Biochemical and Biophysical Research Communications 411 (2011) 88–95

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Enzymatic characterization of Catalase from *Bacillus anthracis* and prediction of critical residues using information theoretic measure of Relative Entropy

Amit Rahi [a,1], Mohd Rehan [b,1], Rajni Garg [a], Deeksha Tripathi [a], Andrew M. Lynn [b,*], Rakesh Bhatnagar [a,*]

[a] Laboratory of Molecular Biology and Genetic Engineering, School of Biotechnology, Jawaharlal Nehru University, New Delhi 110067, India
[b] School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi 110067, India

ARTICLE INFO

Article history:
Received 8 June 2011
Available online 24 June 2011

Keywords:
Catalase
Bacillus anthracis
Sequence analysis
Relative Entropy
ROS
Activity
Mutagenesis

ABSTRACT

In order to cope up with the reactive oxygen species (ROS) generated by host innate immune response, most of the intracellular organisms express Catalase for the enzymatic destruction/detoxification of hydrogen peroxide, to combat its deleterious effects. Catalase thus, scavenges ROS thereby playing a pivotal role in facilitating the survival of the pathogen within the host, and thus contributes to its pathogenesis. *Bacillus anthracis* harbors five isoforms of Catalase, but none of them has been studied so far. Thus, this study is the first attempt to delineate the biochemical and functional characteristics of one of the isoforms of Catalase (Cat1.4) of *B. anthracis*, followed by identification of residues critical for catalysis. The general strategy used, so far for mutational analysis in Catalases is structure based, i.e. the residues in the vicinity of heme were mutated to decipher the enzymatic mechanism. However, in the present study, protein sequence analysis was used for the prediction of catalytically important residues of Catalase. Essential measures were adopted to ensure the accuracy of predictions like after retrieval of well-annotated sequences from the database with EC 1.11.1.6, preprocessing was done to remove irrelevant sequences. The method used for multiple alignment of sequences, was guided by structural alignment and thereafter, an information theoretic measure, Relative Entropy was used for the critical residue prediction. By exploiting this strategy, we identified two previously known essential residues, H55 and Y338 in the active site which were demonstrated to be crucial for the activity. We also identified six novel crucial residues (Q332, Y117, H215, W257, N376 and H146) located distantly from the active site. Thus, the present study highlights the significance of this methodology to identify not only those crucial residues which lie in the active site of Catalase, but also the residues located distantly.

1. Introduction

Bacillus anthracis, a Gram-positive, non-motile, spore-forming organism is the etiological agent of anthrax. Owing to the emergence of multidrug resistance and ineffective vaccinations against anthrax, the inhalation of anthrax spores can prove fatal. Although, the major anti-anthrax drugs/vaccines primarily remain focused on the tripartite exotoxin complex (PA, EF and LF); the other factors underlying the anthrax pathogenesis remain poorly understood. To successfully thrive and confront the hostile microenvironment within the host, most of the pathogens including *B. anthracis*, exploit oxidative stress defense mechanism to subvert and evade the host innate immune system. It has been demonstrated that *katB* or *Catalases* are up-regulated during the growth of *B. anthracis* within the macrophages [1]. The pathogens are endowed with protective mechanisms against reactive oxygen species (ROSs) generated within the host, such as ones generated by professional phagocytic cells [2]. Antioxidant enzymes, like Catalases and superoxide dismutases (SODs) play a pivotal role in facilitating the survival of pathogen within the host and thus, indirectly contribute to the pathogenesis of bacterium. The study of the effect of different organic solvents on the conformation and activity of Catalase could provide clue to understand the mechanism of enzyme action [3]. Catalase, a ubiquitous metalloenzyme guards aerobic organisms, as an intracellular H_2O_2-scavenger. The pathogen-encoded nitric oxide synthase has been shown to activate Catalase and subsequently, suppress the Fenton reaction [4]. Four SODs (including two on the exosporium) provide redundant and combinatorial protection to the bacterium from oxidative stress in the macrophages [5]. The ability of anthrax spores to effectively scavenge free radicals, further indicates that these enzymes are functionally active on the exosporium. On the contrary, the significance of Catalases in combating oxidative stress has not been

* Corresponding authors. Fax: +91 11 26742040 26 (R. Bhatnagar).
E-mail addresses: andrew@mail.jnu.ac.in (A.M. Lynn), rakeshbhatnagar@mail.jnu.ac.in (R. Bhatnagar).
[1] These authors contributed equally to this work.

0006-291X/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2011.06.099

studied in detail, until now. Moreover, there is no evidence for its presence on the exosporium or on the surface of the vegetative cells; besides the presence of an identical $CotJ_c$ protein in the exosporium of *B. subtilis* [6]. However, the presence of five ORFs corresponding to different isoforms of Catalases in the genome of *B. anthracis* prompted us to question the significance of their existence. Thus, the present study was primarily designed to delineate the biochemical and functional characteristics of one of the five isoforms of Catalase (abbreviated as rCat1.4 hereinafter) of *B. anthracis*, followed by identification of residues critical for catalysis. The general strategy, used so far for mutational analysis in Catalases is structure based, i.e. the residues in the vicinity of heme were selected and mutated to decipher the mechanism of enzyme action [7,8]. Once a protein has evolved functionally, the majority of mutations are neutral at the molecular level and do not affect the function and folding of the protein, whereas those mutations which are deleterious (loss of function or misfolding) provide selection pressure for residue conservation [9]. Thus, the residue conservation in a multiple alignment of a protein and its homologs indicate the importance of the residue in maintaining the structure and function of the protein. Traditional methods of measuring conservation are based on the physico-chemical properties of amino acids, and have evolved to the increased use of information theoretic measures [10]. Relative Entropy (RE) is an information theoretic measure of the difference between two probability distributions and has been increasingly applied in bioinformatics for identification of functional residues [11,12]. Conventional use of RE exploits background frequencies from the natural abundance of the amino acid [11]. Further improvements in the prediction of functional residues are achieved by the use of a context-specific background frequency [13–15]. In this study, background frequency in context of Catalase protein family is calculated as observed frequencies of amino acids from the alignment. The residues predicted using RE scoring scheme were validated by site-directed mutagenesis, to assess their role in the catalytic mechanism. These residues are speculated, to provide a valuable insight, into the mechanism of enzyme action that can have far-reaching implications.

