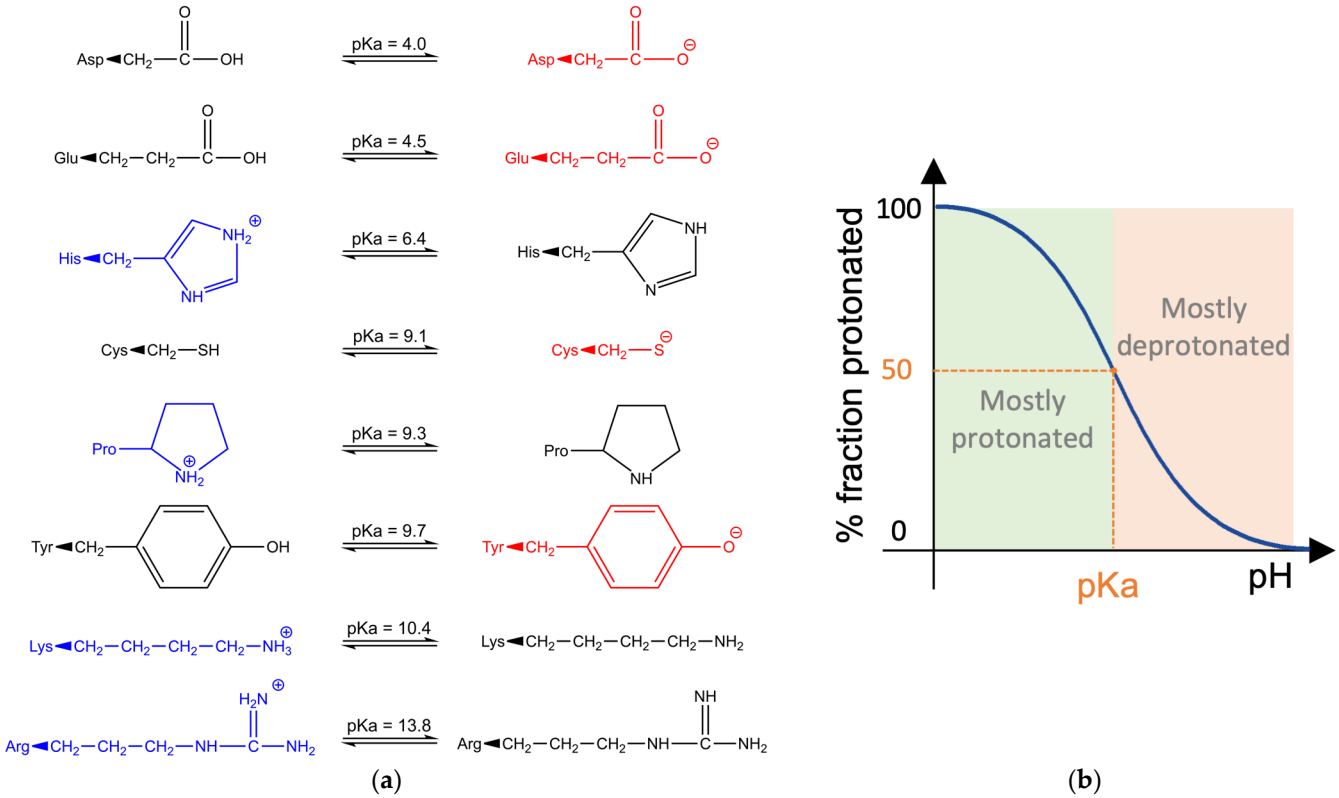

Figure 1. pH effects on amino acids: (**a**) structures and values of amino acid side chains involved in ionization (pK_a values may vary by ± 0.5) [41,42]; (**b**) schematic representation of a titration curve showing the relationship between pH and the fraction of ionizable species that are protonated [43]. When $pH = pK_a$, half of the species will be protonated.

However, these intrinsic values can be significantly perturbed in the active site of enzymes, markedly affecting catalysis, as titratable amino acids adopt a functional or apparent pK_a within the protein environment. This transformation arises from electrostatic interactions with charged or partially charged proximal groups, which generate a localized

microenvironment that modulates their behavior. The polarity and dielectrocharacteristics of the surrounding environment also critically influence the magnitude of this pK_a perturbation by impacting the stability of the associated charged states. All these factors combined fine-tune the pK_as of the titratable amino acids of a protein [41].

For example, in the serine protease mechanism, peptide bond hydrolysis is catalyzed by a catalytic triad composed of serine, histidine, and aspartate, as well as an oxyanion hole that stabilizes the intermediate complex (Figure 2). Initially, the substrate's scissile peptide bond is oriented so that its carbonyl carbon is positioned adjacent to the nucleophilic serine residue. When the reaction begins, the histidine ring nitrogen, which has a pK_a near 7.5 in the free enzyme [44], acts as a general-base, converting the serine into a strongly nucleophilic alkoxide-like species by deprotonating its hydroxyl, thus increasing its nucleophilicity and enabling it to attack the peptide carbonyl (nucleophilic attack). This generates a tetrahedral intermediate, stabilized in the oxyanion hole, and the histidine's pK_a is shifted significantly upward to between 10 and 12. Indicative pK_as are revealed by nuclear magnetic resonance studies of chymotrypsin complexes with peptidyl trifluoroketones, analogs of the tetrahedral intermediate [44]. Subsequently, the histidine acts as a general acid, being the proton donor, thus facilitating the departure of the cleaved amine segment and shifting its pK_a back to near 7.5. Water then enters the active site, once again activated by histidine's elevated basicity, and performs a nucleophilic attack on the carbonyl carbon of the acyl-enzyme intermediate, generating another tetrahedral intermediate, which subsequently collapses by donating the hydrogen from the protonated histidine back to the serine, releasing the cleaved peptide C-terminus and regenerating the free enzyme [41].

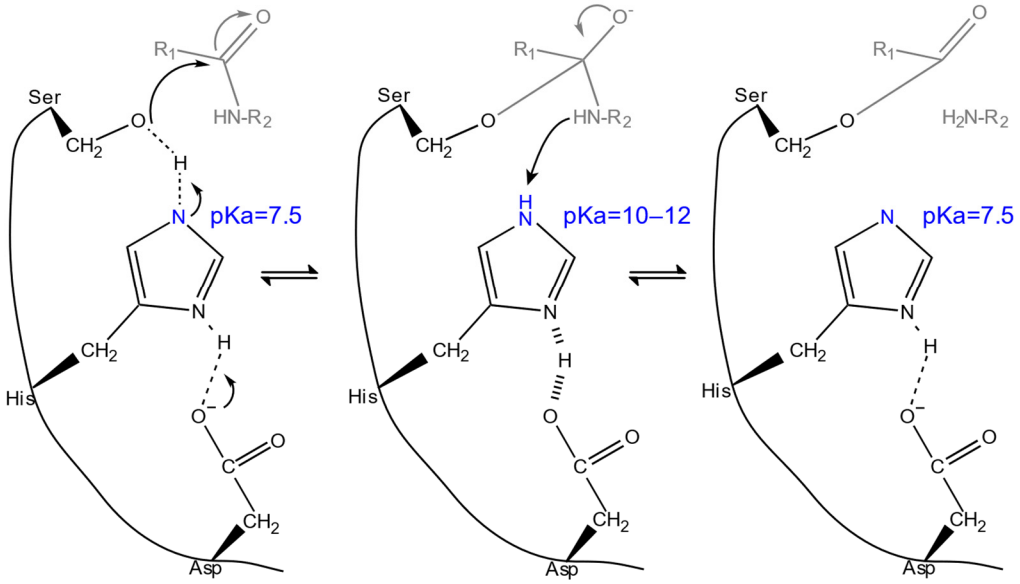

Figure 2. Schematic representation of the α/β hydrolase catalytic mechanism probed on a serine protease. Histidine deprotonates the serine hydroxyl, enhancing its nucleophilicity. During catalysis, the histidine undergoes significant $pK\alpha$ shifts—ranging from near neutral (~7.5) in its general-base form to as high as 10–12 in its protonated, general-acid form—enabling it to toggle between these roles. UniProt accession number: P00767 [44].

Throughout these steps, the substantial shifts in histidine's pK_a are critical for modulating its dual role as acid and base catalyst (Figure 2). The enzyme's optimal activity within a pH range of 8 to 9 [45–47] can be attributed to the ionization states of the histidine residue, which, at this pH, is mostly deprotonated, facilitating its role as a general base in

deprotonating the serine and, thereafter, as a general acid that protonates the departing amine [47].

Over the years, many tools have been created to predict protein pK_a values based on their structures using empirical rules from experiments [48], the Poisson–Boltzmann (PB) equation [49,50], Density Functional Theory (DFT) [51], and ML [52,53], among others [54–56]. Based on these methods, the discrete constant pH framework combines molecular dynamics (MD) with Monte Carlo simulations, allowing a more accurate approximation of the pK_as. In this approach, an MD simulation, which provides enzyme conformational sampling, is occasionally paused to allow resampling of the residues' protonation states [57]. By engineering the pK_as of the catalytic amino acids, it is possible to change an enzyme's pH_{opt} [58,59].

2.2.2. pH Stability

Under extreme pH conditions, the primary mechanism driving protein unfolding is the electrostatic repulsion between like-charged groups within the protein structure, which may subsequently lead to aggregation or irreversible denaturation that is distinctly different from that of thermal denaturation [60]. From an engineering perspective, by modeling these unfolding events with MD simulations or experimental techniques and identifying strategies to mitigate destabilizing interactions, it is possible to rationally engineer pH-dependent stability. Such observation was seen in the stabilization of the 37-residue α/β protein CHABII by a rational single-point mutation (H21F), in which the protonation of the histidine at low pH induced unfolding occurred by destabilizing the hydrophobic core [61]. In this context, however, directed evolution often provides a more practical approach, bypassing the need for exhaustive computational predictions. This is exemplified by the engineered xylanase XynHBN188A, where two amino acid substitutions led to increased specific activity and pH stability [62].

2.3. Structure-Function Relationships

2.3.1. Substrate Affinity and Specificity

Substrate-enzyme complementarity, akin to an "induced-fit" [63,64], or the more recent but not mutually exclusive "conformational-selection" model [65] determines binding efficiency and substrate specificity. The binding energy of the enzyme-ligand complex is also utilized for catalytic turnover [66].

Engineering enzyme-substrate affinity involves tailoring enzymes' interactions with their substrates and products, thereby enhancing enzyme specificity, promiscuity [67], or product release, even altering the final product's composition [68]. Key determinants of substrate recognition are the shape of the active site and binding pocket, the conformations they can adopt, the combination of electrostatic, hydrophobic, hydrogen bonding, and van der Waals interactions [69–71], and the entrance tunnels [72].

A comparative analysis of homologous proteins with varying substrate specificities facilitates the identification of these key structural determinants [66]. In addition, energy-based docking solutions, such as AutoDock 4 (slower, more interpretable) [73] and AutoDock Vina (faster, superior results) [74], quantify protein-ligand binding affinities in silico, relatively accurately. The Molecular Mechanics Generalized Born Surface Area (MMGBSA), or the more rigorous Molecular Mechanics PB Surface Area (MMPBSA), approach provides an even better prediction of binding energy and can dissect interactions using per-residue free energy decomposition or alanine scanning [75–77]. The CAVER, CAVER Analyst and CaverDock suites provide powerful tools for analyzing enzymes with buried active sites to identify bottlenecks in substrate binding or product release and quantify binding energy between the bound and surface state in static and dynamic protein structures [78–80].

2.3.2. Stabilizing Mutations

Thermodynamics provides a useful framework for interpreting the stabilizing effects of mutations. Stabilizing mutations aim to maximize the Gibbs free energy difference (ΔG_f) between folded and unfolded states of a protein, thereby favoring the more thermodynamically stable folded conformation [21,81]. Equation (10) stipulates that protein stabilization can be principally achieved through adjustments to either the enthalpic (ΔH_f) or the entropic (ΔS_{conf}) components of the system, although deconvoluting the effect of a structural modification on each term is difficult [21]:

$$\Delta G_f = \Delta H_f - T\Delta S_{conf} \quad (10)$$

For instance, mutations such as serine to proline in surface loops limit the conformational entropy of the unfolded state, thus increasing ΔS_{conf}, and are unlikely to be interpreted as adding or removing any atomic interactions that contribute to ΔH_f [21,81]. Notwithstanding, by measuring the effect on ΔH_f experimentally, the results may differ from the oversimplified interpretation [21].

Advances in biomolecular force fields, thermodynamic cycle analyses, and ML have facilitated the comprehensive assessment of mutational impacts by employing computational tools to systematically evaluate the effects of single amino acid substitutions. Such tools are Rosetta ddg_monomer [82], FoldX [83], ERIS [84], or DeepDDG [85] which predict the $\Delta\Delta G_f$ of a protein induced by a point mutation, either given by the difference in energy between the WT structure and the point mutant structure, calculated by a forcefield, or, in the last case, via DL. By leveraging these predictors and implementing in silico site-saturated mutagenesis, FireProt's energy-based approach scores suggest single-point mutations by the predicted $\Delta\Delta G_f$ from Rosetta and FoldX with 100% precision and 0% false-positive rate on experimental datasets (after filtering with conservation analysis by the Rate4Site tool [86]), albeit at the expense of omitting some [87].

2.3.3. Flexibility

Enzymes are inherently dynamic and flexible macromolecules, characterized by internal conformational motions essential for substrate binding, product release, and potentially the catalytic mechanism itself [20,88,89]. Flexibility is often evaluated using B-factors (Debye-Waller factors, temperature factors, or atomic displacement parameters) derived from X-ray crystallography [90]. These factors quantify atomic displacement or mobility within the crystalline structure, providing a detailed view of the enzyme's dynamic properties at an atomic scale. Regions characterized by lower B-factors exhibit structural rigidity, whereas higher B-factors indicate flexible domains that are often integral to substrate binding and conformational transitions [91]. B-factors can be visualized by PyMOL's B-factor putty mode (Figure 3), which illustrates atomic mobility by varying thickness and color based on flexibility [92,93]. Variability in B-factors across structures of the same molecule often arises from external influences unrelated to intrinsic molecular properties, such as experimental conditions or computational methodologies, and, therefore, careful consideration when drawing comparisons between different structures is necessary [90]. Aside from X-ray structures, B-factors can be computationally estimated using MD simulations [94], normal mode analysis [95,96], elastic network models [97], and ML models [98].

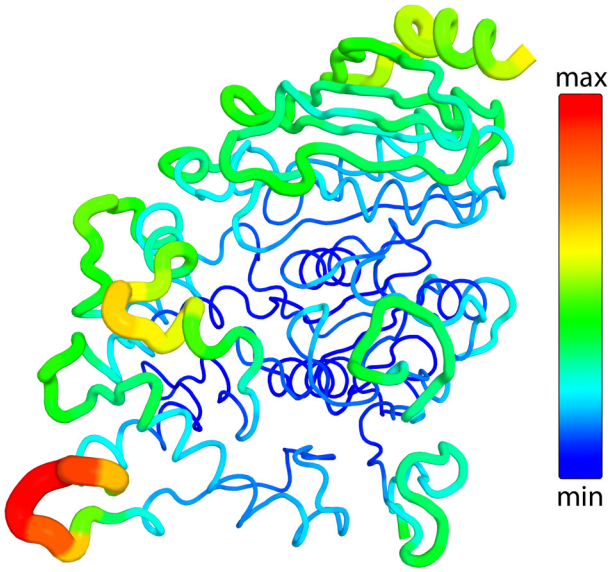

Figure 3. B-factor putty representation created in PyMOL [93] of a feruloyl esterase C from *Fusarium oxysporum* (*Fo*FaeC). Values range from 22.56 to 82.91 Å^2, representing the mean square displacement of atoms from their average positions due to thermal motion [99]. PDB ID: 6FAT.

2.3.4. Activity-Stability Trade-Off

Enzyme flexibility is a critical determinant of substrate specificity, enabling the dynamic structural transitions required for substrate binding, catalytic turnover, and product release through the "induced-fit" [63] or the "conformational-selection" models [65]. Traditionally, excessive flexibility has been thought of as undermining structural stability and being a trade-off with protein activity [100]. Rational protein engineering strategies have introduced targeted modifications, e.g., disulfide bonds, proline residues, and hydrogen bonds, to reduce protein flexibility away from the active site and fine-tune the activity-stability balance, to great acclaim [21].

The hypothesized trade-off has remained a cornerstone in protein engineering with numerous studies documenting evidence supporting its existence [20,101–103]. However, this assumption consequently suggests that a more thermostable enzyme operates as a slower catalyst compared to a less stable homolog under low-temperature conditions; that does not typically occur, neither in engineered systems [20] nor when comparing WT thermophilic and mesophilic homologous enzymes [104]. In fact, thermal stability can be enhanced by increasing conformational entropy upon folding (ΔS_{conf}), thereby generating a more thermostable enzyme while preserving flexibility (Equation (10)). In this case, a more thermostable enzyme may exhibit greater flexibility compared to a less stable homolog, if the increase in ΔS_{conf} results from an increase in the conformational entropy of the native state (S_{conf_F}) [20,105]:

$$\Delta S_{conf} = S_{conf_F} - S_{conf_U} \tag{11}$$

Directed evolution studies further illuminate this dynamic, as engineered enzymes frequently achieve enhanced stability or activity without corresponding losses, showcasing the possibility of decoupling these properties [20], as is clearly the case with all reported directed evolution studies of the PET-degrading PETase from *Ideonella sakaiensis* (*Is*PETase) [106]. Nevertheless, strategies such as directed evolution and rational design are employed to navigate this balance, aiming to enhance both stability and activity without compromising either [4,107,108].

2.3.5. Structure-Function Engineering Insights

To enhance the activity, specificity, and stability of enzymes, a variety of engineering approaches have been developed (see Section 3). Table 1 provides an overview of key modifications used in protein engineering—regardless of their discovery strategy—outlining their mechanisms and typical effects in terms of activity, substrate affinity, and stability while acknowledging that any modification inherently impacts multiple attributes. Figure 4 showcases specific applications of these modifications. Less common enhancements such as allosteric site engineering [109] and helix capping also contribute to enzyme optimization [110].

