Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Integrated genomic, structural, and biochemical characterization of a thermostable manganese superoxide dismutase from Geobacillus thermoleovorans GBRC-MXT211 isolated from Tapovan hot spring

Ruchita Shelat
Gujarat Biotechnology Research Centre
Dhruvi Bhatt
Gujarat Biotechnology Research Centre
Niraj Kumar Singh
National Forensic Sciences University
Chaitanya Joshi
Gujarat Biotechnology Research Centre
Snehal Bagatharia
Gujarat Biotechnology Research Centre
Vikas Patidar
Gujarat Biotechnology Research Centre
Satyamitra Shekh
Gujarat Biotechnology Research Centre
Amrutlal Patel
Gujarat Biotechnology Research Centre
Krishna Bharwad
krishna.scib@gbrc.res.in

Gujarat Biotechnology Research Centre

Article

Keywords: Superoxide dismutase, Geobacillus spp., thermostability, Purification, Whole genome sequencing, Manganese type

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11655694/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Integrated genomic, structural, and biochemical characterization of a thermostable manganese superoxide dismutase from *Geobacillus thermoleovorans* GBRC-MXT211 isolated from Tapovan hot spring

Ruchita Shelat#, Dhruvi Bhatt#, Niraj Kumar Singh, Chaitanya Joshi, Snehal Bagatharia, Vikas Patidar, Satyamitra Shekh*, Amrutlal Patel*, Krishna Bharwad*

Affiliations: Gujarat Biotechnology Research Centre (GBRC), Department of Science and Technology, Government of Gujarat, Gandhinagar-382010, Gujarat, India.

*Corresponding authors

#Authors share equal contributions

Correspondence email: krishna.scib@gbrc.res.in

Abstract

Thermophilic microorganisms are a promising source of stable antioxidant enzymes for industrial and biomedical applications. In this study, we isolated a superoxide dismutase (SOD)-producing bacterium, *Geobacillus thermoleovorans* GBRC-MXT211, from the Tapovan hot spring; the isolate grows at up to 80°C. Annotation of its genome identified an SOD gene, providing a genomic basis for sequence and structural analyses. Multiple sequence alignment of 19 SOD sequences revealed a conserved 3-His/1-Asp metal-coordination motif characteristic of the Fe/Mn-SOD superfamily, and AlphaFold2 modelling placed these residues in proximity within the predicted active site. Ammonium sulphate fractionation was used to partially purify the enzyme, with the 80% saturation fraction showing the highest activity; the active protein was supported as an Mn-SOD by MALDI-TOF MS/MS, with a top Mascot ion score of 62 and 51 in two independent digests, against the identity threshold of 41 ($p < 0.05$). The partially purified enzyme was active over a broad pH range, with similarly high activity from pH 6 to 10, and retained substantial activity after one hour of pre-incubation at temperatures up to 95°C. Its activity was not markedly reduced by hydrogen peroxide or sodium azide, a resistance profile consistent with a manganese-dependent SOD (Mn-SOD). Together, these results indicate that GBRC-MXT211 produces a short-form Mn/Fe-type SOD, most likely Mn-dependent, with notable thermal stability and broad pH tolerance, supporting further investigation for applications that require antioxidant enzymes functioning at elevated temperature or variable pH.

Keywords: Superoxide dismutase, *Geobacillus* spp., thermostability, Purification, Whole genome sequencing, Manganese type

Paper highlights

- Thermophilic *Geobacillus thermoleovorans* GBRC-MXT211 isolated from Tapovan hot spring
- Hybrid 3.48 Mb genome confirms the species and encodes three SOD-family proteins
- Alignment and AlphaFold2 model support a canonical Fe/Mn-SOD active site
- Extracellular SOD stays active at pH 6–10 and after 1 h pre-incubation at 95 °C
- Peroxide/azide resistance and mass spectrometry point to an Mn-SOD

1. Introduction

Superoxide dismutase (SOD) is a vital metalloenzyme that catalyses the dismutation of superoxide radicals (O_{2^-}) into molecular oxygen (O_2) and hydrogen peroxide (H_2O_2), thereby serving as a key enzymatic component of defence against reactive oxygen species (ROS) in living systems [1]. Depending on their metal co-factors, SODs are divided into four major groups: Cu/Zn-SODs, which are mainly found in eukaryotes (although they have also been reported from several bacterial species), Mn-SODs, which are predominantly present in mitochondria and bacteria, Fe-SODs, which are commonly found in plant chloroplasts and prokaryotes, and Ni-SODs, which are limited to certain prokaryotic lineages [2]. Owing to their potential antioxidative properties, SODs have gained widespread interest in the biotechnological field and are widely studied for therapeutic purposes, nutraceuticals, and the cosmetic and food industries [3].

The increasing demand for industrially important thermozymes (such as thermostable SODs) has led researchers to investigate extremophiles. Extremophiles are microorganisms that thrive under extreme conditions such as extreme pH, high or low temperature, salinity or pressure and are a rich source of novel biocatalysts [4]. Among extremophiles, thermophiles and hyperthermophiles are mainly found in geothermal environments and are important reservoirs of thermostable enzymes, also referred to as thermozymes. Thermozymes maintain their structural stability and enzyme activity even at higher temperatures in contrast to their mesophilic counterparts, which usually denature under such conditions [5]. Owing to this, thermozymes are suitable for many industrial applications.

Hot springs are among the most studied geothermal ecosystems worldwide. The Indian Himalayan Region (IHR) contains numerous geothermal hot springs with significant potential for exploring their microbiome, which is still underexplored in many places [6]. One such compelling site is the Tapovan hot spring, situated in Chamoli district of Uttarakhand, India which has not yet been fully investigated. This site is known to harbour a diverse thermophilic microbial community due to high temperatures (reaching up to 95°C), slightly alkaline pH and high sulphur content [7]. Previous studies conducted at this hot spring have reported various thermophilic genera, including *Geobacillus*, *Anoxybacillus*, *Thermus*, *Bacillus*, and *Symbiobacterium* with enzyme capabilities including amylases, lipases, cellulases and proteases [6]. However, thermostable SODs and their associated antioxidant properties from this hot spring remain understudied.

Many Gram-positive, endospore-forming thermophilic bacteria have been well documented from the Tapovan hot spring [8]. Among these, members of the genus *Geobacillus* (part of *Bacillaceae* family) have gained widespread attention due to their ability to produce thermozymes. This is attributed to their ability to grow optimally between 55°C and 80°C while demonstrating broad pH and temperature stability. Species such as *Geobacillus thermocatenulatus*, *G. stearothermophilus*, and *G. thermoleovorans* have been reported from Tapovan as well as from other hot springs in the northern states across Uttarakhand [6, 9, 10, 11]. Even though several thermozymes from *Geobacillus* spp. have been well-characterized, studies focusing on thermostable SODs from this genus remain limited, especially from Tapovan hot spring.

Recent advances in next-generation sequencing (NGS) technologies have enabled comprehensive genome level analysis of extremophiles. Whole-genome sequencing (WGS) provides a more complete picture of the genetic content of an organism than 16S rRNA gene sequencing or targeted amplicon sequencing [12]. WGS reveals SOD gene copy numbers and their arrangement in the genome, provides insights into sequence features associated with enzyme stability and enables identification of genes associated with oxidative stress management, all of which are difficult to obtain using biochemical assays alone. Integration of genomic and enzymatic approaches provides a thorough understanding of molecular mechanisms that facilitate their survival in extreme environments and aid in the production of thermostable enzymes.

Conventionally, large-scale enzyme production has heavily relied on recombinant DNA technology and expression systems [13]. Although effective, this approach requires expertise in molecular cloning, vector design, and host strain optimization, making it a time-consuming and resource-intensive process. In contrast, many microbes naturally secrete enzymes into their culture medium, offering an alternative approach for enzyme recovery and purification [14, 15] This approach provides several key advantages: it eliminates the cell lysis step, which may introduce contaminants (cell debris or intracellular proteins), circumvents the challenges of recombinant expression systems, and is more amenable to scale-up under native conditions. Despite these advantages, extracellular production and purification of thermostable antioxidant enzymes such as SOD remain comparatively underexplored.

Considering these aspects, this study focuses on the isolation of thermostable SOD from *G. thermoleovorans* GBRC-MXT211 (NCBI Accession No:GCF_059611665.1) isolated from the Tapovan hot spring, Uttarakhand, India. In our previous study, two *Geobacillus* spp. were reported and taxonomically characterized from Tapovan hot spring [16]. The present study focuses on *G. thermoleovorans* (GBRC-MXT211) for the purification and characterization of SOD from the cell-free supernatant (extracellular SOD activity). To complement the biochemical characterization, whole-genome sequencing was conducted to investigate the genetic basis of oxidative stress adaptation and SOD production under extreme environmental conditions. The study further describes a cloning-free approach in which SOD activity is recovered directly from the cell-free culture supernatant by ammonium sulphate fractionation, which may simplify future production under native conditions. Together with whole-genome sequencing, sequence and structural analyses, and a preliminary biochemical characterization, this work provides an initial genomic and biochemical description of a thermostable, extracellular SOD activity from *G. thermoleovorans* GBRC-MXT211 isolated from the Tapovan hot spring.

2. Materials and Methods

2.1 Sample collection, isolation and identification of thermophilic bacteria

Soil, water and substrate-enriched samples were collected from the Tapovan hot spring situated in the Chamoli district of Uttarakhand, India (30°29′15″N 79°39′25.2″E). A calibrated probe was used to measure the pH and temperature of the hot spring. Thereafter, the samples were aseptically transferred in sterile screw-capped tubes before transport to the laboratory. For isolation of thermophilic bacteria, samples were serially diluted in sterile phosphate-buffered

saline (PBS, pH 7.4) and aliquots (100 µL) of each dilution were spread plated on nutrient agar (NA) (HiMedia, India) and incubated at 70°C for 24 h. Gram staining was performed to observe the cell morphology, and the colonies were identified using MALDI-TOF (Bruker, USA). Briefly, a single isolated colony was transferred onto a steel target plate, overlaid with α-cyano-4-hydroxycinnamic acid (HCCA) matrix and allowed to air-dry. Spectra were acquired and analysed using the MALDI Biotyper® (MBT) system and software.

2.2 Whole Genome Sequencing and Bioinformatic Analysis

During the isolation and screening study, soil, water, and substrate-bag samples taken from the hot spring yielded a total of 214 thermophilic bacterial isolates. A previous study from our lab reported the details of the isolation, characterisation, and screening of these thermophilic isolates [16]. *Geobacillus thermoleovorans* GBRC-MXT211 was chosen from the isolates for additional whole-genome sequencing and subsequent bioinformatic analysis due to its promising biotechnological potential and its ability to produce thermostable superoxide dismutase (SOD).

2.2.1 DNA extraction and sequencing

Genomic DNA was extracted from GBRC-MXT211 using a modified CTAB-SDS method. DNA was isolated from bacterial cell pellets in which cells were treated with RNase after lysis and DNA was subsequently extracted using 10% SDS and CTAB extraction buffer. Later, DNA was purified using phenol:chloroform:isoamyl alcohol (25:24:1, v/v/v) and chloroform: isoamyl alcohol (24:1, v/v) and precipitated using isopropanol. The resulting DNA pellet was treated with 70% ethanol (twice), air-dried, and eluted in Tris-EDTA (TE) buffer. DNA concentration was measured using Qubit 4 fluorometer (Thermo Fisher Scientific, USA) and purity was assessed using BioTek Cytation 5 (Agilent Technologies). Libraries were prepared using the QIAseq FX DNA Library Kit and the Ligation Sequencing gDNA - Native Barcoding Kit and were sequenced on the Illumina NovaSeq 6000 platform using 2 × 251 bp paired-end chemistry and the Oxford Nanopore GridION platform, respectively.