2. Materials and methods

2.1. Bacterial strains and chemicals

Escherichia coli strains, DH5α and BL21 (λDE3) were used as hosts for cloning and expression, respectively. Plasmid pET-28a (+) for heterologous gene expression, was procured from Novagen (Madison, USA). *B. anthracis* Sterne strain 34F2 (pXO1+, pXO2−) was used. Restriction enzymes were from Fermentas GmbH (Germany). Ni^{2+}–NTA agarose resin was from Qiagen (Hilden, Germany). Oligonucleotides employed in the study were synthesized by Sigma (St. Louis, USA).

2.2. Expression and purification of the rCat1.4

The open reading frame corresponding to the Cat1.4 isoform from *B. anthracis* {(UniProt), YP_027348.1 (488 aa)} was identified. PCR amplification of *cat1.4* (1467 bp) was carried out with primers (Supplementary Table S1), containing *BamH*I and *Sal*I employing genomic DNA of *B. anthracis*. The amplified DNA fragments were ligated to pET-28a (+) expression vector. The recombinant protein (rCat1.4) with N-terminal 6×-His tag was expressed in *E. coli* BL21 (λDE3), by inducing the cells grown till $OD_{600} = 0.6$ with 1 mM IPTG for 4–5 h at 37 °C. The rCat1.4 was purified from the soluble fraction under native conditions using Ni^{2+}–NTA affinity chromatography. The purified protein was dialyzed against 10 mM sodium phosphate buffer, pH 7.4.

2.3. Kinetic parameters of the recombinant Catalase isoform, rCat1.4

2.3.1. Catalase (rCat1.4) activity

The enzymatic activity was measured spectrophotometrically at 25 °C by monitoring the decrease in absorbance at 240 nm (initial linear rate), resulting from the elimination of H_2O_2, using a Cary 100 BIO UV Spectrophotometer (Sunnyvale, CA). The ε for H_2O_2 at 240 nm was taken as 43.6 $M^{-1}cm^{-1}$. The standard assay reaction contained 50 mM sodium phosphate buffer (pH 7.0), 10 mM H_2O_2 and 300 ng of rCat1.4 in a total volume of 1.0 ml.

2.3.2. Determination of optimum pH, thermal stability and catalytic properties of rCat1.4

The optimal pH required for the activity of rCat1.4 was determined as described above except using buffers of varying pH; 50 mM sodium acetate (pH: 3.0, 3.5, 4.0, 4.5 and 5.0), 50 mM sodium phosphate (pH: 6.0, 6.5, 7.0, 7.5 and 8.0) and 50 mM Tris–HCl buffer (pH: 8.5, 9.0, 9.5 and 10.0). The thermal stability of rCat1.4 was determined by incubating the enzyme in 50 mM phosphate buffer (pH 7.0) for 30 min in a broad temperature range of 4–70 °C prior to the activity measurement. The K_m of rCat1.4 for H_2O_2 was determined by employing a wide range of H_2O_2 concentrations (1–90 mM).

2.4. Sequence retrieval

Catalase sequences (470) with EC 1.11.1.6 were sourced from Enzyme Database Release of 05-Feb-2008 [16]. *B. anthracis* Cat1.4 sequence was retrieved from KEGG database and Catalase structures (74) were retrieved from PDB (2009-06-16).

2.5. Pre-processing of sequences

The Markov cluster (MCL) algorithm, based on the precomputed sequence similarity information for the assignment of proteins into subfamilies, has been used previously to cluster protein sequences. Removal of redundant sequences was achieved using BLASTclust, by setting the identity threshold S to 90 keeping all other parameters at their default values [17].

2.6. Structurally guided sequence alignment

The PDB structure sequences were aligned using python script salign.py by Modeller9v7 [18]. The structural alignment thus obtained, was taken as the seed for aligning the protein sequences from enzyme database using MAFFT v6.707 with FFT-NS-i (slow; iterative refinement method) [19].

2.7. Relative Entropy

Relative Entropy at a site '*i*' of MSA was calculated as the divergence of probability distribution of residues from background probability distribution. Background probability is defined, as the probability of occurrence of particular residue in the protein family.

$$RE_i = \sum_{x=1\cdots 20} p_i(x) \log \frac{p_i(x)}{f(x)}$$

wherein; the $p_i(x)$ is the probability of the occurrence of amino acid '*x*' in column '*i*' of the MSA. The $f(x)$ is the background probability of amino acid '*x*' in the protein family and is estimated from the alignment.

2.8. Homology modeling

Homology model of Cat1.4 sequence was generated using Modeller9v7. The PDB id: 2J2M from *Exiguobacterium oxidotolerans*, was used as the template for modeling Cat1.4 sequence [20]. The model was evaluated and validated by PROCHECK v.3.5.4 [21]. The coordinates of heme were taken from the template and the energy was minimized using Gromacs version 4.0.5 [22].

