

Comparative In-Silico Sequence and Structural Analysis of Catalase from *Escherichia coli* and Humans

*¹Sadiq Salisu Yunusa, ¹Muhammad Ali Dikwa, ²Abubakar Yusuf Abubakar, ²Yasir Ahmad, ³Sani Shuaibu Kafin Hausa and ⁴Usman Yahaya

¹Department of Microbiology and Biotechnology, Federal University Dutse, Jigawa State, Nigeria.

²Department of Science Laboratory and Technology, Jigawa State Polytechnic Dutse, Jigawa State.

³Department of Biology, Jigawa State University of Medical and Allied Health Sciences Majia, Jigawa State.

⁴Department of Computer Science, Federal University Dutse, Jigawa State.

*Corresponding authors' email: sadiqsalisu3636@gmail.com

ABSTRACT

Hydrogen peroxide (H₂O₂) is a significant reactive oxygen species that can lead to cellular damage, particularly in eukaryotic organisms. Catalase, an enzyme, plays a crucial role in detoxifying H₂O₂ and sustaining cellular functions. Although substantial research has been conducted on catalase, limited comparative analyses examining the catalase sequences between *Escherichia coli* and humans, particularly in the context of active sites, exist. This study retrieved and analyzed the amino acid sequences and three-dimensional structures of catalase from both organisms using NCBI and PDB, alongside tools such as Clustal Omega, BLAST, MEGA, and PyMOL. The multiple sequence alignment identified 200 conserved, 122 physiochemically similar, and 163 variable amino acid residues, corresponding to an overall sequence identity of 45.17%. Phylogenetic analysis of a small representative dataset illustrated the evolutionary relationship between *E. coli* and human catalase. Structural superimposition of the active-site regions yielded a root mean square deviation (RMSD) of 0.295 Å, indicating close structural correspondence in this region despite the divergence in overall sequence and length. Essential functional residues, including active sites, domains, and motifs, were identified and compared. This computational analysis offers preliminary insights into the evolutionary and structural relationships of catalase between the two organisms, which may help direct future experimental work on oxidative-stress-related enzyme function.

Received: 05 Sept 2026

Accepted: 12 Sept. 2026

Published: 06 Oct. 2026

Keywords: In-silico analysis, Catalase, *Escherichia coli*, Sequence alignment, Oxidative stress

INTRODUCTION

Catalase (EC 1.11.1.6) is a vital antioxidant enzyme found in nearly all aerobic organisms, primarily responsible for regulating hydrogen peroxide (H₂O₂) levels and other cytotoxic oxygen derivatives (Baker et al., 2023). It is an important factor in various biological processes, including inflammation, mutagenesis, prevention of programmed cell death, and inhibition of tumor growth (Nandi et al., 2019; Tian et al., 2026). The enzyme was first crystallized from bovine liver in 1937 (Glorieux and Calderon, 2017), while the first prokaryotic catalase was isolated from the aerobic bacterium *Micrococcus lysodeikticus* in 1948. The gene coding for human catalase is located on chromosome 11, whereas *Escherichia coli* synthesizes two distinct catalases, hydroperoxidase I and II, which differ in physical structure and catalytic capabilities (Scholz et al., 2026).

Numerous studies have been conducted on prokaryotic and lower eukaryotic catalases. Catalase is employed as a therapeutic agent due to its ability to decompose hydrogen peroxide into water and oxygen, effectively combating oxidative stress-related diseases (Nandi et al., 2019). It is increasingly viewed as a promising therapeutic target for diseases linked to oxidative stress, including cancer, neurodegenerative conditions, and cardiovascular diseases (Rasheed, 2024). Treatments based on catalase are believed to offer neuroprotective benefits by preserving neuron integrity and functionality, thus opening paths for the treatment or prevention of neurodegenerative diseases (Olufunmilayo et al., 2023). Despite extensive investigation,

the comparative analysis of catalase sequences between prokaryotes and eukaryotes, specifically *E. coli* and humans, remains underexplored. This study aims to address this deficiency by conducting a comparative analysis of the catalase enzyme to identify conserved regions with functional significance and therapeutic implications. Such comparative studies can enhance the understanding of the enzyme's functional and evolutionary relationships across species, bolstering its potential applications in biotechnology and medicine. This research involved retrieving amino acid sequences for alignment and analyzing the evolutionary relationships followed by a structural comparison to examine the similarities and differences between the two species.

