Vol.01 No.01 (2025) 47-56

Sustainable in Energy Science and Technology

https://ojs.polmed.ac.id/index.php/SiEST

Literature Review: Computational Methods for Designing Thermostable Efficient and Cost-Effective Enzymes for Industrial Applications

Prisca Caesa Moneteringtyas[1*], Nahzim Rahmat[1], Inten Pangestika[1], Sri Rahayu Widya Ningrum[2], Annisa Fillaeli[3], Aliyah Aliyah[4]

[1] *Industrial Chemical Engineering Technology, Department of Mechanical Engineering, Politeknik Negeri Medan, 20155, Medan, Indonesia*

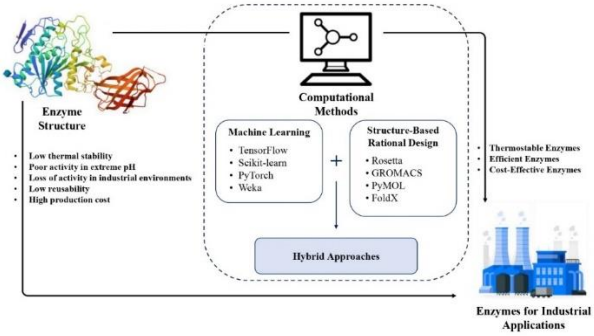
[2] *Study Program of Industrial Chemical Engineering, Politeknik Negeri Lampung, 35144, , Bandar Lampung, Indonesia*

[3] *Department of Chemistry Education, Faculty of Mathematics and Natural Sciences, Universitas Negeri Yogyakarta, Indonesia, Jl. Colombo 1, Yogyakarta 55281, Indonesia*

[4] *Department of Chemical Science and Engineering, Graduate School of Engineering, Kobe Univesity, 1-1 Rokkodai-cho, Nada-Ku, Kobe, Hyogo, 657-8501,Japan*

*Corresponding Author: priscacaesa@polmed.ac.id

Graphical Abstract



Highlights

- *This review explores recent computational methods for enhancing enzyme thermostability for industrial applications.*
- *Structure-based design and machine learning are key approaches, each with unique strengths and limitations.*
- *Hybrid models integrating these techniques show improved predictive accuracy and enzyme performance.*
- *A comparison of leading software tools is provided to guide method selection based on research goals.*

ARTICLE INFO

Keywords:
Computational methods, enzyme, thermostability

Article history:
Received May 26, 2025
Revised June 19, 2025
Accepted July 10, 2025
Available online July 19,2025

https://doi.orgxxxxxxxxxxxxxxxxxxx

ABSTRACT

Enzymes play a vital role as biocatalysts in various industrial applications due to their high specificity and efficiency under mild conditions. However, their limited thermostability significantly constrains their operational lifespan and effectiveness at elevated temperatures. This review examines recent advancements in computational methods aimed at enhancing enzyme thermostability, focusing on structure-based rational design, machine learning, and hybrid approaches. Key findings highlight the effectiveness of structure-based methods, in optimizing enzyme structures, while machine learning approaches demonstrated potential in predicting stabilizing mutations. This review identifies key research gaps and proposes directions for future studies to facilitate the industrial adoption of thermostable enzymes.

1. Introduction

Enzymes are key biocatalysts in industry because of their high specificity as well as efficiency [1]. However, their use in practice is frequently restricted. This is because they lack sufficient thermal stability, which makes the enzyme not effective enough when temperatures are elevated [2]. Improving upon the thermal stability for enzymes, without any sacrifice of their catalytic efficiency, has therefore become a focus inside enzyme engineering.

Current improvements in advanced computing methods present truly hopeful answers to these specific difficulties. To predict and add mutations that make enzymes more thermally stable, was created structure-based rational design, machine learning, and hybrid methods [3].

The purpose of this review is to carefully review the current computational strategies, and to further assess their effectiveness, for improvements regarding the thermal stability of enzymes. A thorough review of nearly all of these approaches should offer some guidance on the directions in which further studies can go in order to help thermostable enzymes find use inside of industry.