2.9. Site-directed mutagenesis of rCat1.4 for validation of predicted functionally critical residues

Point mutations were generated using the Quik-change site-directed mutagenesis kit (Stratagene, USA). Briefly, the pET-28a (+) plasmid harboring the wildtype rCat1.4 isoform was used as a template for PCR based site-directed mutagenesis. The transformants obtained were subjected to dideoxy sequencing to confirm the incorporation of the specific mutations, at the desired site. The oligonucleotides used for specific H55A, Y117A, Y338A, H146A, H215A, W257A, Q332A, N376A mutations are shown in Supplementary Table S1. Moreover, in order to test whether the loss of activity was only restricted to the residues having a higher Relative Entropy score, six additional mutants were created by choosing randomly from low RE positions, having no more than 10% gaps in the alignment column serving as negative controls. The oligonucleotides used for R6A, N82A, V210A, D270A, P391A and Y397A are also shown in Supplementary Table S1. All the mutant proteins were purified as described above and their catalytic potential was compared with the wild-type rCat1.4.

2.10. Secondary structure prediction of rCat1.4 and its point mutants by circular dichroism (CD) spectroscopy

The CD spectra of rCat1.4 and its point mutants (0.2 mg/ml in 10 mM phosphate buffer, pH 7.4) were obtained using Jasco Corp., J-710 Spectropolarimeter at 25 °C using a 1 mm cell, a wavelength scan from 190 to 240 nm at the rate of 20 nm/min. The average of 10 successive spectra was calculated in each case followed by the baseline correction which was done by subtracting the spectrum obtained with the buffer alone. The CD signals were converted to molar ellipticity by using the following relationship: $\theta_m = \theta_o \cdot 100/lc$, where, θ_m is the molar ellipticity (in degrees per square centimeter per decimole), θ_o is the observed ellipticity (in degrees), l is the path length (in centimeters), and c is the molar concentration. The fraction of each secondary structure in the recombinant protein was predicted by K2d software.

3. Results

3.1. Expression and purification of the rCat1.4 of B. anthracis

The rCat1.4 was expressed in BL21 (λDE3) cells post-induction with 1 mM IPTG and an intense band at 58 kDa was observed corresponding to rCat1.4 (data not shown). The recombinant protein was purified to near homogeneity as shown in Fig. 1A.

3.2. Enzymatic characterization rCat1.4 of B. anthracis

The specific activity of rCat1.4 was found to be 5132 mM $min^{-1}mg^{-1}$ under standard assay conditions. The K_m and V_{max} were observed to be 13.744 ± 1.204 mM (for H_2O_2) and 5.904 ± 0.207 mM per minute, respectively. The representative Michaelis–Menten and Lineweaver–Burk plots are shown in panels I and II of Fig. 1B, respectively. The rCat1.4 showed enzymatic activity over a wide pH range of 6–10, with a maximal V_0 (6116.2 mM/min/mg) at pH 7.0. It is stable over a temperature range of 10–35 °C and showed highest activity at 25 °C (data not shown).

3.3. Preprocessing of all Catalase sequences

Clustering of Catalase sequences (including *B. anthracis* Catalases) using software MCL version 06–058 was done as described [23]. Three clusters of sequences (including sequences of PDB structures), first with 377 sequences, second with 193 sequences and third with 3 sequences and 1 protein fragment CAT1_COMTR (57 amino acids) were obtained (Fig. 2A). Second cluster (193 sequences) was chosen for further analysis, since, it included Cat1.4 sequence. These 193 sequences were culled to 90% identity, to give 88 sequences by retaining the maximum representation of PDB structure sequences. These 88 sequences consist of 73 sequences (from enzyme database) and 15 sequences (from PDB structures). The other clusters were of Catalase-peroxidases (KatG) and mitochondrion Catalases (Mcat) and are different from Catalases.

3.4. Multiple sequence alignment containing known motifs

The distal heme ligand binding domain ^{46}RERIPERVVHAKG^{58} encompassing the essential distal histidine residue, and the proximal heme ligand binding domain ^{334}R-F-Y-D^{340} harboring the essential proximal tyrosine residue [24] were evident in our alignment (Fig. 2B). Other Catalase specific motifs like ^{126}VGNNTP^{131} were also found in our alignment [25]. Two other motifs ^{107}RDXRGFAXKFYT^{118} and ^{92}RFSTV^{96} were identified. Tyr117 from ^{107}RDXRGFAXKFYT^{118} was selected for mutation and found to be crucial for activity.

3.5. Homology modeling of B. anthracis Cat1.4

The Catalase of *E. oxidotolerans*, PDB id: 2J2M was used as the template for modeling Cat1.4 sequence since it displayed 65% identity with the Cat1.4 of *B. anthracis*. The model generated (Fig. 3A), when validated by Ramachandran plot revealed the presence of only one residue in the disallowed region, similar to that of the template (Supplementary Table S2). The G-factor was calculated to be >−0.5 and the number of labeled residues deduced from Ramachandran and Chi^1–chi^2 plots were also less. Thus, it can be inferred that the model generated is free from any possible errors and can serve as a good model.

3.6. Prediction of catalytically essential residues for Cat1.4 using RE score

We attempted to predict the critical residues required for the catalytic activity of Cat1.4, using RE based approach. The structurally guided sequence alignment was used for calculating the Relative Entropy and then high-ranking residues (Table 1) were mapped onto the Cat1.4 modeled structure (Fig. 3A). The complete alignment and the RE scores for all the columns are provided as *Supplementary* file 1 and 2. The catalytic residues of bovine liver Catalase (BLC hereinafter), H74, Y357 and N147 correspond to structurally equivalent residues, H128, Y415 and N201 in *E. coli* HPII, respectively [26–27,7]. The H74 and Y357 of BLC are known to play an important role in the catalytic function [26]. These residues displayed high RE scores and thus, were chosen for mutational analysis. The distribution curve generated on the basis of RE scores from the MSA is shown in Fig. 3B.