2.2.2 Genome Assembly and Bioinformatic Analysis

FastQC v0.11.5 and MinKNOW v6.8.7 were used to evaluate the quality of the raw Illumina and Nanopore reads, respectively. fastp v1.3.4 [17], Porechop v0.2.4 [18], and Filtlong v0.3.1 [19] were used for adapter removal, quality filtering, and read trimming. Unicycler v0.5.1 [20] was used to generate a de novo hybrid genome assembly, which was then assessed using

QUAST v5.3.0 [21], BUSCO v6.1.0 [22], and CheckM2 v1.1.0 [23]. Average nucleotide identity (ANI) analysis was carried out against closely related reference genomes for species-level validation after the assembled genome was taxonomically classified using GTDB-Tk [24]. The NCBI Prokaryotic Genome Annotation Pipeline (PGAP) was used for functional annotation.

The genome of *Geobacillus thermoleovorans* GBRC-MXT211 has been submitted to the NCBI GenBank database with accession number GCF_059611665.1. The genome and its characterization are described in a separate manuscript that is under editorial consideration. The assembled genome of GBRC-MXT211 and the publicly accessible genomes of *G. thermoleovorans* from NCBI were used for whole-genome comparative phylogenetic analysis. Using *G. thermoleovorans* KCTC 3570 (GCF_001610955.1) as the reference genome, core-genome SNPs were identified using Snippy [25]. Gubbins was used to identify and remove recombinant regions [26], and IQ-TREE 2 [27] was used to construct a maximum-likelihood phylogenetic tree from the recombination-filtered core SNP alignment.

2.3 SOD gene identification and comparative analysis

The annotated genome of GBRC-MXT211 was screened for SOD-associated genes after genome assembly, quality evaluation, taxonomic confirmation, and functional annotation [28]. Three SOD-family proteins were identified in the genome: a canonical Fe/Mn-SOD-type protein, an N-terminally extended Fe/Mn-SOD-type protein, and a Cu/Zn-SOD. The physicochemical characteristics, conserved sequence features, and sequence similarity of the canonical Mn/Fe-SOD (SodA-type) protein with representative SOD proteins from various bacterial and eukaryotic organisms were further examined. SignalP 6.0 and PSORTb were used for signal peptide and subcellular localization prediction, respectively.

2.3.1 Sequence dataset

A total of 19 SOD protein sequences representing Fe/Mn-SOD family, all from bacterial and eukaryotic sources were compiled in FASTA format (Table 1). The dataset included sequences from thermophilic and mesophilic *Geobacillus*, *Bacillus*, *Anoxybacillus*, *Aeribacillus*, and *Thermus* species, together with the human mitochondrial manganese SOD (NP_000627.2) included as a reference eukaryotic sequence.

Table 1: Accession numbers, source organisms, and full-length sequence lengths of the nineteen SOD sequences used in this study

Accession	Organism	Length (aa, full-length)
WP_011231937.1	*Geobacillus thermoleovorans (GBRC MXT211)*	204
OQP13510.1	*Geobacillus thermoleovorans*	437
WP_434407099.1	*Geobacillus thermoleovorans*	418
WP_069304620.1	*Geobacillus sp.*	437
UBT21395.1	*Geobacillus sp.*	204
WP_429908981.1	*Geobacillus thermocatenulatus*	437
WP_044736957.1	*Geobacillus kaustophilus*	437
KYD16580.1	*Anoxybacillus flavithermus*	303
BAW01276.1	*Thermus thermophilus*	204
WP_004398583.1	*Bacillus sp.*	202
NP_000627.2	*Homo sapiens (Mn-SOD, mitochondrial)*	222
WP_193442426.1	*Geobacillus stearothermophilus*	418
ANZ31612.1	*Parageobacillus thermoglucosidasius*	204
WP_429650695.1	*Aeribacillus pallidus*	292
WP_429771673.1	*Bacillus smithii*	296
SPY11823.1	*Bacillus subtilis*	281
WP_150825824.1	*Flavobacterium thermophilum* (NCTC 11526)	204
BAA02655.1	*Thermus aquaticus*	204
WP_066230298.1	*Geobacillus jurassicus*	437

***** *Genus assignment is provisional in the NCBI database (bracketed genus); the protein record is placed in the Firmicutes*

2.3.2 Multiple Sequence Alignment

Multiple sequence alignment (MSA) of the 19 SOD sequences was performed using MAFFT v7.505 [29] with the auto option, which selected the L-INS-i strategy (iterative refinement with local pairwise alignment information) using the BLOSUM62 scoring matrix, as appropriate for this dataset size (<200 sequences). The resulting alignment was 446 columns in length.

Because the dataset combined canonical short-form Mn/Fe-SOD sequences (~200–225 aa) with an N-terminally extended Mn-SOD variant found in several *Geobacillus* sequences (~435–436 aa), column-wise sequence occupancy was calculated across the full alignment to

identify the shared conserved catalytic domain; a contiguous block of 211 columns (positions 236–446 of the full alignment) in which all 19 sequences were represented was extracted as the region of interest for downstream conservation analysis, visualization, and phylogenetic reconstruction, thereby avoiding the bias that the poorly aligned, gap-rich N-terminal extensions would otherwise introduce into branch-length estimation. The trimmed alignment was visualized in an ESPript-style format, with residues shaded by conservation level (black, dark gray, and light gray denoting fully, strongly, and weakly conserved columns, respectively) and annotated with standard Clustal consensus symbols. Columns bearing an invariant histidine or aspartate residue across all sequences were flagged as candidate metal-coordinating active-site positions based on sequence conservation alone, and their positions were then examined in the predicted structure.

2.3.3 Phylogenetic tree construction

Phylogenetic relationships among the 19 SOD sequences were inferred from the trimmed 211-column alignment using IQ-TREE v2.0.7 [27]. The best-fit substitution model was selected using ModelFinder [30] under the Bayesian Information Criterion (BIC), which identified the WAG+G4 model (Whelan and Goldman substitution matrix with a discrete gamma-distributed rate heterogeneity, 4 categories; gamma shape parameter $\alpha = 0.92$) as optimal. Maximum-likelihood tree search was performed with default IQ-TREE heuristics (nearest-neighbor interchange and stochastic perturbation), and clade support was assessed using the ultrafast bootstrap approximation (UFBoot) with 1000 replicates [31]. The final consensus tree was midpoint-rooted, ladderized, and visualized using Biopython and matplotlib.

2.3.4 Structure modelling and active-site mapping

The three-dimensional (3D) structure of the SOD protein from the Tapovan isolate (GBRC-MXT211) was predicted using AlphaFold2 [32] as implemented in ColabFold [33], which generates multiple sequence alignments via MMseqs2 for input to the AlphaFold2 network. The top-ranked model by predicted local distance difference test (pLDDT) score (mean pLDDT = 97.98; 98.5% of residues scoring >90, indicating very high per-residue confidence) was selected for downstream analysis. Of the eight invariant His/Asp positions identified in the alignment, four (His-27, His-82, Asp-164, His-168) corresponded to the canonical Fe/Mn-SOD metal ligands and were mapped onto the predicted structure and visualised using UCSF Chimera [34] to assess their spatial arrangement.

2.4 SOD purification and characterization

2.4.1 Qualitative SOD Activity Screening by NBT Plate Assay

For preliminary screening of extracellular superoxide dismutase (SOD) producing bacteria, a plate assay was performed based on nitroblue tetrazolium (NBT) reduction. A single bacterial colony was streaked onto the NA plate and incubated at 70°C for 24 h. After incubation, plates were gently flooded with a solution containing 0.3 mM NBT, 28 µM riboflavin, and 28 mM TEMED and exposed to uniform UV/visible light for 20 minutes. SOD positive activity was identified by the appearance of colourless or light-coloured halo zones against a dark purple stained background. Isolates exhibiting prominent halo zones were further shortlisted for enzyme characterisation [35]. To validate the thermostable property of SOD producing bacteria, different growth temperatures were utilized starting from 37–80°C and simultaneously, their morphological characteristics were observed.

2.4.2 Preparation of Crude Enzyme Extract using Ammonium sulphate precipitation

The selected isolate was inoculated with 1% inoculum into 2000 mL Erlenmeyer flasks containing 1000 mL of nutrient broth and incubated at 70°C with continuous shaking at 150 rpm for 24 h. Next day, the culture was centrifuged at 10,000 × g for 15 minutes at 4°C to remove bacterial cells and debris. The remaining supernatant was filtered using a 0.22 µm filter and stored as crude enzyme extract. This was labelled as cell-free supernatant (CFS) which contains the extracellular protein fraction.

The collected CFS was precipitated using ammonium sulphate ($(NH_4)_2SO_4$) for partial purification. Precipitation was performed stepwise increasing the saturation levels to 20, 40, 60 and 80%. At each step, ammonium sulphate ($(NH_4)_2SO_4$) was added gradually with gentle continuous stirring at 4°C for 12 to 16 h at each step. Fractions were collected after each precipitation level (20, 40, 60 and 80%) by centrifuging the solution at 10,000 × g for 10 minutes at 4°C. The collected pellet from each fraction was carefully separated from the supernatant (which was carried forward for the next precipitation step). The protein pellets of each fractions were resuspended in a minimal volume of 50 mM potassium phosphate buffer (pH 7.4) and allowed to undergo a salt exchange using Thermo Scientific Snakeskin dialysis tubing (10K MWCO) for removal of excess residual salt and debris. The dialysed 80% fraction was used directly for biochemical characterisation. For protein identification, a portion of the dialysed fraction was additionally concentrated using a 10 kDa MWCO centrifugal filter and

precipitated with acetone (sample-to-acetone ratio 1:4); the resulting pellet was resuspended in 100 µL potassium phosphate buffer.

2.4.3 Protein Quantification and Native Polyacrylamide Gel Electrophoresis

Total protein concentration at each step was determined by the Bradford assay using bovine serum albumin (BSA) as a standard. Absorbance was measured at 595 nm, using BioTek Cytation 5.

Subsequently, PAGE was performed for in-gel SOD activity detection. 12% Native polyacrylamide gel was prepared and initially run to remove any excess unbound ammonium persulphate (APS) which may otherwise hinder in SOD activity detection. A 1X non-reducing loading dye (without SDS) was used to track the progress of the pre-run. The gel buffer was replaced with a fresh one after the pre-run was over. Samples were loaded and the gel was then re-run at 100 V. Protein bands were visualised by Coomassie Brilliant Blue (CBB) R-250 stain for molecular weight determination [validated using a Page Ruler Prestained Protein Ladder (Thermo Scientific) (10–250 kDa)] and simultaneously using NBT reduction method for SOD activity detection. Briefly, the gel was incubated for 30 minutes in a freshly prepared 50 mM potassium phosphate buffer and rinsed thoroughly. It was then treated with 0.48 mM nitroblue tetrazolium (NBT) prepared in a potassium phosphate buffer for another 30 minutes. Following a gentle rinse to remove excess NBT, 30 µM riboflavin (in phosphate buffer) was added and incubated for 30 minutes. The gel was exposed to UV/Vis light until colourless bands of SOD activity appeared against a uniformly dark purple background [35].

2.4.4 Protein identification using MALDI-TOF

The SOD activity corresponding protein bands were excised from activity-stained native PAGE gel using a sterilized razor blade (1 mm × 1 mm pieces) and transferred to a sterile microcentrifuge tube. Gel slices were destained by 100 mM ammonium bicarbonate/50% acetonitrile (ACN) solution at 37°C for 30 min incubation. This step was repeated until all the staining reagent was removed. Next, the gel pieces were incubated with 5 mM Tris(2-carboxyethyl) phosphine (TCEP) at 60°C for 10 min followed by incubation (in shaking condition) for 15 min at 37°C in freshly prepared 100 mM iodoacetamide (IAA) in the dark. Excess IAA was removed by rinsing the gel slices with 100 mM ammonium bicarbonate/50% ACN for seven to eight times. Gel pieces were dehydrated by adding 50 µL of ACN and incubated for 15 min at room temperature, followed by brief air drying for 5–10 min. For trypsin digestion, lyophilized Thermo Scientific™ Pierce™ Trypsin Protease was reconstituted

in 50 mM acetic acid to a final concentration of 0.01 mg/mL in 100 mM ammonium bicarbonate (pH 8.0). 50 µL of the diluted trypsin solution was added to the dried gel pieces and incubated at 37°C for 8 to 10 h. Next day, the digested solution was vortex mixed for 10-15 min for complete mixing. Peptide mixture was resuspended in 0.1% TFA, spotted onto a MALDI steel target plate with matrix, and analysed by MALDI-TOF mass spectrometry (Bruker). Protein identification was performed using Mascot (Matrix Science) using Swiss-Prot database.