2. Methods

This literature review illustrated the potential role that computers could play in improving the industrial stability of enzymes. The review reviewed journal articles published between 2015 on-wards for the reason that those articles are considered the latest breakthrough in developing computational tools and methods. Some journals in 2005 was also included due to its foundational contributions to structure-based rational design.

The study evaluated the literature focused on computer tools, namely, Rosetta, GROMACS, PyMOL, TensorFlow, Scikit-learn. Review papers and works that relied only on laboratory studies were not included. Terms such as "thermostable enzymes computational design" and "machine learning for stability in enzyme" were used to carry out the search.

Searches using database sites, such as ScienceDirect, Springer, Wiley Online Library, and IEEE Xplore, were carried out to retrieve the articles. The search was able to cover a wide range of computer techniques, which were all practically applied in industries.

Thus, it was possible to analyze the collected data through computational methods and main findings derived from the research gaps studies. A particular emphasis was therefore placed on studies that employed various computer techniques since this method is very significant in industrial applicances.

This method also allowed a comprehensive overview of current advancements and challenges in computational enzyme design, as it balanced the views on existing methods' potential and limitations.

3. Results and Discussion

There has been significant progress in computational methods to improve enzyme thermostability, especially through the use of structure-based rational design, machine learning, and hybrid approaches. Each method has its own specificities as outlined in Table 1.

Table 1 Review

Title	Authors and Year	Research Topic	Method/ Software	Key Findings	Research Gaps
Computational Thermostabilization of an Enzyme	Korkegian et al. (2005) [4]	Enhancing enzyme thermostability	**Rosetta** (structure-based design)	Successfully improved enzyme thermostability through structural modifications.	Limited exploration of machine learning approaches.
Protein Thermostability Engineering	Modarres et al., (2016) [5]	Principles of protein engineering for thermostability	**PyMOL, FoldX** (rational design and directed evolution)	Identified key factors influencing protein thermostability.	Lack of integration between computational and experimental approaches.
Thermostability Engineering of Industrial Enzymes through Structure-Based Rational Design	Nezhad et al. (2022) [6]	Rational design for industrial enzyme thermostability	**Rosetta, GROMACS** (structure-based rational design)	Enhanced stability without compromising catalytic efficiency.	Limited exploration of data-driven methods.
Data-Driven Strategies for the Computational Design of Enzyme Thermal Stability	Dou et al. (2023) [7]	Data-driven approaches for enzyme design	**Scikit-learn, TensorFlow** (machine learning and data mining)	Demonstrated effectiveness of data-driven strategies in predicting stability.	Need for hybrid approaches combining data-driven and rational design.
Rational-Design Engineering to Improve Enzyme Thermostability	Pongsupasa et al., (2023) [8]	Improving enzyme stability through	**PyRosetta, GROMACS** (site-directed mutagenesis	Achieved enhanced stability by modifying specific	Insufficient analysis of long-term stability in

Title	Authors and Year	Research Topic	Method/ Software	Key Findings	Research Gaps
		rational design	and molecular dynamics simulations)	amino acid sites.	industrial conditions.
Achieving Thermostability of a Phytase with Resistance up to 100 °C	Tu et al. (2024) [9]	Enhancing phytase stability	**Schrödinger Suite, NAMD** (directed evolution and rational design)	Successfully stabilized phytase at high temperatures.	Limited application to other types of enzymes.
Computational Advances in Protein Engineering and Enzyme Design	Derat & Kamerlin. (2022) [10]	Computational methods in enzyme engineering	**AlphaFold, PyMOL, MODELLER** (AI-based tools and molecular simulations)	Showcased the role of AI in accelerating enzyme design.	Insufficient focus on cost-effectiveness for industrial applications.
Current Advances in Design and Engineering Strategies of Industrial Enzymes	Dinmukhamed et al. (2021) [11]	Strategies for industrial enzyme design	**Rosetta, Scikit-learn** (hybrid approaches: rational design + machine learning)	Highlighted benefits of combining multiple approaches.	Lack of detailed analysis on implementation challenges.
Computational design of an efficient and thermostable esterase for polylactic acid depolymerization	Xie et al. (2024) [12]	Enzyme design for plastic degradation	**Rosetta, GROMACS, AutoDock** (computational protein design and molecular dynamics)	Enhanced degradation efficiency and stability of esterase.	Limited scalability studies for industrial applications.
A general Temperature-Guided Language model to engineer enhanced Stability and Activity in Proteins	Jiang et al., (2024) [13]	AI-based models for protein engineering	**PyTorch, Transformers Library** (temperature-guided language models)	Improved both stability and activity of proteins using AI.	Lack of comparative studies with traditional methods.