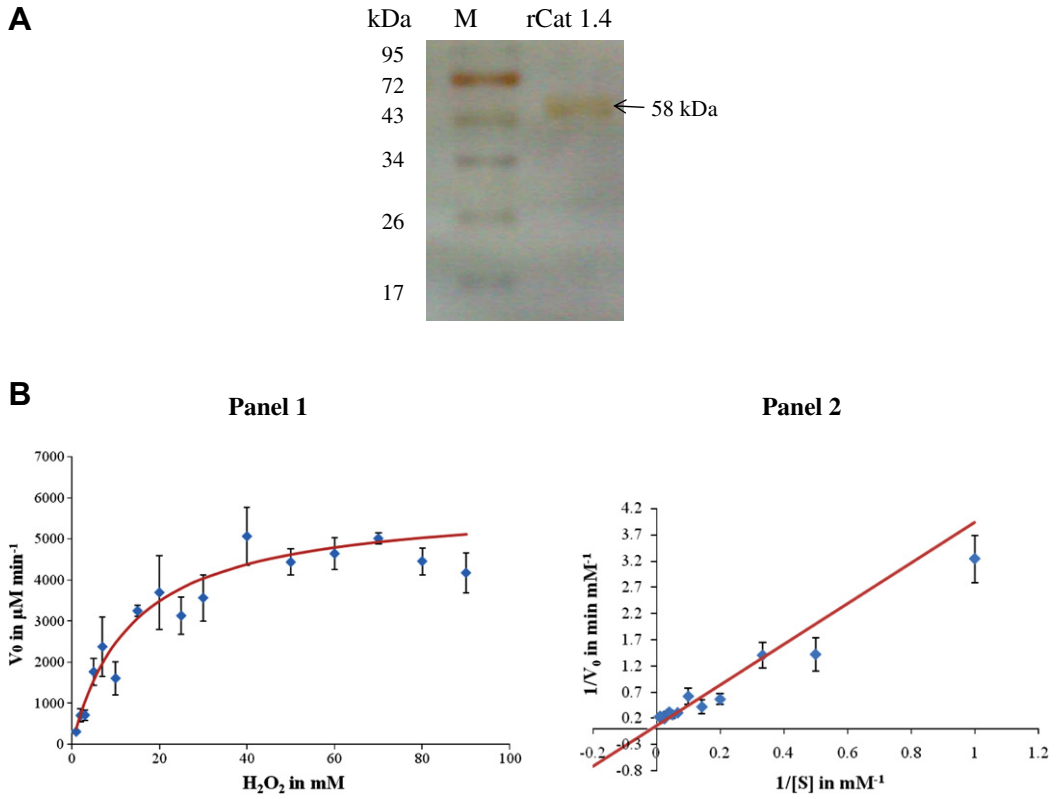

Fig. 1. Enzymatic activity of rCat1.4 of *Bacillus anthracis* (A) silver stained PAG showing purified rCat1.4. (B) Elucidation of enzymatic activity of rCat1.4, (I) represents Michaelis–Menten plot and (II) Lineweaver–Burk plot. Points represent mean ± S.D. for three independent experiments done in triplicates.

3.7. Mutational analysis of the residues that demonstrated high RE scores for their ability to participate in enzyme catalysis

Based on high RE score, the top ranking residues (H55, Y117, H146, H215, W257, Q332, Y338 and N376) were mutated to alanine by site-directed mutagenesis. The yields of the recombinant mutant proteins were comparable to the wild type rCat1.4 obviating the possibility of any toxic effects conferred to the protein upon alanine substitution (data not shown). The enzymatic activity of the mutant proteins was investigated (Fig. 4A). Except Q332A and N376A, all other mutants exhibited a total loss of the enzymatic activity ($\geqslant$96%). The Q332A mutant showed 60% of the activity while the N376A was marginally active ($\approx$30%). There was no significant change in the enzymatic activity of all the negative control mutants (possessing low RE score) from that of rCat1.4 as anticipated (Fig. 4A).

The far-UV CD spectrum of rCat1.4 showed negative peaks at 222 and 209 nm and a positive peak at 195 nm, hallmark of high α-helical and low β-sheet content. The mutants Y117A, H146A and W257A exhibited substantial perturbation in their secondary structure content relative to the wildtype whereas H55A did not (Fig. 4B). There was no significant change in the CD spectra of D270A and V210A (negative controls) mutants from that of rCat1.4 (Fig. 4B). This clearly suggests that the selected residues on the basis of high RE score do lead to loss of activity either by affecting the mechanism of catalysis or changing the native fold of the enzyme.

4. Discussion

Catalases relieve almost all the organisms from the deleterious peroxide stress by catalyzing the same to free oxygen and water. *B. anthracis spores* reside within the alveolar macrophages and germinate into vegetative bacilli, leading to pulmonary anthrax. One of the major protective mechanisms employed by the host against an intracellular pathogen, is generation of reactive oxygen intermediates (ROIs) like H_2O_2. To cope up with this host response, the pathogen relies upon enzymes like Catalases and superoxide dismutases (SODs). It has been demonstrated that actively growing *B. subtilis* expresses two major Catalases, *KatA* and *KatB*; wherein, KatA is the major Catalase isoform while *katB* gene is placed under the stringent control of stress-regulated sigma factor [28–30]. Interestingly, a third Catalase, encoded by *katX* was also identified to function in the dormant spores [28]. In *Enterococcus faecalis* V583 strain, although *katA* gene is under the control of its native promoter, no KatA protein could be detected in heme deficient medium [31]. Therefore, the expression of different isoforms of Catalase might not dependent only on the presence and amount of ROIs, but also on their co-factors like heme. In *B. anthracis*, the expression of different Catalase isoforms is expected to be location and stage-dependent, i.e. it varies *in vitro* culture filtrates, during sporulation, within the macrophages and toxin-producing vegetative bacilli during systemic infection. We thus believe, that multiple isoforms for a single protein in *B. anthracis*, is not only a mere gene duplication event in evolution; rather is a tightly controlled intricate phenomenon wherein, the expression and localization of each Catalase isoform is temporally and spatially distributed to perform a destined stage-/location-specific function.