MATERIALS AND METHODS

Data Retrieval

Amino acid sequences of catalase from *Escherichia coli* and human sources were collected in FASTA format from the National Center for Biotechnology Information (NCBI) databases (Sayers et al., 2022). *E. coli* encodes two distinct catalase isoenzymes, hydroperoxidase I (HPI, encoded by katG) and hydroperoxidase II (HPII, encoded by katE), which differ in structure, regulation, and catalytic behaviour (Scholz et al., 2026). The monofunctional, large-subunit catalase KatE (HPII) was selected for this comparison because, unlike the bifunctional catalase-peroxidase KatG, it is structurally and mechanistically more directly comparable to the monofunctional human catalase (CAT), making it the more biologically appropriate reference isoenzyme for this

analysis. Conclusions drawn here therefore pertain specifically to the KatE (HP1I) isoform of *E. coli* catalase and should not be generalised to *E. coli* catalase activity as a whole. The accession numbers and PDB identifiers used for both sequences and structures (Table 1) were cross-checked against the NCBI and RCSB PDB databases prior to analysis.

Multiple Sequence Alignment (MSA)

The collected sequences were aligned using Clustal Omega (version 1.2.4, EMBL-EBI web server, accessed 2025) with default parameters to identify conserved and variable regions between the two species (Sievers and Higgins, 2018). Residues were classified as conserved (identical residue at an aligned position in both sequences), similar (non-identical residues sharing comparable physicochemical properties such as size, charge, or polarity, based on the Clustal similarity-scoring groups), or variable (aligned positions with no conservation of identity or physicochemical property).

Sequence Identity and Similarity

Sequence identity and similarity were assessed using BLASTp (NCBI) with the BLOSUM45 scoring matrix, selected because its lower substitution penalties are better suited to detecting distant homology between evolutionarily divergent, low-identity sequences such as prokaryotic and human catalase, and an E-value threshold of 0.05, following the approach of Shahid et al. (2021) and Timsina et al. (2025). The alignment length, percentage identity, percentage similarity, number and location of gaps, and query/subject coverage returned by BLASTp were recorded, and identical residues (percentage identity) were distinguished from chemically conservative substitutions (percentage similarity) throughout the Results.

Phylogenetic Tree Construction

A phylogenetic tree was constructed using the maximum likelihood method with the JTT (Jones-Taylor-Thornton) substitution model and 100 bootstrap replicates in MEGA version 11 (Tamura et al., 2021), following sequence

alignment in Clustal Omega. Given the limited number of taxa included, this tree is intended only as a preliminary illustration of the relationship between *E. coli* and human catalase rather than a robust phylogenetic reconstruction; a taxonomically broader dataset and $\geq 1,000$ bootstrap replicates, with node-specific bootstrap support explicitly reported, would be required to draw firmer evolutionary conclusions.

2.5 Structural Comparison

Three-dimensional structures of *E. coli* KatE (PDB: 1IPH) and human catalase (PDB: 1QQW) were obtained from the RCSB PDB and superimposed using the PyMOL "align" command (PyMOL Molecular Graphics System; Jones, 2021), restricted to the active-site residues listed in Table 2 rather than the full-length structures, given the substantial difference in overall sequence length and identity between the two proteins. The number and percentage of atoms/residues included in the final alignment after PyMOL's outlier-rejection cycles, together with the alignment coverage, were recorded alongside the resulting root mean square deviation (RMSD).

Active Site and Functional Motif Identification

Active sites, motifs, and conserved domains across the two sequences were identified and analyzed using Pfam, UniProt, and relevant literature (Putnam et al., 2000; Mistry et al., 2021). Residues were further classified by functional role, distinguishing catalytic residues (directly involved in H₂O₂ turnover), substrate-binding residues (involved in positioning H₂O₂ within the active site), and structurally important residues (maintaining active-site architecture without direct catalytic or binding roles), based on the structural alignment between the two proteins.