2.4.5 Size exclusion chromatography (SEC)

For SEC, 80% ammonium sulphate fraction was dialysed against potassium phosphate buffer (50 mM, pH 7.4). This sample was further dialysed using 150 mM sodium phosphate buffer and filtered using 0.22 µm filter. It was then loaded onto an AdvanceBio SEC 300Å protein standard column (Agilent Technologies) (4.6×300 mm) pre-equilibrated with the same buffer and eluted at 0.5 mL/min. Fractions (0.5 mL) were collected and screened for SOD activity by NBT reduction assay and protein content by Bradford assay.

2.5 Biochemical characterisation of partially purified SOD fraction

SOD activity assays and enzyme activity was calculated using the dialysed fraction that retained sufficient enzyme activity and protein concentration for reliable biochemical characterization.

2.5.1 SOD activity using standard curve

Quantitative enzyme activity assay for SOD was performed using the Invitrogen™ Superoxide Dismutase (SOD) Colorimetric Activity Assay Kit. Following the manufacturer's instructions, SOD standard (provided in the kit) was reconstituted in 250 µL of assay buffer to a concentration of 4 U/mL, and eight serial two-fold dilutions (4, 2, 1, 0.5, 0.25, 0.125, 0.0625, and 0 U/mL) were prepared in the same buffer for standard curve. The enzyme sample was appropriately diluted in an assay buffer to fall within this linear range. Working reagents were freshly prepared and added sequentially to each well of a clear 96-well plate. After incubation at room temperature, absorbance was measured at 450 nm using BioTek Cytation 5. SOD activity was calculated from the standard curve.

2.5.2 Effect of pH on SOD activity

The effect of pH on SOD activity was determined by assaying the enzyme across a pH range of 2.0 to 10.0. The following buffer systems were used at a concentration of 50 mM. Glycine-

HCl buffer was used for pH 2.0, sodium acetate buffer for pH 4.0 and 5.0, sodium phosphate buffer for pH 6.0 and 7.0, Tris-HCl buffer for pH 8.0, and glycine buffer for pH 10.0. The enzyme assay was conducted using SOD activity kit (Invitrogen) at each respective pH and expressed as percentage inhibition of WST reduction.

2.5.3 pH stability of SOD

pH stability assay was performed by pre-incubating the dialysed fraction in prepared buffers ranging from pH 2 to 10 at 37°C for 60 min in the absence of substrate. Residual SOD activity was measured under standard assay conditions and, as for the pH activity assay (Section 2.5.2), is expressed as percentage inhibition, where a higher value indicates higher SOD activity. The untreated control (enzyme in standard assay buffer) was assayed in parallel for comparison.

2.5.4 Effect of Temperature on SOD activity

The effect of temperature on SOD activity was determined by conducting the enzyme assay at temperatures ranging from 30 to 95°C (30, 40, 50, 60, 70, 80 and 95°C). The reaction was initiated at each respective temperature, and SOD activity was measured spectrophotometrically and expressed as percentage inhibition of WST reduction.

2.5.5 Thermostability of SOD

Temperature stability was assessed by pre-incubating the enzyme sample at 30, 40, 50, 60, 70, 80, and 95°C for 60 min in the absence of substrate. After incubation, samples were cooled to room temperature and residual SOD activity was measured under standard assay conditions, expressed as percentage inhibition (Section 2.5.4). An unincubated sample maintained at 4°C was assayed in parallel as the control.

2.5.6 Effect of Inhibitors on SOD activity

The influence of inhibitors on SOD activity was investigated by pre-incubating the enzyme sample with 5 mM hydrogen peroxide (H_2O_2) and 5 mM sodium azide (NaN_3) at 37°C for 30 min. Residual SOD activity was measured under standard conditions using SOD activity kit (Invitrogen) and expressed as a percentage activity relative to control (enzyme incubated without any inhibitor).

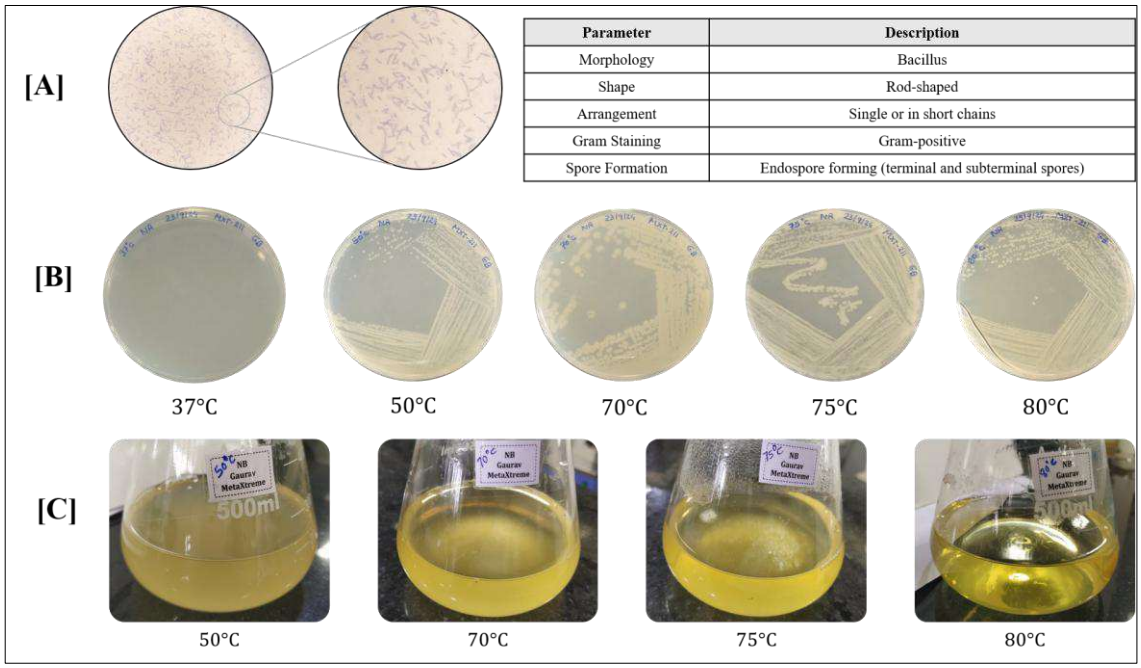
2.6 Statistical analysis

All the experiments were performed in triplicate, and results are expressed as mean ± standard deviation (SD). Statistical analysis was performed using GraphPad Prism v8.0.

Results

3.1 Sample collection, Isolation and initial identification of thermophilic bacteria

Water, soil and substrate (bags) samples were collected from Tapovan hot spring, Uttarakhand as discussed earlier [16]. The pH recorded at the site was 7.4 and temperature up to 95°C. Bacterial colonies were differentiated based on their colony morphologies and the organisms that grew consistently on subculture were identified using MALDI-TOF. Among the notable organisms identified was *Geobacillus thermoleovorans* capable of growing up to 80°C. Figure 1 illustrates the microscopic and colony morphology of the organism. The Gram staining image shows purple rods (arranged singly or in chains) suggesting Gram-positive nature. In addition, endospores can also be observed towards the terminal end of the rods which is a characteristic feature of *Bacillus* spp. The table in the figure 1 depicts the morphological characteristics of the organism. A variety of growth temperatures (37, 50, 60, 70, 75 and 80°C) were utilized to understand the thermotolerance property of the organism. It was validated on nutrient agar plates where growth was observed up to 80°C. Similarly, broth cultures revealed turbidity and yellowish pigmentation, which are indicative of growth.



Parameter	Description
Morphology	Bacillus
Shape	Rod-shaped
Arrangement	Single or in short chains
Gram Staining	Gram-positive
Spore Formation	Endospore forming (terminal and subterminal spores)

Figure 1. Morphological and growth characteristics of *Geobacillus thermoleovorans* MXT211. (A) Gram staining showing purple, rod-shaped cells and the accompanying table summarises the morphological features. (B) Observation of colony growth on nutrient agar after incubation at 37, 50, 70, 75 and 80 °C. (C) Observation of growth in nutrient broth at 50, 70, 75 and 80 °C.

3.2 Genome annotation and Bioinformatics analysis

Genomic DNA was extracted and quantified using Qubit 4 fluorometer (Thermo Fisher Scientific, USA) yielding a concentration of 37.6 ng/µL. In addition, DNA purity was assessed by measuring the A260:A280 ratio using BioTek Cytation 5 (Agilent Technologies), which yielded a value of 1.84. The extracted DNA was then sequenced using the Illumina NovaSeq platform (2 × 251 bp paired end) and the Oxford Nanopore GridION platform. After quality filtering and adapter trimming, 842,686 reads passed the Illumina quality check (416.34 Mbp) and 155,530 reads passed the Nanopore quality check (256.6 Mbp). *De novo* hybrid assembly generated using Unicycler provided two circular contigs of 3,426,319 bp and 56,688 bp, giving a total genome size of 3.48 Mb with a GC content of 52.43% and an N50 of 3.43 Mb. Assembly completeness assessed using BUSCO and CheckM2, was 99.1% and 99.99%, respectively, indicating a near-complete genome assembly. Finally, the assembled genome sequence was submitted to the NCBI database with accession number GCF_059611665.1.

Taxonomic identification based on the assembled genome was performed using GTDB-Tk with the GTDB reference database, which placed the isolate (GBRC-MXT211) within *Geobacillus thermoleovorans*, showing the highest average nucleotide identity (ANI) of 99.01% to the type strain KCTC 3570 (RefSeq accession GCF_001610955.1). A whole-genome-based phylogenetic analysis was carried out using 19 publicly available *Geobacillus thermoleovorans* genomes together with strain GBRC-MXT211. Core-genome single nucleotide polymorphisms (SNPs) were identified using Snippy, recombinant regions were removed using Gubbins, and a maximum-likelihood tree was constructed from the recombination-filtered core-genome SNP alignment using IQ-TREE 2. The resulting phylogeny placed GBRC-MXT211 (unpublished data) close to the reference strain KCTC 3570 while forming a distinct branch within the species, consistent with the ANI-based identification.

Genome annotation using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) identified a total of 3,555 genes, comprising 3,435 coding sequences (CDS), 3,359 predicted protein-coding genes, 120 RNA genes (88 tRNAs and 27 rRNAs), 76 pseudogenes, 5 ncRNAs, and 3 CRISPR arrays. The SOD coding gene investigated in this study is among these predicted protein-coding genes, which provided the genomic context for the sequence and structural analyses (unpublished data).

3.3 SOD sequence identification and analysis

The annotated GBRC-MXT211 genome was screened for SOD-associated genes, identifying three predicted SOD-family proteins, comprising a canonical Fe/Mn-SOD-type protein, a Cu/Zn-SOD, and an N-terminally extended Fe/Mn-SOD-type protein. The predicted Mn/Fe-SOD (SodA-type) protein was selected as the primary candidate for further sequence and comparative analysis because its predicted molecular mass (~22.9 kDa) is consistent with the monomer size of bacterial Mn-SODs, which was selected as the primary candidate because its predicted molecular mass (~22.9 kDa) is consistent with the monomer size of bacterial Mn-SODs and because the SOD-active band in the cell-free supernatant was tentatively identified as Mn-SOD by MALDI-TOF MS/MS.

SignalP 6.0 analysis predicted no classical signal peptide in the Mn-SOD or Exosporium SOD sequences (Figure 2). In contrast, the Cu/Zn-SOD was predicted to contain a Sec-dependent lipoprotein signal peptide (Sec/SPII), with a predicted cleavage site between residues 17 and 18 (probability = 0.9976), suggesting potential export and membrane-associated localization. Subcellular localization analysis using PSORTb predicted the Mn-SOD and Exosporium SOD as extracellular proteins, with localization scores of 9.72 and 9.60, respectively, despite the absence of detectable classical signal peptides by SignalP 6.0. In contrast, the Cu/Zn-SOD was predicted to be cytoplasmic-membrane associated (score 9.81), consistent with the Sec/SPII lipoprotein signal peptide predicted by SignalP (Supplementary Table 1).