Information on studies regarding computational methods for improving enzyme thermostability is furthersummarized in Table 1. It contains the title, names of authors, date of publication, methodology or software used, important findings, and identified research gaps. The outcome suggests that hybrid strategies combining structure-based

design and machine learning would counterbalance the limitations of independent approaches and would be a strong strategy to improve enzyme thermostability.

With vivid differences with respect to efficacy and feasibility, almost all the reviewed studies in Table 1 have successfully conducted enzyme thermal stabilization.For example, in a work done by Tu et al. [9], extraordinary enhancement of phytase stability was reported at 100 °C, using an approach involving a combination of Schrödinger Suite with NAMD. However, the limitations of this method apply to a few groups of enzymes, whereas others have never been validated in industry yet.

Furthermore, Derat & Kamerlin [10] recount the importance of AI in fast-tracking enzyme design with software such as AlphaFold and PyMOL. Their results show AI much fast-forwards enzyme structure prediction and fruitful alterations.

Different studies show that indeed almost all the computational techniques are successfully applied to enhance enzyme thermostability, but those differ on their efficiency, scalability, and applicability in industry. For instance, Tu et al. [9] reported outstanding enhanced stability of phytase using Schrödinger Suite and NAMD, which supports the promise of combining advanced force fields with high parallelization for enzyme design. However, such approaches would be questionable when considering their industrial applications due to high computation costs and the need for extensive experimental verifications.

Similarly, the application of AI software including AlphaFold has taken to task the reliability of predictions in protein structure as stated by Derat & Kamerlin [10]. Nevertheless, the major drawback remains high computational expenses along with hardly any attention to cost-effectiveness, making these processes untenable for industrial application. Solving such problems implies further resource-poor AI models and integrating cost evaluation frameworks into computational protocols.

3. 1 New Methods to Enhance Enzyme Heat Stability

Software such as Rosetta and GROMACS are very useful in sharpening enzyme temperature forays with specific techniques. They assist in the design of enzyme structures by proposing alterations that stabilize the molecule, such as enhancing hydrogen bonds and salt bridges. In 2005, Korkegian et al. [4] showed that Rosetta could be used to modify three specific sections of an enzyme, this is also supported by [14] which also uses Rosetta to enzyme design and activity enhancement so that it makes enzyme was more stable and efficient when heated.

Machine learning is also playing an important role in making enzymes more temperature-resistant. Though less widely adopted to date, it has immense potential. Machine learning algorithms can also help researchers predict the chemical changes that make enzymes more stable. TensorFlow and Scikit-learn also assist with creating models that efficiently process complex data with picking up relations between system changes in the enzyme sequence and stability. In 2023, Dou et al. [7] successfully showed that

these models are capable of predicting enzyme stability with few experiments. The trouble is, these models can be difficult to interpret, particularly in understanding how the changes improve the stability of enzymes. Explainable AI models can be useful in this case since they are transparent when they point out why specific changes act in the intended way.

3. 2 Alternatives Sofware and Their Roles

Part of running successful computational research is the decision on what software to use. Table 2 presents a comparison of key software alternatives for molecular dynamics simulations, machine learning, protein structure modeling, molecular visualization, and protein programming, emphasizing their unique strengths and suitable contexts for application.