Besides characterizing the enzymatic potential of rCat1.4 of *B. anthracis*, we have elucidated its functionally important residues using sequence analysis by employing the information theoretical measure of Relative Entropy, instead of general structure- based approaches. To ensure the accuracy of predictions, well-annotated sequences of Catalase with EC 1.11.1.6 were taken and clustered

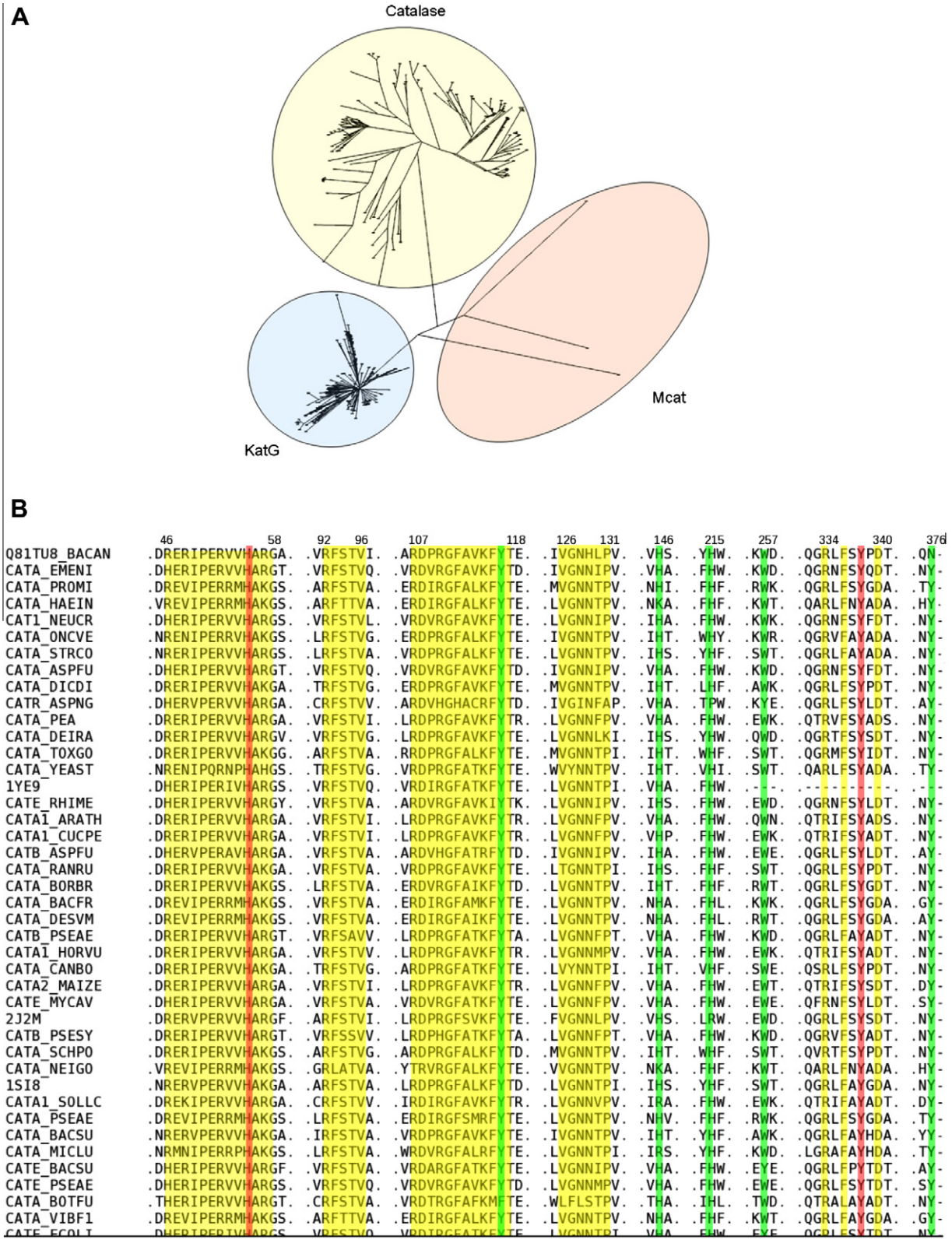

Fig. 2. (A) Clustering of Catalase sequences. (B) Multiple sequence alignment showing conserved motifs and crucial residues. The motifs are highlighted in yellow. Columns showing essential known residues (Distal His and Proximal Tyr) are highlighted in red whereas the identified crucial residues are highlighted in green. The motif identified in our alignment, ^{107}RDXRGFAXKFYT^{118} contains Y116 which is selected for mutation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

using MCL. Of the resultant three clusters (Fig. 2A), two were found to be Catalase-peroxidases (KatG) and mitochondrial Catalases (Mcat), whereas the other constituted the remaining Catalase sequences. The Catalase-peroxidases (KatG) are bifunctional enzymes that exhibit both Catalase and peroxidase activity and they do not share sequence similarity with mono-functional Catalases