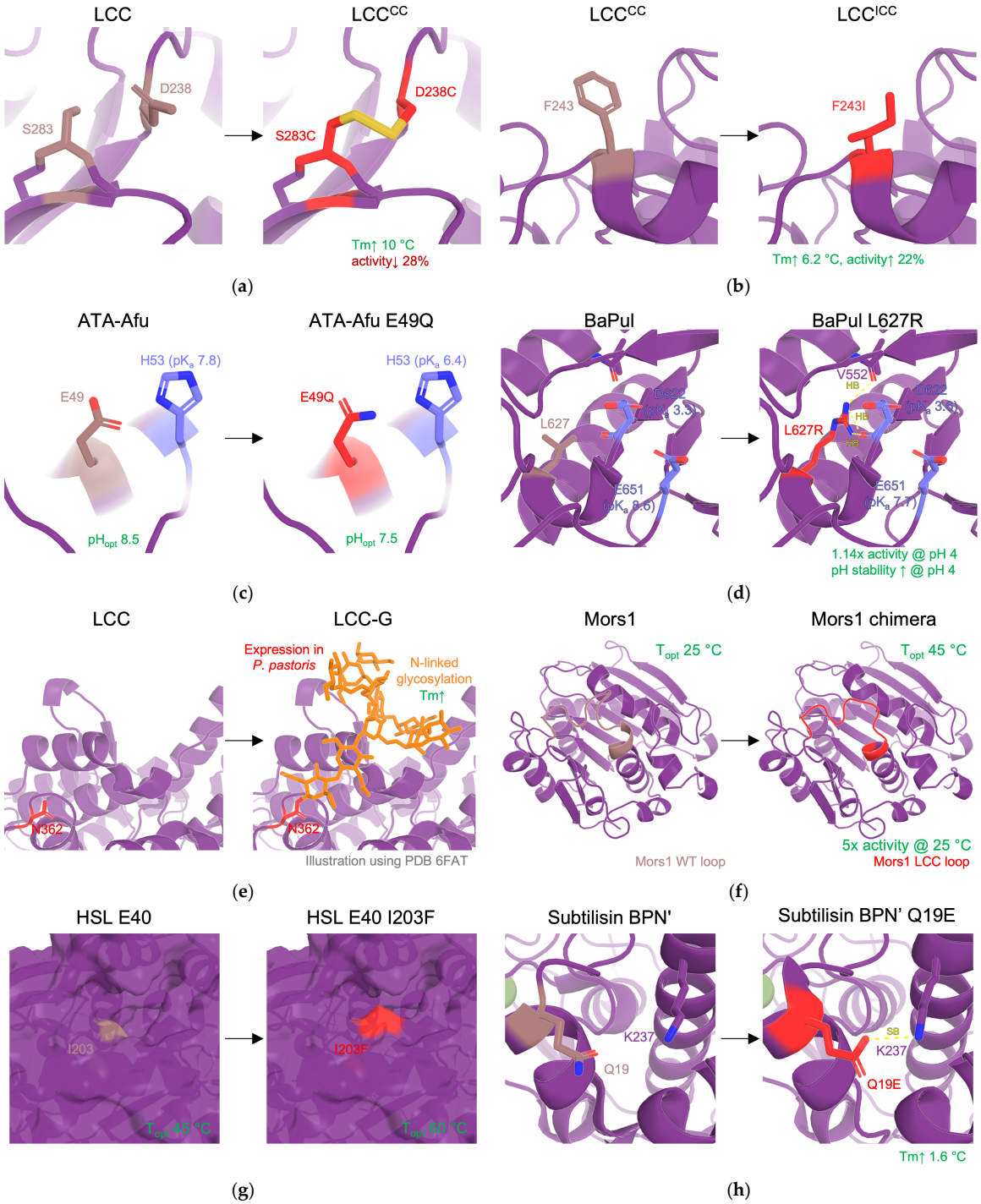

Figure 4. Examples of key modifications in protein engineering: (**a**) Engineering a disulfide bridge (D238C and S283C) in LCC increases its thermal stability but decreases activity by 28% [4]; (**b**) Single-point mutation F243I increases the activity of LCC^{CC} to 22% higher than WT; (**c**) Engineering pKa values

of catalytic amino acids by single-point mutation E49Q decreases pH_{opt} in ATA-Afu [58]; (**d**) L627R in *Ba*Pul introduces hydrogen bonds, leading to pK_a shifts for D622 and E651 and increased activity and stability at pH 4.0 [121]; (**e**) Expressing WT LCC in *Pichia pastoris* introduces N-glycosylation and increases thermal stability [117]; (**f**) Grafting an active site loop from LCC to Mors1 increases T_{opt} and activity [122]; (**g**) I203F in HSL_E40 improves thermal stability through hydrophobic interactions [123]; (**h**) Q19E in subtilisin BPN′ introduces a salt bridge, increasing T_m [124]. LCC: leaf and branch compost cutinase; ATA-Afu: amine transaminase from *Aspergillus fumigatus*; BaPul: *Bacillus acidopullulyticus* pullulanase; Mors1: *Moraxella sp.* TA144 cutinase; HSL: hormone-sensitive lipase.

Table 1. Key modifications in protein engineering and approaches that enable them.

Modification	Mechanism	Primary Effects	Engineering Approach	Refs.
Single-point mutations	Selective single point mutations with energetically favorable residues (minimizing $\Delta\Delta G_f$), residues important for substrate binding (surface electrostatics/hydrophobicity) or enhancing activity.	activity ↑↓** affinity ↑↓ stability ↑↓	RD, SRD, DE, ML *	[4,12,21,106, 111]
Disulfide bridges	Covalent linkage of cysteine residues to rigidify the protein backbone and constrain conformational freedom.	stability ↑ activity ↓ (typically)	RD, ML	[4,12,21,112]
Shifting pK_a values	Modifying electrostatic microenvironment with single-point mutations to fine-tune the pK_as of catalytic residues.	activity ↑ (in different pH)	RD, DE	[11,58,59,113]
Hydrogen bond network optimization	Modification of hydrogen-bonding network to reinforce affinity with substrate and active site stability.	activity ↑ affinity ↑	RD	[114]
Salt bridges	Introducing salt bridges to reinforce protein structure.	stability ↑	RD	[12,21,115]
Glycosylation	Introducing sites for post translational modifications.	stability↑	RD	[116,117]
Surface loop engineering	Rational remodeling of surface loops via shortening or loop grafting.	activity ↑↓ affinity ↑↓ stability ↑↓	RD, SRD	[118,119]
Hydrophobic core packing	Optimized distribution of hydrophobic residues to eliminate internal cavities.	stability↑	RD	[21,120]

* DE: directed evolution, ML: machine learning, RD: rational design, SRD: semi-rational design. ** ↑: increased, ↓: decreased.

3. Protein Engineering Approaches and Strategies

Protein engineering strategies can be broadly categorized into rational design, directed evolution, semi-rational design, and, more recently, ML/DL approaches [12]. Figure 5 provides an overview of these schemes.

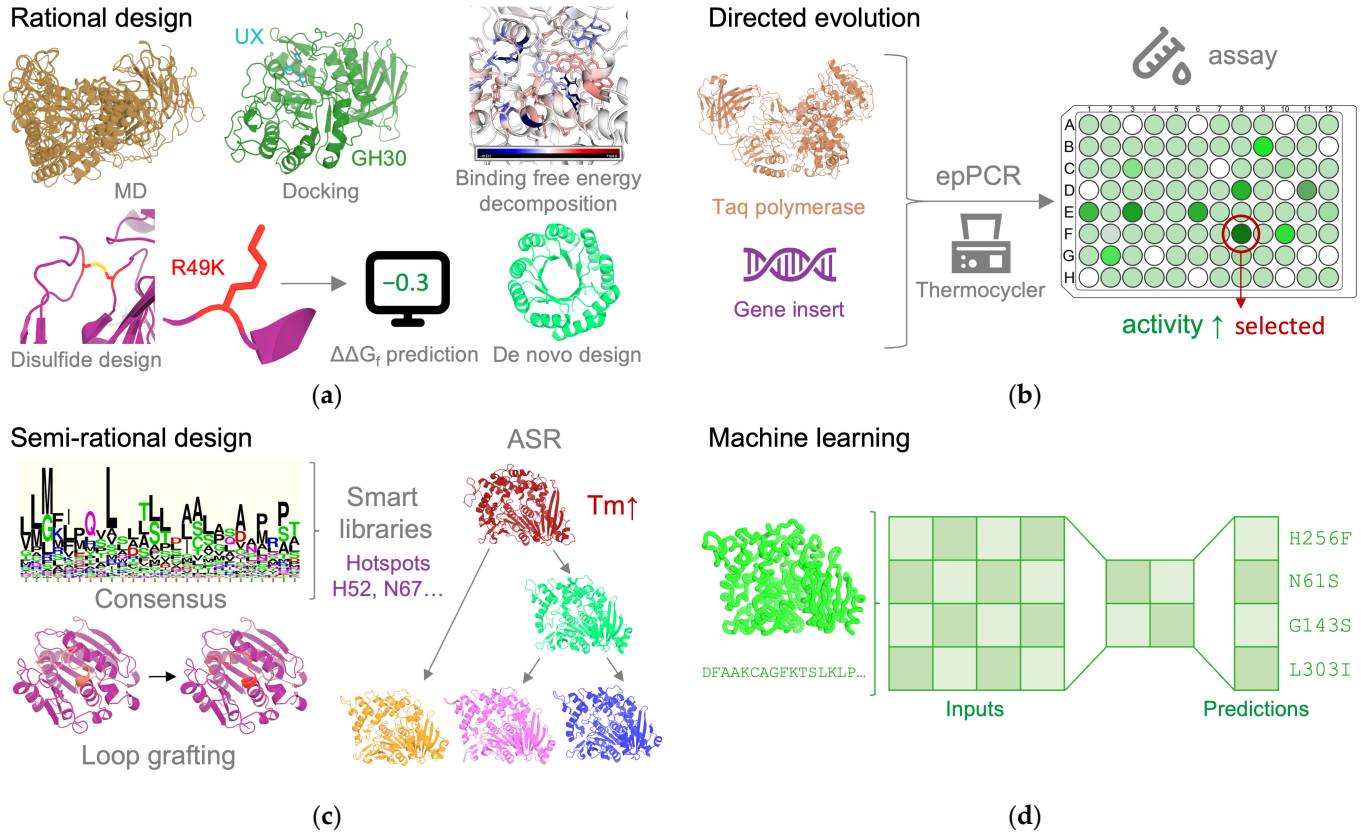

Figure 5. Schematic representation of protein engineering approaches. (**a**) Rational design approach, including molecular dynamics (MD), docking, binding free energy decompositions, disulfide bond design, $\Delta\Delta G_f$ predictions, and de novo design; (**b**) directed evolution approach, using epPCR and activity-based selection assays; (**c**) semi-rational design, with consensus analysis for hotspots identification, loop grafting, and ancestral sequence reconstruction (ASR); (**d**) machine learning (ML) in protein engineering, showing input sequence and structural data to predict specific mutations. The consensus sequence logo has been created using Weblogo 3 [125].

3.1. Rational Design

Rational design mainly refers to leveraging knowledge of a target enzyme's structural and functional attributes and using computational modeling and simulation frameworks to predict mutations, insertions, or deletions aimed at augmenting enzymatic performance. Most recent studies on improving plastic-degrading enzymes primarily adopt a structure-based approach, exploiting the extensive structural and functional data available for these enzymes [12]. In return, these enzymes also function as benchmark platforms for advancing and refining protein engineering methodologies. Another domain of rational design involves the de novo synthesis of novel enzymes, by incorporating active sites and substrate-binding pockets predicted to catalyze a reaction of interest into geometrically compatible native scaffolds [126].

3.1.1. Structure-Based Design

Structure-based rational design relies on computational tools that analyze and manipulate protein structures. These structures may be obtained through experimental methods (e.g., X-ray crystallography, cryo-electron microscopy, and nuclear magnetic resonance), homology modeling (e.g., SWISS-MODEL and MODELLER) leveraging information from similar solved structures, or, lately, DL-based predictors (e.g., AlphaFold 2.0 and ESMFold), which achieve unprecedented accuracy in structural prediction while maintaining high computational efficiency [127–129]. Docking software, such as AutoDock Vina, predict

protein-ligand binding modes [74]. Visualization platforms such as PyMOL [130] and ChimeraX [131] enable detailed analysis of enzymatic structures, facilitating the identification of protein regions critical for activity, substrate affinity, and stability. These tools help pinpoint key protein-ligand interactions (i.e., hydrogen bonds, hydrophobic contacts, electrostatic interactions, steric hindrances, etc.) and evaluate flexibility by analyzing B-factors, which highlight potential engineering hotspots [93]. Integrated mutagenesis tools allow for the visualization and evaluation of potential mutations. Energy-based methods (e.g., Rosetta ddg_monomer [82] and FoldX [83], both integrated into FireProt [87]) systematically evaluate the $\Delta\Delta G_f$ associated with point mutations, and specialized programs such as Disulfide by Design 2.0 enable the introduction of disulfide bonds [132].

Similarly, MD simulations using software such as GROMACS [133], AMBER [134], and CHARMM [135] enable the exploration of protein and protein-ligand dynamics. MD simulations can extract B-factors [94], estimate binding free energies with MMPBSA/MMGBSA, and decompose the contributions of each amino acid to these energies [75,77]. Additionally, they provide insights into the frequency and time-dependency of interactions, revealing how often and for how long contacts (e.g., hydrogen bonds) or conformational states occur [136]. They can also calculate the activation energy barrier of a reaction using quantum mechanics/molecular mechanics and umbrella sampling [137], as well as model reaction mechanisms at an atomic level [138]. The simulated trajectories can be analyzed to reveal conformational clusters representing accessible states [139], while Markov State Models may be used to characterize these states by their populations and transition kinetics [140]. In addition, Principal Component Analysis reduces the dimensionality of the configurational space by unveiling the dominant modes of motions [141] and also allows plotting of the free energy landscape projected onto principal components [142]. Furthermore, Constant-pH MD addresses the limitation of fixed protonation states by allowing titratable residues to dynamically switch between protonated and deprotonated forms, capturing pH-dependent behaviors and allowing the determination of their pK_a values [57]. Moreover, MD simulations can explore the denaturation of proteins by simulating them under diverse thermal or environmental conditions [143,144].

3.1.2. De Novo Design

The overarching goal of protein engineering is the de novo development of enzymatic functions derived from first principles [66], once considered to be impossible [145]. Leveraging computational methods to create structures and functionalities absent in nature, "true" de novo design aims to construct entirely new proteins from scratch, relying purely on computationally generated backbones, while the "minimalist" approach utilizes known stable protein folds as the foundation for introducing new functional sites, aiming to establish the feasibility of a catalytic reaction [146]. Computational tools, such as Rosetta [147,148], play a pivotal role in designing scaffolds and optimizing sequences for functional and structural refinement [149]. Rosetta Match [150] and AsiteDesign [151] extend these capabilities by enabling the precise redesigning or grafting of active sites onto protein scaffolds and engineering the stability of the transition state [152].

Although de novo-designed enzymes frequently exhibit limited catalytic efficiencies, iterative optimization through directed evolution, discussed in the following section, or other engineering approaches can substantially enhance their performance [153]. Despite the inherent challenges, advances in computational frameworks continue to highlight the potential of this approach [154]. Prominent achievements include the development of catalysts for Diels-Alder reactions [155], ester hydrolysis [156], and retro-aldol reactions [157]. In the context of plastic-degrading enzymes, a polycarbonate hydrolase has

been successfully designed de novo by introducing a catalytic site into a thermostable scaffold [158].

3.2. Directed Evolution

Directed evolution has accelerated the optimization of polymer-degrading enzymes, providing robust complementarity to rational design. In contrast to rational design, which demands detailed structural and functional insights to guide enzyme engineering, directed evolution operates independent of prior knowledge regarding structure-function relationships, establishing it as an exceptionally versatile approach, particularly valuable for enzymes with limited structural characterization [159].

By simulating natural selection, directed evolution iteratively generates extensive libraries of enzyme variants and screens them to identify enhanced traits. In vitro Polymerase Chain Reaction (PCR)-based methodologies, including error-prone PCR (epPCR), Site Saturation Mutagenesis (SSM) (focused mutagenesis), and recombination-driven DNA shuffling, represent the principal strategies for exploring the mutational landscape [160], although SSM could be considered part of all engineering approaches after identifying hotspots.

However, despite its transformative potential, directed evolution presents inherent challenges. The primary bottleneck in directed evolution experiments lies in identifying improved variants, which heavily depends on high-throughput screening platforms tailored to the specific enzymatic activity under investigation [160,161]. Furthermore, inherent biases introduced by the experimental methodologies (e.g., the preference of Taq polymerase for AT $\rightarrow$ GC transitions and AT $\rightarrow$ TA transversions in epPCR) further limit exploration of the mutational landscape. Lastly, even protein libraries containing millions of variants can only probe an infinitesimal fraction of the immense sequence space theoretically available for an average protein [160,162].

3.3. Semi-Rational Design

Semi-rational design leverages computational tools to extract evolutionary insights from homologous proteins based on conserved sequences, structures, and functional data with a focus on creating small, high-quality mutational libraries, allowing for more efficient sampling of sequence space [13]. Databases such as the Protein Data Bank (PDB), UniProt [163], CAZy (Carbohydrate-Active enZYmes) database [164], and PAZy (Plastics-Active enZYmes) database [165] catalog comprehensive protein data including sequences, structures, functional information, and the organism of origin. To capitalize on that, the Basic Local Alignment Search Tool (BLAST) [166] and Many-against-Many sequence searching (MMseqs2) [167] identify homologous proteins through sequence analysis. Recently, advancements in computational structure prediction, as illustrated by AlphaFold 2.0 [127], have enabled tools such as Foldseek to perform rapid and precise structural homolog searches across extensive protein databases [168], often achieving higher sensitivity than sequence-based methods since structural cores evolve slower than sequences [169].

Back-to-consensus design is a semi-rational protein engineering strategy that exploits evolutionary information to enhance protein stability [13,170]. By aligning homologous sequences, conserved residues are identified, reflecting evolutionary pressure for functional importance. Target proteins are mutated to match the consensus sequence, increasing melting temperatures by 10–32 °C. However, not all conserved residues contribute positively to stability, as approximately 50% are stabilizing, 10% are neutral, and 40% are destabilizing, thus requiring precise selection. Full-length de novo sequences using the most frequent residues can also be constructed with some success [13].

Ancestral Sequence Reconstruction (ASR) infers the sequences of ancient proteins by analyzing and comparing the sequences of their modern descendants using phylogenetic

methods [13]. The functional and structural properties of ancient proteins often reveal enhanced stability, elevated promiscuity, or functionalities that have been lost or modified in contemporary proteins [171]. ASR has been successfully used by Pfizer Inc. to engineer ene-reductases with enhanced thermostability [172], utilizing the FireProt^{ASR} tool [173].

