Insertion of PrpoD_rpoS Fragment Enhances Expression of Recombinant Protein By Dps Auto-Inducible Promoter in Escherichia Coli

Dina Mulyanti (✉ dina.mulyanti@unisba.ac.id)
  Bandung Institute of Technology; Bandung Islamic University

Sundani Nurono Soewandhi
  Bandung Institute of Technology

Catur Riani
  Bandung Institute of Technology   https://orcid.org/0000-0001-7082-0264

Original Article

Abstract

Background: Nowadays, recombinant therapeutic proteins have been widely produced and consumed. For the safety and effectiveness of the protein production, an auto-inducible expression vector is required to replace inducer interference, which is uneconomic and could be harmful. In this research, an auto-inducible expression plasmid, pCAD$_2$_sod, which was under *dps* (RpoS-dependent gene) promoter control, was modified to provide RpoS at earlier phase, hence accumulate more target protein by generated a new plasmid, pCAD$_2^+$_sod. pCAD$_2$_sod had been constructed to automatically induce the expression of recombinant superoxide dismutase (SOD) from *Staphylococcus equorum* (rMnSODSeq) in the stationary growth phase of *Escherichia coli*. This work aimed to obtain pCAD$_2^+$_sod and determined the expression level of rMnSODSeq.

Method and Results: A synthetic *rpoS* coding region under *rpoD* promoter control (p*rpoD_rpoS*) was inserted to pCAD$_2$_sod and generated pCAD$_2^+$_sod. The rMnSODSeq (24.3 kDa) produced from pCAD$_2^+$_sod was ~1.5 fold higher at 37 °C and more intense at 43 °C compared to that produced from pCAD$_2$_sod, likewise shifted to earlier phase (after 1 h of incubation), as shown in the SDS-PAGE. The dismutase activity was also retained after zymography assay. The mRNA level of *rMnSODSeq* from pCAD$_2^+$_sod, in all growth phases, gave a consistently lower Cq values compared to the one carried pCAD$_2$_sod and so gave higher quantification values relative to *rho* reference gene as well.

Conclusions: The p*rpoD_rpoS* insertion shifts and increases the rMnSODSeq production from stationary to exponential phase, either in the RNA or protein level. The pCAD$_2^+$_sod plasmid has good potential for further recombinant protein productions.

Introduction

The usage of recombinant protein as therapeutic agent has annually increased approximately 35 per cent since 2001. It is a good economical development indication for biopharmaceutical company [1]. In 2018, there are 316 biopharmaceutical products consist of recombinant therapeutic protein and nucleic acid approved in USA and Europe [2]. The major challenges in recombinant protein production are to reduce the production cost, improve the productivity both in upstream and downstream, and obtain high titer while maintaining the quality of the recombinant protein products [1].

Escherichia coli had attested its versatility and economic potential in production [3]. There is a wealth of knowledge and comprehensive tools for *E. coli* systems, such as expression vectors, production strains, protein folding and fermentation technologies that are well tailored for industrial applications. Hence, with its recent advancements, the use of *E. coli* has been a preferred choice and a workhorse not only for expression of non-glycosylated proteins in the biotech industry, but also a complex protein production, bacterial N-linked glycosylation, novel strain engineering and creation of *E. coli* cell-free systems [4]. Varieties of promoters are used in the production of recombinant proteins. Some of the strong promoters

are induced by isopropyl-β-D-thiogalacto-pyranoside (IPTG) (i.e. *tac, trc,* and *lac*), which however has been reported to the limitation of its toxicity for the host cells and also causing an ineffective production cost due to its expensiveness [1, 4–6].