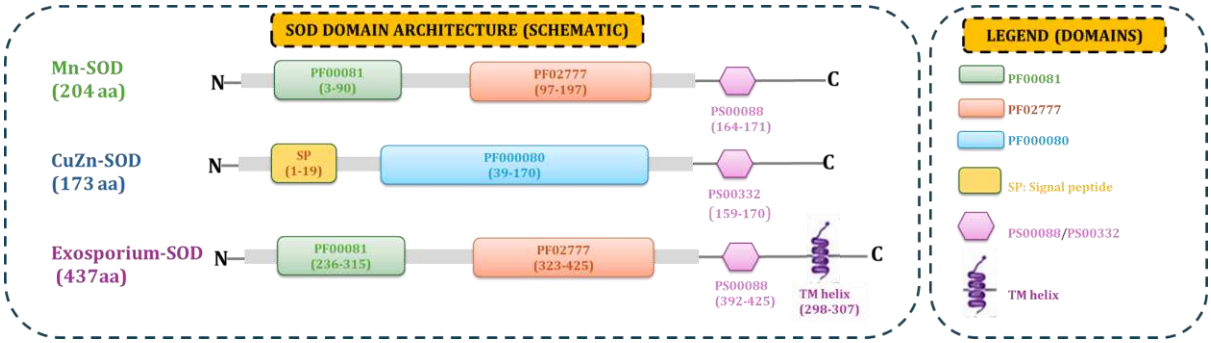


Figure 2. Schematic representation of the domain architecture of the three MXT211 SOD proteins based on DeepTMHMM prediction and InterProScan/domain analysis. Mn-SOD and Exosporium-SOD contain the characteristic PF00081 and PF02777 Mn/Fe-SOD domains, whereas Cu/Zn-SOD contains the PF00080 Cu/Zn-SOD domain. The conserved PS00088/PS00332 motifs, N-terminal signal peptide, and predicted transmembrane helix are also indicated.

Domain analysis was performed using DeepTMHMM prediction and InterProScan/domain analysis, revealing distinct but related architectures among the three SOD proteins. Mn-SOD

(204 aa) showed the characteristic Mn/Fe-SOD domains PF00081 (3–90) and PF02777 (97–197), along with the conserved PS00088 motif (164–171), supporting its classification as a Mn/Fe-SOD. Cu/Zn-SOD (173 aa) displayed a distinct Cu/Zn-SOD architecture, with PF00080 (39–170) and the conserved PS00332 motif (159–170), together with an N-terminal signal peptide (1–17 aa). Exosporium-SOD (437 aa) also contained the Mn/Fe-SOD domains PF00081 (236–315) and PF02777 (323–425) and the PS00088 motif (392–399), along with an extended protein region and a predicted transmembrane segment. Overall, the domain profiles support Mn-SOD and Exosporium-SOD as Mn/Fe-SOD-related proteins, while Cu/Zn-SOD represents a distinct Cu/Zn-SOD type. This approximate correspondence between the predicted protein size and the observed SOD activity supported the selection of Mn-SOD for subsequent sequence and structural analyses although the native gel does not allow molecular mass to be determined directly.

3.3.1 Multiple sequence alignment reveals a conserved catalytic core

MAFFT alignment of the 19 SOD sequences produced a 446-column alignment, of which 56 columns (12.6%) were fully conserved across all sequences. The alignment length substantially exceeded the length of any individual sequence (202–437 aa) due to the presence of long, non-homologous N-terminal extensions unique to the Fe/Mn-SOD-type sequences from *Geobacillus* and related genera, which introduced extensive gap regions when aligned against the shorter, canonical short-form Mn/Fe-SOD sequences. Column-occupancy analysis identified a 211-residue block (alignment positions 236–446) that was populated in all 19 sequences, corresponding to the conserved SOD catalytic domain. Within this trimmed region, 55 columns (26.1%) were fully conserved, an additional 33 columns (15.6%) were strongly similar, and 20 columns (9.5%) were weakly similar, together indicating that just over half (51.2%) of the catalytic-domain positions are conserved or semi-conserved across this taxonomically diverse sequence set (Supplementary figure 1). Pairwise identity within the trimmed region ranged from 36.6% to 100%, with the highest identities observed among closely related *Geobacillus* sequences (e.g., OQP13510.1 and WP_069304620.1, 100% identity) and the lowest identities involving the more divergent *Bacillus subtilis* (SPY11823.1) and *Aeribacillus pallidus* (WP_429650695.1) sequences. Eight alignment columns were invariant histidine or aspartate residues across all 19 sequences (His-27, His-28, His-31, His-32, His-79, His-82, Asp-164 and His-168 in the GBRC-MXT211 numbering) and are marked in the alignment figure. Four of these (His-27, His-82, Asp-164 and His-168) correspond to the positions expected to form the 3-His/1-Asp metal-binding site of Fe/Mn-SODs; Asp-164 and

His-168 fall within the PS00088 signature motif (DVWEHAYY). The other four conserved histidines (His-28, His-31, His-32 and His-79) are in the vicinity of the active site but are not expected to be direct metal ligands.

3.3.2 Phylogenetic analysis

Maximum-likelihood phylogenetic analysis under the WAG+G4 model resolved the nineteen SOD sequences into several well-supported clades (Figure 3). A large clade comprising all *Geobacillus* Fe/Mn-SOD-type sequences (OQP13510.1, WP_069304620.1, WP_434407099.1, WP_429908981.1, WP_044736957.1, WP_066230298.1, WP_193442426.1) together with *Anoxybacillus flavithermus* (KYD16580.1) was recovered with moderate-to-strong support (bootstrap = 52–97%), reflecting the close taxonomic relationship among these thermophilic *Bacillaceae*. Within this clade, *G. kaustophilus* (WP_044736957.1) and *G. jurassicus* (WP_066230298.1) formed a tight, well-supported subclade (bootstrap = 83%), consistent with their high pairwise sequence identity (99.0%). A second clade grouped *Aeribacillus pallidus* (WP_429650695.1), *Bacillus smithii* (WP_429771673.1), and *Bacillus subtilis* (SPY11823.1) with weak-to-moderate support (bootstrap = 44–57%), reflecting their greater sequence divergence from the *Geobacillus* clade. The canonical short-form Mn-SOD sequences did not form a single monophyletic group with the N-terminally extended *Geobacillus* Mn-SOD sequences. GBRC-MXT211, UBT21395.1, and WP_150825824.1 formed a well-supported, near-identical cluster (bootstrap = 73–75%; pairwise identity 99.0%), which grouped with WP_004398583.1 (*Bacillus* sp.) and ANZ31612.1 (*Parageobacillus thermoglucosidasius*). WP_150825824.1 is annotated as [*Flavobacterium*] *thermophilum* NCTC 11526; the bracketed genus denotes a provisional assignment, and NCBI places the record in the Firmicutes, so its close relationship to the *Geobacillus*-type sequences is not unexpected. The two *Thermus* sequences (BAW01276.1, *T. thermophilus*; BAA02655.1, *T. aquaticus*) formed a distinct, strongly supported clade (bootstrap = 100%), consistent with their shared genus. The human mitochondrial Mn-SOD sequence (NP_000627.2) occupied a relatively divergent position from the bacterial sequences. Overall, the tree topology broadly recapitulated known taxonomic relationships among the *Bacillaceae*, with strongly supported terminal clades (bootstrap ≥ 70%) corresponding to congeneric or closely related species pairs, and comparatively weaker support (bootstrap 44–57%) at several deeper internal nodes reflecting the limited phylogenetic signal available from a single, relatively short protein-coding locus.

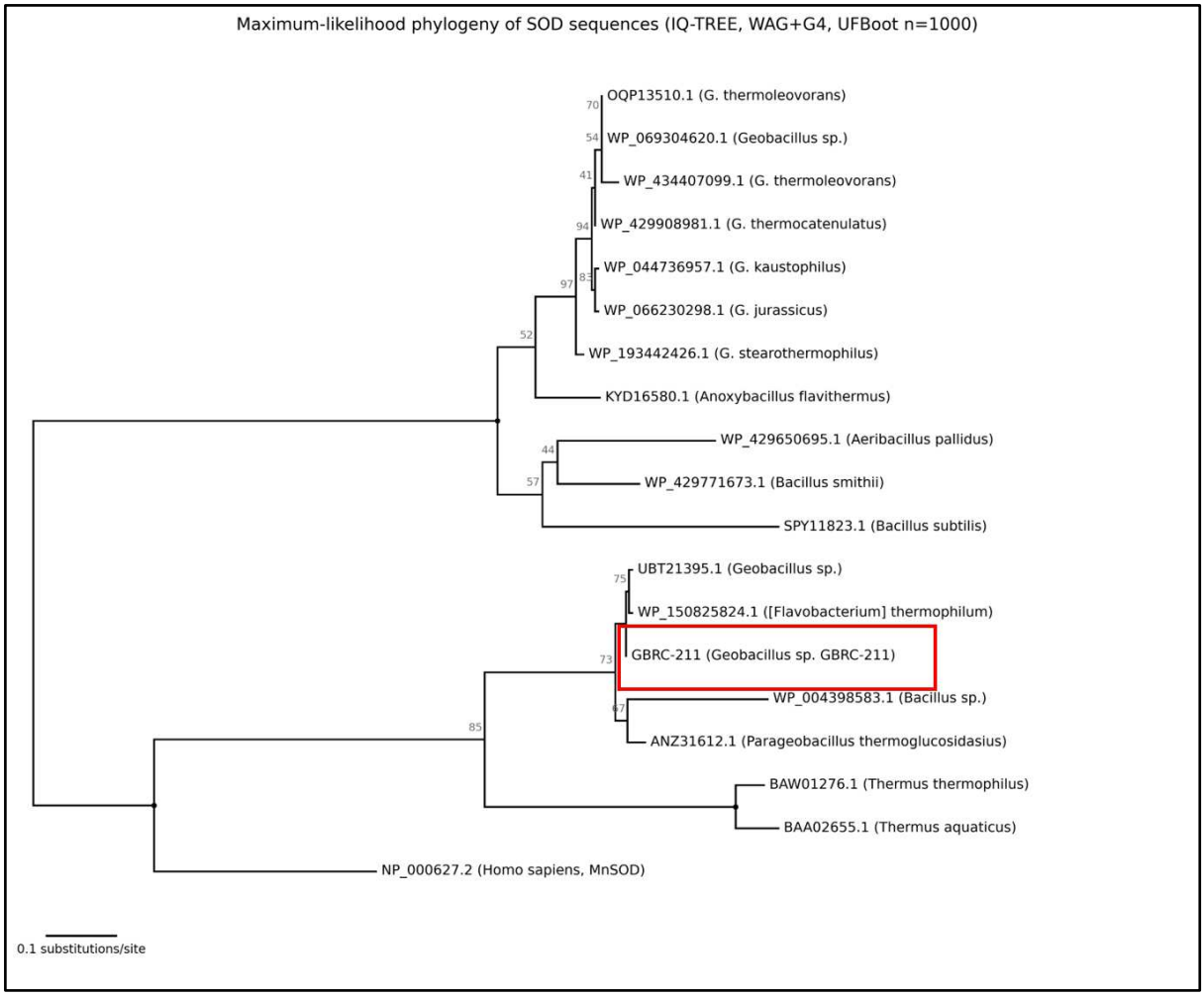


Figure 3: Maximum-likelihood phylogenetic tree of 19 SOD sequences inferred from the trimmed catalytic-domain alignment (IQ-TREE, WAG+G4 model, 1000 ultrafast bootstrap replicates). The tree is midpoint-rooted; numbers at internal nodes indicate bootstrap support (%); nodes with 100% support are marked with a filled circle. Scale bar indicates substitutions per site.