The persistence of specific amino acids within proteins of the same family across evolutionary timelines illustrates the need to maintain the protein's biological activity. ConSurf analyses evolutionary conservation at the residue level, identifying sites under intense selective pressure that are likely important for maintaining structural or functional integrity. The results are presented as a color-coded visualization superimposed on the protein structure, ranging from variable to highly conserved residues [174]. Utilizing this concept, HotSpot Wizard identifies and evaluates evolutionary variable amino acids in the active site or along the access tunnels as key targets for mutagenesis [175]. Concurrently, residue-residue coevolution can be uncovered using GREMLIN [176] and EVcouplings [177] for uncovering epistatic interactions (i.e., where consequences of a mutation in one residue are dependent on the state of another) and has been proposed as a strategy to design smart mutational libraries [178].

Structure-based recombination approaches, such as SCHEMA [179], decompose proteins into structurally compatible fragments, facilitating the generation of chimeric libraries that preserve core folds and minimize misfolding by choosing the least disruptive crossover locations, generating functional chimeric proteins [180]. Similarly, LoopGrafter has demonstrated remarkable efficacy, achieving a 40,000-fold enhancement in the bioluminescence efficiency of the grafted variant relative to the scaffold enzyme of an ancestral dehalogenase-luciferase [118,181].

3.4. Machine Learning and Deep Learning

The integration of ML into protein engineering represents a revolutionary shift in the field, offering an unprecedented capacity to predict structures from sequences with exceptional speed and accuracy, navigate high-dimensional sequence-function landscapes, and generate optimized protein variants that surpass the limitations of traditional protein engineering approaches. Expansive databases, modern model architectures, and present-day hardware have culminated in successful uses of this technology in enzyme engineering, showcasing the transformative potential of ML in streamlining the development of more efficient and stable enzymes [182]. ML encompasses a diverse set of algorithms and techniques for predictive tasks and pattern recognition, including decision trees, support vector machines, k-nearest neighbors, linear regression models, and neural networks (NNs) [183,184].

NNs, inspired by biological neural structures, become "deep" upon incorporating two or more hidden layers [183]. DL, a subset of ML, has transformed data science by enabling breakthroughs in areas such as computer vision, natural language processing (NLP), autonomous systems, and, of course, protein engineering [185]. Deep neural networks construct hierarchical feature representations, wherein layers proximal to the input data capture elementary patterns while successive, deeper layers deduce increasingly abstract and complex features. These architectures inherently demand access to extensive datasets and substantial computational resources to achieve effective model performance [183].

Advancements in accessible hardware and community-driven tools such as ColabFold [186], supported by Google Colab, have greatly enhanced the reach of ML applications. Hugging Face (huggingface.co) offers a comprehensive suite of DL resources, including a vast repository of pre-trained models and datasets spanning diverse domains, including emerging applications such as protein engineering. Their open-source libraries, such as Transformers [187], enable efficient integration with widely used DL frameworks, includ-

ing PyTorch and TensorFlow, streamlining the construction, training, and deployment of ML models.

3.4.1. ML Paradigms

Supervised learning involves predicting specific properties from labeled datasets by establishing a mapping between inputs (e.g., enzyme sequences or structures) and outputs (e.g., enzyme structures or properties, respectively). This is achieved by minimizing the error between the model's predictions and the provided labels [188]. For example, a training dataset might include enzyme sequences (inputs), annotated/labeled with experimentally determined characteristics such as structures, or properties such as thermal stability, substrate specificity, or catalytic efficiency (outputs). AlphaFold 2.0 serves as a prominent example of a supervised learning-based approach by utilizing sequences as inputs and experimental structures as labels, with the data originating from PDB, and demonstrates exceptional accuracy in predicting protein structures [127].

Unsupervised learning enables the identification of hidden patterns and relationships in unlabeled data by generating outputs that replicate the given input data and using the errors in these outputs to refine the model. It is commonly applied in tasks such as clustering and representation learning. In representation learning, models are trained to encode data in a format that is beneficial for subsequent tasks by mapping high-dimensional inputs to low-dimensional spaces [189], often as part of pre-training [190]. Pre-trained models can then be adapted or fine-tuned for specific downstream applications through transfer learning. This approach is frequently used with large Protein Language Models (PLMs) to minimize further training requirements and effectively leverage smaller labeled datasets [191]. A prime example is the ESM model from Meta's Fundamental AI Research Protein Team, which is trained to predict the identity of randomly masked amino acids in protein sequences [192]. These models can be fine-tuned/trained [193], modified [194], or used as feature extractors/encoders [128] for subsequent supervised learning tasks.

Semi-supervised learning is a hybrid approach that combines a small set of labeled data with a larger set of unlabeled data. By leveraging the structure of the unlabeled data, semi-supervised models improve prediction accuracy by learning a better representation of the input, while reducing reliance on extensive experimental annotations, particularly advantageous in protein engineering, where experimental datasets are often limited [184,195].

Reinforcement Learning (RL) emerges as a powerful tool for exploring protein sequence-function landscapes. In the RL setting, an agent proposes actions (e.g., amino acid substitutions or entire protein sequences) and observes feedback (a "reward") that reflects how well those designs perform (e.g., measured experimentally or predicted by a model). RL methods continually update their strategy (policy) to propose better designs without necessitating a predefined labeled dataset, as the necessary annotations are generated dynamically throughout the experimental process [184]. This paradigm has been successful in simulating directed evolution, either in silico or in combination with in vitro experiments [196,197], and generating compounds predicted to interact effectively with biological targets [198].

3.4.2. Training Datasets

A robust foundation for ML training is provided by accurate and well-structured enzyme databases. PDB, UniProt [163], and other standardized datasets, such as CASP [199], provide sequence and structural data that can be used for representation learning [200], pre-training [192], or training structure predictors [127,128]. NCBI [201], JGI IMG [202], and BacDive [203] databases collectively provide taxonomic, sequence, and cultivation data that have been used in the training of ThermoProt [204] and Preoptem [205], enabling

them to classify enzymes as thermophilic or mesophilic. CAZy and PAZy can be used to train classification models for natural and synthetic polymer-degrading enzymes or organisms based on their substrates [206,207]. BRENDA [208] and Sabio-RK [209] catalog kinetic parameters and are used for training kinetic predictors, such as DLKcat [210] and TurNuP [211], while EnzymeML provides a framework for standardizing kinetic data exchange [212]. For stability, SAPPHIRE [213], ProThermDB [214], and FireProtDB [215] offer sequence and mutation-specific thermostability data, enabling tools such as DeepDDG [85] and ThermoMPNN [216]. BindingDB [217] and PDBbind [218] catalog protein-ligand affinity data, SoluProtMutDB [219] curates mutational data on soluble expression, the SignalP 6.0 [220] dataset supports signal peptide detection, and MutaDescribe [221] provides rich textual annotations for the effects of mutations on proteins. Additionally, databases such as GotEnzymes [222] and AlphaFoldDB [223] have initiated the systematic organization of predictions derived from AI tools, thereby streamlining access to experimentalists and facilitating deeper engagement with the field.

3.4.3. Model Architectures

Traditional ML models, such as random forests [224] and gradient boosting [225], remain valuable for tasks where well-labeled and moderately sized datasets prevail. However, recent breakthroughs in DL model architectures have largely amplified the impact of such databases. Convolutional Neural Networks (CNNs) are a type of DL architecture characterized by hidden layers that are locally connected to subsequent layers through convolutional filters (also called kernels), traditionally used for computer vision. This local connectivity enables CNNs to efficiently extract local features, which are then hierarchically combined into more complex representations [226]. In the field of protein engineering, CNNs have demonstrated great predictive power and capability in learning the fitness landscape of proteins [227]. CNNs have been applied in binding site detection [228], optimal amino acid prediction [111], thermostability estimation [205], and de novo protein design [229], addressing the inverse folding problem.

In recent years, borrowing techniques from NLP, pre-trained PLMs, such as ProtBERT [191] and ESM-2 [128], have leveraged transformer-based architectures with multi-head attention mechanisms. The attention mechanism enables these models to selectively focus on the most relevant aspects of the input by assigning varying levels of importance, or "weights", to different parts of the sequence, thus allowing the models to effectively capture both local and global dependencies across the entire input sequence through the use of multiple attention heads [230]. By learning these patterns, PLMs have become highly effective for predicting protein structures, annotating functions, and assessing the effects of mutations. For example, ProtBERT has been applied in tools such as SignalP 6.0 [220] and BertThermo [231], ESM-2 powers applications such as ESMFold [128] and PepMLM [232], ProtGPT2 enables de novo protein design [233] and, while not technically a PLM, AlphaFold 2.0 is a prominent example of a model that applies attention mechanisms within its Evoformer block to accurately predict protein structures from their sequences [127]. Transformer-based PLMs have largely replaced CNNs in popularity due to their often superior performance. However, it is important to note that CNNs can still be competitive and, in specific contexts, outperform transformer-based architectures [234–236].

In parallel with these developments, diffusion models have recently emerged as promising generative approaches that iteratively refine random noise into structured and coherent outputs and were originally popularized for generating images from text prompts [237,238]. Their applications span docking [237], de novo design with RFdiffusion [238], NNPs [239], and structure prediction with AlphaFold 3 [240]. Graph Neural Networks (GNNs), on the other hand, treat proteins as graphs, where amino acid residues

(or atoms) serve as nodes and edges represent interactions or spatial proximity [241]. As a result, GNN-based methods have shown promise in tasks such as predicting protein-protein interactions [241] and protein solubility [242] in NNPs [243] and de novo sequence design from structure in ProteinMPNN [126,244].

3.4.4. Interpreting ML Models

Interpreting the predictions of ML models in protein engineering is important for uncovering insights into fundamental protein sequence-function or structure-function relationships. Traditional ML models, such as decision trees, random forests, and linear regression, are inherently interpretable. These models clearly show how input features influence predictions, whether through detailed decision pathways, feature importance metrics, or coefficients, making them highly useful for exploratory research [245]. However, understanding the predictions of DL models, which are generally more accurate, can be difficult because of their large size and complex architecture.

CNNs process input data by extracting features using filters, focusing on key regions that significantly impact the output. During training, these filters identify patterns from the input dataset important to predictions, thereby elucidating the relationship between input features, such as structure, and predictive outcomes, such as function [246]. Likewise, attention mechanisms in transformer-based models also enable the identification of critical input regions influencing predictions, by analyzing the attention scores given to each input token, providing insights into the relationship between the input data and the prediction, thereby also enabling the visualization of sequence-function relationships [247]. Unsupervised techniques, such as Sparse Autoencoders, aid in extracting latent representations by compressing high-dimensional data into lower-dimensional, interpretable forms, uncovering patterns across datasets. For example, InterPLM extracted 143 biological concepts (e.g., functional domains and structural motifs) learned by the PLM ESM-2 from its unsupervised training [248]. Additionally, recent advancements and the incorporation of Chain-of-Thought (CoT), prompting strategies in systems such as MutaPLM [221], have demonstrated progress in providing human-readable step-by-step explanations for mutational effects.

4. Lessons from the Industrial Application of Engineered PETases

The escalating issue of plastic waste accumulation has driven significant research into enzymatic strategies for plastic degradation, as traditional waste management techniques, such as incineration and landfill disposal, not only fail to address the scale of this issue but also contribute to additional environmental concerns, including greenhouse gas emissions and microplastic soil contamination [249,250]. Furthermore, traditional thermomechanical recycling methods degrade plastics, such as PET, due to chain breakage, crystallinity increase, and chemical degradation of their building blocks [16]. In this context, enzymes capable of catalyzing the breakdown of polymeric materials offer a sustainable and eco-friendly alternative for addressing the plastic waste crisis [12]. PETases stand out as efficient and potentially transformative biocatalysts for tackling PET pollution and supporting sustainable material reuse in a circular economy [251], with an already industrialized application made possible by protein engineering of LCC [3,4,18].

4.1. Biocatalysis of PET

Enzymes responsible for PET depolymerization are part of the Enzyme Commission number (EC) 3.1.1 class of carboxylic ester hydrolases [208] and feature a catalytic triad (Ser-His-Asp) typical of the α/β hydrolase superfamily. PET degradation occurs in a hydrophobic cleft on the surface of PETases, which facilitates interaction with the polymer.

The hydrolysis of ester bonds begins with the catalytic serine's oxygen atom attacking the carbonyl carbon of the scissile ester bond, leading to bond cleavage (Figure 2) [18]. PETases, either alone or synergistically with mono(2-hydroxyethyl) terephthalate (MHET) hydrolases (MHETases), fully depolymerize PET to its monomers, i.e., ethylene glycol and terephthalic acid [252].

The first highly efficient PET depolymerase, *Thermobifida fusca* cutinase (*Tf*Cut), was reported in 2005 [18,253], bringing significant attention to cutinases and lipases for PET bio-recycling. Since then, other PET-degrading enzymes have been isolated from various taxonomic groups, with benchmark LCC [254] and *Is*PETase [252] being the most widely studied and cited in the field. LCC is recognized as the most effective WT PETase at temperatures exceeding the glass transition temperature of PET (~70 °C), while *Is*PETase demonstrates the best performance at moderate temperatures (<45 °C) [255], with melting temperatures (T_m) of ~84.7 °C and ~46.4 °C, respectively [4].

Protein engineering has played a pivotal role in transitioning PETases from an academic curiosity focus to industrially relevant targets. WT PETases, while effective at degrading PET under laboratory conditions, exhibited limitations such as suboptimal catalytic efficiency and low thermostability at industrially relevant temperatures, as well as poor activity on semi-crystalline PET, rendering them unsuited for industrial applications. Through iterative rounds of rational design, directed evolution, semi-rational design, and ML, these challenges are being systematically addressed [256].

4.2. Protein Engineering of PETases

Engineered variants of benchmark enzymes such as *Is*PETase and LCC have demonstrated outstanding performance in terms of catalytic activity and durability, making them viable candidates for industrial deployment, highlighting the transformative potential of the aforementioned protein engineering approaches in addressing diverse modern challenges through optimized or completely novel enzymatic systems [3]. Table 2 showcases notable applications of protein engineering in PETase optimization. In column 3 (targeted properties), the first property mentioned corresponds to the main target of each study, while the rest were of secondary consideration to the researchers; this emphasizes the synergistic integration of multiple protein engineering strategies for optimal results. Figure 6 highlights all residues of *Is*PETase and LCC modified in the studies presented in Table 2, demonstrating their diverse positions, not solely concentrated close to the active site but all over the enzyme, inside and outside.

Rational design examples are plentiful in PETase engineering, mostly aiming for increased stability and modulating PET binding affinity. The most influential PETase engineering study is the engineering of LCC to LCC^{ICCG} (Table 2, study 10), which was later industrialized for PET bio-recycling. In this study, a disulfide bond was introduced at sites D238C and S283C to increase stability, which resulted in a ΔT_m = +10 °C, albeit with a 28% decreased activity (LCC^{CC}). Point mutation F243I on LCC^{CC} regained activity 22% higher than WT but decreased ΔT_m to +6.6 °C (LCC^{ICC}). The latter was proposed by identifying hotspots through in-silico docking studies, SSM, and experimental screening. Finally, Y127G from the same docking and SSM study did not affect activity but increased melting temperature, resulting in LCC^{ICCG} and demonstrating ΔT_m = +9.3 °C and 82% PET conversion in 20 h compared to 53% of WT at 72 °C [4]. Study 11 (LCC^{ICCG}-RIP) focused on (i) introducing proline residues, (ii) introducing hydrophilic interactions on the surface, and iii) increasing internal hydrophobic interactions to further increase the stability of LCC^{ICCG} [257]. The GRAPE strategy, which includes $\Delta\Delta G_f$ calculations with FoldX and Rosetta, among others, for identification of stabilizing mutations, was utilized to engineer DuraPETase (Table 2 study 6), an *Is*PETase variant, and TurboPETase (Table 2 study 16), a

*Bhr*PETase variant [258,259]. For engineering substrate affinity, Studies 8 (CombiPETase), 9 (TS-Δ*Is*PET), and 14 (LCC-A2) all employed docking methodologies to identify hotspot residues in the active site of the respective enzymes, significantly increasing activity in all cases [257,260,261].

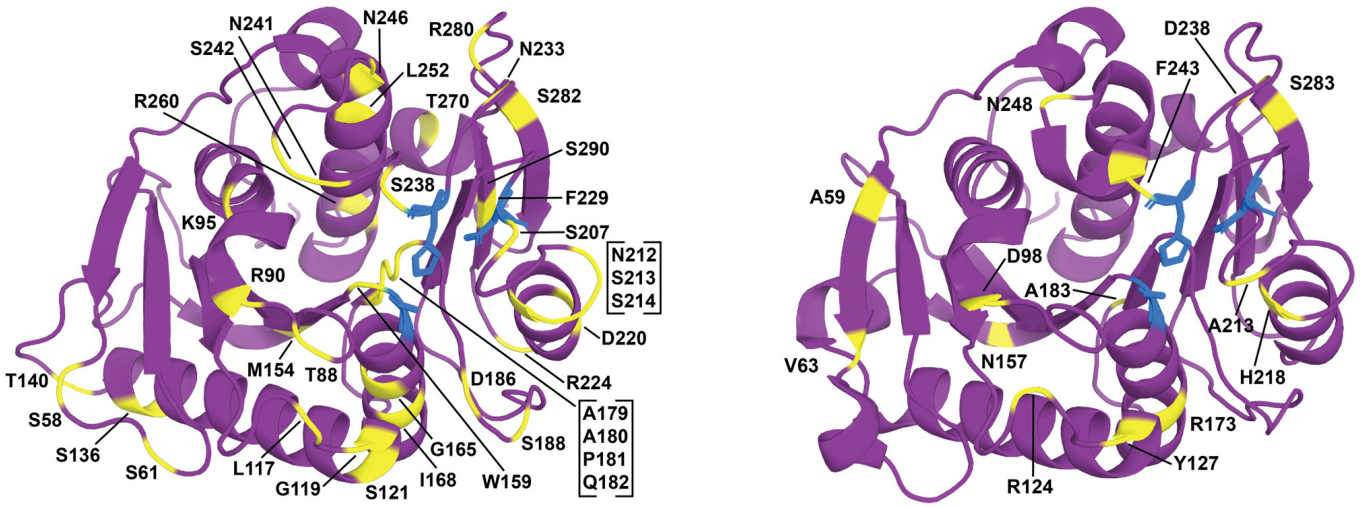

Figure 6. Cartoon representation of the (**a**) *Ideonella sakaiensis* PETase (*Is*PETase) and the (**b**) leaf and branch compost cutinase (LCC). WT residues are shown in purple, while modified residues are shown in yellow and are also labeled (see Table 2 for reference). The catalytic triad residues are shown as blue sticks. PDB IDs: 6IWL for *Is*PETase and 4EB0 for LCC.