Table 2. Software Alternatives for Designing Thermostable, Efficient, and Cost-Effective Enzymes

Category	Software & Alternatives	Detailed Explanation
Molecular Dynamics Simulation	GROMACS ↔ NAMD ↔ AMBER	GROMACS excels in speed and is open-source, suitable for small to medium-scale simulations; NAMD is superior for large-scale simulations due to high parallelization; AMBER offers higher accuracy for complex proteins [15].
Machine Learning (ML)	Scikit-learn ↔ TensorFlow ↔ PyTorch ↔ Weka	Scikit-learn is ideal for traditional models (regression, SVM) with simple syntax; TensorFlow is strong for large-scale production with TPU support; PyTorch is flexible for research; Weka provides a GUI for non-coding analysis [16].
Protein Structure Modeling	MODELLER ↔ Rosetta ↔ AlphaFold ↔ AutoDock Vina	MODELLER focuses on homology modeling with templates, Rosetta supports de novo modeling and protein design, AlphaFold achieves the highest accuracy in structure prediction, and AutoDock Vina specializes in small-molecule docking [17]
Molecular Visualization	PyMOL ↔ VMD	PyMOL is preferred for high-quality publication images with detailed display control; VMD is tailored for analyzing molecular dynamics data with robust scripting capabilities.
Protein Programming	PyRosetta ↔ Rosetta	Rosetta offers diverse protocols for protein design and prediction, whereas PyRosetta allows for Python scripting to automate complex tasks with greater flexibility.
Protein Design	MODELLER ↔ Swiss-Model	MODELLER provides more comprehensive features for homology modeling with better parameter control, while Swiss-Model is more user-friendly for beginners seeking quick results without complex configurations.

Table 2 summarizes several software tools that are used for activities such as molecular dynamics simulations, machine learning, protein structure prediction, molecular visualization, and protein design. When selecting which software is best, it is important to consider the research requirements, i.e., accuracy, scalability, and computation costs. In some cases, merging and aligning different software tools will enhance the computation performance [18].

This side-by-side comparison illustrates the advantage of selecting your software according to the particular demands of your research environment. To illustrate, when choosing between NAMD and GROMACS for molecular dynamics computation, you would need to consider the size as well as the complexity of the system that is under study. GROMACS would most suitably be applicable to small systems, while NAMD is ideally suited for large systems due to its scalability architecture [19].

3. 3 Synergistic Potential of Hybrid Approaches

Hybrid approaches, where structure-based design and machine learning are combined, have been suggested as a remedy to surpass the drawback of using one approach in silico. A recent study by Dinmukhamed et al. [11] showed that hybrid models, where structural information is utilized for training machine learning models, performed significantly better than single approaches for predicting stabilizing mutations. This fusion brings together the predictiveness of machine learning and the structural precision of structure-based methods to enable possibilities of mutations to be explored even further. Using the incorporation of Rosetta's structural modeling into TensorFlow's machine learning framework as an example, this hybrid approach resulted in the creation of enzymes that not only performed better but even proved more stable in character, actually bridging the accuracy-scalability gap.

3. 4 Effectiveness and Limitations of Methods

All these computational approaches discussed here have both strengths and weaknesses. Structure-based design is excellent at giving very high-level control over structural modification but is very time-consuming and typically limited to small enzyme sizes [5]. Machine learning, on the other hand, while able to process big data very effectively, is not very transparent and requires a lot of training data [12].

The combination of rational design and machine learning offers a possible solution to such complexities. For example, Dinmukhamed et al. [11] established that the combination of these approaches not only enhanced the stability of enzymes but also accelerated the designing process and ensured more accurate results. In this method, structure simulation data are employed to train machine learning models, which then aid in designing improved mutation designs.