In the previous research, an auto-inducible expression plasmid with *dps* promoter (pCAD*sod*) had been constructed to express *sodA* from *Staphylococcus equorum* encoding a Mn superoxide dismutase (rMnSODSeq) in stationary phase. The *dps* promoter in a high stability medium copy number plasmid, pCAD*sod*, had given a high level expression of rMnSODSeq when bacterial cells entered stationary phase in *E. coli* BL21(DE3) [5]. By using an auto-inducible expression plasmid, the recombinant protein expression will be free of chemical impurity originated from the induction and offers a low-cost production system [5]. Some strategies have been developed to make an auto-induction expression system by modified the growth medium components [7], and using the Self-InducibLe Expression (SILEX) system [6]. Other strategy that offered by this study is by modified the presence of sigma factor needed by the promoter to earlier phase. Hence, promoter can be activated faster, and then target protein accumulates more. pCAD*sod* plasmid carries *dps* promoter which has been known as an RNA polymerase sigma factor subunit S (σ^s, RpoS)-dependent [8, 9].

In *E. coli,* the σ^S or σ^{38} controls the regulon of starvation, while other sigma factors evolved to respond different stressors [10]. Some of the genes induced as part of the RpoS response are truly RpoS specific, while others are also expressed by the vegetative sigma factor, RpoD (σ^{70}), under some specific growth conditions [11]. RpoD is an essential main vegetative sigma factor that controlled approximately more than 2000 genes in *E. coli* that activated by high growth rate [12]. Modifying both sigma factors in one expression system could become a new invention in the field of study regarding recombinant protein overproduction in *E. coli*. The presence of RpoS could automatically induce the activation of *dps* promoter in the constructed expression vector and make the recombinant protein production become more cost-effective. This research aimed to obtain synthetic gene encoding RpoS fused with *rpoD* promoter (p*rpoD_rpoS*) to be inserted into pCAD$_2$_*sod*, generated a modified auto-inducible expression vector, pCAD$_2$$^+$_*sod*, and be determined for the expression of its recombinant gene, *rMnSODSeq*.

Materials And Methods

Bacterial Strain, Plasmid and Culture Medium

Bacteria *E. coli* strain TOP10 was used for cloning and gene expression. Plasmid pCAD$_2$_*sod* is available at the Laboratory of Pharmaceutical Biotechnology, School of Pharmacy, Institut Teknologi Bandung, Bandung, Indonesia, while plasmid pUC*rpoS* were prepared by GenScript, USA with the gene of interest sequences were constructed synthetically. Selective Luria Bertani (LB) agar and liquid medium containing 100 µg ml^{-1} of ampicillin was used for selection, bacterial growth, and protein production.

Plasmids Constructions

Two plasmids, pUC*rpoS* and pCAD$_2$$^+$_*sod*, were constructed and the diagram for the construction is displayed in Fig.1. RpoS sequences consist of 330 amino acids was obtained from NCBI (Acc. Nr: WP_000081550.1) with preferred *E. coli* codons optimization for its highly expression was performed using online program Optimizer [13]. Meanwhile, the terminator sequence of *rpoS* was predicted using online program ARNold [14]. The nucleotide sequences for *rpoD* promoter and its transcription initiation site was taken prior to the *rpoD* Open Reading Frame (ORF) from NCBI (Acc. Nr: NC_012971.2 [3079129..3080970]) which was analyzed using online program BPROM (Softberry inc.) and Fruitfly promoter prediction program [15]. Analyses were done altogether with optimization of the Translation Initiation Rates (TIR) by designing a synthetic Ribosome Binding Site (RBS) using Genome Compiler program (Genome Compiler Corp.). The *rpoS* mRNA secondary structure was predicted using Kinefold [16]. Those complete fusion sequences were flanked by *Hpa*I restriction site sequence in both 3' and 5' ends. The complete synthetic gene sequence carried *rpoS* ORF under *rpoD* promoter control (p*rpoD*_*rpoS*) was inserted into pUC57 plasmid, generated pUC*rpoS* (GenScript, USA).

The pCAD$_2$$^+$_*sod* plasmid was prepared by introducing the p*rpoD*_*rpoS* gene in the pUC*rpoS* into pCAD$_2$_sod using QuickStep PCR cloning method [17] with modification. pCAD$_2$_*sod* was derived from pCAD*sod* that carries *rMnSODSeq* ORF under *dps* promoter control [5]. The pCAD$_2$_*sod* plasmid was also carrying *cer* fragment and *dapD* coding gene (not discussed in this paper). The *Hpa*I recognition site sequences were present in the plasmid in two places flanking *bla* coding gene. The pCAD$_2$$^+$_*sod* plasmid was designed to carry p*rpoD*_*rpoS* gene which was inserted into the *Hpa*I restriction enzyme recognition site near the *dps* promoter sequence.