Directed evolution approaches for evolving *Is*PETase stability are represented by studies 1 and 2 of Table 2, which resulted in a significant increase in T_m with multiple mutations [262,263]. In the case of DepoPETase, semi-rational design was also used to further identify mutation R260Y via focused SSM of positively charged amino acids on the opposite side of the substrate-binding pocket, and D186H and N233K were obtained from the literature, further refining the obtained variant [263].

Semi-rational design approaches were also used for engineering *Is*PETase (Table 2 studies 5, 7, 8, and 9) and LCC (Table 2 study 12). Study 3 identified mutations S121D/D186H for *Is*PETase based on structural comparison with *Tf*Cut, where these residues generate a hydrogen bond that stabilizes the β6-β7 loop [264]. In study 8, an ASR approach was used to identify *Is*PETase mutation K95N, which exhibited an increase in thermal stability and activity [260]. In study 5, eight hotspots were identified in the binding site of *Is*PETase by examining four homologs (smart library) and deeming them variable. After focused mutagenesis, a combinatorial approach and insights from the literature yielded the final variant that exhibited 58-fold increased activity at 37 °C compared to WT [265]. Studies 7 and 9 employed similar concepts, utilizing *Is*PETase's homologs to calculate the likelihood and conservation of amino acids in specific positions, respectively. In the first case (study 7), the Premuse tool was developed to calculate position-specific amino acid probabilities from a library of homologs, guiding mutation selection and generating a variant with $\Delta T_m = +10.4$ °C and 40-fold activity increase at 40 °C in 24 h [266]. In the later study (study 9), conservation analysis was used to avoid the substitution of highly conserved residues using Rate4Site [86,261]. Additionally, study 12 identified hotspots on LCC^{ICCG} based on two approaches. First, a conservation scheme categorized 4203 homologous proteins into high and low-temperature datasets, based on scoring from the DL-based tool Preoptem, and determined the probability of amino acids at each position, identifying 18 candidate mutations that were not only conserved in the high-temperature dataset but also absent

from the low-temperature dataset and the target protein. The second approach followed a coevolutionary scheme utilizing EVcouplings to identify hotspots, which were subsequently screened through EVmutation and Preoptem scoring functions, further identifying 18 additional mutations. Experimental validation independently screened six beneficial mutations and LCC^{ICCG}_I6M. Incorporating all six demonstrated an increased T_{opt} from 65 °C to 75–80 °C for 39% crystalline PET [267].

Machine learning has been used to improve *Bhr*PETase (Table 2, study 16), *Is*PETase (Table 2, studies 4, 6, and 7), and LCC (Table 2, study 12). In the case of *Bhr*PETase, a PLM, trained with a masked language modeling objective on ~26,000 homologous sequences to predict the real amino acid at the masked position, was used to suggest mutations that improved activity but resulted in decreased thermal stability. After the rational design of stabilizing mutations, the final variant, with a $T_m = 84$ °C and a 3.4-fold improvement in specific activity towards PET films compared to WT, was obtained [259]. In the case of *Is*PETase variant FAST-PETase (study 4), the CNN-based MūtCompute, trained to predict the masked amino acid at the center of a chemical environment extracted from a protein structure [268], was used to obtain a discrete probability distribution for the structural fit of all 20 canonical amino acids at every position, identifying mutations S121E, T140D, R224Q, and N233K and combinatorically assembling them across *Is*PETase, Thermo-PETase, and DuraPETase to obtain the best variant containing N233K and R224Q on top of Thermo-PETase [111].

Other non-typical approaches include: (i) study 15, in which LCC was expressed in *Pichia pastoris,* increasing thermal stability and activity through the introduction of N-glycosylation [117], (ii) study 18, in which an active site loop from LCC was grafted to Mors1, resulting in a shift in optimal temperature from 25 °C to 45 °C and a 5-fold increase in PET hydrolysis compared with WT at 25 °C [122], (iii) study 19, in which a PETase was designed de novo [269], and (iv) study 20, in which *Is*PETase was fused with *Is*MHETase, separated with a linker, improving turnover relative to the free enzymes [270]. An unconventional but notable filter in study 14 (LCC-YGA) incorporated a correlation-based accumulated mutagenesis (CAM) strategy that accounts for the amino acids exhibiting highly correlated or anti-correlated motions. Through MD simulations, the correlation is quantified by the covariance between the fluctuations of two atoms. Mutations are then introduced in regions with little cross-correlated dynamics [271].

Noteworthily, optimizing both the enzyme's active site and distal regions is crucial for enhancing catalytic efficiency and overall stability (Figure 6). For instance, incorporation of disulfide bonds, proline residues, and other stability-enhancing mutations consistently raises the melting temperature and broadens the operational range of PETases, typically without overly compromising enzymatic activity. Additionally, rational and semi-rational approaches synergize effectively with directed evolution and ML-driven approaches to pinpoint beneficial substitutions at a scale and speed not achievable through trial-and-error alone. The most comprehensive studies integrate multiple methodologies and systematically rationalize obtained results, even when they initially manifest as random changes.

Table 2. Protein engineering examples for enhancing PETases.

#	Enzyme	Targeted Properties	Engineering Strategies	Modifications	Results	Ref
1	HotPETase	stability, activity	DE	*Is*PETase variant: S121E, D186H, R280A, P181V, S207R, S214Y, Q119K, S213E, N233C, S282C, R90T, Q182M, N212K, R224L, S58A, S61V, K95N, M154G, N241C, K252M, T270Q	ΔT_m = +35.5 °C	[262]
2	DepoPETase	stability, affinity, activity	DE, SRD (focused surface charge mutations with SSM), literature	*Is*PETase variant: T88I, D186H, D220N, N233K, N246D, R260Y, S290P	ΔT_m = +23.3 °C 1407-fold more products towards amorphous PET film at 50 °C	[263]
3	Thermo-PETase	stability, activity	RD (structure-based approach), SRD (adopting features from homolog TfCUT), literature	*Is*PETase variant: S121E, D186H, R280A	ΔT_m = +8.81 °C activity enhanced by 14-fold at 40 °C	[264]
4	FAST-PETase	stability	ML	Thermo-PETase variant: N233K, R224Q	2.4- and 38- fold higher activity at 40 and 50 °C, respectively, compared to ThermoPETase	[111]
5	*Is*PETase variant	affinity, stability	SRD (smart libraries from homologs), literature	*Is*PETase variant: S121E, D186H, S242T, N246D (based on Thermo-PETase)	ΔT_m = +12 °C 58-fold increased activity at 37 °C	[265]
6	DuraPETase	stability, activity	RD ($\Delta\Delta G_f$)	*Is*PETase variant: S214H, I168R, W159H, S188Q, R280A, A180I, G165A, Q119Y, L117F, T140D	Enhanced degradation performance (300-fold) on semicrystalline PET films at 40 °C	[258]
7	*Is*PETase variant	stability, activity	SRD (position-specific amino acid probabilities)	*Is*PETase variant: W159H, F229Y	ΔT_m = +10.4 °C 40-fold activity increase at 40 °C in 24 h	[266]
8	CombiPETase	affinity, stability, activity	RD (MD, engineering flexibility engineering, disulfides, hydrophobic core packing, hydrogen bond breaking), SRD (ASR), literature	*Is*PETase variant: K95N, S136E, A179C, D186A, S214T, N233C, S282C	ΔT_m = +27.2 °C 4.25-fold increased activity when compared to WT at their respective T_{opt} 24.6-fold increased protein yield	[260]
9	TS-Δ*Is*PET	activity, affinity, stability	RD (identifying hotspots through protein-ligand interaction analysis with MD, rational mutations, salt bridge), SRD (conservation analysis followed by SSM), literature	*Is*PETase variant: S121E, W159H, D186H, F238A	ΔT_m = +4.9 °C Increased catalytic activity on PET	[261]
10	LCC^{ICCG}	stability	RD (docking to identify hotspots followed by SSM, disulfide design)	LCC variant: F243I, D238C, S283C, Y127G	ΔT_m = +9.3 °C 82% PET conversion in 20 h compared to 53% of WT at 72 °C	[4]
11	LCC^{ICCG}_RIP	stability	RD (proline residues, hydrophilic surface, hydrophobic core)	LCC^{ICCG} variant: A59R, V63I, N248P	More products at 85 °C	[272]

Table 2. *Cont.*

#	Enzyme	Targeted Properties	Engineering Strategies	Modifications	Results	Ref
12	LCC^{ICCG}_I6M	activity, stability	ML, SRD (coevolutionary analysis)	LCC^{ICCG} variant: S32L, D18T, S98R, T157P, E173Q, N213P	T_{opt} for 39% crystalline PET increased from 65 °C to 75–80 °C	[267]
13	LCC-A2	affinity	RD (docking)	LCC^{ICCG} variant: H218Y, N248D	ΔT_m = +1.11 °C Increased relative activity by 80.1% at 78 °C compared to LCCICCG	[257]
14	LCC-YGA	affinity, activity	RD (remodeling hydrophilicity of binding site, correlation based accumulated mutagenesis strategy), SRD (homolog information), literature	LCC^{ICCG} variant: H183Y, L124G, S29A	2.07-fold hydrolytic activity of LCCICCG	[271]
15	LCC-G	stability	RD (glycosylation)	Introduction of N-linked glycosylation at sites N197, N239, and N266 by expressing WT LCC in *Pichia pastoris*	Increased T_m, at 70 and 75 °C, 1.6- and 1.2-fold more active, respectively	[117]
16	TurboPETase	stability, activity	ML, RD ($\Delta\Delta G_f$), literature	*Bhr*PETase variant: H218S, F222I, W104L, F243T, A209R, D238K, A251C, A281C	T_m = 84 °C and a 3.4-fold improvement in specific activity towards GF-PET films	[259]
17	Est1 variant	stability	SRD (consensus design)	Est1 variant A68V, T253P	Increased T_m and activity	[273]
18	Mors1 chimera	activity	SRD (loop exchange)	Loop exchange of an active site loop from LCC	Shift in optimal temperature from 25 °C to 45 °C, increase 5x in PET hydrolysis when compared with WT at 25 °C.	[122]
19	HSH-25	De novo PETase activity	RD (de novo)	De novo design of a 25 amino acid thermostable peptide capable of depolymerizing PET	Confirmed degradation of PET	[269]
20	*Is*PETase-*Is*MHETase chimera	activity	RD (fusion with linker to achieve synergistic action)	Construction of a bifunctional chimeric enzyme fusion of *Is*PETase with *Is*MHETase	Chimeric proteins of varying linker lengths all exhibit improved turnover relative to the free enzymes	[270]

Note: DE: directed evolution, ASR: ancestral sequence reconstruction, *Bhr*PETase: bacterium HR29 polyethylene terephthalate hydrolase, Est1: *Thermobifida alba* AHK119 cutinase, *Is*MHETase: *Ideonella sakaiensis* mono(2-hydroxyethyl) terephthalate hydrolase, *Is*PETase: *Ideonella sakaiensis* PETase; LCC: leaf and branch compost cutinase, ML: machine learning, Mors1: *Moraxella sp.* TA144 cutinase, RD: rational design, SRD: semi-rational design, SSM: site saturation mutagenesis, *Tf*CUT: *Thermobifida fusca* cutinase.

5. Conclusions and Prospects

This paper presents a comprehensive review of protein engineering strategies aimed at enhancing enzyme performance for industrial applications, focusing on key principles, methods, and lessons learned from industrialized PETases. These efforts illustrate the transformative potential of protein engineering in addressing industrial and environmental challenges.

Interdisciplinary collaborations that combine computational enzyme design groups with experimentalists and dedicated AI researchers hold the potential to unlock new frontiers in protein engineering. Looking forward, the integration of advanced DL frameworks, such as diffusion models and PLMs, offers exciting prospects for accelerating enzyme design and bridging the gap between laboratory innovation and industrial implementation, potentially rendering the protein engineering problem trivial through text prompts to fully functional designs, just like recent text-to-image models.

Despite their successes, the strategies reviewed exhibit notable limitations. Rational design, while highly effective, relies on detailed structural and mechanistic data, which are not always available or straightforward to obtain. Directed evolution, although powerful, is constrained due to the vast number of potential sequence combinations, experimental biases, and the need for efficient, high-throughput screening methods. Semi-rational design remains reliant on comprehensive evolutionary insights which may not always be available or translate effectively into smart library designs. ML approaches show substantial promise but often require large, high-quality datasets and considerable computational resources, which can limit their practical application.

Scaling these strategies for industrial applications presents additional hurdles. All of these approaches require significant investments in time, expertise, and infrastructure, which may not always align with the cost-sensitive nature of industrial biotechnology. While engineered enzymes may demonstrate significantly enhanced activity, stability, and specificity, their overall production costs, including research, design, scale-up, and the expected return on investment from societal, environmental, and financial perspectives must be carefully evaluated to ensure economic viability. Future efforts should prioritize addressing scalability and cost-effectiveness, alongside a more holistic evaluation of engineered enzymes' lifecycle impacts. Addressing these factors will help establish protein engineering as a foundation for versatile and innovative industrial biotechnology.

Author Contributions: K.G.: Methodology, Validation, Investigation, Writing—original draft, Writing—review and editing, Visualization. C.F.: Validation, Writing—review and editing. E.T.: Conceptualization, Writing—review and editing, Supervision, Funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: The research work was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) through the research project EnZyReMix (Project Number: 015024), implemented under H.F.R.I.'s call "Basic Research Financing (Horizontal Support of all Sciences)" within the framework of the National Recovery and Resilience Plan Greece 2.0, funded by the European Union–NextGenerationEU (Implementation body: HFRI).

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

ASR	Ancestral Sequence Reconstruction
BLAST	Basic Local Alignment Search Tool
CAM	Correlation-based Accumulated Mutagenesis
CAZy	Carbohydrate-Active enZYmes
CNNs	Convolutional Neural Networks
CoT	Chain-of-Thought
DFT	Density Functional Theory
DL	Deep Learning
epPCR	error-prone PCR
GNNs	Graph Neural Networks
LCC	Leaf Compost Cutinase
MD	Molecular Dynamics
MHET	Mono(2-hydroxyethyl) terephthalate
MHETases	MHET hydrolases
ML	Machine Learning
MMGBSA	Molecular Mechanics Generalized Born Surface Area
MMPBSA	Molecular Mechanics Poisson–Boltzmann Surface Area
MMRT	Macromolecular Rate Theory
MMseq2	Many-against-Many sequence searching
nanoDSF	nano-Differential Scanning Fluorimetry
NLP	Natural Language Processing
NNs	Neural Networks
PAZy	Plastics-Active enZYmes
PB	Poisson–Boltzmann
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PET	Polyethylene terephthalate
PETases	PET hydrolases
PLMs	Protein Language Models
RL	Reinforcement Learning
SSM	Site Saturation Mutagenesis
WT	Wild-type

References

1. Dietrich, O.; Heun, M.; Notroff, J.; Schmidt, K.; Zarnkow, M. The Role of Cult and Feasting in the Emergence of Neolithic Communities. New Evidence from Göbekli Tepe, South-Eastern Turkey. *Antiquity* **2012**, *86*, 674–695. [CrossRef]
2. Leek, F.F. Teeth and Bread in Ancient Egypt. *J. Egypt. Archaeol.* **1972**, *58*, 126–132. [CrossRef]
3. Arnal, G.; Anglade, J.; Gavalda, S.; Tournier, V.; Chabot, N.; Bornscheuer, U.T.; Weber, G.; Marty, A. Assessment of Four Engineered PET Degrading Enzymes Considering Large-Scale Industrial Applications. *ACS Catal.* **2023**, *13*, 13156–13166. [CrossRef] [PubMed]
4. Tournier, V.; Topham, C.M.; Gilles, A.; David, B.; Folgoas, C.; Moya-Leclair, E.; Kamionka, E.; Desrousseaux, M.L.; Texier, H.; Gavalda, S.; et al. An Engineered PET Depolymerase to Break down and Recycle Plastic Bottles. *Nature* **2020**, *580*, 216–219. [CrossRef]
5. Xia, W.; Xu, X.; Qian, L.; Shi, P.; Bai, Y.; Luo, H.; Ma, R.; Yao, B. Engineering a Highly Active Thermophilic β-Glucosidase to Enhance Its PH Stability and Saccharification Performance. *Biotechnol. Biofuels* **2016**, *9*, 147. [CrossRef] [PubMed]
6. Cui, H.; Stadtmüller, T.H.J.; Jiang, Q.; Jaeger, K.E.; Schwaneberg, U.; Davari, M.D. How to Engineer Organic Solvent Resistant Enzymes: Insights from Combined Molecular Dynamics and Directed Evolution Study. *ChemCatChem* **2020**, *12*, 4073–4083. [CrossRef]
7. McDonald, A.D.; Higgins, P.M.; Buller, A.R. Substrate Multiplexed Protein Engineering Facilitates Promiscuous Biocatalytic Synthesis. *Nat. Commun.* **2022**, *13*, 5242. [CrossRef] [PubMed]