3. 5 Cost-Effectiveness of Computational and Hybrid Methods in Enzyme Design

In comparison to traditional experimental approaches, computer-aided and hybrid designs reduce the costs by incurring lower trial-and-error testing and cheaper reagent, equipment, and labor inputs [20]. Computer programming environments like Rosetta, GROMACS, TensorFlow, and Scikit-learn offer predictive accuracy through which scientists filter and rank more viable enzyme mutations to minimize expensive experimental testing. The computer assistances accelerate enzyme design workflow in terms of shorter development timelines and fewer overall expenses [21].

Also, these computational approaches can reduce failure rates through intentional design and predictive computation, which equate to fewer failed experiments and their costs. Their scalability, enabled by software and cloud computing, enable one to handle large volumes of data without spending much more money, which makes them suitable for industrial use. While the initial investment in computing hardware may be high, the long-term return on investment includes reduced dependence on trial-and-error approaches, compressed time-to-market for thermostable enzymes, and increased return on investment (ROI).

4. Conclusion

The review highlights the way computational approaches have compressed the design process for thermostable enzymes, through the combination of machine learning and structure-based rational design. The construction of hybrid strategies that blend these methods provides a promising route for enhancing enzyme stability without loss of catalytic function. Yet, there are a number of fundamental issues to be addressed, such as scalability, cost, and the need for thorough experimental validation of computational design predictions. These issues will be addressed through concerted research efforts, the establishment of standard experimental protocols, and calibration of calculation devices for wider industrial application. Further emphasis must be placed on the integration of data-driven methods and experimental back-loops to iteratively refine the design process and also to establish the practical viability of thermostable enzymes. Having overcome systematically these challenges, we can now employ the full power of computational tools to revolutionize enzyme engineering in industry.

Acknowledgements
The authors would like to express their sincere gratitude to the Industrial Chemical Engineering Technology program, Department of Mechanical Engineering, Politeknik Negeri Medan, for providing access to literature resources and continuous academic support throughout the preparation of this review. The authors also appreciate the valuable feedback from colleagues, which contributed to improving the clarity and depth of this manuscript.

CRediT Authorship Contribution Statement
Prisca Caesa Moneteringtyas: Conceptualization, Supervision, Writing – review & editing, Project administration. **Nahzim Rahmat:** Investigation, Formal analysis, Visualization. **Inten Pangestika:** Methodology, Writing – original draft, Data curation, **Sri Rahayu Widya Ningrum:** Software, Literature search, **Annisa Fillaeli:** Resources, Validation, Writing – review & editing. **Aliyah Aliyah:**Writing – review & editing, translating.

Conflicts of Interest
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Choi JM, Han SS, Kim HS. Industrial applications of enzyme biocatalysis: Current status and future aspects. Bioechnol Adv 2015;33(7):1443-1454. https://doi.org/10.1016/j.biotechadv.2015.02.014

[2] Rahban M, Zolghadri S, Salehi N, Ahmad F, Haertlé T, Rezaei-Ghaleh N, et al. Thermal stability enhancement: Fundamental concepts of protein engineering strategies to manipulate the flexible structure. Int J Biol Macromol 2022;214:642–654. https://doi.org/10.1016/j.ijbiomac.2022.06.154

[3] Zheng N, Cai Y, Zhang Z, Zhou H, Deng Y, Du S, et al. Tailoring industrial enzymes for thermostability and activity evolution by the machine learning-based iCASE strategy. Nat Commun 2025;16:604. https://doi.org/10.1038/s41467-025-55944-5

[4] Korkegian A, Black ME, Baker D, Stoddard BL. Computational Thermostabilization of an Enzyme. Science (1979) 2005;308:857–860. https://doi.org/10.1126/science.1107387

[5] Modarres HP, Mofrad MR, Sanati-Nezhad A. Protein thermostability engineering. RSC Adv 2016;6(116):115252–115270. **https://doi.org/10.1039/C6RA16992A**

[6] Nezhad NG, Rahman RNZRA, Normi YM, Oslan SN, Shariff FM, Leow TC. Thermostability engineering of industrial enzymes through structure modification. Appl Microbiol Biotechnol 2022;106:4845–4866. https://doi.org/10.1007/s00253-022-12067-x