The modifications introduced on the QuickStep PCR cloning method were on the primer sequences and their molar ratios used to optimize the PCR condition. The first step of PCR to create megaprimers was conducted with following condition: one minute of pre-denaturation step at 95 °C, 30 cycles of a sequence of 10 s at 95 °C, 30 s at 65 °C, 40 s at 72 °C, and finally 2 minutes of post-elongation step at 72 °C. Each primer pair was used in separated PCR mixture with the same condition as mentioned. The PCR reaction mixture consisted of 2.5 µl of 2mM dNTPs, 1.5 µl of 25mM MgSO$_4$, 0.5 µl of KOD-Plus-Neo (Toyobo), 2.5 µL of 10× PCR buffer for KOD-Plus-Neo, 2 ng of pUC*rpoS* as the template, and 10 µM for each primer pair of ForQS/Rev_rpoQ (Asym 1) or RevQS/For_rpoQ (Asym 2), with molar ratio of 10:1 and 25:1, respectively, after optimization. The volume was adjusted to 25 µl with nuclease-free water. After PCR product was confirmed to have an ssDNA form, it was purified and prepared to be used in the second PCR step as the megaprimer. The second step of PCR used the first step PCR product with following condition: 1 minute of pre-denaturation step at 95 °C, 30 cycles of a sequence of 15 s at 95 °C, 1 minute at 62 °C, 3 minutes at 72 °C, and finally 5 minutes of post-elongation step at 72 °C. The PCR reaction mixture consisted of 5 µl of 2mM dNTPs, 3 µl of 25mM MgSO$_4$, 1 µl of KOD-Plus-Neo (Toyobo), 5 µL of 10× PCR buffer for KOD-Plus-Neo, 20 ng of pCAD$_2$_*sod* as the template, and 200 ng of each megaprimer from previous step.

The second step PCR product was directly digested with *Dpn*I (Thermo Scientific) and incubated at 37 °C. After 72 hours incubation, the PCR product was introduced into competent *E. coli* TOP10 using heat-shock method (42 °C; 90 s) and the transformants were selected on an LB agar plate containing ampicillin, overnight at 37 °C. The transformants which grew after incubation were characterized by isolating their plasmids. The characterizations included three different methods: PCR, restriction, and DNA sequencing analyses. In PCR characterization, the isolated plasmid was amplified using pair of primers (F$_{AanIbla}$ and R$_{NotIsod}$) [5] which differentiated the length of PCR product as larger if p*rpoD_rpoS* fragment was successfully inserted into the pCAD$_2$_*sod* plasmid compared to the one from the original plasmid (no insertion). The isolated plasmid was also digested by *Hpa*I (Thermo Scientific) restriction enzyme and the DNA fragment formed after digestion was visualized using agarose gel electrophoresis. The last characterization was DNA sequencing towards particular sequence in the isolated plasmid using primer F$_{Konf}$ and R$_{Konf}$ (Macrogen Inc., Singapore) to confirm the presence of p*rpoD_rpoS* nucleotide sequences in the plasmid. The primers sequences are listed in Table.1.

Overproduction and Zymography Assay of rMnSODSeq

rMnSODSeq was overproduced in each of *E. coli* TOP10 harboring pCAD$_2$_*sod* and pCAD$_2$$^+$_*sod* as a fusion protein with 6×His-tag at its amino terminus. The recombinant bacteria were grown in an Erlenmeyer flask containing LB broth at 37 °C on a rotary shaker at an agitation speed of 150 rpm with a starting OD$_{600}$ around 0.05. To study the auto-inducible expression of rMnSODSeq, the cells were harvested in some time points (after 1, 2, 3, 4, 5, 6, and 24 hours of incubation). The protein expression was checked from 200 ml of culture on each point and cold-centrifuged for 15 minutes at 2851×*g*. The cell pellets were then resuspended, disrupted by ultra-sonication, and characterized using Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) standard protocol as previously reported [18], which were carried out in triplicate. The activity of rMnSODSeq was observed on a 10% native PAGE followed by negative staining zymography analysis with modifications [18]. The presence of a clear zone at the expected band after exposure in the gel is an indication of rMnSODSeq activity.