8. Victorino da Silva Amatto, I.; Gonsales da Rosa-Garzon, N.; Antônio de Oliveira Simões, F.; Santiago, F.; Pereira da Silva Leite, N.; Raspante Martins, J.; Cabral, H. Enzyme Engineering and Its Industrial Applications. *Biotechnol. Appl. Biochem.* **2022**, *69*, 389–409. [CrossRef]
9. Tang, W.L.; Zhao, H. Industrial Biotechnology: Tools and Applications. *Biotechnol. J.* **2009**, *4*, 1725–1739. [CrossRef] [PubMed]
10. Han, S.S.; Kyeong, H.H.; Choi, J.M.; Sohn, Y.K.; Lee, J.H.; Kim, H.S. Engineering of the Conformational Dynamics of an Enzyme for Relieving the Product Inhibition. *ACS Catal.* **2016**, *6*, 8440–8445. [CrossRef]
11. Borges, P.T.; Silva, D.; Silva, T.F.D.; Brissos, V.; Cañellas, M.; Lucas, M.F.; Masgrau, L.; Melo, E.P.; Machuqueiro, M.; Frazão, C.; et al. Unveiling Molecular Details behind Improved Activity at Neutral to Alkaline pH of an Engineered DyP-Type Peroxidase. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 3899–3910. [CrossRef] [PubMed]
12. Zhu, B.; Wang, D.; Wei, N. Enzyme Discovery and Engineering for Sustainable Plastic Recycling. *Trends Biotechnol.* **2022**, *40*, 22–37. [CrossRef] [PubMed]
13. Porebski, B.T.; Buckle, A.M.; By, E.; Daggett, V. Consensus Protein Design. *PEDS* **2016**, *29*, 245–251. [CrossRef]
14. OECD Global Plastics Outlook-Plastics Use by Polymer. Available online: https://ourworldindata.org/grapher/plastic-production-polymer (accessed on 29 January 2025).
15. Sarda, P.; Hanan, J.C.; Lawrence, J.G.; Allahkarami, M. Sustainability Performance of Polyethylene Terephthalate, Clarifying Challenges and Opportunities. *J. Polym. Sci.* **2022**, *60*, 7–31. [CrossRef]
16. Del Mar Castro López, M.; Ares Pernas, A.I.; Abad López, M.J.; Latorre, A.L.; López Vilariño, J.M.; González Rodríguez, M.V. Assessing Changes on Poly(Ethylene Terephthalate) Properties after Recycling: Mechanical Recycling in Laboratory versus Postconsumer Recycled Material. *Mater. Chem. Phys.* **2014**, *147*, 884–894. [CrossRef]
17. Kaabel, S.; Daniel Therien, J.P.; Deschênes, C.E.; Duncan, D.; Friščic, T.; Auclair, K. Enzymatic Depolymerization of Highly Crystalline Polyethylene Terephthalate Enabled in Moist-Solid Reaction Mixtures. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2026452118. [CrossRef] [PubMed]
18. Tournier, V.; Duquesne, S.; Guillamot, F.; Cramail, H.; Taton, D.; Marty, A.; André, I. Enzymes' Power for Plastics Degradation. *Chem. Rev.* **2023**, *123*, 5612–5701. [CrossRef] [PubMed]
19. Robinson, P.K. Enzymes: Principles and Biotechnological Applications. *Essays Biochem.* **2015**, *59*, 1. [CrossRef] [PubMed]
20. Miller, S.R. An Appraisal of the Enzyme Stability-Activity Trade-Off. *Evolution* **2017**, *71*, 1876–1887. [CrossRef] [PubMed]
21. Eijsink, V.G.H.; Bjørk, A.; Gåseidnes, S.; Sirevåg, R.; Synstad, B.; Burg, B.V.D.; Vriend, G. Rational Engineering of Enzyme Stability. *J. Biotechnol.* **2004**, *113*, 105–120. [CrossRef] [PubMed]
22. Wolfenden, R.; Snider, M.; Ridgway, C.; Miller, B. The Temperature Dependence of Enzyme Rate Enhancements. *J. Am. Chem. Soc.* **1999**, *121*, 7419–7420. [CrossRef]
23. Kavanau, J.L. Enzyme Kinetics and the Rate of Biological Processes. *J. Gen. Physiol.* **1950**, *34*, 193–209. [CrossRef]
24. Thomas, T.M.; Scopes, R.K. The Effects of Temperature on the Kinetics and Stability of Mesophilic and Thermophilic 3-Phosphoglycerate Kinases. *Biochem. J.* **1998**, *330*, 1087–1095. [CrossRef] [PubMed]
25. Daniel, R.M.; Danson, M.J. A New Understanding of How Temperature Affects the Catalytic Activity of Enzymes. *Trends Biochem. Sci.* **2010**, *35*, 584–591. [CrossRef] [PubMed]
26. Arcus, V.L.; Prentice, E.J.; Hobbs, J.K.; Mulholland, A.J.; Van Der Kamp, M.W.; Pudney, C.R.; Parker, E.J.; Schipper, L.A. On the Temperature Dependence of Enzyme-Catalyzed Rates. *Biochemistry* **2016**, *55*, 1681–1688. [CrossRef] [PubMed]
27. Eyring, H. The Activated Complex in Chemical Reactions. *J. Chem. Phys.* **1935**, *3*, 107–115. [CrossRef]
28. Daniel, R.M.; Danson, M.J.; Eisenthal, R. The Temperature Optima of Enzymes: A New Perspective on an Old Phenomenon. *Trends Biochem. Sci.* **2001**, *26*, 223–225. [CrossRef] [PubMed]
29. Hei, D.J.; Clark, D.S. Estimation of Melting Curves from Enzymatic Activity–Temperature Profiles. *Biotechnol. Bioeng.* **1993**, *42*, 1245–1251. [CrossRef]
30. Daniel, R.M.; Peterson, M.E.; Danson, M.J.; Price, N.C.; Kelly, S.M.; Monk, C.R.; Weinberg, C.S.; Oudshoorn, M.L.; Lee, C.K. The Molecular Basis of the Effect of Temperature on Enzyme Activity. *Biochem. J.* **2010**, *425*, 353–360. [CrossRef]
31. Klotz, I.M.; Robert, M. Rosenberg The Second Law of Thermodynamics. In *Chemical Thermodynamics: Basic Concepts and Methods*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2008; pp. 111–157. ISBN 978-0-471-78015-1.
32. Klotz, I.M.; Robert, M. Rosenberg Applications of the First Law to Gases. In *Chemical Thermodynamics: Basic Concepts and Methods*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2008; pp. 81–109. ISBN 978-0-471-78015-1.
33. Klotz, I.M.; Robert, M. Rosenberg Equilibrium and Spontaneity for Systems at Constant Temperature. In *Chemical Thermodynamics: Basic Concepts and Methods*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2008; pp. 159–191. ISBN 978-0-471-78015-1.
34. Vieille, C.; Zeikus, G.J. Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability. *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 1–43. [CrossRef]
35. Johnson, C.M. Differential Scanning Calorimetry as a Tool for Protein Folding and Stability. *Arch. Biochem. Biophys.* **2013**, *531*, 100–109. [CrossRef] [PubMed]

36. Greenfield, N.J. Using Circular Dichroism Collected as a Function of Temperature to Determine the Thermodynamics of Protein Unfolding and Binding Interactions. *Nat. Protoc.* **2006**, *1*, 2527. [CrossRef]
37. Santiago, P.S.; Moura, F.; Moreira, L.M.; Domingues, M.M.; Santos, N.C.; Tabak, M. Dynamic Light Scattering and Optical Absorption Spectroscopy Study of PH and Temperature Stabilities of the Extracellular Hemoglobin of Glossoscolex Paulistus. *Biophys. J.* **2007**, *94*, 2228. [CrossRef] [PubMed]
38. Gao, K.; Oerlemans, R.; Groves, M.R. Theory and Applications of Differential Scanning Fluorimetry in Early-Stage Drug Discovery. *Biophys. Rev.* **2020**, *12*, 85. [CrossRef] [PubMed]
39. Fitter, J. Structural and Dynamical Features Contributing to Thermostability in α-Amylases. *Cell. Mol. Life Sci.* **2005**, *62*, 1925–1937. [CrossRef] [PubMed]
40. Sharma, A.; Gupta, G.; Ahmad, T.; Mansoor, S.; Kaur, B. Enzyme Engineering: Current Trends and Future Perspectives. *Food Rev. Int.* **2021**, *37*, 121–154. [CrossRef]
41. Harris, T.K.; Turner, G.J. Structural Basis of Perturbed pKa Values of Catalytic Groups in Enzyme Active Sites. *IUBMB Life* **2002**, *53*, 85–98. [CrossRef]
42. Fitch, C.A.; Platzer, G.; Okon, M.; Garcia-Moreno, B.E.; McIntosh, L.P. Arginine: Its pKa Value Revisited. *Protein Sci.* **2015**, *24*, 752–761. [CrossRef] [PubMed]
43. Dobrev, P.; Vemulapalli, S.P.B.; Nath, N.; Griesinger, C.; Grubmüller, H. Probing the Accuracy of Explicit Solvent Constant PH Molecular Dynamics Simulations for Peptides. *J. Chem. Theory Comput.* **2020**, *16*, 2561–2569. [CrossRef]
44. Lin, J.; Cassidy, C.S.; Frey, P.A. Correlations of the Basicity of His 57 with Transition State Analogue Binding, Substrate Reactivity, and the Strength of the Low-Barrier Hydrogen Bond in Chymotrypsin. *Biochemistry* **1998**, *37*, 11940–11948. [CrossRef]
45. Bender, M.L.; Killheffer, J.V.; Cohen, S. Chymotrypsin. *Crit. Rev. Biochem. Mol. Biol.* **1973**, *1*, 149–199. [CrossRef]
46. Hess, G.P.; McConn, J.; Ku, E.; McConkey, G. Studies of the Activity of Chymotrypsin. *Philos. Trans. R. Soc. London. B Biol. Sci.* **1970**, *257*, 89–104. [CrossRef] [PubMed]
47. Hofer, F.; Kraml, J.; Kahler, U.; Kamenik, A.S.; Liedl, K.R. Catalytic Site pKa Values of Aspartic, Cysteine, and Serine Proteases: Constant PH MD Simulations. *J. Chem. Inf. Model.* **2020**, *60*, 3030–3042. [CrossRef] [PubMed]
48. Li, H.; Robertson, A.D.; Jensen, J.H. Very Fast Empirical Prediction and Rationalization of Protein pKa Values. *Proteins* **2005**, *61*, 704–721. [CrossRef] [PubMed]
49. Reis, P.B.P.S.; Vila-Vicosa, D.; Rocchia, W.; Machuqueiro, M. PypKA: A Flexible Python Module for Poisson-Boltzmann-Based pKa Calculations. *J. Chem. Inf. Model.* **2020**, *60*, 4442–4448. [CrossRef] [PubMed]
50. Havranek, J.J.; Harbury, P.B. Tanford-Kirkwood Electrostatics for Protein Modeling. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11145–11150. [CrossRef]
51. Thapa, B.; Schlegel, H.B. Density Functional Theory Calculation of pKa's of Thiols in Aqueous Solution Using Explicit Water Molecules and the Polarizable Continuum Model. *J. Phys. Chem. A* **2016**, *120*, 5726–5735. [CrossRef] [PubMed]
52. Cai, Z.; Liu, T.; Lin, Q.; He, J.; Lei, X.; Luo, F.; Huang, Y. Basis for Accurate Protein pKa Prediction with Machine Learning. *J. Chem. Inf. Model.* **2023**, *63*, 2936–2947. [CrossRef] [PubMed]
53. Johnston, R.C.; Yao, K.; Kaplan, Z.; Chelliah, M.; Leswing, K.; Seekins, S.; Watts, S.; Calkins, D.; Chief Elk, J.; Jerome, S.V.; et al. Epik: pKa and Protonation State Prediction through Machine Learning. *J. Chem. Theory Comput.* **2023**, *19*, 2380–2388. [CrossRef] [PubMed]
54. Ho, J.; Coote, M.L. First-Principles Prediction of Acidities in the Gas and Solution Phase. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 649–660. [CrossRef]
55. Alongi, K.S.; Shields, G.C. Theoretical Calculations of Acid Dissociation Constants: A Review Article. *Annu. Rep. Comput. Chem.* **2010**, *6*, 113–138. [CrossRef]
56. Seybold, P.G.; Shields, G.C. Computational Estimation of pKa Values. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2015**, *5*, 290–297. [CrossRef]
57. Martins de Oliveira, V.; Liu, R.; Shen, J. Constant PH Molecular Dynamics Simulations: Current Status and Recent Applications. *Curr. Opin. Struct. Biol.* **2022**, *77*, 102498. [CrossRef] [PubMed]
58. Xiang, C.; Ao, Y.F.; Höhne, M.; Bornscheuer, U.T. Shifting the PH Optima of (R)-Selective Transaminases by Protein Engineering. *Int. J. Mol. Sci.* **2022**, *23*, 15347. [CrossRef]
59. Thomas, P.G.; Russell, A.J.; Fersht, A.R. Tailoring the PH Dependence of Enzyme Catalysis Using Protein Engineering. *Nature* **1985**, *318*, 375–376. [CrossRef]
60. Kishore, D.; Kundu, S.; Kayastha, A.M. Thermal, Chemical and PH Induced Denaturation of a Multimeric β-Galactosidase Reveals Multiple Unfolding Pathways. *PLoS ONE* **2012**, *7*, e50380. [CrossRef] [PubMed]
61. Wei, Z.; Song, J. Molecular Mechanism Underlying the Thermal Stability and PH-Induced Unfolding of CHABII. *J. Mol. Biol.* **2005**, *348*, 205–218. [CrossRef]
62. Xiang, L.; Lu, Y.; Wang, H.; Wang, M.; Zhang, G. Improving the Specific Activity and pH Stability of Xylanase XynHBN188A by Directed Evolution. *Bioresour. Bioprocess.* **2019**, *6*, 25. [CrossRef]

63. Selvaraj, C.; Rudhra, O.; Alothaim, A.S.; Alkhanani, M.; Singh, S.K. Structure and Chemistry of Enzymatic Active Sites That Play a Role in the Switch and Conformation Mechanism. *Adv. Protein Chem. Struct. Biol.* **2022**, *130*, 59–83. [CrossRef] [PubMed]
64. Koshland, D.E. The Key–Lock Theory and the Induced Fit Theory. *Angew. Chem. Int. Ed. Engl.* **1995**, *33*, 2375–2378. [CrossRef]
65. Paul, F.; Weikl, T.R. How to Distinguish Conformational Selection and Induced Fit Based on Chemical Relaxation Rates. *PLoS Comput. Biol.* **2016**, *12*, e1005067. [CrossRef]
66. Harris, J.L.; Craik, C.S. Engineering Enzyme Specificity. *Curr. Opin. Chem. Biol.* **1998**, *2*, 127–132. [CrossRef]
67. De Groeve, M.R.M.; Remmery, L.; Van Hoorebeke, A.; Stout, J.; Desmet, T.; Savvides, S.N.; Soetaert, W. Construction of Cellobiose Phosphorylase Variants with Broadened Acceptor Specificity towards Anomerically Substituted Glucosides. *Biotechnol. Bioeng.* **2010**, *107*, 413–420. [CrossRef] [PubMed]
68. Li, J.X.; Fang, X.; Zhao, Q.; Ruan, J.X.; Yang, C.Q.; Wang, L.J.; Miller, D.J.; Faraldos, J.A.; Allemann, R.K.; Chen, X.Y.; et al. Rational Engineering of Plasticity Residues of Sesquiterpene Synthases from *Artemisia annua*: Product Specificity and Catalytic Efficiency. *Biochem. J.* **2013**, *451*, 417–426. [CrossRef]
69. Ringe, D.; Petsko, G.A. How Enzymes Work. *Science* **2008**, *320*, 1428–1429. [CrossRef] [PubMed]
70. Warshel, A.; Sharma, P.K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M.H.M. Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **2006**, *106*, 3210–3235. [CrossRef] [PubMed]
71. Pedersen, J.N.; Zhou, Y.; Guo, Z.; Pérez, B. Genetic and Chemical Approaches for Surface Charge Engineering of Enzymes and Their Applicability in Biocatalysis: A Review. *Biotechnol. Bioeng.* **2019**, *116*, 1795–1812. [CrossRef] [PubMed]
72. Chaloupková, R.; Sýkorová, J.; Prokop, Z.; Jesenská, A.; Monincová, M.; Pavlová, M.; Tsuda, M.; Nagata, Y.; Damborský, J. Modification of Activity and Specificity of Haloalkane Dehalogenase from *Sphingomonas paucimobilis* UT26 by Engineering of Its Entrance Tunnel. *J. Biol. Chem.* **2003**, *278*, 52622–52628. [CrossRef]
73. Morris, G.M.; Ruth, H.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791. [CrossRef] [PubMed]
74. Trott, O.; Olson, A.J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization and Multithreading. *J. Comput. Chem.* **2010**, *31*, 455. [CrossRef] [PubMed]
75. Miller, B.R.; McGee, T.D.; Swails, J.M.; Homeyer, N.; Gohlke, H.; Roitberg, A.E. MMPBSA.Py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321. [CrossRef] [PubMed]
76. Ylilauri, M.; Pentikäinen, O.T. MMGBSA as a Tool to Understand the Binding Affinities of Filamin-Peptide Interactions. *J. Chem. Inf. Model.* **2013**, *53*, 2626–2633. [CrossRef]
77. Kumari, R.; Kumar, R.; Lynn, A. G_mmpbsa-A GROMACS Tool for High-Throughput MM-PBSA Calculations. *J. Chem. Inf. Model.* **2014**, *54*, 1951–1962. [CrossRef] [PubMed]
78. Klvana, M.; Pavlova, M.; Koudelakova, T.; Chaloupkova, R.; Dvorak, P.; Prokop, Z.; Stsiapanava, A.; Kuty, M.; Kuta-Smatanova, I.; Dohnalek, J.; et al. Pathways and Mechanisms for Product Release in the Engineered Haloalkane Dehalogenases Explored Using Classical and Random Acceleration Molecular Dynamics Simulations. *J. Mol. Biol.* **2009**, *392*, 1339–1356. [CrossRef] [PubMed]
79. Vavra, O.; Filipovic, J.; Plhak, J.; Bednar, D.; Marques, S.M.; Brezovsky, J.; Stourac, J.; Matyska, L.; Damborsky, J. CaverDock: A Molecular Docking-Based Tool to Analyse Ligand Transport through Protein Tunnels and Channels. *Bioinformatics* **2019**, *35*, 4986–4993. [CrossRef] [PubMed]
80. Chovancova, E.; Pavelka, A.; Benes, P.; Strnad, O.; Brezovsky, J.; Kozlikova, B.; Gora, A.; Sustr, V.; Klvana, M.; Medek, P.; et al. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput. Biol.* **2012**, *8*, e1002708. [CrossRef]
81. Matthews, B.W.; Nicholson, H.; Becktel, W.J. Enhanced Protein Thermostability from Site-Directed Mutations That Decrease the Entropy of Unfolding. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 6663–6667. [CrossRef] [PubMed]
82. Kellogg, E.H.; Leaver-Fay, A.; Baker, D. Role of Conformational Sampling in Computing Mutation-Induced Changes in Protein Structure and Stability. *Proteins* **2011**, *79*, 830–838. [CrossRef]
83. Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; Serrano, L. The FoldX Web Server: An Online Force Field. *Nucleic Acids Res.* **2005**, *33*, W382. [CrossRef] [PubMed]
84. Yin, S.; Ding, F.; Dokholyan, N.V. Eris: An Automated Estimator of Protein Stability. *Nat. Methods* **2007**, *4*, 466–467. [CrossRef]
85. Cao, H.; Wang, J.; He, L.; Qi, Y.; Zhang, J.Z. DeepDDG: Predicting the Stability Change of Protein Point Mutations Using Neural Networks. *J. Chem. Inf. Model.* **2019**, *59*, 1508–1514. [CrossRef] [PubMed]
86. Pupko, T.; Bell, R.E.; Mayrose, I.; Glaser, F.; Ben-Tal, N. Rate4Site: An Algorithmic Tool for the Identification of Functional Regions in Proteins by Surface Mapping of Evolutionary Determinants within Their Homologues. *Bioinformatics* **2002**, *18*, S71–S77. [CrossRef] [PubMed]
87. Bednar, D.; Beerens, K.; Sebestova, E.; Bendl, J.; Khare, S.; Chaloupkova, R.; Prokop, Z.; Brezovsky, J.; Baker, D.; Damborsky, J. FireProt: Energy- and Evolution-Based Computational Design of Thermostable Multiple-Point Mutants. *PLoS Comput. Biol.* **2015**, *11*, e1004556. [CrossRef]