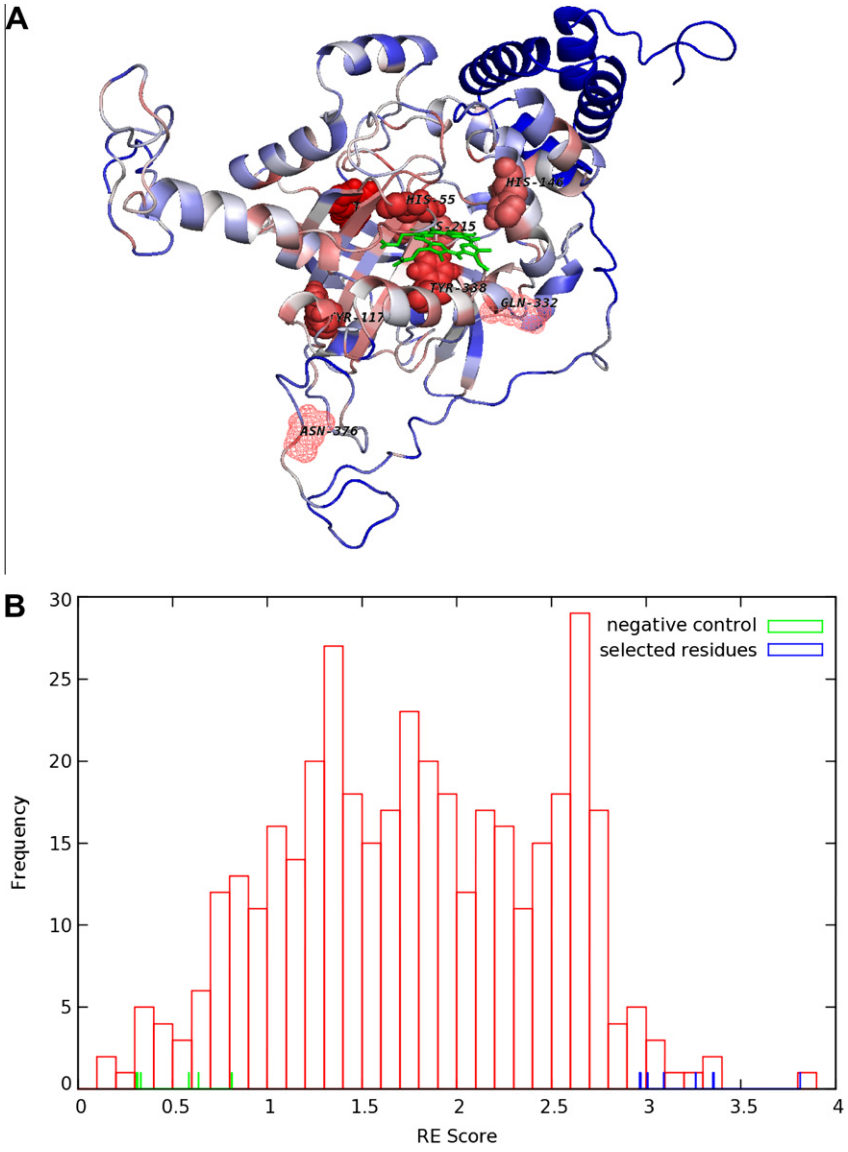

Fig. 3. (A) Cartoon representation of 3D homology model of Cat1.4. The Relative Entropy scores for the Catalase family sequences are projected onto the 3D structure. The residues for which the mutated protein retained residual activity out of the residues selected for mutation are shown as mesh with labeling. The rest other selected residues are shown as labeled spheres on the model. Heme is shown as sticks in green color. The structure is viewed using Pymol v0.99 (http://www.pymol.org). (B) Histogram of RE score of complete MSA. The high RE scoring predicted functional residues selected for mutation are shown at right tail in blue color whereas the low RE scoring residues serving as negative controls are shown in green color at left tail of the histogram. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Twelve top ranking residues based on Relative Entropy score.

BAT Cat1.4 residue	Structurally equivalent residue in BLC	Relative Entropy	Loss in activity upon mutation (%)	Location of residue	Distance from heme (Å)
H55[a]	H74	3.36	100	Distal	3.03
Y177	Y136	3.26	100	Proximal	>10
H146	H165	2.97	100	Distal	>10
K149	K168	2.93	No expression	Distal	9.56
M192	M211	3.13	No expression	Distal	8.24
H215	H234	3.01	100	Proximal	>10
W257	W276	3.81	100	Proximal	>10
Q262	Q281	2.98	No expression	Proximal	>10
W283	W302	3.06	No expression		>10
Q332	Q351	2.97	40	Proximal	6.44
Y338[a]	Y357	3.35	100	Proximal	1.93
N376	Y403	3.09	70		>10

[a] Known essential residues.