RESULTS AND DISCUSSION

Results

The catalase sequences and their three-dimensional structures were retrieved from various databases, and their accession numbers are listed in Table 1.

Table 1: Retrieved Sequence Accession numbers of Catalase in *Escherichia coli* and Humans from Different Protein Databases

S/N	Organism	Gene	NCBI	PDB	References
1	<i>Escherichia coli</i>	KatE	CAK0673004	1IPH	Bravo <i>et al</i> , 1995
2	<i>Homo sapiens</i>	CAT	CAA27721	1QQW	Putnam <i>et al</i> , 2000

Note: All accession numbers and PDB identifiers listed above were verified against the NCBI and RCSB PDB databases prior to analysis. The *E. coli* entry corresponds specifically to the KatE (HP1I) isoenzyme; conclusions regarding "E. coli catalase" in this study should be read as pertaining to this isoenzyme and not generalised to the katG-encoded HPI catalase

Sequence Analysis

The alignment of the two catalase sequences using the Clustal Omega tool comprised 485 aligned positions, of which 200 residues were identical (conserved), 122 were non-identical but shared similar physicochemical properties such as size, charge, or polarity (similar), and 163 showed no conservation of identity or property (variable); these figures refer to individual conserved amino acid residues rather than contiguous conserved regions. The complementary BLASTp

analysis (Section 2.3) returned an overall percentage identity of 45.17% across the aligned region, together with the corresponding percentage similarity, alignment length, number of gaps, and query/subject coverage; the modest numerical difference between this value and the proportion of conserved positions from the Clustal Omega alignment ($200/485 = 41.2\%$) reflects differences in gap handling and scoring between the two methods, as shown in fig 1 below:

Tool Output	Alignments	Guide Tree	Phylogenetic Tree	Results Viewers	Result Files	Submission
Tool output	CLUSTAL O(1.2.4) multiple sequence alignment					
Download						
CAA27721 CAK0673004	-----MADSRDP MSQQNEKNPQHQSPLHDSSEAKPGMDSLAPEDGSHRPAAEPTTPGAQPTAPGSLKAPDT . : *					7 60
CAA27721 CAK0673004	ASDOMQHWKEQRAAQKADVLITGAGNPVGDKLVNITVGRPLLVQDVVITDEMAHFDRE RNEKLNLEDVRKGSSENYALTNOGVRIADDONSLRAGNRGPTLLDFILREKITHFDE . :					