88. Ramanathan, A.; Agarwal, P.K. Evolutionarily Conserved Linkage between Enzyme Fold, Flexibility, and Catalysis. *PLoS Biol.* **2011**, *9*, e1001193. [CrossRef]
89. Teilum, K.; Olsen, J.G.; Kragelund, B.B. Functional Aspects of Protein Flexibility. *Cell Mol. Life Sci.* **2009**, *66*, 2231–2247. [CrossRef]
90. Mlynek, G.; Djinović-Carugo, K.; Carugo, O. B-Factor Rescaling for Protein Crystal Structure Analyses. *Crystals* **2024**, *14*, 443. [CrossRef]
91. Sun, Z.; Liu, Q.; Qu, G.; Feng, Y.; Reetz, M.T. Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. *Chem. Rev.* **2019**, *119*, 1626–1665. [CrossRef] [PubMed]
92. Ru, W.J.; Xia, B.B.; Zhang, Y.X.; Yang, J.W.; Zhang, H.B.; Hu, X.Q. Development of Thermostable Dextranase from *Streptococcus mutans* (*Sm*dexTM) through in Silico Design Employing B-Factor and Cartesian-ΔΔG. *J. Biotechnol.* **2022**, *360*, 142–151. [CrossRef]
93. Mura, C. Development & Implementation of a PyMOL "putty" Representation. *arXiv* **2014**, arXiv:1407.5211. [CrossRef]
94. Roe, D.R.; Cheatham, T.E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095. [CrossRef]
95. Alexandrov, V.; Lehnert, U.; Echols, N.; Milburn, D.; Engelman, D.; Gerstein, M. Normal Modes for Predicting Protein Motions: A Comprehensive Database Assessment and Associated Web Tool. *Protein Sci.* **2005**, *14*, 633. [CrossRef] [PubMed]
96. Zhou, L.; Liu, Q. Aligning Experimental and Theoretical Anisotropic B-Factors: Water Models, Normal-Mode Analysis Methods, and Metrics. *J. Phys. Chem. B* **2014**, *118*, 4069–4079. [CrossRef] [PubMed]
97. Fuglebakk, E.; Reuter, N.; Hinsen, K. Evaluation of Protein Elastic Network Models Based on an Analysis of Collective Motions. *J. Chem. Theory Comput.* **2013**, *9*, 5618–5628. [CrossRef]
98. Pandey, A.; Liu, E.; Graham, J.; Chen, W.; Keten, S. B-Factor Prediction in Proteins Using a Sequence-Based Deep Learning Model. *Patterns* **2023**, *4*, 100805. [CrossRef] [PubMed]
99. Dimarogona, M.; Topakas, E.; Christakopoulos, P.; Chrysina, E.D. The Crystal Structure of a *Fusarium oxysporum* Feruloyl Esterase That Belongs to the Tannase Family. *FEBS Lett.* **2020**, *594*, 1738–1749. [CrossRef] [PubMed]
100. Somero, G.N. Proteins and Temperature. *Annu. Rev. Physiol.* **1995**, *57*, 43–68. [CrossRef] [PubMed]
101. Svingor, Á.; Kardos, J.; Hajdú, I.; Németh, A.; Závodszky, P. A Better Enzyme to Cope with Cold. Comparative Flexibility Studies on Psychrotrophic, Mesophilic, and Thermophilic IPMDHs. *J. Biol. Chem.* **2001**, *276*, 28121–28125. [CrossRef]
102. Hou, Q.; Rooman, M.; Pucci, F. Enzyme Stability-Activity Trade-Off: New Insights from Protein Stability Weaknesses and Evolutionary Conservation. *J. Chem. Theory Comput.* **2023**, *19*, 3664–3671. [CrossRef]
103. Vanella, R.; Küng, C.; Schoepfer, A.A.; Doffini, V.; Ren, J.; Nash, M.A. Understanding Activity-Stability Tradeoffs in Biocatalysts by Enzyme Proximity Sequencing. *Nat. Commun.* **2024**, *15*, 1807. [CrossRef]
104. Hernández, G.; Jenney, F.E.; Adams, M.W.W.; LeMaster, D.M. Millisecond Time Scale Conformational Flexibility in a Hyperthermophile Protein at Ambient Temperature. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3166–3170. [CrossRef] [PubMed]
105. Karshikoff, A.; Nilsson, L.; Ladenstein, R. Rigidity versus Flexibility: The Dilemma of Understanding Protein Thermal Stability. *FEBS J.* **2015**, *282*, 3899–3917. [CrossRef]
106. Joho, Y.; Vongsouthi, V.; Gomez, C.; Larsen, J.S.; Ardevol, A.; Jackson, C.J. Improving Plastic Degrading Enzymes via Directed Evolution. *Protein Eng. Des. Sel.* **2024**, *37*, gzae009. [CrossRef] [PubMed]
107. Stimple, S.D.; Smith, M.D.; Tessier, P.M. Directed Evolution Methods for Overcoming Trade-Offs between Protein Activity and Stability. *AIChE J.* **2020**, *66*, e16814. [CrossRef] [PubMed]
108. Xie, Y.; An, J.; Yang, G.; Wu, G.; Zhang, Y.; Cui, L.; Feng, Y. Enhanced Enzyme Kinetic Stability by Increasing Rigidity within the Active Site. *J. Biol. Chem.* **2014**, *289*, 7994–8006. [CrossRef] [PubMed]
109. Yang, J.S.; Seo, S.W.; Jang, S.; Jung, G.Y.; Kim, S. Rational Engineering of Enzyme Allosteric Regulation through Sequence Evolution Analysis. *PLoS Comput. Biol.* **2012**, *8*, e1002612. [CrossRef] [PubMed]
110. Parker, M.H.; Hefford, M.A. Introduction of Potential Helix-Capping Residues into an Engineered Helical Protein. *Biotechnol. Appl. Biochem.* **1998**, *28*, 69–76. [CrossRef] [PubMed]
111. Lu, H.; Diaz, D.J.; Czarnecki, N.J.; Zhu, C.; Kim, W.; Shroff, R.; Acosta, D.J.; Alexander, B.R.; Cole, H.O.; Zhang, Y.; et al. Machine Learning-Aided Engineering of Hydrolases for PET Depolymerization. *Nature* **2022**, *604*, 662–667. [CrossRef] [PubMed]
112. Gao, X.; Dong, X.; Li, X.; Liu, Z.; Liu, H. Prediction of Disulfide Bond Engineering Sites Using a Machine Learning Method. *Sci. Rep.* **2020**, *10*, 10330. [CrossRef] [PubMed]
113. Mary, B.; And, T.-C.; Nielsen, J.E. Redesigning Protein pKa Values. *Protein Sci.* **2007**, *16*, 239. [CrossRef]
114. Liu, H.; Ding, Y.; Mazurkewich, S.; Pei, W.; Wei, X.; Larsbrink, J.; Chipot, C.; Hong, Z.; Cai, W.; Zong, Z. Boosting Enzyme Activity in Biomass Conversion by Modulating the Hydrolysis Process of Cellobiohydrolases. *ACS Catal.* **2024**, *14*, 16044–16054. [CrossRef]
115. Kordes, S.; Romero-Romero, S.; Lutz, L.; Höcker, B. A Newly Introduced Salt Bridge Cluster Improves Structural and Biophysical Properties of de Novo TIM Barrels. *Protein Sci.* **2021**, *31*, 513. [CrossRef]
116. Fonseca-Maldonado, R.; Vieira, D.S.; Alponti, J.S.; Bonneil, E.; Thibault, P.; Ward, R.J. Engineering the Pattern of Protein Glycosylation Modulates the Thermostability of a GH11 Xylanase. *J. Biol. Chem.* **2013**, *288*, 25522–25534. [CrossRef] [PubMed]

117. Shirke, A.N.; White, C.; Englaender, J.A.; Zwarycz, A.; Butterfoss, G.L.; Linhardt, R.J.; Gross, R.A. Stabilizing Leaf and Branch Compost Cutinase (LCC) with Glycosylation: Mechanism and Effect on PET Hydrolysis. *Biochemistry* **2018**, *57*, 1190–1200. [CrossRef] [PubMed]
118. Planas-Iglesias, J.; Opaleny, F.; Ulbrich, P.; Stourac, J.; Sanusi, Z.; Pinto, G.P.; Schenkmayerova, A.; Byska, J.; Damborsky, J.; Kozlikova, B.; et al. LoopGrafter: A Web Tool for Transplanting Dynamical Loops for Protein Engineering. *Nucleic Acids Res.* **2022**, *50*, W465–W473. [CrossRef] [PubMed]
119. Wang, J.; You, S.; Zha, Z.; Li, J.; Zhang, W.; Bai, Z.; Hu, Y.; Wang, X.; Chen, Y.; Chen, Z.; et al. Loop Engineering of a Thermostable GH10 Xylanase to Improve Low-Temperature Catalytic Performance for Better Synergistic Biomass-Degrading Abilities. *Bioresour. Technol.* **2021**, *342*, 125962. [CrossRef]
120. Yang, J.K.; Park, M.S.; Waldo, G.S.; Suh, S.W. Directed Evolution Approach to a Structural Genomics Project: Rv2002 from *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 455–460. [CrossRef] [PubMed]
121. Chen, A.; Xu, T.; Ge, Y.; Wang, L.; Tang, W.; Li, S. Hydrogen-Bond-Based Protein Engineering for the Acidic Adaptation of *Bacillus acidopullulyticus* Pullulanase. *Enzyme Microb. Technol.* **2019**, *124*, 79–83. [CrossRef] [PubMed]
122. Blázquez-Sánchez, P.; Vargas, J.A.; Furtado, A.A.; Griñen, A.; Leonardo, D.A.; Sculaccio, S.A.; Pereira, H.D.M.; Sonnendecker, C.; Zimmermann, W.; Díez, B.; et al. Engineering the Catalytic Activity of an Antarctic PET-Degrading Enzyme by Loop Exchange. *Protein Sci.* **2023**, *32*, e4757. [CrossRef] [PubMed]
123. Li, P.Y.; Chen, X.L.; Ji, P.; Li, C.Y.; Wang, P.; Zhang, Y.; Xie, B.B.; Qin, Q.L.; Su, H.N.; Zhou, B.C.; et al. Interdomain Hydrophobic Interactions Modulate the Thermostability of Microbial Esterases from the Hormone-Sensitive Lipase Family. *J. Biol. Chem.* **2015**, *290*, 11188–11198. [CrossRef] [PubMed]
124. Erwin, C.R.; Barnett, B.L.; Oliver, J.D.; Sullivan, J.F. Effects of Engineered Salt Bridges on the Stability of Subtilisin BPN'. *Protein Eng. Des. Sel.* **1990**, *4*, 87–97. [CrossRef] [PubMed]
125. Crooks, G.E.; Hon, G.; Chandonia, J.M.; Brenner, S.E. WebLogo: A Sequence Logo Generator. *Genome Res.* **2004**, *14*, 1188–1190. [CrossRef] [PubMed]
126. Yeh, A.H.W.; Norn, C.; Kipnis, Y.; Tischer, D.; Pellock, S.J.; Evans, D.; Ma, P.; Lee, G.R.; Zhang, J.Z.; Anishchenko, I.; et al. De Novo Design of Luciferases Using Deep Learning. *Nature* **2023**, *614*, 774–780. [CrossRef] [PubMed]
127. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [CrossRef] [PubMed]
128. Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; et al. Evolutionary-Scale Prediction of Atomic-Level Protein Structure with a Language Model. *Science* **2023**, *379*, 1123–1130. [CrossRef]
129. Jisna, V.A.; Jayaraj, P.B. Protein Structure Prediction: Conventional and Deep Learning Perspectives. *Protein J.* **2021**, *40*, 522–544. [CrossRef] [PubMed]
130. WL, D. The PyMOL Molecular Graphics System (DeLano Scientific, San Carlos, CA) 2002. Available online: https://www.pymol.org/ (accessed on 29 January 2025).
131. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Meng, E.C.; Couch, G.S.; Croll, T.I.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Structure Visualization for Researchers, Educators, and Developers. *Protein Sci.* **2020**, *30*, 70. [CrossRef]
132. Craig, D.B.; Dombkowski, A.A. Disulfide by Design 2.0: A Web-Based Tool for Disulfide Engineering in Proteins. *BMC Bioinform.* **2013**, *14*, 346. [CrossRef] [PubMed]
133. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M.R.; Smith, J.C.; Kasson, P.M.; Van Der Spoel, D.; et al. GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29*, 845–854. [CrossRef]
134. Case, D.A.; Aktulga, H.M.; Belfon, K.; Cerutti, D.S.; Cisneros, G.A.; Cruzeiro, V.W.D.; Forouzesh, N.; Giese, T.J.; Götz, A.W.; Gohlke, H.; et al. AmberTools. *J. Chem. Inf. Model.* **2023**, *63*, 6183–6191. [CrossRef] [PubMed]
135. Brooks, B.R.; Brooks, C.L.; Mackerell, A.D.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614. [CrossRef] [PubMed]
136. Morales-Quintana, L.; Carrasco-Orellana, C.; Beltrán, D.; Moya-León, M.A.; Herrera, R. Molecular Insights of a Xyloglucan Endo-Transglycosylase/Hydrolase of Radiata pine (*Pr*XTH1) Expressed in Response to Inclination: Kinetics and Computational Study. *Plant Physiol. Bioch* **2019**, *136*, 155–161. [CrossRef] [PubMed]
137. Jäckering, A.; Kamp, M.v.d.; Strodel, B.; Zinovjev, K. Influence of Wobbling Tryptophan and Mutations on PET Degradation Explored by QM/MM Free Energy Calculations. *J. Chem. Inf. Model.* **2024**, *64*, 7544–7554. [CrossRef] [PubMed]
138. Berraud-Pache, R.; Garcia-Iriepa, C.; Navizet, I. Modeling Chemical Reactions by QM/MM Calculations: The Case of the Tautomerization in Fireflies Bioluminescent Systems. *Front. Chem.* **2018**, *6*, 358639. [CrossRef]
139. Peng, J.H.; Wang, W.; Yu, Y.Q.; Gu, H.L.; Huang, X. Clustering Algorithms to Analyze Molecular Dynamics Simulation Trajectories for Complex Chemical and Biological Systems. *CJCP* **2018**, *31*, 404–420. [CrossRef]
140. Husic, B.E.; Pande, V.S. Markov State Models: From an Art to a Science. *J. Am. Chem. Soc.* **2018**, *140*, 2386–2396. [CrossRef] [PubMed]