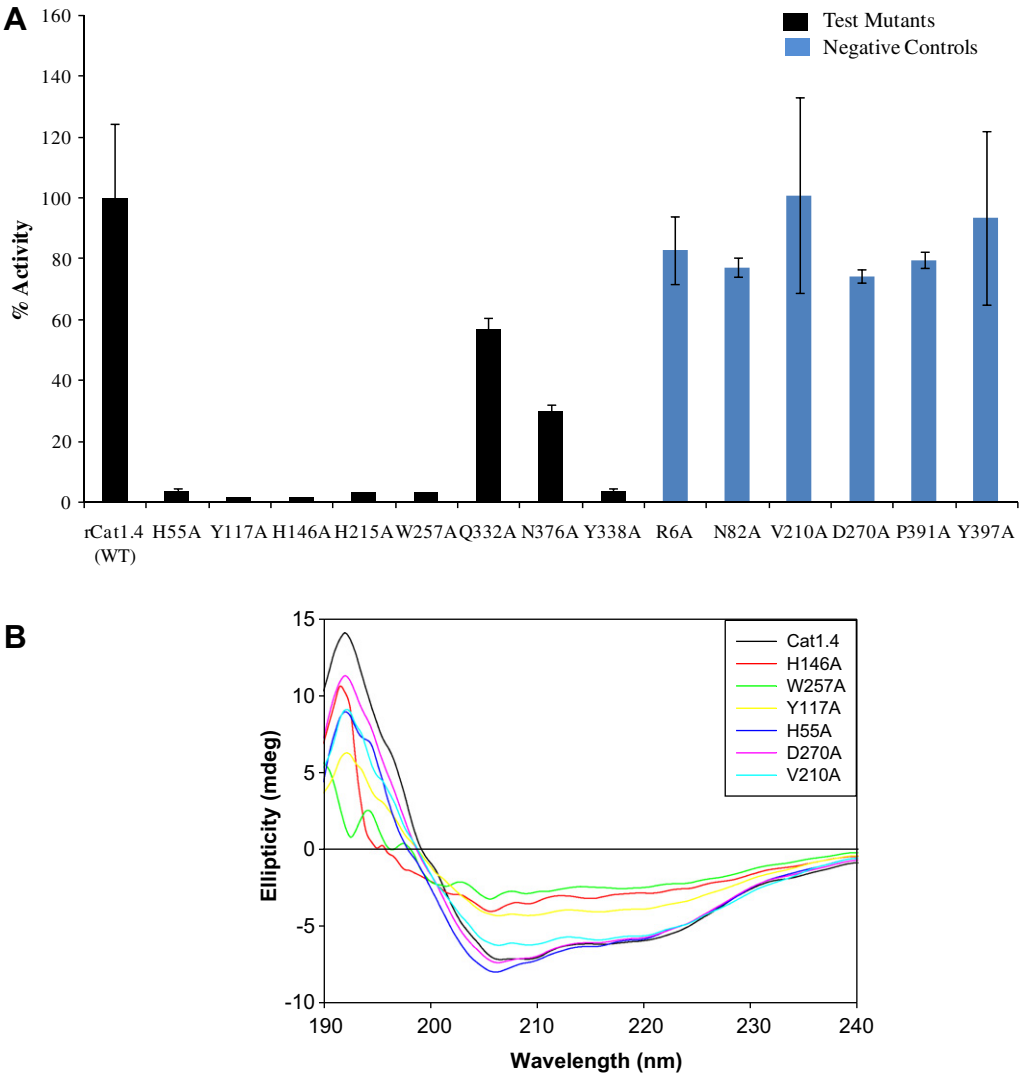

Fig. 4. (A) Enzymatic activity of catalytic variants of Catalase. Bars represent mean ± S.D. for three independent experiments done in triplicates. (B) CD spectra for wild type rCat1.4, H55A, Y117A, H146A, W257A, D270A and V210A.

[32,33]. The clustering removed KatG (Catalase-peroxidases) and Mcat (mitochondrion Catalases) sequences from Catalases to avoid deterioration in the alignment quality. To ensure further improvement in alignment quality, a strategy is employed wherein structural alignment is used as a seed to build the alignment of sequences. The structurally guided sequence alignment thus obtained, contained the known motifs and has better alignment quality. The columns in the alignment were ranked for its importance in the activity by calculating the RE score. RE with appropriate background frequencies, provides a better measure, as conserved positions are under significant evolutionary pressure, which are expected to have amino acid distributions very different from those of columns under no pressure [13]. Thus, this measure increases the scores for aligned columns containing "rare" residues, which are often functionally important. Top 12 high scoring residues were selected from Cat1.4 for mutational analysis (Table 1). We observed that two of the identified catalytic residues, H55 (corresponds to H74 of BLC) and Y338 (Y357 of BLC), lie in the active site (Fig. 3A) at distal and proximal side of heme, respectively, whereas other functional residues were located distantly (>10 Å) from the heme and were not present in the active site (Table 1). The efficiency of RE scoring methodology was also validated by evaluating the potential of low RE scoring residues as negative controls. None of the low RE scoring residues demonstrated appreciable role in catalysis.

Thus, the present study, for the first time, describes the characterization of Catalase (Cat1.4) from *B. anthracis* and identification of crucial residues for Cat1.4 activity. We have identified six novel residues (Q332, Y117, H215, W257, N376 and H146), which were previously not known to be important for the enzyme activity. Upon mutations in identified known crucial residues, lying in the active site (H55 and Y338), there was no change in the secondary structure compared to wildtype whereas identified 3 crucial residues away from active site (Y117, W257 and H146) on mutation perturbed the CD spectra. Despite being distant from the active site and showing loss of activity, H215 mutant did not show significant change in secondary structure. This observation suggests that the residue might be involved in the catalytic mechanism unknown so far. Partial loss in activity of Q332 and N376 and no significant change in secondary structure suggest that the residues might be involved in maintaining the native fold of the protein and a slight change in native fold might lead to partial loss in activity (*Supplementary* Fig. 1). For the negative control mutants (possessing low RE score), neither the activity nor the structure was perturbed as expected (Fig. 4A and B, CD Spectra for all not shown). The current approach, thus, facilitates selection of crucial residues, ones present

in the active site playing role in interaction of substrate directly and others, which are away from active site, but required for acquiring native fold of functional protein. Considering the inter-species variation amongst bacterial and human Catalases, understanding the enzyme mechanism and identification of crucial residues can have far-reaching implications for designing novel, specific and precise therapeutics directed against *B. anthracis*.

Acknowledgments

The financial assistance provided by Indian Council of Medical Research (ICMR), India is duly acknowledged. M. Rehan acknowledges Council of Scientific and Industrial Research, India, for Senior Research fellowship. We also acknowledge Mr. S. Matta, Dr. S. Agarwal and Dr. Shivani Agarwal for critical reading of the manuscript. The assistance extended by Mr. S. Kumar Suman and Mr. K. Gopal in CD experiments is deeply appreciated as well.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2011.06.099.