rMnSODSeq was overproduced as previous experiment above with slight differences on the growth temperature. At first, the recombinant *E. coli* was grown in an Erlenmeyer flask containing LB broth at 37 °C, and then switched to 43 °C after 2 h (OD$_{600}$ ~ 0.1-0.2) to observe the rMnSODSeq expression as an impact of RpoS involvement at the exponential and stationary phases due to heat environmental change.

Quantitative Real-Time PCR (qPCR)

Overnight culture of *E. coli* TOP10 harboring pCAD$_2$_*sod* and pCAD$_2$$^+$_*sod* was each sub-cultured to three new LB broths in Erlenmeyer flasks with ratio of 1:5. Each flask differentiates for the incubation times of 2 (OD$_{600}$ 0.1-0.3), 5 (OD$_{600}$ 0.8-1.0), and 24 h (OD$_{600}$ 1.8-2.0) at 37 °C on a rotary shaker at an agitation speed of 150 rpm with a starting OD$_{600}$ around 0.05, represented *E. coli* growth phase of mid-exponential,

late-exponential, and stationary phase, respectively. The cells were harvested and total RNA from 0.05 gram of the cell pellets were extracted and purified using TRIzol® reagent (Invitrogen) according to the manufacturer's instructions. RNA concentration and purity were confirmed by measuring the ratio of absorbance at 260nm/280nm and 260nm/230nm using µDrop Multiscan GO Microplate spectrophotometer (Thermo Scientific) before and after treated with RNase-free DNase I (Thermo Scientific). One microgram of total RNA was reverse transcribed using random primer in the ReverTra Ace-α® (Toyobo) master mix of 20 µl total volume following manufacturer's instructions.

The qPCR was performed on the CFX96 Thermal Cycler (BioRad) using SensiFAST SYBR NO-ROX master mix (Bioline). Each reaction mixture consisted of 10 µl SensiFAST SYBR NO-ROX master mix, 20µM of each forward and reverse primer, and 1 µl of cDNA template in a total volume of 20 µl. Cycling was performed in following condition: 2 min at 95 °C, followed by 40 cycles of 10 s at 95 °C and 20 s at 60 °C. Melt curve analysis program (65 to 95 °C with an increment of 0.5 °C for 5 s) was added in the end of amplification program to verify the primer specificity. All qPCR assays were carried out using two technical replication and non-template control (NTC), as well as two independent cDNA syntheses. The amplification efficiency of each primer was determined using 2.5-fold dilution series of cDNA synthesized from the isolated RNA of *E. coli* TOP10 pCAD$_2^+$_*sod* in various growth phases. The efficiency was calculated as E = $10^{(-1/B)}$ with B was a linear regression slope in the Y = Bx + A equation of logarithmic cDNA amounts towards the Cq values. The Cq values of the target genes, *rMnSODSeq* and *rpoS,* were measured as well as the reference gene for each experiment. Analysis for changes in the gene expression was quantified by relative quantification (RQ) assay for both target genes which was calculated through $2^{-\Delta\Delta Cq}$ method [19].

Results

Construction of pCAD$_2^+$_*sod*

The pCAD$_2^+$_*sod* was constructed by QuickStep PCR cloning method. In our work, the ratio of the primers was modified and optimized to form an ssDNA carrying a fusion DNA fragment, p*rpoD_rpoS*, which was subsequently used as megaprimers. The presence of p*rpoD_rpoS* DNA fragment was showed by an amplification product using a F$_{AanIbla}$ and R$_{NotIsod}$ primer pair. Five, out of seven, isolated plasmids from the *E. coli* TOP10 transformants gave a PCR product of 2957 basepairs (bp) in size, showed that the DNA fragment (1127 bp) was inserted. Meanwhile, *E. coli* transformants that gave PCR product with smaller size (1830 bp) did not contain modified plasmid. The digestion of the isolated modified plasmid using *Hpa*I restriction enzyme produced three DNA fragments i.e the plasmid backbone, p*rpoD_rpoS* fragment, and *bla* gene (3296, 1121, and 1009bp, respectively) and the sequencing of the plasmid using a F$_{Konf}$ and R$_{Konf}$ primer pair showed that the plasmid carried a correct sequence of the p*rpoD_rpoS* fragment. These data confirmed that pCAD$_2^+$_*sod* was correctly constructed (Suppl.1).