3.3.3 Three-dimensional structure prediction

The AlphaFold2 structural model of the GBRC-MXT211 SOD protein (mean pLDDT = 97.98) predicted a compact, single-chain, two-domain fold consistent with the canonical Fe/Mn-SOD architecture, comprising an N-terminal α-helical domain and a C-terminal mixed α/β domain (Figure 4). The four residues identified from the alignment as invariant histidine/aspartate positions: His-27, His-82, Asp-164 and His-168 converge spatially in the modelled structure to form a single pocket at the interface of the two domains (Figure 4, inset), consistent with the expected 3-His/1-Asp metal-coordination geometry of the Fe/Mn-SOD active site. The high pLDDT scores indicate high confidence in the predicted local structure surrounding the putative active-site residues. The predicted structure is consistent with the canonical Fe/Mn-

SOD fold and places the conserved metal-coordinating residues within a common active-site region.

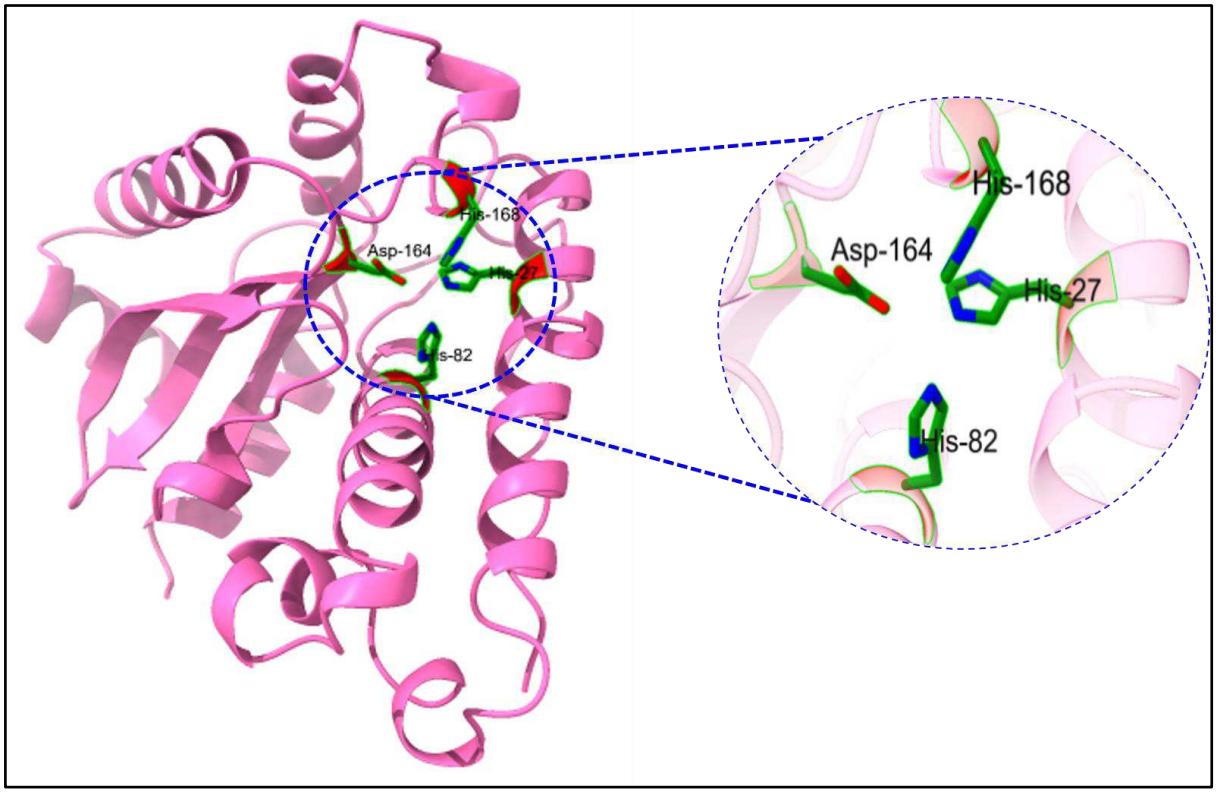


Figure 4. Structure prediction of the thermostable superoxide dismutase (SOD) from *Geobacillus thermoleovorans* MXT211. AlphaFold2 (ColabFold) model of the full-length protein, shown as a cartoon representation. The dashed oval marks the predicted active-site region along with Close-up view of the active-site pocket, with the side chains of His-27, His-82, His-168 and Asp-164 shown as green sticks. These four residues, which were found to be invariant across the multiple sequence alignment.

3.4 SOD purification and characterization

3.4.1 SOD Activity Screening by NBT Plate Assay

Thermophilic bacterial isolates were qualitatively screened for superoxide dismutase activity using the NBT plate assay. In this assay, superoxide radicals generated *in situ* reduce nitroblue tetrazolium (NBT) to blue-purple formazan, whereas SOD-mediated removal of superoxide results in reduced NBT reduction and the formation of a clear zone around SOD-producing colonies. The colonies of the selected isolate GBRC-MXT211 consistently exhibited distinct clear halos surrounding their growth against the purple background, indicating SOD activity. The formation of clear halos was observed across the colonies in the assay and was reproducible in triplicate petri-dishes, supporting the consistent SOD-positive phenotype of GBRC-MXT211 (representative image shown here). Based on this activity, GBRC-MXT211

was selected for subsequent genomic and biochemical characterization (Supplementary figure 2).

3.4.2 Ammonium sulphate precipitation

The selected isolate was cultivated in nutrient broth for 24 h at 70°C, after which the culture was centrifuged at 10,000 × g for 15 min at 4°C. The resulting cell-free supernatant was collected for partial purification of extracellular proteins by sequential ammonium sulphate precipitation at 20, 40, 60, and 80% saturation. Protein concentration was determined for each precipitated fraction using the Bradford assay (Supplementary table 2). No visible precipitate was obtained at 40% ammonium sulphate saturation, whereas the 60% fraction showed the highest protein concentration among the recovered fractions. In contrast, SOD activity staining indicated that the 80% ammonium sulphate fraction exhibited the most prominent SOD-active band (Figure 5). Thus, although the highest protein concentration was observed in the 60% fraction, the 80% fraction showed the greatest relative SOD activity. Based on the activity profile, the 80% ammonium sulphate fraction was selected for subsequent purification. The selected fraction was dialysed against 50 mM potassium phosphate buffer (pH 7.4) using a 10-kDa molecular-weight-cutoff membrane and subsequently concentrated using a 10-kDa MWCO membrane. A portion was concentrated and acetone-precipitated for protein identification (Section 3.4.4).

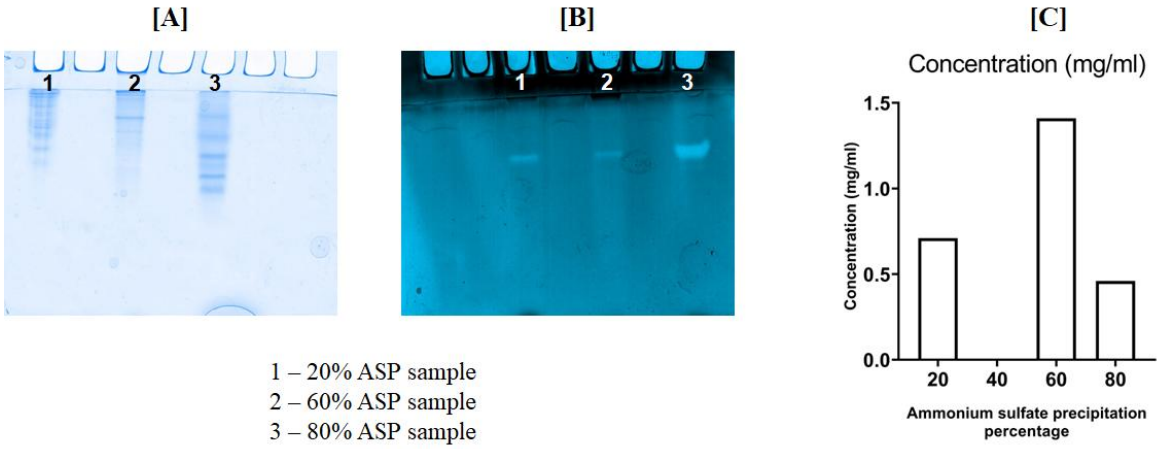
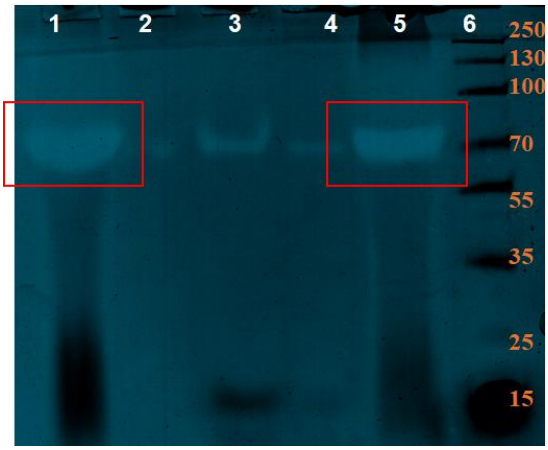


Figure 5. Ammonium sulfate precipitation (ASP) fractionation of superoxide dismutase (SOD) from *Geobacillus thermoleovorans* MXT211. (A) Native PAGE of the protein fractions precipitated at 20% (lane 1), 60% (lane 2) and 80% (lane 3), stained with Coomassie Brilliant Blue. (B) In-gel SOD activity staining of the same fractions were clear bands against dark blue background shows SOD activity. (C) Protein concentration (mg/mL) of each fraction.

3.4.3 SOD activity staining and protein profile after initial purification

The protein concentration after dialysis was calculated using Bradford reagent and the absorbance was measured at 595 nm. The concentration of different fractions is provided in figure 6. The dialysed fractions were loaded onto Native PAGE for activity staining. The SOD activity was calculated using standard curve from the commercial SOD. The dialysed fraction (partially purified) showed an activity of 0.24 U/mL and specific activity of 10.2 U/mg.

[A]



[B]

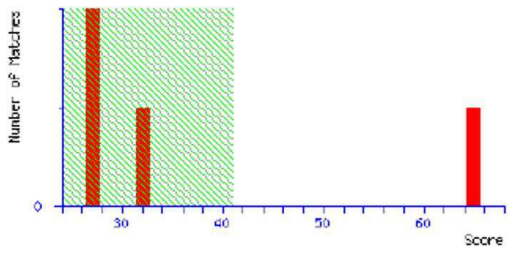
Lane No	Sample position	Protein conc. (mg/mL)
1	80% dialysed, concentrated and acetone precipitated	0.981
2	Empty	-
3	80% dialysed	0.242
4	80% dialysed and flow through	0.002
5	80% dialysed and concentrated	0.853
6	Ladder	-

Figure 6. Native PAGE staining for SOD activity screening (A) In-gel SOD activity staining of different fractions after 80% ammonium sulphate precipitation. (B) Inset table shows the lane number and their respective fractions along with their protein concentration. Red highlighted boxes displayed the highest in-gel SOD activity

3.4.4 Protein identification using MALDI-TOF MS/MS

The dialysed fraction and the dialysed, concentrated and acetone-precipitated fraction, both derived from the same 80% ammonium sulphate fraction, were resolved on native PAGE and activity-stained. The SOD-active band from each preparation (highlighted in red boxes in Fig. 6) was excised for in-gel trypsin digestion. Both the bands migrated at the same position, indicating that the SOD activity was retained after dialysis, concentration and acetone precipitation. The peptide fragments were analysed by mass spectrometry and searched against the Mascot database. Of the 17 queries submitted to the Mascot database (obtained from dialysed, concentrated and acetone-precipitated sample, lane 1) one peptide match exceeded the identity threshold, and one additional match exceeded the homology threshold. The highest ion score was 62, exceeding the Mascot identity threshold of 41 at $p < 0.05$ (Figure 7). The band from lane 5 (dialysed fraction) was also digested and out of 9 queries submitted, one

peptide match exceeded the identity threshold of 41 with the ion score of 51 (Supplementary fig 3). Both the bands were supported as Mn-SOD by Mascot-based peptide matching and organism identified as *Bacillus caldotenax*; however, the limited number of confidently matched peptides means that this identification should be considered supportive rather than definitive.