References

[1] N.H. Bergman, E.C. Anderson, E.E. Swenson, B.K. Janes, N. Fisher, M.M. Niemeyer, A.D. Miyoshi, P.C. Hanna, Transcriptional profiling of *Bacillus anthracis* during infection of host macrophages, Infect Immun. 75 (2007) 3434–3444.
[2] B. Halliwell, Phagocyte-derived reactive species: salvation or suicide?, Trends Biochem Sci. 31 (2006) 509–515.
[3] M. Rehan, H. Younus, Effect of organic solvents on the conformation and interaction of Catalase and antiCatalase antibodies, Int. J. Biol. Macromol. 38 (2006) 289–295.
[4] K. Shatalin, I. Gusarov, E. Avetissova, Y. Shatalina, L.E. McQuade, S.J. Lippard, E. Nudler, *Bacillus anthracis*-derived nitric oxide is essential for pathogen virulence and survival in macrophages, Proc. Natl. Acad. Sci. USA 105 (2008) 1009–1013.
[5] R.J. Cybulski Jr., P. Sanz, F. Alem, S. Stibitz, R.L. Bull, A.D. O'Brien, Four superoxide dismutases contribute to *Bacillus anthracis* virulence and provide spores with redundant protection from oxidative stress, Infect. Immun. 77 (2009) 274–285.
[6] R.W. Seyler Jr., A.O. Henriques, A.J. Ozin, C.P. Moran Jr., Assembly and interactions of cotJ-encoded proteins, constituents of the inner layers of the *Bacillus subtilis* spore coat, Mol. Microbiol. 25 (1997) 955–966.
[7] M.J. Maté, M.S. Sevinc, B. Hu, J. Bujons, J. Bravo, J. Switala, W. Ens, P.C. Loewen, I. Fita, Mutants that alter the covalent structure of Catalase hydroperoxidase II from *Escherichia coli*, J. Biol. Chem. 274 (1999) 27717–27725.
[8] P. Chelikani, I. Fita, P.C. Loewen, Diversity of structures and properties among Catalases, Cell Mol. Life Sci. 6 (2004) 192–208.
[9] P.S. Shenkin, B. Erman, L.D. Mastrandrea, Information-theoretical entropy as a measure of sequence variability, Proteins 11 (1991) 297–313.
[10] W.S. Valdar, Scoring residue conservation, Proteins 48 (2002) 227–241.
[11] T. Cover, J. Thomas, Elements of Information Theory, 2009.
[12] C.E. Shannon, The mathematical theory of communication, 1963, MD Comput. 14 (1997) 306–317.
[13] K. Wang, R. Samudrala, Incorporating background frequency improves entropy-based residue conservation measures, BMC Bioinform. 7 (2006) 385.
[14] K. Kapoor, M. Rehan, A. Kaushiki, R. Pasrija, A.M. Lynn, R. Prasad, Rational mutational analysis of a multidrug MFS transporter CaMdr1p of *Candida albicans* by employing a membrane environment based computational approach, PLoS Comput. Biol. 5 (2009) e1000624.
[15] K. Kapoor, M. Rehan, A.M. Lynn, R. Prasad, Employing information theoretic measures and mutagenesis to identify residues critical for drug-proton antiport function in Mdr1p of *Candida albicans*, PLoS ONE 5 (2010) e11041.
[16] A. Bairoch, The ENZYME database in 2000, Nucleic Acids Res. 28 (2000) 304–305.
[17] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.
[18] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[19] K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, Nucleic Acids Res. 30 (2002) 3059–3066.
[20] Hara, N. Ichise, K. Kojima, H. Kondo, S. Ohgiya, H. Matsuyama, I. Yumoto, Relationship between the size of the bottleneck 15 Å from iron in the main channel and the reactivity of Catalase corresponding to the molecular size of substrates, Biochemistry 46 (2007) 11–22.
[21] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK – a program to check the stereochemical quality of protein structures, J. Appl. Cryst. 26 (1993) 283–291.
[22] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: a message-passing parallel molecular dynamics implementation, Comput. Phys. Commun. 91 (1995) 43–56.
[23] S. Van Dongen, Graph Clustering by Flow Simulation. PhD thesis, University of Utrecht, 2000.
[24] E.J. Alyamani, P. Brandt, J.A. Pena, A.M. Major, J.G. Fox, S. Suerbaum, J. Versalovic, Helicobacter hepaticus Catalase shares surface-predicted epitopes with mammalian Catalases, Microbiology 153 (2007) 1006–1016.
[25] L. Vera-Cabrera, W.M. Johnson, O. Welsh, F.L. Resendiz-Uresti, M.C. Salinas-Carmona, Distribution of a *Nocardia brasiliensis* Catalase gene fragment in members of the genera, Nocardia, Gordona and Rhodococcus, J. Clin. Microbiol. 37 (1999) 1971–1976.
[26] T.J. Reid 3rd, M.R. Murthy, A. Sicignano, N. Tanaka, W.D. Musick, M.G. Rossmann, Structure heme environment of beef liver Catalase at 2.5 Å resolution, Proc. Natl. Acad. Sci. USA 78 (1981) 4767–4771.
[27] P. Loewen, Probing the structure of Catalase HPII of *Escherichia coli* – a review, Gene 179 (1996) 39–44.
[28] Bagyan, L. Casillas-Martinez, P. Setlow, The katX gene, which codes for the Catalase in spores of *Bacillus subtilis*, is a forespore-specific gene controlled by sigmaF, and KatX is essential for hydrogen peroxide resistance of the germinating spore, J. Bacteriol. 180 (1998) 2057–2062.
[29] D.K. Bol, R.E. Yasbin, The isolation, cloning and identification of a vegetative Catalase gene from *Bacillus subtilis*, Gene 109 (1991) 31–37.
[30] S. Engelmann, C. Lindner, M. Hecker, Cloning, nucleotide, sequence and regulation of katE encoding a sigma B-dependent Catalase in *Bacillus subtilis*, J. Bacteriol. 177 (1995) 5598–5605.
[31] L. Frankenberg, M. Brugna, L. Hederstedt, *Enterococcus faecalis* heme-dependent Catalase, J. Bacteriol. 184 (2002) 6351–6356.
[32] M. Bernroitner, M. Zamocky, P.G. Furtmuller, G.A. Peschek, C. Obinger, Occurrence, phylogeny, structure, and function of Catalases and peroxidases in cyanobacteria, J. Exp. Bot. 60 (2009) 423–440.
[33] J. Godocikova, M. Zamocky, M. Buckova, C. Obinger, B. Polek, Molecular diversity of katG genes in the soil bacteria Comamonas, Arch. Microbiol. 192 (2010) 175–184.