Shift of rMnSODSeq Production into Exponential Phase

The level of rMnSODSeq of 24.3 kDa produced from *E. coli* TOP10 pCAD$_2$$^+$_*sod* after 24 h of incubation at 37 °C was higher about ~1.5 fold compared to that produced from *E. coli* TOP10 pCAD$_2$_*sod* in the SDS-PAGE electrophoregram (Fig. 2A) after analyzed using ImageJ software [20]. The intensity of the protein band was based on a comparison to an *E. coli* protein of 20 kDa band (à in Fig. 2) which was constitutively produced in all conditions under study. The rMnSODSeq produced from *E. coli* TOP10 pCAD$_2$$^+$_*sod* retained its dismutase activity demonstrated by the formation of clear zone on the zymogram (Fig. 2B). In the study of the role of RpoS during exponential phase, the rMnSODSeq protein band has been produced in the exponential phase from *E. coli* TOP10 pCAD$_2$$^+$_*sod* after 1 h incubation, while the one from pCAD$_2$_*sod* after 4 h incubation at 37 °C (Fig. 2C).

Effect of Growth Temperature on rMnSODSeq Overproduction

In the exponential phase (4 h) at 43 °C, the intensity of rMnSODSeq protein band produced from *E. coli* Top10 pCAD$_2$$^+$_*sod* was more intense compared to that from *E. coli* Top10 pCAD$_2$_*sod*. The protein band was also more intense compared to the band produced at 37 °C. As previously mentioned, a protein of 20 kDa was used in the comparison of protein band intensity. Meanwhile, the comparison of the rMnSODSeq protein band in the stationary phase (18 h) showed that the intensity of the protein band from both recombinant clones was quite similar. Though based on comparing to the 20 kDa protein band (à in Fig. 2), the intensity of protein bands at 43 °C tended to be more intense compared to that at 37 °C (Fig. 2D).

Quantitative Real-Time PCR (qPCR)

The Cq values trend of *rMnSODSeq* synthesis cDNA from *E. coli* TOP10 pCAD$_2$$^+$_*sod*, in all phases, were found consistently lower to the one carried pCAD$_2$_*sod*, therefore gave higher values of the relative quantification (RQ) as well (Fig. 3). Unfortunately, through the biological variations of the samples, the standard deviation (SD) of the RQ value was large among every experiment even though the enhancement of the data trend was consistent (Table. 2). A huge differentiation on the RQ value between the duplication independent cDNA was also depending on the relative calculation towards the reference gene.

In the calculation of amplification efficiency for each reference gene candidate, *recA* and *rho* gave E = 1.98 and 1.97, respectively, while *gyrA* gave E = 2.72 which was higher than the acceptable amplification efficiencies in range of 1.85-2.05 [21] (Table. 1). The Cq data trend of *recA* was inconsistent in the stationary phase among three different samples, in contrast to the Cq value of *rho*. Despite the consistent Cq of *gyrA* in various growth phase, too high efficiency showed that the use of *gyrA* as reference gene needed more optimization in the experimental condition applied. Thus, *rho* was chosen as the reference gene in this study (Suppl. 2). Using 2$^{-\Delta\Delta Cq}$ method, the expression of *rMnSODSeq* relative to *rho* gene from *E. coli* TOP10 pCAD$_2$$^+$_*sod* showed that the gene was expressed in 752.2; 3206.8; and 21334.8 fold

towards the non-recombinant one as well as higher compared to the one harboring pCAD$_2$_*sod* in mid-exponential, late-exponential, and stationary phase, respectively (Table. 2).