Peptide score distribution. Ions score is $-10\log(P)$, where P is the probability that the observed match is a random event.
There is **1** peptide match above identity threshold and **1** match above homology threshold for **17** queries. On average, individual ions scores **> 41** (beyond green shading) indicate **identity or extensive homology** ($p<0.05$).

Type	Score	Mass	Peptide sequence
SODM_BACCA Superoxide dismutase [Mn] OS=Bacillus caldotenax	65	22875	K.SLEELLSNLEALPESIR.T

Figure 7. Identification of the purified SOD from *Geobacillus thermoleovorans* MXT211 by peptide mass fingerprinting. (Above) Mascot score distribution for 17 queries, where scores above 41 (outside the green shaded area) indicate a significant match ($p < 0.05$). (Tbale below) The top hit was manganese-dependent superoxide dismutase from *Bacillus caldotenax* (SODM_BACCA), with a score of 65 and a matched peptide of K.SLEELLSNLEALPESIR.T.

3.4.5 Size exclusion chromatography of dialysed sample

The dialysed protein fraction was further analysed by size-exclusion chromatography using a PDA detector at 220 nm (Figure 8). Two major chromatographic peaks were observed, with retention times of approximately 4.4 and 5.2 min. The earlier peak at 4.4 min represented the predominant higher-apparent-molecular-size species, whereas the later peak at 5.2 min represented a lower-apparent-molecular-size species. Based on the expected oligomeric organization of bacterial Mn-SODs, these peaks may correspond to different oligomeric states of the enzyme, with the earlier-eluting species being consistent with a higher-order oligomer and the later-eluting species with a lower-order oligomer. However, assignment of the peaks specifically as tetrameric and dimeric forms requires molecular-mass calibration of the SEC column or independent determination of the molecular mass of the purified protein. Therefore,

the chromatographic profile provides evidence for the presence of different molecular-size species but does not, by itself, definitively establish their oligomeric states.

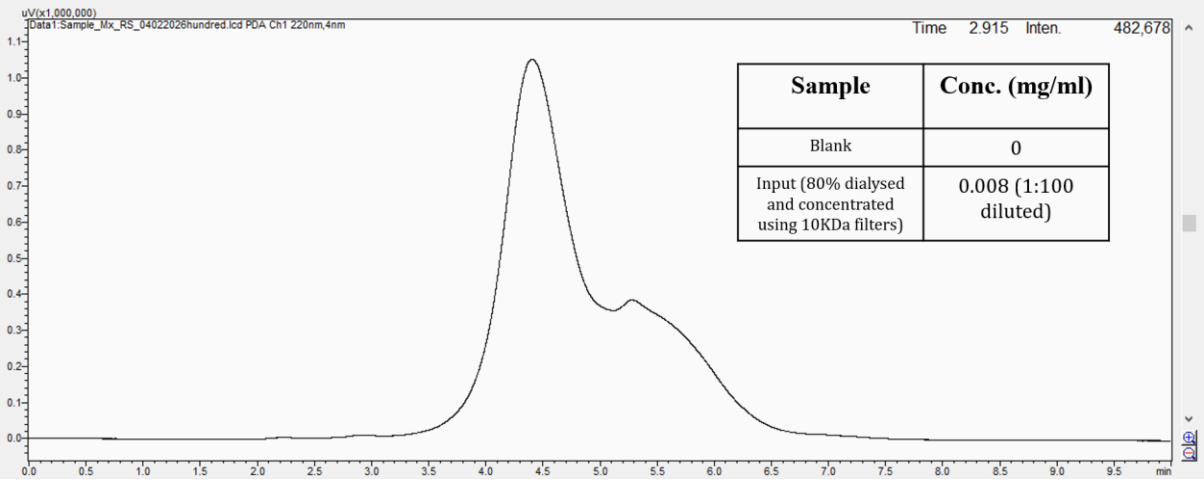


Figure 8: Size exclusion chromatography for partial purification. The dialysed fraction was loaded into the SEC column and monitored at 220 nm. Two major protein peaks were observed at approximately 4.4 and 5.2 min. (Inset table shows the protein concentration of the input fraction)

3.5 Biochemical characterisation of partially purified SOD fraction

3.5.1 Effect of pH & pH stability on SOD activity

The effect of pH on SOD activity was evaluated across a pH range of 2 – 10 using Invitrogen SOD assay activity kit that utilizes WST based inhibition approach. Apparent SOD activity was detected across pH 2–10, with unusually high inhibition values at pH 2 and 4 (Figure 9). Because the assay system contains xanthine oxidase, which contributes to superoxide generation, assay performance may be affected under strongly acidic conditions. Therefore, the unusually high inhibition observed at pH 2 and 4 was interpreted cautiously. Within the more reliable pH range, comparatively similar SOD activity was observed under near neutral to mildly alkaline conditions, particularly between pH 6 to 10. Figure 9 also shows a pH stability graph that is consistent with the earlier one. The stability study was conducted by pre-incubating the enzyme in different pH buffers (pH 2-10) for one hour. Residual activity after pre-incubation is presented as percentage inhibition, with the untreated control shown for comparison. The graph pattern matches closely with the activity profile where similar high activity (% inhibition) can be seen between pH 6 to 10. Extreme acidic pH alters the xanthine oxidase system and hence does not represent the true nature of the SOD enzyme. The observed retention of activity across neutral to mildly alkaline conditions indicates broad pH tolerance of the partially purified SOD preparation.

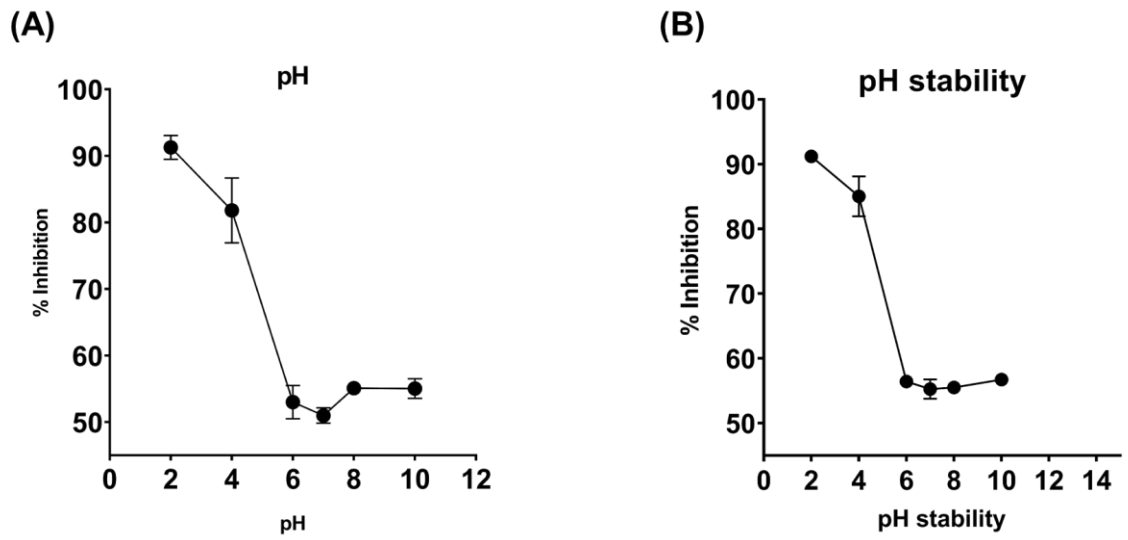


Figure 9: Effect of pH on the activity and stability of the purified SOD from *Geobacillus thermoleovorans* MXT211. (A) SOD activity (% inhibition) measured at pH 2-10. (B) Residual activity after incubation at pH 2-10. Data are mean ± SD.

3.5.2 Effect of temperature and temperature stability on SOD activity

The effect of temperature on SOD activity was assessed at various temperatures ranging from 30°C to 95°C (Figure 10). High apparent SOD activity was detected across the tested temperature range; however, these measurements were interpreted cautiously because assay components may be affected by elevated temperature. The slight decline in % inhibition at higher temperatures could be due to temperature sensitivity of components like xanthine oxidase which is mostly inactivated above 55°C and the WST that undergoes auto reduction leading to lowered % inhibition values. This provides stronger evidence for intrinsic thermal stability than the direct activity measurements performed at elevated assay temperatures. Thermal stability was measured by incubating the enzyme at different temperatures for one hour. Residual activity after pre-incubation is presented as percentage inhibition, with the untreated control shown for comparison. From the figure 10, it can be observed that the enzyme retains thermostability properties even after exposure to extreme temperatures. This provides evidence for the earlier result and further supports the enzyme as thermostable.

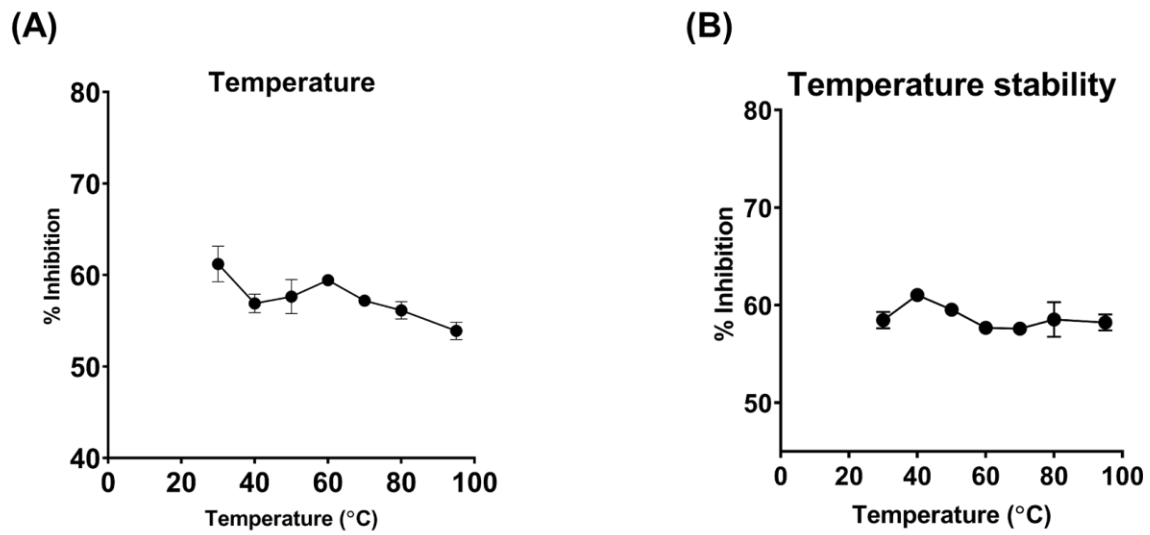


Figure 10. Effect of temperature on the activity and stability of the purified SOD from *Geobacillus thermoleovorans* MXT211. (A) SOD activity (% inhibition) measured at 30-95 °C. (B) Residual activity after incubation at 30-95 °C. Activity stayed at about 55-60% across the whole range, showing that the enzyme is thermostable. Data are mean ± SD.

3.5.3 Inhibitor profiling to infer the metal cofactor

Inhibitor profiling was performed to obtain supporting evidence for the likely metal cofactor of the SOD. Inhibitors such as hydrogen peroxide [(H_2O_2), 5 mM] and sodium azide [(NaN_3), 5 mM] are well studied and reported [36] for distinguishing between metal cofactors. Figure 11 shows the bar graph of SOD activity against these inhibitors. The enzyme was insensitive to both agents. It showed resistance to H_2O_2 and NaN_3, which is characteristic of Mn-SOD. Fe-SOD is partly resistant to NaN_3 and inhibited by H_2O_2 whereas Cu/Zn SOD is inhibited by both the agents. The observed resistance to H_2O_2 and NaN_3 was consistent with the Mn-SOD assignment and provided a supporting result for the MALDI-based identification.