141. Palma, J.; Pierdominici-Sottile, G. On the Uses of PCA to Characterise Molecular Dynamics Simulations of Biological Macromolecules: Basics and Tips for an Effective Use. *ChemPhysChem* **2023**, *24*, e202200491. [CrossRef]
142. Lemkul, J.A. Introductory Tutorials for Simulating Protein Dynamics with GROMACS. *J. Phys. Chem. B* **2024**, *128*, 9418–9435. [CrossRef] [PubMed]
143. Benrezkallah, D. Molecular Dynamics Simulations at High Temperatures of the *Aeropyrum pernix* L7Ae Thermostable Protein: Insight into the Unfolding Pathway. *J. Mol. Graph. Model.* **2024**, *127*, 108700. [CrossRef]
144. Gattin, Z.; Riniker, S.; Hore, P.J.; Mok, K.H.; Van Gunsteren, W.F. Temperature and Urea Induced Denaturation of the TRP-Cage Mini Protein TC5b: A Simulation Study Consistent with Experimental Observations. *Protein Sci.* **2009**, *18*, 2090–2099. [CrossRef] [PubMed]
145. Korendovych, I.V.; DeGrado, W.F. De Novo Protein Design, a Retrospective. *Q. Rev. Biophys.* **2020**, *53*, e3. [CrossRef]
146. Marshall, L.R.; Zozulia, O.; Lengyel-Zhand, Z.; Korendovych, I.V. Minimalist de Novo Design of Protein Catalysts. *ACS Catal.* **2019**, *9*, 9265. [CrossRef] [PubMed]
147. Leman, J.K.; Weitzner, B.D.; Lewis, S.M.; Adolf-Bryfogle, J.; Alam, N.; Alford, R.F.; Aprahamian, M.; Baker, D.; Barlow, K.A.; Barth, P.; et al. Macromolecular Modeling and Design in Rosetta: Recent Methods and Frameworks. *Nat. Methods* **2020**, *17*, 665–680. [CrossRef]
148. Richter, F.; Leaver-Fay, A.; Khare, S.D.; Bjelic, S.; Baker, D. De Novo Enzyme Design Using Rosetta3. *PLoS ONE* **2011**, *6*, e19230. [CrossRef] [PubMed]
149. Kaufmann, K.W.; Lemmon, G.H.; Deluca, S.L.; Sheehan, J.H.; Meiler, J. Practically Useful: What the Rosetta Protein Modeling Suite Can Do for You. *Biochemistry* **2010**, *49*, 2987–2998. [CrossRef]
150. Zanghellini, A.; Jiang, L.; Wollacott, A.M.; Cheng, G.; Meiler, J.; Althoff, E.A.; Röthlisberger, D.; Röthlisberger, R.; Baker, D. New Algorithms and an in Silico Benchmark for Computational Enzyme Design. *Protein Sci.* **2006**, *15*, 2785. [CrossRef] [PubMed]
151. Roda, S.; Terholsen, H.; Meyer, J.R.H.; Cañellas-Solé, A.; Guallar, V.; Bornscheuer, U.; Kazemi, M. AsiteDesign: A Semirational Algorithm for an Automated Enzyme Design. *J. Phys. Chem.* **2023**, *127*, 2661–2670. [CrossRef] [PubMed]
152. Weitzner, B.D.; Kipnis, Y.; Daniel, A.G.; Hilvert, D.; Baker, D. A Computational Method for Design of Connected Catalytic Networks in Proteins. *Protein Sci.* **2019**, *28*, 2036–2041. [CrossRef]
153. Khersonsky, O.; Röthlisberger, D.; Wollacott, A.M.; Murphy, P.; Dym, O.; Albeck, S.; Kiss, G.; Houk, K.N.; Baker, D.; Tawfik, D.S. Optimization of the In-Silico-Designed Kemp Eliminase KE70 by Computational Design and Directed Evolution. *J. Mol. Biol.* **2011**, *407*, 391–412. [CrossRef]
154. Pan, X.; Kortemme, T. Recent Advances in de Novo Protein Design: Principles, Methods, and Applications. *J. Biol. Chem.* **2021**, *296*, 100558. [CrossRef] [PubMed]
155. Siegel, J.B.; Zanghellini, A.; Lovick, H.M.; Kiss, G.; Lambert, A.R.; St. Clair, J.L.; Gallaher, J.L.; Hilvert, D.; Gelb, M.H.; Stoddard, B.L.; et al. Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction. *Science* **2010**, *329*, 309–313. [CrossRef]
156. Li, G.; Xu, L.; Zhang, H.; Liu, J.; Yan, J.; Yan, Y. A De Novo Designed Esterase with P-Nitrophenyl Acetate Hydrolysis Activity. *Molecules* **2020**, *25*, 4658. [CrossRef] [PubMed]
157. Jiang, L.; Althoff, E.A.; Clemente, F.R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J.L.; Betker, J.L.; Tanaka, F.; Barbas, C.F.; et al. De Novo Computational Design of Retro-Aldol Enzymes. *Science* **2008**, *319*, 1387–1391. [CrossRef] [PubMed]
158. Holst, L.H.; Madsen, N.G.; Toftgård, F.T.; Rønne, F.; Moise, I.M.; Petersen, E.I.; Fojan, P. De Novo Design of a Polycarbonate Hydrolase. *Protein Eng. Des. Sel.* **2023**, *36*, gzad022. [CrossRef] [PubMed]
159. Satta, A.; Zampieri, G.; Loprete, G.; Campanaro, S.; Treu, L.; Bergantino, E.; Satta, A.; Zampieri, G.; Loprete, G.; Campanaro, S.; et al. Metabolic and Enzymatic Engineering Strategies for Polyethylene Terephthalate Degradation and Valorization. *Rev. Environ. Sci. Biotechnol.* **2024**, *23*, 351–383. [CrossRef]
160. Wang, Y.; Xue, P.; Cao, M.; Yu, T.; Lane, S.T.; Zhao, H. Directed Evolution: Methodologies and Applications. *Chem. Rev.* **2021**, *121*, 12384–12444. [CrossRef] [PubMed]
161. Xiao, H.; Bao, Z.; Zhao, H. High Throughput Screening and Selection Methods for Directed Enzyme Evolution. *Ind. Eng. Chem. Res.* **2015**, *54*, 4011–4020. [CrossRef]
162. Lutz, S. Beyond Directed Evolution-Semi-Rational Protein Engineering and Design. *Curr. Opin. Biotechnol.* **2010**, *21*, 734. [CrossRef] [PubMed]
163. de Almeida Paiva, V.; de Souza Gomes, I.; Monteiro, C.R.; Mendonça, M.V.; Martins, P.M.; Santana, C.A.; Gonçalves-Almeida, V.; Izidoro, S.C.; de Melo-Minardi, R.C.; de Azevedo Silveira, S. Protein Structural Bioinformatics: An Overview. *Comput. Biol. Med.* **2022**, *147*, 105695. [CrossRef]
164. Cantarel, B.I.; Coutinho, P.M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes Database (CAZy): An Expert Resource for Glycogenomics. *Nucleic Acids Res.* **2008**, *37*, D233. [CrossRef] [PubMed]
165. Buchholz, P.C.F.; Feuerriegel, G.; Zhang, H.; Perez-Garcia, P.; Nover, L.L.; Chow, J.; Streit, W.R.; Pleiss, J. Plastics Degradation by Hydrolytic Enzymes: The Plastics-Active Enzymes Database—PAZy. *Proteins* **2022**, *90*, 1443–1456. [CrossRef] [PubMed]

166. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and Applications. *BMC Bioinform.* **2009**, *10*, 421. [CrossRef]
167. Steinegger, M.; Söding, J. MMseqs2 Enables Sensitive Protein Sequence Searching for the Analysis of Massive Data Sets. *Nat. Biotechnol.* **2017**, *35*, 1026–1028. [CrossRef] [PubMed]
168. van Kempen, M.; Kim, S.S.; Tumescheit, C.; Mirdita, M.; Lee, J.; Gilchrist, C.L.M.; Söding, J.; Steinegger, M. Fast and Accurate Protein Structure Search with Foldseek. *Nat. Biotechnol.* **2023**, *42*, 243–246. [CrossRef]
169. Illergård, K.; Ardell, D.H.; Elofsson, A. Structure Is Three to Ten Times More Conserved than Sequence—A Study of Structural Response in Protein Cores. *Proteins* **2009**, *77*, 499–508. [CrossRef] [PubMed]
170. Musil, M.; Jezik, A.; Horackova, J.; Borko, S.; Kabourek, P.; Damborsky, J.; Bednar, D. FireProt 2.0: Web-Based Platform for the Fully Automated Design of Thermostable Proteins. *Brief. Bioinform.* **2023**, *25*, bbad425. [CrossRef]
171. Risso, V.A.; Sanchez-Ruiz, J.M.; Ozkan, S.B. Biotechnological and Protein-Engineering Implications of Ancestral Protein Resurrection. *Curr. Opin. Struct. Biol.* **2018**, *51*, 106–115. [CrossRef] [PubMed]
172. Livada, J.; Vargas, A.M.; Martinez, C.A.; Lewis, R.D. Ancestral Sequence Reconstruction Enhances Gene Mining Efforts for Industrial Ene Reductases by Expanding Enzyme Panels with Thermostable Catalysts. *ACS Catal.* **2023**, *13*, 2576–2585. [CrossRef]
173. Musil, M.; Khan, R.T.; Beier, A.; Stourac, J.; Konegger, H.; Damborsky, J.; Bednar, D. FireProtASR: A Web Server for Fully Automated Ancestral Sequence Reconstruction. *Brief. Bioinform.* **2021**, *22*, bbaa337. [CrossRef]
174. Ashkenazy, H.; Abadi, S.; Martz, E.; Chay, O.; Mayrose, I.; Pupko, T.; Ben-Tal, N. ConSurf 2016: An Improved Methodology to Estimate and Visualize Evolutionary Conservation in Macromolecules. *Nucleic Acids Res* **2016**, *44*, W344–W350. [CrossRef]
175. Pavelka, A.; Chovancova, E.; Damborsky, J. HotSpot Wizard: A Web Server for Identification of Hot Spots in Protein Engineering. *Nucleic Acids Res.* **2009**, *37*, W376–W383. [CrossRef] [PubMed]
176. Kamisetty, H.; Ovchinnikov, S.; Baker, D. Assessing the Utility of Coevolution-Based Residue-Residue Contact Predictions in a Sequence- and Structure-Rich Era. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 15674–15679. [CrossRef] [PubMed]
177. Hopf, T.A.; Green, A.G.; Schubert, B.; Mersmann, S.; Schärfe, C.P.I.; Ingraham, J.B.; Toth-Petroczy, A.; Brock, K.; Riesselman, A.J.; Palmedo, P.; et al. The EVcouplings Python Framework for Coevolutionary Sequence Analysis. *Bioinformatics* **2019**, *35*, 1582–1584. [CrossRef]
178. Hopf, T.A.; Ingraham, J.B.; Poelwijk, F.J.; Schärfe, C.P.I.; Springer, M.; Sander, C.; Marks, D.S. Mutation Effects Predicted from Sequence Co-Variation. *Nat. Biotechnol.* **2017**, *35*, 128–135. [CrossRef] [PubMed]
179. Voigt, C.A.; Martinez, C.; Wang, Z.G.; Mayo, S.L.; Arnold, F.H. Protein Building Blocks Preserved by Recombination. *Nat. Struct. Biol.* **2002**, *9*, 553–558. [CrossRef] [PubMed]
180. Meyer, M.M.; Hochrein, L.; Arnold, F.H. Structure-Guided SCHEMA Recombination of Distantly Related β-Lactamases. *PEDS* **2006**, *19*, 563–570. [CrossRef]
181. Schenkmayerova, A.; Pinto, G.P.; Toul, M.; Marek, M.; Hernychova, L.; Planas-Iglesias, J.; Daniel Liskova, V.; Pluskal, D.; Vasina, M.; Emond, S.; et al. Engineering the Protein Dynamics of an Ancestral Luciferase. *Nat. Commun.* **2021**, *12*, 3616. [CrossRef] [PubMed]
182. Notin, P.; Rollins, N.; Gal, Y.; Sander, C.; Marks, D. Machine Learning for Functional Protein Design. *Nat. Biotechnol.* **2024**, *42*, 216–228. [CrossRef] [PubMed]
183. Shinde, P.P.; Shah, S. A Review of Machine Learning and Deep Learning Applications. In Proceedings of the 2018 4th International Conference on Computing, Communication Control and Automation, ICCUBEA, Pune, India, 16–18 August 2018. [CrossRef]
184. Narayanan, H.; Dingfelder, F.; Butté, A.; Lorenzen, N.; Sokolov, M.; Arosio, P. Machine Learning for Biologics: Opportunities for Protein Engineering, Developability, and Formulation. *Trends Pharmacol. Sci.* **2021**, *42*, 151–165. [CrossRef] [PubMed]
185. Hu, B.; Tan, C.; Wu, L.; Zheng, J.; Xia, J.; Gao, Z.; Liu, Z.; Wu, F.; Zhang, G.; Li, S.Z. Advances of Deep Learning in Protein Science: A Comprehensive Survey. *arXiv* **2024**, arXiv:2403.05314.
186. Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: Making Protein Folding Accessible to All. *Nat. Methods* **2022**, *19*, 679–682. [CrossRef] [PubMed]
187. Wolf, T.; Debut, L.; Sanh, V.; Chaumond, J.; Delangue, C.; Moi, A.; Cistac, P.; Rault, T.; Louf, R.; Funtowicz, M.; et al. Transformers: State-of-the-Art Natural Language Processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*; Association for Computational Linguistics: Kerrville, TX, USA, 2020; pp. 38–45. [CrossRef]
188. Crisci, C.; Ghattas, B.; Perera, G. A Review of Supervised Machine Learning Algorithms and Their Applications to Ecological Data. *Ecol. Modell.* **2012**, *240*, 113–122. [CrossRef]
189. Lampropoulos, A.S.; Tsihrintzis, G.A. The Learning Problem. In *Machine Learning Paradigms, Intelligent Systems Reference Library 92*; Springer: Cham, Switzerland, 2015; pp. 31–61. ISBN 9783319191348.
190. Li, H.; Zhang, R.; Min, Y.; Ma, D.; Zhao, D.; Zeng, J. A Knowledge-Guided Pre-Training Framework for Improving Molecular Representation Learning. *Nat. Commun.* **2023**, *14*, 7568. [CrossRef] [PubMed]

191. Elnaggar, A.; Heinzinger, M.; Dallago, C.; Rehawi, G.; Wang, Y.; Jones, L.; Gibbs, T.; Feher, T.; Angerer, C.; Steinegger, M.; et al. ProtTrans: Toward Understanding the Language of Life Through Self-Supervised Learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 7112–7127. [CrossRef]
192. Rives, A.; Meier, J.; Sercu, T.; Goyal, S.; Lin, Z.; Liu, J.; Guo, D.; Ott, M.; Zitnick, C.L.; Ma, J.; et al. Biological Structure and Function Emerge from Scaling Unsupervised Learning to 250 Million Protein Sequences. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2016239118. [CrossRef] [PubMed]
193. Chu, S.K.S.; Narang, K.; Siegel, J.B. Protein Stability Prediction by Fine-Tuning a Protein Language Model on a Mega-Scale Dataset. *PLoS Comput. Biol.* **2024**, *20*, e1012248. [CrossRef]
194. Kroll, A.; Ranjan, S.; Engqvist, M.K.M.; Lercher, M.J. A General Model to Predict Small Molecule Substrates of Enzymes Based on Machine and Deep Learning. *Nat. Commun.* **2023**, *14*, 2787. [CrossRef] [PubMed]
195. Wang, Y.; Mei, C.; Zhou, Y.; Wang, Y.; Zheng, C.; Zhen, X.; Xiong, Y.; Chen, P.; Zhang, J.; Wang, B. Semi-Supervised Prediction of Protein Interaction Sites from Unlabeled Sample Information. *BMC Bioinform.* **2019**, *20*, 699. [CrossRef] [PubMed]
196. Angermueller, C.; Dohan, D.; Belanger, D.; Deshpande, R.; Murphy, K.; Colwell, L. Model-Based Reinforcement Learning for Biological Sequence Design. In Proceedings of the International Conference on Learning Representations, New Orleans, LA, USA, 6–9 May 2019.
197. Sun, H.; He, L.; Deng, P.; Liu, G.; Liu, H.; Cao, C.; Ju, F.; Wu, L.; Qin, T.; Liu, T.-Y. Accelerating Protein Engineering with Fitness Landscape Modeling and Reinforcement Learning. *bioRxiv* **2024**. [CrossRef]
198. Olivecrona, M.; Blaschke, T.; Engkvist, O.; Chen, H. Molecular De-Novo Design through Deep Reinforcement Learning. *J. Cheminform* **2017**, *9*, 48. [CrossRef]
199. AlOmari, M.; AlOmari, A.; Alsmadi, I. CASP Dataset and Protein Structures Prediction. *SSRN Electron. J.* **2022**. [CrossRef]
200. Harding-Larsen, D.; Funk, J.; Madsen, N.G.; Gharabli, H.; Acevedo-Rocha, C.G.; Mazurenko, S.; Welner, D.H. Protein Representations: Encoding Biological Information for Machine Learning in Biocatalysis. *Biotechnol. Adv.* **2024**, *77*, 108459. [CrossRef] [PubMed]
201. Sayers, E.W.; Beck, J.; Bolton, E.E.; Bourexis, D.; Brister, J.R.; Canese, K.; Comeau, D.C.; Funk, K.; Kim, S.; Klimke, W.; et al. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **2021**, *49*, D10–D17. [CrossRef] [PubMed]
202. Chen, I.M.A.; Chu, K.; Palaniappan, K.; Pillay, M.; Ratner, A.; Huang, J.; Huntemann, M.; Varghese, N.; White, J.R.; Seshadri, R.; et al. IMG/M v.5.0: An Integrated Data Management and Comparative Analysis System for Microbial Genomes and Microbiomes. *Nucleic Acids Res.* **2019**, *47*, D666–D677. [CrossRef]
203. Reimer, L.C.; Sardà Carbasse, J.; Koblitz, J.; Ebeling, C.; Podstawka, A.; Overmann, J. BacDive in 2022: The Knowledge Base for Standardized Bacterial and Archaeal Data. *Nucleic Acids Res* **2021**, *50*, D741. [CrossRef] [PubMed]
204. Erickson, E.; Gado, J.E.; Avilán, L.; Bratti, F.; Brizendine, R.K.; Cox, P.A.; Gill, R.; Graham, R.; Kim, D.J.; König, G.; et al. Sourcing Thermotolerant Poly(Ethylene Terephthalate) Hydrolase Scaffolds from Natural Diversity. *Nat. Commun.* **2022**, *13*, 7850. [CrossRef]
205. Zhang, Y.; Guan, F.; Xu, G.; Liu, X.; Zhang, Y.; Sun, J.; Yao, B.; Huang, H.; Wu, N.; Tian, J. A Novel Thermophilic Chitinase Directly Mined from the Marine Metagenome Using the Deep Learning Tool Preoptem. *Bioresour. Bioprocess.* **2022**, *9*, 54. [CrossRef]
206. Hasegawa, N.; Sugiyama, M.; Igarashi, K. Random Forest Machine-Learning Algorithm Classifies White- and Brown-Rot Fungi According to the Number of the Genes Encoding Carbohydrate-Active EnZyme Families. *Appl. Environ. Microbiol.* **2024**, *90*, e0048224. [CrossRef]
207. Jiang, R.; Yue, Z.; Shang, L.; Wang, D.; Wei, N. PEZy-Miner: An Artificial Intelligence Driven Approach for the Discovery of Plastic-Degrading Enzyme Candidates. *Metab. Eng. Commun.* **2024**, *19*, e00248. [CrossRef] [PubMed]
208. Chang, A.; Jeske, L.; Ulbrich, S.; Hofmann, J.; Koblitz, J.; Schomburg, I.; Neumann-Schaal, M.; Jahn, D.; Schomburg, D. BRENDA, the ELIXIR Core Data Resource in 2021: New Developments and Updates. *Nucleic Acids Res* **2021**, *49*, D498–D508. [CrossRef] [PubMed]
209. Wittig, U.; Rey, M.; Weidemann, A.; Kania, R.; Müller, W. SABIO-RK: An Updated Resource for Manually Curated Biochemical Reaction Kinetics. *Nucleic Acids Res.* **2018**, *46*, D656–D660. [CrossRef]
210. Li, F.; Yuan, L.; Lu, H.; Li, G.; Chen, Y.; Engqvist, M.K.M.; Kerkhoven, E.J.; Nielsen, J. Deep Learning-Based Kcat Prediction Enables Improved Enzyme-Constrained Model Reconstruction. *Nat. Catal.* **2022**, *5*, 662–672. [CrossRef]
211. Kroll, A.; Rousset, Y.; Hu, X.P.; Liebrand, N.A.; Lercher, M.J. Turnover Number Predictions for Kinetically Uncharacterized Enzymes Using Machine and Deep Learning. *Nat. Commun.* **2023**, *14*, 4139. [CrossRef] [PubMed]
212. Lauterbach, S.; Dienhart, H.; Range, J.; Malzacher, S.; Spöring, J.D.; Rother, D.; Pinto, M.F.; Martins, P.; Lagerman, C.E.; Bommarius, A.S.; et al. EnzymeML: Seamless Data Flow and Modeling of Enzymatic Data. *Nat. Methods* **2023**, *20*, 400–402. [CrossRef]
213. Charoenkwan, P.; Schaduangrat, N.; Moni, M.A.; Lio', P.; Manavalan, B.; Shoombuatong, W. SAPPHIRE: A Stacking-Based Ensemble Learning Framework for Accurate Prediction of Thermophilic Proteins. *Comput. Biol. Med.* **2022**, *146*, 105704. [CrossRef]