Discussion

In our previous work, we constructed a recombinant plasmid, pCAD$_2$_*sod*, where the expression of *rMnSODSeq* gene was regulated under *dps* promoter, hence the rMnSODSeq was produced at the stationary phase [5]. In stationary phase, cells enter a period of no growth since the metabolism-linked genes are turned off [22], though it provides high cell densities for increased product formation. However, low growth rates and protease activity brought on by depleted nutrient levels can reduce the yield of foreign protein in stationary phase [23].

Plasmid pCAD$_2$_*sod* is equipped with RpoS-dependent *dps* promoter to express rMnSODSeq in the stationary phase. The presence of RpoS in exponential phase by default is low, since its translation is shut off even if its transcription occurs, and during normal conditions the RpoS is rapidly degraded [11]. The changes in RpoS level during growth phase play an important role in providing differential expression of RpoS-dependent gene expression [24]. The use of *dps* promoter in achieving high level protein production in *E. coli* has been shown by involving H$_2$O$_2$ induction [25]. By inclusion of additional RpoS from the pCAD$_2$$^+$_*sod* plasmid at the exponential growth phase in this work, it was as expected that activation of *dps* promoter and improvement of the auto-inducible production of rMnSODSeq occurred at earlier phase.

RpoD and RpoS are both sigma factor subunits of RNA polymerase that are able to be effectively competed [11]. RpoD is known to be a constitutive RNA polymerase works in the early phase of bacterial growth [26], while RpoS is an alternative sigma factor produced by many Gram-negative bacteria and primarily controls genes which are expressed in stationary phase in response to nutrient deprivation [27]. The addition of the p*rpoD*_*rpoS* fragment was intended to provide recombinant RpoS, which hypothetically shifts the expression of rMnSODSeq coding region in the logarithmic phase by affecting the activity of the stationary phase *dps* promoter, in the overproduction process. Moreover, the alteration of initial sequence of the *rpoS* will reduce the involvement of negative regulators on the transcription level in *rpoS* expression at the exponential phase.

In our work, the production of additional recombinant RpoS from the pCAD$_2$$^+$_*sod* plasmid in the system switched the expression of rMnSODSeq coding region to exponential phase through the earlier activation of stationary *dps* promoter. To study the role of RpoS in the current expression system, the overproduction incubation temperature was switched from 37 °C to 43 °C after the culture reach the mid-exponential phase. The temperature was maintained at 37 °C from the beginning to allow the activation of the *rpoD* promoter. The result showed that higher incubation temperature affected the expression of rMnSODSeq in both exponential and stationary phase. This present work reports the role of synthetic *rpoS* coding region under *rpoD* promoter regulation in pCAD$_2$_*sod* on the shift of gene expression from

stationary to exponential phase and the level expression of gene product. Nevertheless, further study for H_2O_2 induction in pCAD$_2^+$_sod plasmid is worth to be done.

Battesti, et., al (2011) described several RpoS regulons regarding stress inductions and each of its effector genes. Stresses related to dps include DNA damage by UV, high temperature, low pH, high pH, and oxidative stress; but only the first two stresses were related to the regulation level of mRNA stability, translation, and degradation of RpoS. Related to heat-shock, in a high temperature at 43.5 °C, the level of RpoS protein was found increase 20 fold associated with the inactivation of the essential endonuclease RNase E along with a detection of 3 fold increasing in rpoS mRNA levels as well [28]. The condition relate to the effect of higher incubation temperature at 43 °C in this research. The production of RpoS is highly regulated at the transcriptional, translational, and posttranslational levels [29]. The transcriptional level could be set aside, since the native rpoS promoter was replaced by rpoD in prpoD_rpoS sequence. Meanwhile, in translational regulation, the RBS of native rpoS needs some sRNAs, such as DsrA, RprA, and ArcZ, as positive regulators [30, 31]. Each of these sRNAs requires RNA chaperone Hfq to base-pairing with the 5'UTR of the rpoS mRNA and opening a stem-loop that represses translation initiation in various conditions [29, 31]. This complexity of translational regulation was simplified in the recombinant rpoS mRNA secondary structure by using synthetic RBS sequences, which have the same TIR as rpoD promoter.