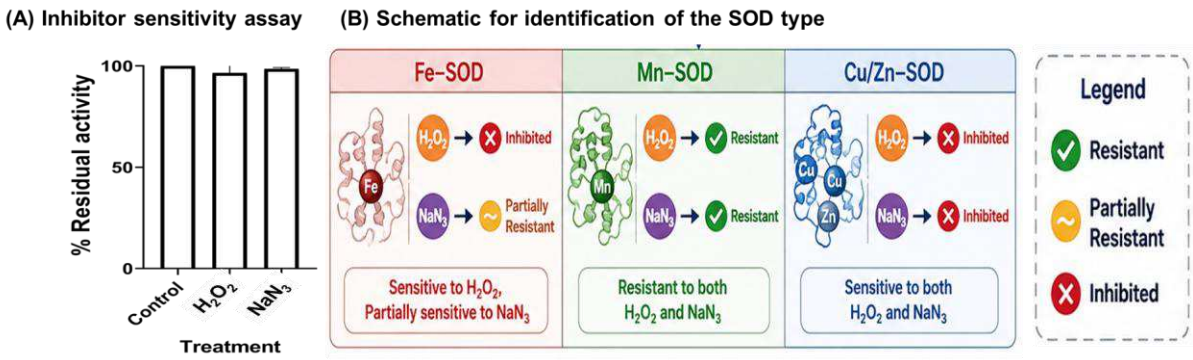


Figure 11. Identification of SOD type based on inhibitor sensitivity. (A) Experimental determination of SOD activity in the presence of H_2O_2 and NaN_3 to assess the inhibitor

sensitivity of the partially purified SOD enzyme. (B) Schematic for identification of the SOD type based on the observed inhibition pattern: Fe-SOD showed inhibition by H_2O_2 and partial resistance to NaN_3, Mn-SOD was resistant to both inhibitors, whereas Cu/Zn-SOD was inhibited by both H_2O_2 and NaN_3.

4. Discussion

This study aimed to isolate, identify, and characterize a superoxide dismutase (SOD) enzyme from a thermophilic bacterium recovered from the Tapovan hot spring, Uttarakhand, India. Through a combination of computational and experimental approaches, this study investigated the enzyme type, structural architecture, and functional properties of the SOD. Together, these complementary approaches (phenotypic characterization, sequence and structural bioinformatics, protein purification, mass-spectrometry-based identification, and biochemical characterization) provide preliminary evidence that GBRC-MXT211 produces a manganese-dependent superoxide dismutase (Mn-SOD) with substantial thermal stability and broad pH tolerance.

GBRC-MXT211 showed Gram-positive, rod-shaped, endospore-forming cells, yellow pigmentation, and growth up to 80 °C. These characteristics are consistent with *Geobacillus* and with thermophilic bacteria previously reported from the Tapovan/Soldhar region [8, 37, 38]. Earlier isolates from this region were mainly identified using phenotypic characteristics and 16S rRNA analysis, which has limited resolution for closely related *Geobacillus* species [39]. The genome-based identification of GBRC-MXT211 therefore provides stronger support for species assignment.

The GBRC-MXT211 genome was assembled into two circular contigs with a total size of 3.48 Mb and showed high completeness based on BUSCO and CheckM2 analyses [22,23]. Its genome size was comparable to previously reported *G. thermoleovorans* strains from India and other geothermal environments [40-43]. GTDB-Tk placed GBRC-MXT211 within *G. thermoleovorans*, with 99.01% ANI to the type strain KCTC 3570, well above the generally accepted species boundary (approximately 95–96% ANI) [24, 44, 45]. The strain also formed a distinct branch within *G. thermoleovorans* in the core-genome SNP analysis. The presence

of two circular contigs may indicate an additional replicon (possibly a plasmid), although further analysis is required to confirm this.

Table 3. Genome features of GBRC-MXT211 compared with published *G. thermoleovorans* genomes.

Strain	Source and conditions	Genome	GC (%)	Genes	Reference
GBRC-MXT211	Tapovan hot spring, Uttarakhand, India	3.48 Mb; 2 circular contigs (hybrid assembly)	52.43	3,555 (3,435 CDS, of which 3,359 protein-coding), PGAP	This study
CCB_US3_UF5	Ulu Slim hot spring, Malaysia (92.4 °C, pH 7)	3.60 Mb; complete circular chromosome	52.3	3,887 protein genes (KEGG)	[42]
N7	Paniphala hot spring, West Bengal, India (63 °C, pH 7.6)	3.40 Mb; draft, 226 contigs	52.4	3,530 (3,403 CDS), PGAP	[41]
RL	Manikaran hot spring, Himalaya, India	3.39 Mb; draft, 138 contigs	52.3	NR	[40]
ARTRW1	Armutlu hot spring, Turkey (74.9 °C, pH 7.1)	3.60 Mb; circular genome	NR	NR	[43]
SGAir0734	Singapore air	Chromosome plus plasmid	NR	NR	[46]

Three SOD-family proteins were identified in the GBRC-MXT211 genome: a canonical Mn/Fe-SOD, an N-terminally extended Fe/Mn-SOD, and a Cu/Zn-SOD. The presence of multiple SODs has also been reported in *Geobacillus* and other spore-forming bacteria [47, 48]. Such redundancy may provide protection against oxidative stress under different cellular conditions. Among the three proteins, the canonical Mn-SOD was selected because the SOD-active band was identified as Mn-SOD by Mascot-based peptide matching (using Swiss-Prot database) and its predicted mass (~22.9 kDa) is consistent with the monomer size of bacterial Mn-SODs.

SignalP and PSORTb predicted a Sec-dependent signal peptide and membrane-associated localization for the Cu/Zn-SOD, whereas the Mn-SOD and exosporium-associated SOD lacked classical signal peptides but were predicted to be extracellular [49, 50]. The detection of SOD activity in the culture supernatant without a conventional signal peptide is not unprecedented. An extracellular and cytosolic Fe-SOD from *Mycobacterium bovis* BCG was reported without evidence of a cleaved N-terminal signal sequence [51], suggesting that non-classical export or

release mechanisms may occur. Although SOD activity was detected in the cell-free supernatant, the present data cannot distinguish active secretion from release associated with cell lysis or sporulation, and no intracellular fraction was analysed for comparison. The extracellular origin of this SOD therefore requires further experimental confirmation.

Sequence alignment of 19 SOD proteins showed conservation of the characteristic His/Asp residues associated with the Fe/Mn-SOD catalytic centre. The AlphaFold2 model further showed that His-27, His-82, Asp-164, and His-168 clustered within the predicted active-site region, supporting the assignment of the protein to the Fe/Mn-SOD family [32, 52-54]. The GBRC-MXT211 protein belongs to the compact, short-form group of approximately 200–225 residues. Similar compact Mn-SODs have been reported from *Bacillus* and *Geobacillus*, whereas longer Fe/Mn-SODs containing N-terminal extensions have also been described [49, 55, 56]. These observations support the structural classification of GBRC-MXT211 as a canonical Fe/Mn-SOD-type protein. However, because Fe- and Mn-SODs can be structurally similar and some enzymes show cambialistic behaviour [57], direct determination of the bound metal would provide stronger confirmation of Mn specificity.

The 80% ammonium sulphate fraction showed the strongest SOD activity even though the 60% fraction contained more total protein. This indicates that protein abundance did not directly correspond to enrichment of the target enzyme and supports the use of activity-guided fraction selection. The specific activity of the dialysed fraction obtained was 10.2 U/mg. This value should not be directly compared with specific activities reported for purified recombinant SODs because the present preparation was only partially purified and the activity units may have been determined using a different assay system.

MALDI-TOF MS/MS provided additional support for the identification of the active protein as Mn-SOD. A single peptide exceeded the identity threshold in each of two independent digests (ion scores 62 and 51) exceeding the identity threshold of 41 ($p < 0.05$) [58]. Both the bands were identified from the same organism. Since the Swiss-Prot database does not contain the GBRC-MXT211 proteome, the best match was a homologous Mn-SOD from *Bacillus caldotenax*. Therefore, the MS result should be considered supportive rather than conclusive. The agreement between the MS result, predicted protein size, conserved catalytic residues, and biochemical properties is consistent with the identification of the enzyme. Re-analysis of the MS/MS spectra against the GBRC-MXT211 predicted proteome would provide further confirmation.

Size-exclusion chromatography produced two major protein peaks, suggesting the presence of proteins with different apparent molecular sizes. Different oligomeric states have been reported for bacterial Mn-SODs, including dimers and tetramers [48, 56]. As the peaks were detected by absorbance at 220 nm in a partially purified preparation and were not tested for SOD activity, they cannot be attributed to the SOD.

The enzyme showed comparatively high and similar activity from pH 6 to 10, with the separate pH stability experiment supporting a preference for neutral to alkaline conditions. Reliable activity was observed from pH 6 to 10; the unusually high inhibition values at pH 2 and 4 were attributed to assay artefacts and were not interpreted as enzyme activity. Similar pH tolerance has been reported for other thermophilic SODs [59].

The enzyme also showed substantial activity across the tested temperature range of 30–95 °C. Because the assay components can themselves be affected by temperature, the pre-incubation experiment provides stronger evidence for intrinsic thermostability. GBRC-MXT211 showed a thermal behaviour like other reported thermophilic Mn-SODs, including the *Geobacillus* EPT3 enzyme and the Mn-SOD reported from a thermophilic *Geobacillus* isolate [48,56]. A hyper-thermostable Mn-SOD has also been reported from *Thermus thermophilus* HB27 [60]. More quantitative parameters such as melting temperature and enzyme half-life would be useful for future comparison.

The inhibitor profile also supported the Mn-SOD assignment. GBRC-MXT211 showed no marked loss of activity in the presence of hydrogen peroxide or azide. Resistance to hydrogen peroxide is consistent with Mn-SODs, although inhibitor behaviour alone cannot definitively distinguish all Fe/Mn-SOD variants [48,56,57]. Direct measurement of the bound metal would therefore provide stronger confirmation of the enzyme type.

Overall, GBRC-MXT211 produces an SOD with good thermal tolerance and activity over a broad pH range. These properties indicate potential for further evaluation in applications where antioxidant enzymes are required under elevated temperature or variable pH conditions. However, the present enzyme preparation is only partially purified and showed a relatively modest activity of 0.24 U/mL. Further purification, recombinant expression, and detailed stability studies are therefore needed before its application potential can be assessed.

5. Conclusion

This study describes *Geobacillus thermoleovorans* GBRC-MXT211, a thermophile isolated from the Tapovan hot spring that grows up to 80 °C and produces SOD activity in its culture supernatant. Hybrid genome sequencing confirmed its species identity (99.01% ANI to the type strain) and revealed three SOD-family genes: a canonical Mn/Fe-SOD, an N-terminally extended Fe/Mn-SOD and a Cu/Zn-SOD. Sequence alignment and AlphaFold2 modelling of the canonical protein supported a conserved 3-His/1-Asp metal-binding site. The partially purified SOD obtained by ammonium sulphate fractionation was active from pH 6 to 10 and remained active after one hour of pre-incubation at temperatures up to 95 °C. Its resistance to hydrogen peroxide and sodium azide, together with peptide matching by mass spectrometry, is consistent with a manganese-dependent SOD

6. Limitations and future directions

This study has several limitations. First, the SOD gene identified in the genome has not been directly linked to the protein that was purified and characterised; this link is assumed from the mass-spectrometry match and the predicted properties of the protein. The SOD-active band was excised from native gels of partially purified preparations, so co-migrating proteins, including the exosporium-associated Fe/Mn-SOD that is also predicted to be extracellular, cannot be excluded. Second, the mass-spectrometry identification rests on a single peptide exceeding the identity threshold in each of two digests, which supports the assignment but does not establish it. Third, the enzyme was found in the cell-free supernatant, but no intracellular fraction was analysed, so active secretion cannot be distinguished from release by cell lysis or sporulation. Fourth, the two peaks seen by size-exclusion chromatography were detected at 220 nm on an uncalibrated column and were not tested for SOD activity, so they cannot be assigned to specific oligomeric states of the enzyme. Fifth, the pH and temperature assays rely on a coupled xanthine oxidase/WST system, which is itself affected at extreme pH and temperature, so some of the apparent activity values may reflect assay artefacts rather than enzyme behaviour. Finally, the bound metal was inferred from inhibitor sensitivity and sequence features rather than measured directly, and because Fe- and Mn-SODs can be structurally similar, and some are cambialistic, the Mn assignment remains provisional.