[7] Dou Z, Sun Y, Jiang X, Wu X, Li Y, Gong B, et al. Data-driven strategies for the computational design of enzyme thermal stability: trends, perspectives, and prospects. Acta Biochim Biophys Sin (Shanghai) 2023;55:343–355. https://doi.org/10.3724/abbs.2023033

[8] Pongsupasa V, Anuwan P, Maenpuen S, Wongnate T. Rational-Design Engineering to Improve Enzyme Thermostability. In: Magnani F, Marabelli C, Paradisi F, editors. Enzyme Engineering: Methods and Protocols, New York, NY: Springer US; 2022, p. 159–178. https://doi.org/10.1007/978-1-0716-1826-4_9

[9] Tu T, Wang Q, Dong R, Liu X, Penttinen L, Hakulinen N, et al. Achieving thermostability of a phytase with resistance up to 100 oC. Journal of Biological Chemistry 2024;300(12). https://doi.org/10.1016/j.jbc.2024.107992

[10] Derat E, Kamerlin SCL. Computational Advances in Protein Engineering and Enzyme Design. J Phys Chem B 2022;126(13):2449–2451. https://doi.org/10.1021/acs.jpcb.2c01198

[11] Dinmukhamed T, Huang Z, Liu Y, Lv X, Li J, Du G, et al. Current advances in design and engineering strategies of industrial enzymes. Systems Microbiology and Biomanufacturing 2021;1(1):15–23. https://doi.org/10.1007/s43393-020-00005-9

[12] Xie B, Zhang J, Sun H, Bai R, Lu D, Zhu Y, et al. Computational design of an efficient and thermostable esterase for polylactic acid depolymerization. Green Chemistry 2024;26(12):7268–7279. https://doi.org/10.1039/D3GC04888H

[13] Jiang F, Li M, Dong J, Yu Y, Sun X, Wu B, et al. A general temperature-guided language model to design proteins of enhanced stability and activity. Sci Adv 2024;10:eadr2641. https://doi.org/10.1126/sciadv.adr2641

[14] Sun R, Wu D, Chen P, Zheng P. Cutting-edge computational approaches in enzyme design and activity enhancement. Biochem Eng J 2024;212:109510. https://doi.org/10.1016/j.bej.2024.109510

[15] Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K. Scalable molecular dynamics with NAMD. J Comput Chem. 2005 Dec;26(16):1781-1802. https://doi.org/10.1002/jcc.20289

[16] Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al. TensorFlow: A system for large-scale machine learning 2016.https://doi.org/10.48550/arXiv.1605.08695

[17] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021;596:583–589. https://doi.org/10.1038/s41586-021-03819-2

[18] Krylov A, Windus TL, Barnes T, Marin-Rimoldi E, Nash JA, Pritchard B, et al. Perspective: Computational chemistry software and its advancement as illustrated through three grand challenge cases for molecular science. J Chem Phys 2018;149:180901. https://doi.org/10.1063/1.5052551

[19] A. Poghosyan, A. ARTASHES Shahinyan, A. H. Poghosyan, G. A. Yeghiazaryan, H. H. Gharabekyan, and A. A. Shahinyan, "The GROMACS and NAMD Software Packages Comparison," *Commun Comput Phys*, vol. 1, no. 4, pp. 736–743, Aug. 2006, [Online]. Available: http://www.global-sci.com/

[20] Landes SJ, McBain SA, Curran GM. An introduction to effectiveness-implementation hybrid designs. Psychiatry Res 2019;280:112513. https://doi.org/10.1016/j.psychres.2019.112513

[21] Landwehr GM, Bogart JW, Magalhaes C, Hammarlund EG, Karim AS, Jewett MC. Accelerated enzyme engineering by machine-learning guided cell-free expression. Nat Commun 2025;16:865. https://doi.org/10.1038/s41467-024-55399-0