CAA27721

RIPERVVHAKGAGAFGYFEVTHDITKYSKAKVFEHIGKKTPIAVRFSTVAGESGADTVR

127

CAK0673004

RIPERTVHARGSAAHGYFQPKSLSDITKADFLSDPNKITPVFVRFSTVQGGAGSADTVR

180

CAA27721

DPRGFAVKFYTEDGNWDLVGNNTPIFFIRDPILFSPFIHSOKRNPQTHL---

183

CAK0673004

DIRGFATKFTYEEGIFDLVGNNTPIFFIQDAHKFPDFVHAVKPEPHWAIPOGQSAHDTFW

240

CAA27721

DFWLSRPESLHVQVSLFSDRGIPDGRHMMNGYGSHTFKLVNANGEAVYCKFHYKTQGIK

243

CAK0673004

DYVSLQPETLHNVMWMSDRGIPRSYRTMEGFIHTFRLINAEKGATFVRHFWKPLAGKA

300

CAA27721

NI SVE DAARI SQEDPDYGI RDLFNATATGKYP SWIT YIQVMTF NQAETFP NPF DL TKVW

303

CAK0673004

SLVWDEAOKLTGRDPDFHRRELWEAIEAGDFPEYELGFLIPEEDEFKDFDLDLP TKLI

360

CAA27721

PHKDYPLIPVCKLVLRNPNVYFAEVEQIAFDPSNMPPGIEASPKHLOGRLFAYPDTHR

363

CAK0673004

PEELVPVQVRGKMLNRNPDNFFAENEQAAPHGHTVPGLDFTNDPLLQGRFLFSYTDQTI

420

CAA27721

HRLG-PNYLHIPVNCYRVARVANYORDGPMCMODNOGGAPNYYPSNFGAP--EEOPSALE

420

CAK0673004

SRLGGPNFHEIPINRPTC-PYHNFORDGMHRMGID-TNPANYEPNSINDNWPRETTPPGPK

478

CAA27721

-HST-QYSG----EVRRENTANDDNVTQVRAFVYVNVINEEQKRLCENIAGH-LKDAQI

472

CAK0673004

RGGEESYQERVEGNKVRERSPSFGEYYSHIRLEWLSQITPEQRHIV-DGFSFELSKVVRP

537

CAA27721

FIQKKA VKNFTEVHPDYGSHI-----QALLDKYNAEKPK-----NAIHTEVQSGSH

518

CAK0673004

YIRERVVQDLAHTDLTAQAVAKNLGIELTDDQLNITPPPDVNLKKPKSLSLYAIPDGD

597

CAA27721

LAAREKANL-----

527

CAK0673004

VKGRVVAILLNDEVRSADLLAILKALKAKGVHAKLLYSRMGEVTADDGTVLPAAATFAGA

657

CAA27721

527

CAK0673004

PSLTVDIVPCGNTADTADNGDANYILMEAYKHLKPIALAGDARKFKATIKVADQGEFG

717

CAA27721

527

CAK0673004

TVVADSADGSEFMDLTLTMAHVRVWSRIPKTDKTPA-----

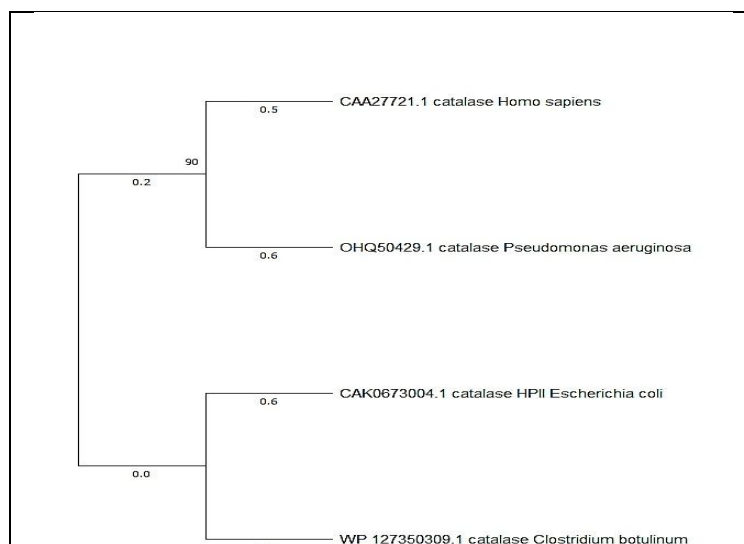
753

Figure 1: Multiple Sequence Alignment of Catalase in *Escherichia coli* and Human using Clustal Omega

Phylogenetic Analysis

Phylogenetic tree using maximum likelihood illustrates the evolutionary relationships of humans, *Escherichia coli* and other two bacteria species namely *Pseudomonas aeruginosa* and *Clostridium botulinum*, with a bootstrap support value of

90% at the key node based on 100 replicates, as shown below in fig 2. Given the small number of taxa (n = 4) included in this analysis, the tree should be interpreted as an illustrative comparison rather than a definitive phylogenetic reconstruction:

Figure 2: Phylogenetic Tree of Catalase in *Escherichia coli*, Human and other Prokaryotes using MEGA

Three-dimensional Structure Analysis

The 3D structural comparison gives an RMSD value of 0.295 Å corresponding coverage reported alongside this value, as shown in the superimposed active-site residues (Table 2), using the PyMOL "align" command as described in Section 2.5, with the

number/percentage of atoms included in the final alignment and the

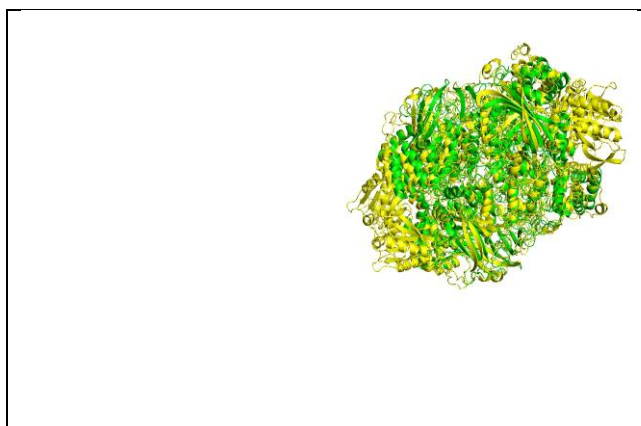


Figure 3: 3D Structure Comparison of Catalase Active Sites in *Escherichia coli* and Human using PyMOL

The key residues within both enzymes' active sites are retrieved from existing literatures and are presented in Table 2.