Future work should address these points. Purifying the enzyme to homogeneity (for example, by ion-exchange followed by calibrated SEC) would allow the active protein to be identified with greater peptide coverage, and re-searching the MS/MS data against the predicted GBRC-

MXT211 proteome would link the protein to its gene. Recombinant expression of the canonical SOD gene would confirm that this gene encodes the activity and remove the contribution of other SODs. Measuring metal content (ICP-MS) and testing activity with Mn or Fe reconstitution would establish the metal cofactor. Comparing intracellular and extracellular fractions would test whether the enzyme is secreted. Finally, assays that do not rely on xanthine oxidase, together with quantitative stability parameters (melting temperature and half-life) and kinetic constants, would give a more reliable picture of the enzyme's pH and thermal behaviour and its suitability for applications.

Data availability

The whole-genome sequence of *Geobacillus thermoleovorans* strain GBRC-MXT211 is available in the NCBI GenBank database under accession number JCAOYO000000000.1. The corresponding manganese-dependent superoxide dismutase (Mn-SOD) protein sequence is available in the NCBI Protein database under accession number MIE8011117.1. Raw sequencing data are available in the NCBI Sequence Read Archive (SRA) under accession numbers SRX34047969 (Oxford Nanopore sequencing) and SRX34047968 (Illumina sequencing). The AlphaFold2-predicted three-dimensional structure of Mn-SOD has been submitted to ModelArchive under deposition accession number ma-2zwyb (https://modelarchive.org/doi/10.5452/ma-2zwyb).

References

1. Wang, Y., Branicky, R., Noë, A. & Hekimi, S. Superoxide dismutases: dual roles in controlling ROS damage and regulating ROS signaling. *J. Cell Biol.* **217**, 1915–1928 (2018).
2. Zheng, M. et al. The applications and mechanisms of superoxide dismutase in medicine, food, and cosmetics. *Antioxidants* **12**, 1675 (2023).
3. Younus, H. Therapeutic potentials of superoxide dismutase. *Int. J. Health Sci. (Qassim)* **12**, 88–93 (2018).
4. van den Burg, B. Extremophiles as a source for novel enzymes. *Curr. Opin. Microbiol.* **6**, 213–218 (2003).
5. Vieille, C. & Zeikus, G. J. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* **65**, 1–43 (2001).
6. Verma, J., Sourirajan, A. & Dev, K. Bacterial diversity in 110 thermal hot springs of Indian Himalayan Region (IHR). *3 Biotech* **12**, 238 (2022).
7. Rawat, N. & Joshi, G. K. Bacterial community structure analysis of a hot spring soil by next generation sequencing of ribosomal RNA. *Genomics* **111**, 1053–1058 (2019).

8. Pandey, A. et al. Thermophilic bacteria that tolerate a wide temperature and pH range colonize the Soldhar (95 °C) and Ringigad (80 °C) hot springs of Uttarakhand, India. *Ann. Microbiol.* **65**, 809–816 (2015).
9. Arya, M. et al. Statistical optimization of cellulase production by a novel thermophilic lignocellulolytic bacterium *Geobacillus stearothermophilus* TP-3 from Tapovan hot spring in India. *J. Pure Appl. Microbiol.* **19**, 1085–1100 (2025).
10. Pandey, A. et al. *Geobacillus stearothermophilus* (GBPI_16): a resilient hyperthermophile isolated from an autoclaved sediment sample. *Proc. Natl. Acad. Sci. India Sect. B Biol. Sci.* **84**, 349–356 (2014).
11. Kumar, A. et al. *Geobacillus thermoleovorans* MTCC 13131: an amide-hydrolyzing thermophilic bacterium isolated from a hot spring of Manikaran. *Indian J. Microbiol.* **62**, 618–626 (2022).
12. Satam, H. et al. Next-generation sequencing technology: current trends and advancements. *Biology* **12**, 997 (2023).
13. Siddikey, F. et al. Enzyme technology in the food industry: molecular mechanisms, applications, and sustainable innovations. *Food Sci. Nutr.* **13**, e70927 (2025).
14. Porta, R., Pandey, A. & Rosell, C. M. Enzymes as additives or processing aids in food biotechnology. *Enzyme Res.* **2010**, 436859 (2011).
15. Devi, M. et al. Advances in microbial enzyme technology for food processing strategies and applications. *Discov. Food* **6**, 56 (2026).
16. Bhatt, D. et al. Investigating the biotechnological potential of thermophiles using integrated culture-dependent and metagenomic approaches from Tapovan hot springs for industrial thermozymes. *BMC Microbiol.* (2026).
17. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
18. Bonenfant, Q., Noé, L. & Touzet, H. Porechop_ABI: discovering unknown adapters in Oxford Nanopore Technology sequencing reads for downstream trimming. *Bioinform. Adv.* **3**, vbac085 (2022).
19. Wick, R. R. Filtlong v0.3.1: quality filtering tool for long reads. *GitHub* https://github.com/rrwick/Filtlong (2025).
20. Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **13**, e1005595 (2017).
21. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).
22. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
23. Chklovski, A., Parks, D. H., Woodcroft, B. J. & Tyson, G. W. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nat. Methods* **20**, 1203–1212 (2023).

24. Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* **38**, 5315–5316 (2022).
25. Seemann, T. Snippy: fast bacterial variant calling from NGS reads. *GitHub* https://github.com/tseemann/snippy (2015).
26. Croucher, N. J. et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* **43**, e15 (2015).
27. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
28. Shukla, N. et al. Genome data associated with *Geobacillus thermoleovorans* strain GBRC-MXT211 [dataset]. *Figshare* https://doi.org/10.6084/m9.figshare.33728434 (2026).
29. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
30. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
31. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
32. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
33. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
34. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
35. Beauchamp, C. & Fridovich, I. Superoxide dismutase: improved assays and an assay applicable to acrylamide gels. *Anal. Biochem.* **44**, 276–287 (1971).
36. Sheng, Y. et al. Superoxide dismutases and superoxide reductases. *Chem. Rev.* **114**, 3854–3918 (2014).
37. Nazina, T. N. et al. Taxonomic study of aerobic thermophilic bacilli: descriptions of *Geobacillus subterraneus* gen. nov., sp. nov. and *Geobacillus uzenensis* sp. nov. from petroleum reservoirs and transfer of *Bacillus stearothermophilus*, *Bacillus thermocatenulatus*, *Bacillus thermoleovorans*, *Bacillus kaustophilus*, *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations. *Int. J. Syst. Evol. Microbiol.* **51**, 433–446 (2001).
38. Arya, M., Joshi, G. K., Gupta, A. K., Kumar, A. & Raturi, A. Isolation and characterization of thermophilic bacterial strains from Soldhar (Tapovan) hot spring in Central Himalayan Region, India. *Ann. Microbiol.* **65**, 1457–1464 (2015).
39. Ticona, A. R. P. et al. Functional characterization, genome assembly, and annotation of *Geobacillus* sp. G4 isolated from a geothermal field in Tacna, Peru. *Microorganisms* **13**, 1374 (2025).

40. Singh, D. N. et al. Genome sequencing revealed the biotechnological potential of an obligate thermophile *Geobacillus thermoleovorans* strain RL isolated from hot water spring. *Indian J. Microbiol.* **59**, 351–355 (2019).
41. Bose, S. et al. Genome sequence of the multiple-protease-producing strain *Geobacillus thermoleovorans* N7, a thermophilic bacterium isolated from Paniphala hot spring, West Bengal, India. *Genome Announc.* **4**, e01202-16 (2016).
42. Muhd Sakaff, M. K. L., Abdul Rahman, A. Y., Saito, J. A., Hou, S. & Alam, M. Complete genome sequence of the thermophilic bacterium *Geobacillus thermoleovorans* CCB_US3_UF5. *J. Bacteriol.* **194**, 1239 (2012).
43. Oztug, M., Cebeci, A., Mumcu Topaloğlu, H., Akgoz, M. & Gül Karagüler, N. Whole-genome sequence of *Geobacillus thermoleovorans* ARTRW1, isolated from Armutlu geothermal spring, Turkey. *Microbiol. Resour. Announc.* **9**, e00269-20 (2020).
44. Richter, M. & Rosselló-Móra, R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. USA* **106**, 19126–19131 (2009).
45. Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114 (2018).
46. Gaultier, N. E. et al. Genome sequence of *Geobacillus thermoleovorans* SGAir0734, isolated from Singapore air. *Genome Announc.* **6**, e00636-18 (2018).
47. Inaoka, T., Matsumura, Y. & Tsuchido, T. Molecular cloning and nucleotide sequence of the superoxide dismutase gene and characterization of its product from *Bacillus subtilis*. *J. Bacteriol.* **180**, 3697–3703 (1998).
48. Zou, Y. et al. Identification of three superoxide dismutase genes from a *Geobacillus* sp. *Protein J.* **30**, 66–71 (2011).
49. Teufel, F. et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat. Biotechnol.* **40**, 1023–1025 (2022).
50. Yu, N. Y. et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26**, 1608–1615 (2010).
51. Kang, S.-K., Jung, Y.-J., Kim, C.-H. & Song, C.-Y. Extracellular and cytosolic iron superoxide dismutase from *Mycobacterium bovis* BCG. *Clin. Diagn. Lab. Immunol.* **5**, 784–789 (1998).
52. Parker, M. W. & Blake, C. C. F. Crystal structure of manganese superoxide dismutase from *Bacillus stearothermophilus* at 2.4 Å resolution. *J. Mol. Biol.* **199**, 649–661 (1988).
53. Ludwig, M. L., Metzger, A. L., Pattridge, K. A. & Stallings, W. C. Manganese superoxide dismutase from *Thermus thermophilus*: a structural model refined at 1.8 Å resolution. *J. Mol. Biol.* **219**, 335–358 (1991).
54. Adams, J. J., Morton, C. J. & Parker, M. W. The crystal structure of the manganese superoxide dismutase from *Geobacillus stearothermophilus*: Parker and Blake (1988) revisited. *Aust. J. Chem.* **73**, 145–150 (2020).

55. Areekit, S. et al. Cloning, expression, and characterization of thermotolerant manganese superoxide dismutase from *Bacillus* sp. MHS47. *Int. J. Mol. Sci.* **12**, 844–856 (2011).
56. Zhu, Y., Wang, G., Ni, H., Xiao, A. & Cai, H. Cloning and characterization of a new manganese superoxide dismutase from deep-sea thermophile *Geobacillus* sp. EPT3. *World J. Microbiol. Biotechnol.* **30**, 1347–1357 (2014).
57. Lancaster, V. L., LoBrutto, R., Selvaraj, F. M. & Blankenship, R. E. A cambialistic superoxide dismutase in the thermophilic photosynthetic bacterium *Chloroflexus aurantiacus*. *J. Bacteriol.* **186**, 3408–3414 (2004).
58. Perkins, D. N., Pappin, D. J. C., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567 (1999).
59. Monsalves, M. T., Amenabar, M. J., Ollivet-Besson, G. P. & Blamey, J. M. Effect of UV radiation on a thermostable superoxide dismutase purified from a thermophilic bacterium isolated from a sterilization drying oven. *Protein Pept. Lett.* **20**, 749–754 (2013).
60. Liu, J. G., Yin, M. M., Zhu, H., Lu, J. R. & Cui, Z. F. Purification and characterization of a hyperthermostable Mn-superoxide dismutase from *Thermus thermophilus* HB27. *Extremophiles* **15**, 221–226 (2011).

Acknowledgements

The authors would like to acknowledge Dr. Ishan Rawal for his assistance in carrying out size exclusion chromatography.

Funding

The authors would like to acknowledge the Department of Science and Technology, Government of Gujarat, India for the project grant (Ref. No: GBRC/GoG/DST/JD-2/BS14/2017-18).

Declaration

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.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementarySOD15102026.docx