214. Nikam, R.; Kulandaisamy, A.; Harini, K.; Sharma, D.; Michael Gromiha, M. ProThermDB: Thermodynamic Database for Proteins and Mutants Revisited after 15 Years. *Nucleic Acids Res.* **2021**, *49*, D420–D424. [CrossRef] [PubMed]
215. Stourac, J.; Dubrava, J.; Musil, M.; Horackova, J.; Damborsky, J.; Mazurenko, S.; Bednar, D. FireProtDB: Database of Manually Curated Protein Stability Data. *Nucleic Acids Res.* **2021**, *49*, D319–D324. [CrossRef]
216. Dieckhaus, H.; Brocidiacono, M.; Randolph, N.Z.; Kuhlman, B. Transfer Learning to Leverage Larger Datasets for Improved Prediction of Protein Stability Changes. *Proc. Natl. Acad. Sci. USA* **2024**, *121*, e2314853121. [CrossRef]
217. Gilson, M.K.; Liu, T.; Baitaluk, M.; Nicola, G.; Hwang, L.; Chong, J. BindingDB in 2015: A Public Database for Medicinal Chemistry, Computational Chemistry and Systems Pharmacology. *Nucleic Acids Res* **2015**, *44*, D1045. [CrossRef]
218. Wang, R.; Fang, X.; Lu, Y.; Yang, C.Y.; Wang, S. The PDBbind Database: Methodologies and Updates. *J. Med. Chem.* **2005**, *48*, 4111–4119. [CrossRef] [PubMed]
219. Velecký, J.; Hamsikova, M.; Stourac, J.; Musil, M.; Damborsky, J.; Bednar, D.; Mazurenko, S. SoluProtMutDB: A Manually Curated Database of Protein Solubility Changes upon Mutations. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 6339. [CrossRef] [PubMed]
220. Teufel, F.; Almagro Armenteros, J.J.; Johansen, A.R.; Gíslason, M.H.; Pihl, S.I.; Tsirigos, K.D.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 6.0 Predicts All Five Types of Signal Peptides Using Protein Language Models. *Nat. Biotechnol.* **2022**, *40*, 1023–1025. [CrossRef] [PubMed]
221. Luo, Y.; Nie, Z.; Hong, M.; Zhao, S.; Zhou, H.; Nie, Z. MutaPLM: Protein Language Modeling for Mutation Explanation and Engineering. *arXiv* **2024**, arXiv:2410.22949 (accessed on 13 January 2025). [CrossRef]
222. Li, F.; Chen, Y.; Anton, M.; Nielsen, J. GotEnzymes: An Extensive Database of Enzyme Parameter Predictions. *Nucleic Acids Res.* **2023**, *51*, D583–D586. [CrossRef]
223. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; et al. AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Res.* **2022**, *50*, D439–D444. [CrossRef]
224. Li, Y.; Fang, J. PROTS-RF: A Robust Model for Predicting Mutation-Induced Protein Stability Changes. *PLoS ONE* **2012**, *7*, e47247. [CrossRef] [PubMed]
225. Rawi, R.; Mall, R.; Kunji, K.; Shen, C.H.; Kwong, P.D.; Chuang, G.Y. PaRSnIP: Sequence-Based Protein Solubility Prediction Using Gradient Boosting Machine. *Bioinformatics* **2018**, *34*, 1092–1098. [CrossRef]
226. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions. *J. Big Data* **2021**, *8*, 53. [CrossRef] [PubMed]
227. Horne, J.; Shukla, D. Recent Advances in Machine Learning Variant Effect Prediction Tools for Protein Engineering. *Ind. Eng. Chem. Res.* **2022**, *61*, 6235–6245. [CrossRef] [PubMed]
228. Petrovski, Ž.H.; Hribar-Lee, B.; Bosnić, Z. CAT-Site: Predicting Protein Binding Sites Using a Convolutional Neural Network. *Pharmaceutics* **2022**, *15*, 119. [CrossRef]
229. Zhang, Y.; Chen, Y.; Wang, C.; Lo, C.C.; Liu, X.; Wu, W.; Zhang, J. ProDCoNN: Protein Design Using a Convolutional Neural Network. *Proteins* **2020**, *88*, 819–829. [CrossRef] [PubMed]
230. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. *arXiv* **2017**, arXiv:1706.03762 (accessed on 13 January 2025). [CrossRef]
231. Pei, H.; Li, J.; Ma, S.; Jiang, J.; Li, M.; Zou, Q.; Lv, Z. Identification of Thermophilic Proteins Based on Sequence-Based Bidirectional Representations from Transformer-Embedding Features. *Appl. Sci.* **2023**, *13*, 2858. [CrossRef]
232. Chen, T.; Dumas, M.; Watson, R.; Vincoff, S.; Peng, C.; Zhao, L.; Hong, L.; Pertsemlidis, S.; Shaepers-Cheu, M.; Wang, T.Z.; et al. PepMLM: Target Sequence-Conditioned Generation of Therapeutic Peptide Binders via Span Masked Language Modeling. *arXiv* **2023**, arXiv:2310.03842v3.
233. Ferruz, N.; Schmidt, S.; Höcker, B. ProtGPT2 Is a Deep Unsupervised Language Model for Protein Design. *Nat. Commun.* **2022**, *13*, 4348. [CrossRef] [PubMed]
234. Yang, K.K.; Fusi, N.; Lu, A.X. Convolutions Are Competitive with Transformers for Protein Sequence Pretraining. *Cell Syst.* **2024**, *15*, 286–294.e2. [CrossRef] [PubMed]
235. Tay, Y.; Dehghani, M.; Gupta, J.; Aribandi, V.; Bahri, D.; Qin, Z.; Metzler, D. Are Pre-Trained Convolutions Better than Pre-Trained Transformers? In Proceedings of the ACL-IJCNLP 2021-59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Proceedings of the Conference, Virtual, 1–6 August 2021; pp. 4349–4359. [CrossRef]
236. Matsoukas, C.; Haslum, J.F.; Söderberg, M.; Smith, K. Is It Time to Replace CNNs with Transformers for Medical Images? *arXiv* **2021**, 9038. [CrossRef]
237. Yim, J.; Stärk, H.; Corso, G.; Jing, B.; Barzilay, R.; Jaakkola, T.S. Diffusion Models in Protein Structure and Docking. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2024**, *14*, e1711. [CrossRef]

238. Watson, J.L.; Juergens, D.; Bennett, N.R.; Trippe, B.L.; Yim, J.; Eisenach, H.E.; Ahern, W.; Borst, A.J.; Ragotte, R.J.; Milles, L.F.; et al. De Novo Design of Protein Structure and Function with RFdiffusion. *Nature* **2023**, *620*, 1089–1100. [CrossRef] [PubMed]
239. Arts, M.; Garcia Satorras, V.; Huang, C.W.; Zügner, D.; Federici, M.; Clementi, C.; Noé, F.; Pinsler, R.; van den Berg, R. Two for One: Diffusion Models and Force Fields for Coarse-Grained Molecular Dynamics. *J. Chem. Theory Comput.* **2023**, *19*, 6151–6159. [CrossRef]
240. Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A.J.; Bambrick, J.; et al. Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3. *Nature* **2024**, *630*, 493–500. [CrossRef]
241. Jha, K.; Saha, S.; Singh, H. Prediction of Protein–Protein Interaction Using Graph Neural Networks. *Sci. Rep.* **2022**, *12*, 8360. [CrossRef] [PubMed]
242. Chen, J.; Zheng, S.; Zhao, H.; Yang, Y. Structure-Aware Protein Solubility Prediction from Sequence through Graph Convolutional Network and Predicted Contact Map. *J. Cheminform* **2021**, *13*, 7. [CrossRef] [PubMed]
243. Busk, J.; Schmidt, M.N.; Winther, O.; Vegge, T.; Jørgensen, P.B. Graph Neural Network Interatomic Potential Ensembles with Calibrated Aleatoric and Epistemic Uncertainty on Energy and Forces. *PCCP* **2023**, *25*, 25828–25837. [CrossRef] [PubMed]
244. Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R.J.; Milles, L.F.; Wicky, B.I.M.; Courbet, A.; de Haas, R.J.; Bethel, N.; et al. Robust Deep Learning–Based Protein Sequence Design Using ProteinMPNN. *Science* **2022**, *378*, 49–56. [CrossRef] [PubMed]
245. Sidak, D.; Schwarzerová, J.; Weckwerth, W.; Waldherr, S. Interpretable Machine Learning Methods for Predictions in Systems Biology from Omics Data. *Front. Mol. Biosci.* **2022**, *9*, 926623. [CrossRef]
246. Hochuli, J.; Helbling, A.; Skaist, T.; Ragoza, M.; Koes, D.R. Visualizing Convolutional Neural Network Protein-Ligand Scoring. *J. Mol. Graph. Model.* **2018**, *84*, 96–108. [CrossRef]
247. Vig, J.; Madani, A.; Varshney, L.R.; Xiong, C.; Socher, R.; Rajani, N.F. BERTology Meets Biology: Interpreting Attention in Protein Language Models. In Proceedings of the ICLR 2021-9th International Conference on Learning Representations, Vienna, Austria, 3–7 May 2021. [CrossRef]
248. Simon, E.; Zou, J. InterPLM: Discovering Interpretable Features in Protein Language Models via Sparse Autoencoders. *bioRxiv* **2024**. [CrossRef]
249. Zhang, H.; Themelis, N.J.; Bourtsalas, A. Environmental Impact Assessment of Emissions from Non-Recycled Plastic-to-Energy Processes. *Waste Dispos. Sustain. Energy* **2021**, *3*, 1–11. [CrossRef]
250. Pratiwi, O.A.; Achmadi, U.F.; Kurniawan, R. Microplastic Pollution in Landfill Soil: Emerging Threats the Environmental and Public Health. *Environ. Anal. Health Toxicol.* **2024**, *39*, e2024009. [CrossRef] [PubMed]
251. Qiu, J.; Chen, Y.; Zhang, L.; Wu, J.; Zeng, X.; Shi, X.; Liu, L.; Chen, J. A Comprehensive Review on Enzymatic Biodegradation of Polyethylene Terephthalate. *Environ. Res.* **2024**, *240*, 117427. [CrossRef] [PubMed]
252. Yoshida, S.; Hiraga, K.; Takehana, T.; Taniguchi, I.; Yamaji, H.; Maeda, Y.; Toyohara, K.; Miyamoto, K.; Kimura, Y.; Oda, K. A Bacterium That Degrades and Assimilates Poly(Ethylene Terephthalate). *Science* **2016**, *351*, 1196–1199. [CrossRef]
253. Müller, R.J.; Schrader, H.; Profe, J.; Dresler, K.; Deckwer, W.D. Enzymatic Degradation of Poly(Ethylene Terephthalate): Rapid Hydrolyse Using a Hydrolase from *T. fusca*. *Macromol. Rapid Commun.* **2005**, *26*, 1400–1405. [CrossRef]
254. Sulaiman, S.; Yamato, S.; Kanaya, E.; Kim, J.J.; Koga, Y.; Takano, K.; Kanaya, S. Isolation of a Novel Cutinase Homolog with Polyethylene Terephthalate-Degrading Activity from Leaf-Branch Compost by Using a Metagenomic Approach. *Appl. Environ. Microbiol.* **2012**, *78*, 1556–1562. [CrossRef]
255. Britton, D.; Liu, C.; Xiao, Y.; Jia, S.; Legocki, J.; Kronenberg, J.; Montclare, J.K. Protein-Engineered Leaf and Branch Compost Cutinase Variants Using Computational Screening and IsPETase Homology. *Catal. Today* **2024**, *433*, 114659. [CrossRef]
256. Liu, F.; Wang, T.; Yang, W.; Zhang, Y.; Gong, Y.; Fan, X.; Wang, G.; Lu, Z.; Wang, J. Current Advances in the Structural Biology and Molecular Engineering of PETase. *Front. Bioeng. Biotechnol.* **2023**, *11*, 1263996. [CrossRef] [PubMed]
257. Zheng, Y.; Li, Q.; Liu, P.; Yuan, Y.; Dian, L.; Wang, Q.; Liang, Q.; Su, T.; Qi, Q. Dynamic Docking-Assisted Engineering of Hydrolases for Efficient PET Depolymerization. *ACS Catal.* **2024**, *14*, 3627–3639. [CrossRef]
258. Cui, Y.; Chen, Y.; Liu, X.; Dong, S.; Tian, Y.; Qiao, Y.; Mitra, R.; Han, J.; Li, C.; Han, X.; et al. Computational Redesign of a PETase for Plastic Biodegradation under Ambient Condition by the GRAPE Strategy. *ACS Catal.* **2021**, *11*, 1340–1350. [CrossRef]
259. Cui, Y.; Chen, Y.; Sun, J.; Zhu, T.; Pang, H.; Li, C.; Geng, W.C.; Wu, B. Computational Redesign of a Hydrolase for Nearly Complete PET Depolymerization at Industrially Relevant High-Solids Loading. *Nat. Commun.* **2024**, *15*, 1–12. [CrossRef] [PubMed]
260. Joho, Y.; Royan, S.; Caputo, A.T.; Newton, S.; Peat, T.S.; Newman, J.; Jackson, C.; Ardevol, A. Enhancing PET Degrading Enzymes: A Combinatory Approach. *ChemBioChem* **2024**, *25*, e202400084. [CrossRef] [PubMed]
261. Pirillo, V.; Orlando, M.; Tessaro, D.; Pollegioni, L.; Molla, G. An Efficient Protein Evolution Workflow for the Improvement of Bacterial PET Hydrolyzing Enzymes. *Int. J. Mol. Sci.* **2022**, *23*, 264. [CrossRef] [PubMed]
262. Bell, E.L.; Smithson, R.; Kilbride, S.; Foster, J.; Hardy, F.J.; Ramachandran, S.; Tedstone, A.A.; Haigh, S.J.; Garforth, A.A.; Day, P.J.R.; et al. Directed Evolution of an Efficient and Thermostable PET Depolymerase. *Nat. Catal.* **2022**, *5*, 673–681. [CrossRef]
263. Shi, L.; Liu, P.; Tan, Z.; Zhao, W.; Gao, J.; Gu, Q.; Ma, H.; Liu, H.; Zhu, L. Complete Depolymerization of PET Wastes by an Evolved PET Hydrolase from Directed Evolution. *Angew. Chem. Int. Ed.* **2023**, *62*, e202218390. [CrossRef]

264. Son, H.F.; Cho, I.J.; Joo, S.; Seo, H.; Sagong, H.Y.; Choi, S.Y.; Lee, S.Y.; Kim, K.J. Rational Protein Engineering of Thermo-Stable PETase from *Ideonella sakaiensis* for Highly Efficient PET Degradation. *ACS Catal.* **2019**, *9*, 3519–3526. [CrossRef]
265. Son, H.F.; Joo, S.; Seo, H.; Sagong, H.Y.; Lee, S.H.; Hong, H.; Kim, K.J. Structural Bioinformatics-Based Protein Engineering of Thermo-Stable PETase from *Ideonella sakaiensis*. *Enzym. Microb. Technol.* **2020**, *141*, 109656. [CrossRef] [PubMed]
266. Meng, X.; Yang, L.; Liu, H.; Li, Q.; Xu, G.; Zhang, Y.; Guan, F.; Zhang, Y.; Zhang, W.; Wu, N.; et al. Protein Engineering of Stable IsPETase for PET Plastic Degradation by Premuse. *Int. J. Biol. Macromol.* **2021**, *180*, 667–676. [CrossRef]
267. Ding, Z.; Xu, G.; Miao, R.; Wu, N.; Zhang, W.; Yao, B.; Guan, F.; Huang, H.; Tian, J. Rational Redesign of Thermophilic PET Hydrolase LCC^{ICCG} to Enhance Hydrolysis of High Crystallinity Polyethylene Terephthalates. *J. Hazard. Mater.* **2023**, *453*, 131386. [CrossRef] [PubMed]
268. Kulikova, A.V.; Diaz, D.J.; Loy, J.M.; Ellington, A.D.; Wilke, C.O. Learning the Local Landscape of Protein Structures with Convolutional Neural Networks. *J. Biol. Phys.* **2021**, *47*, 435–454. [CrossRef]
269. Koch, J.; Hess, Y.; Bak, C.R.; Petersen, E.I.; Fojan, P. Design of a Novel Peptide with Esterolytic Activity toward PET by Mimicking the Catalytic Motif of Serine Hydrolases. *J. Phys. Chem. B* **2024**, *128*, 10363–10372. [CrossRef] [PubMed]
270. Knott, B.C.; Erickson, E.; Allen, M.D.; Gado, J.E.; Graham, R.; Kearns, F.L.; Pardo, I.; Topuzlu, E.; Anderson, J.J.; Austin, H.P.; et al. Characterization and Engineering of a Two-Enzyme System for Plastics Depolymerization. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 25476–25485. [CrossRef] [PubMed]
271. Zheng, Y.; Zhang, J.; You, S.; Lin, W.; Su, R.; Qi, W. Efficient Thermophilic Polyethylene Terephthalate Hydrolase Enhanced by Cross Correlation-Based Accumulated Mutagenesis Strategy. *Bioresour. Technol.* **2024**, *406*, 130929. [CrossRef]
272. Zeng, W.; Li, X.; Yang, Y.; Min, J.; Huang, J.W.; Liu, W.; Niu, D.; Yang, X.; Han, X.; Zhang, L.; et al. Substrate-Binding Mode of a Thermophilic PET Hydrolase and Engineering the Enzyme to Enhance the Hydrolytic Efficacy. *ACS Catal.* **2022**, *12*, 3033–3040. [CrossRef]
273. Thumarat, U.; Kawabata, T.; Nakajima, M.; Nakajima, H.; Sugiyama, A.; Yazaki, K.; Tada, T.; Waku, T.; Tanaka, N.; Kawai, F. Comparison of Genetic Structures and Biochemical Properties of Tandem Cutinase-Type Polyesterases from *Thermobifida alba* AHK119. *J. Biosci. Bioeng.* **2015**, *120*, 491–497. [CrossRef] [PubMed]