Table 2. Retrieved Active Sites Sequences of Catalase in *Escherichia coli* and Human

S/N	Organism	Number of Amino Acids	Amino acid sequences
1	<i>Escherichia coli</i>	5	His-128, Ser-167, Asn-201, Arg-411, Tyr-415
2	<i>Homo sapiens</i>	4	His-75, Asn-148, Arg-354, Tyr-358

References: Bravo et al, 1995, Putnam et al, 2000

Of these residues, the histidine and asparagine/arginine pairs (*E. coli* His-128/Asn-201; human His-75/Asn-148) constitute the core catalytic residues directly involved in H₂O₂ turnover, the arginine and tyrosine residues (*E. coli* Arg-411/Tyr-415; human Arg-354/Tyr-358) are primarily substrate-binding residues that help position H₂O₂ within the active site, and serine-167 in *E. coli* is a structurally important residue with no direct human counterpart in this alignment. The structural alignment used to derive the RMSD in Section 3.3 confirms a positional correspondence between each *E. coli* residue and its listed human counterpart.

Discussion

The aligned sequences (see Figure: 1) analyzed using the Clustal Omega tool revealed notable sequence-based similarities. These results underscore the close relationship and conservation of these enzymes from prokaryotes to eukaryotes, despite the presence of divergent residues. The multiple sequence alignment showed a significant number of conserved and functionally similar residues, with a moderate sequence identity of 45.17%. This indicates that, despite the variations, essential residues necessary for enzyme functionality are largely retained.

Moreover, the highly conserved amino acids: glycine, proline, aspartic acid, arginine, and phenylalanine, across multiple regions suggest that these residues have played a vital role in the evolution, structural stability, and catalytic function of catalase in both prokaryotes (e.g., *Escherichia coli*) and eukaryotes (such as humans). These amino acids are crucial for the enzyme's functional activity, including protein folding, flexibility, and maintaining the integrity of the active site, which is essential for efficiently decomposing hydrogen peroxide. Similar findings were reported by Shahid et al. (2021), based on 45 full-length catalase enzyme sequences from diverse organisms.

The phylogenetic tree (see Figure 2) illustrates two clusters within this limited dataset: one containing human catalase and *Pseudomonas aeruginosa* (NCBI: OHQ50429), and the other comprising *E. coli* catalase and *Clostridium botulinum* (WP_127350309). Because only four taxa were included, this

clustering pattern, and in particular the apparent grouping of human catalase with *P. aeruginosa*, should be treated as preliminary and re-examined using a larger, taxonomically representative dataset before being interpreted as a robust phylogenetic signal. The separation between the *E. coli* and human catalase sequences is nonetheless consistent with the known evolutionary divergence between prokaryotes and eukaryotes, and the conservation of the active-site residues described earlier is consistent with, though not sufficient on its own to demonstrate, descent from a common catalase ancestor and a long-conserved role in oxidative-stress defence.

The structures of *Escherichia coli* and human catalase were elucidated using X-ray diffraction and are available in the PDB database. When aligning the structures with a focus on the active sites, the root mean square deviation (RMSD) was found to be 0.295 Å, demonstrating a high level of similarity in these vital regions despite variations in sequence and length (Jones, 2021). This finding is crucial since structural conservation, especially within active sites, often holds greater importance than sequence similarity in preserving enzymatic activity. The 3D structure comparison (see Figure 3) illustrates the conserved and variable regions of *Escherichia coli* catalase (yellow) and human catalase (green). Key catalytic residues, including asparagine, arginine, and tyrosine, further support the preservation of the catalytic mechanism that detoxifies hydrogen peroxide.

While *Escherichia coli* (1-493) and human (24-490) catalase domains differ in length, they represent the essential domain of mono-functional, heme-containing catalase, contributing to structural stability and enzyme integrity (Blum et al., 2025). Additionally, the human motif (524-527) serves as a peroxisomal targeting signal, allowing catalase to localize to peroxisomes and protect cells from oxidative damage (Purdue and Lazarow, 1996). These regions facilitate the effective function of catalase in both organisms across various cellular environments while preserving its core activity (Baker et al., 2023).

The conservation of active-site residues identified here supports the view that catalase's core mechanism for

mitigating oxidative stress has been maintained across the evolutionary distance separating *E. coli* and humans. As noted by Xu et al. (2025), human catalase has separately been engineered through recombinant DNA technology and site-directed mutagenesis to enhance its reaction rate for industrial applications, illustrating that its catalytic efficacy can be modulated. However, the present sequence and structural comparison does not itself establish any disease-specific or therapeutic relevance; extrapolating these computational findings to specific conditions such as cancer or neurodegenerative disease would go beyond what this analysis can support.

Although these findings indicate potential therapeutic applications, caution is warranted—this study is purely in-silico and does not provide direct evidence of therapeutic efficacy in biological systems. Further research, including molecular docking and experimental validation, is necessary to investigate the feasibility of employing catalase in clinical or biotechnological settings.

CONCLUSION

This research compared the catalase enzyme of *Escherichia coli*, specifically the KatE (HP11) isoenzyme, with human catalase, encoded by the KatE and CAT genes, respectively, through multiple sequence alignment, a preliminary phylogenetic analysis, and active-site structural superimposition. The results revealed moderate overall sequence identity alongside strong conservation of key active-site residues, and, within the limits of the small taxon sample used, a pattern of divergence consistent with the known evolutionary separation of prokaryotes and eukaryotes. These computational findings indicate that catalase's core active-site architecture has been substantially conserved across this evolutionary distance, which is consistent with its long-standing role in detoxifying hydrogen peroxide, a reactive oxygen species implicated in oxidative-stress-related cellular damage in both prokaryotic and eukaryotic organisms. Establishing any specific relevance of these findings to particular diseases or therapeutic applications, or generalising them beyond the KatE isoenzyme of *E. coli*, would require dedicated experimental follow-up beyond the scope of this in-silico comparison.

REFERENCES

- Baker, A., Lin, C. C., Lett, C., Karpinska, B., Wright, M. H., & Foyer, C. H. (2023). Catalase: A critical node in the regulation of cell fate. *Free radical biology & medicine*, 199, 56–66. <https://doi.org/10.1016/j.freeradbiomed.2023.02.009>
- Nandi, A., Yan, L. J., Jana, C. K., & Das, N. (2019). Role of Catalase in Oxidative Stress- and Age-Associated Degenerative Diseases. *Oxidative medicine and cellular longevity*, 2019, 9613090. <https://doi.org/10.1155/2019/9613090>
- Tian, Y., Liu, Y. R., Jin, H. Z., Lin, Q. X., Zhao, W. Y., Song, W. W., Gong, Y. N., Deng, Y. T., Wang, S. S., Wang, K., Tian, L., & Gu, D. N. (2026). Catalase inhibits tumor growth by alleviating oxidative stress and promoting the polarization of tumor-associated macrophage from M2 to M1 phenotype in non-small cell lung cancer. *Biology direct*, 21(1), 24. <https://doi.org/10.1186/s13062-026-00730-6>
- Glorieux, C. & Calderon, P. (2017). Catalase, a remarkable enzyme: targeting the oldest antioxidant enzyme to find a new cancer treatment approach. *Biological Chemistry*, 398(10), 1095–1108. <https://doi.org/10.1515/hsz-2017-0131>
- Scholz, A. T., Coppo, L., Nolan, E., Hernandez, M., Wang, X., Mishra, P., Schnell, R., Pató, Z. A., Chen, Y., Dagnell, M., Andor, A., Cheng, Q., & Arnér, E. S. J. (2026). A katE katG double-knockout *E. coli* strain eliminates the risk of catalase contamination in recombinant proteins. *Applied microbiology and biotechnology*, 110(1), 136. <https://doi.org/10.1007/s00253-026-13820-2>
- Rasheed, Z. (2024). Therapeutic potentials of catalase: Mechanisms, applications, and future perspectives. *International Journal of Health Sciences*, 18(2), 1–6. PMID: 38455600.
- Olufunmilayo, E. O., Gerke-Duncan, M. B., & Holsinger, R. M. D. (2023). Oxidative stress and antioxidants in neurodegenerative disorders. *Antioxidants*, 12(2), 517. <https://doi.org/10.3390/antiox12020517>
- Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., Connor, R., Funk, F., Murphy, T., Phan, L., Skripchenko, Sherry, S. T. (2022). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, 50(D1), D20–D26. <https://doi.org/10.1093/nar/gkab1112>
- Shahid, M., Amjad, M., Ashraf, U., Jamal, A., & Wattoo, J. (2021). Computational analysis of catalase from different source organisms. *Pakistan Journal of Botany*. 54(35). <https://doi.org/10.30848/PJB2022-1>
- Sievers, F., & Higgins, D. G. (2018). Clustal Omega for making accurate alignments of many protein sequences. *Protein science : a publication of the Protein Society*, 27(1), 135–145. <https://doi.org/10.1002/pro.3290>
- Timsina, R., Gora, R. A., & Ferguson, D. J., Jr (2025). Proteomic and metabolomic analysis reveals new insights into quaternary amine metabolism in *Citrobacter amalonaticus* CJ25. *mSphere*, 10(9), e0042125. <https://doi.org/10.1128/msphere.00421-25>
- Jones, D. (2021). Student Tutorial for PyMOL: for educational use Visualizing proteins using molecular graphics. 10.13140/RG.2.2.33676.64641
- Putnam, C. D., Arvai, A. S., Bourne, Y., & Tainer, J. A. (2000). Active and inhibited human catalase structures: Ligand and NADPH binding and catalytic mechanism. *Journal of Molecular Biology*, 296, 295–309
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic acids research*, 49(D1), D412–D419. <https://doi.org/10.1093/nar/gkaa913>
- Bravo, J., Verdaguer, N., Tormo, J., Betzel, C., Switala, J., Loewen, P. C., & Fita, I. (1995). Crystal structure of catalase HP11 from *Escherichia coli*. *Structure (London, England : 1993)*, 3(5), 491–502. [https://doi.org/10.1016/s0969-2126\(01\)00182-4](https://doi.org/10.1016/s0969-2126(01)00182-4)

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA 11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, 38(7), 3022-3027. <https://doi.org/10.1093/molbev/msab120>

Xu, S., Chen, Y. M., Meng, X. Y., Pan, R., Yan, A. X., Li, Z. M., & Li, Z. L. (2025). Computational-assisted protein engineering to develop thermostable and highly active catalase for industrial and biocatalytic applications. *Bioresource technology*, 437, 133081. <https://doi.org/10.1016/j.biortech.2025.133081>

Blum, M., Andreeva, A., Florentino, L. C., Chuguransky, S. R., Grego, T., Hobbs, E., Pinto, B. L., Orr, A., Paysan-Lafosse, T., Ponamareva, I., Salazar, G. A., Bordin, N.,

Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D. H., Letunic, I., Llinares-López, F., Marchler-Bauer, A., Meng-Papaxanthos, L., Mi, H., Natale, D. A., Orengo, C. A., Pandurangan, A. P., Piovesan, D., Rivoire, C., Sigrist, C. J. A., Thanki, N., Thibaud-Nissen, F., Thomas, P. D., Tosatto, S. C. E., Wu, C. H., & Bateman, A. (2025). InterPro: The protein sequence classification resource in 2025. *Nucleic Acids Research*. Advance online publication. <https://doi.org/10.1093/nar/gkae1082>

Purdue, P. E., & Lazarow, P. B. (1996). Targeting of human catalase to peroxisomes is dependent upon a novel COOH-terminal peroxisomal targeting sequence. *The Journal of cell biology*, 134(4), 849–862. <https://doi.org/10.1083/jcb.134.4.849>



©2026 This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license viewed via <https://creativecommons.org/licenses/by/4.0/> which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited appropriately.