Our prediction using in silico study showed that the native rpoS mRNA secondary structure had a high free-energy of -50.2 kcal/mol, while the recombinant version gave less negative free-energy of -10.7 kcal/mol (Fig. 1). In such highly spontaneus reaction which might be formed from the native rpoS mRNA has put the position of RBS and start codon were closed and defined why in nature it needed some components, such as sRNAs, as positive regulators, whilst the recombinant rpoS mRNA has an open formation. Hence, the recombinant RpoS expression could also aside the regulation in the translational level based on the prediction. The fact that rpoS coding sequence was inserted into pCAD$_2$_sod, a medium copy number plasmid, would stabilize the gene expression since the stability of medium copy number plasmid was significantly higher than that of high copy number [5]. Nevertheless, the stability of the rpoS mRNA and the degradation of the RpoS are still a concern at exponential growth; thereby the focus of this research was to minimize the role of negative regulators related to this.

The expression profile of rMnSODSeq mRNA in different growth phases of E. coli from the host harboring each of pCAD$_2$_sod and pCAD$_2^+$_sod described the effect of the prpoD_rpoS insertion in the expression system. A reference gene is critical in normalizing cellular mRNA data. It should be stably expressed and appropriate for normalization under the experimental conditions described [32]. For transcriptional study in E. coli, the choice of reliable reference genes has not been systematically validated [33]. A study from Zhou, et.,al. (2011) identified that cysG was found as one of reliable novel reference gene for transcription analysis in recombinant protein producing E. coli. However, the use of cysG in this research was found unsuitable, since all the Cq values in various amounts of cDNA were approximating the NTC Cq values. This finding most probably caused by the state of cysG expression in the conducted experimental

condition was not fit. The *cysG* gene was found the least stable during stress condition compared to *rssA* and *hcaT* gene [34]. Candidates of reference genes in this research should be suitable upon the experimental design which was stable in expression throughout the growth phase of *E. coli*. Unfortunately, there was lack of information regarding *E. coli* reference gene that suitable with this experimental design. Thus, functional categories of genes such as DNA replication and transcription had become a consideration in the selection of reference gene candidates, i.e *recA* and *gyrA* (DNA replication) and *rho* (transcription). Those genes were considered to had most stable expression across different conditions tested in other bacteria, i.e. *Klebsiella pneumoniae* [35].

The qPCR results indicated that the insertion of p*rpoD_rpoS* into the plasmid has affected the shift and enhancement of rMnSODSeq expression in *E. coli.* The expression of *rMnSODSeq* in the RNA level was in favor to its protein profile in correspondingly host growth curve. Nevertheless, despite the differentiation of the regulation between recombinant RpoS under *rpoD* promoter control to the native has enhanced the *rMnSODSeq* RQ value, which also showed in the protein band on the SDS-PAGE, the RQ value of *rpoS* mRNA was found inconsistent (Fig. 3). Since the primer used in this study to quantify the *rpoS* not specifically quantified the recombinant one, there was presumably another mechanism that causing the *rMnSODSeq* RQ enhancement. To ensure the factors that associated to the rMnSODSeq expression enhancement, further study needs to be conducted, i.e microarray or transcriptome profiling [36–38].

This work concluded that insertion of a coding sequence of *rpoS* gene which expression directed by *rpoD* promoter shifts the rMnSODSeq overproduction from stationary to exponential phase and increase its level of production in *E. coli*, either in the RNA or protein level. This *E. coli* expression system has good potential for further recombinant protein productions. To have an overall study to the gene expression profile of the recombinant and non-recombinant *E. coli* host using microarray or transcriptome profiling would be an advantage to clarify the role of RpoS through rMnSODSeq expression enhancement. The study on other recombinant protein model substituted *rMnSODSeq* ORF in the modified plasmid is our future plan to give additional insight upon the behavior of the constructed expression system to another recombinant protein profile.

Declarations

Ethical Approval

The ethical approval is not applicable, since it did not use animal nor human participants.

Authors Contributions

DM performed experiments, analyzed the data, and wrote the paper.

SNS supervised and analyzed the data.

CR supervised, analyzed the data, and wrote the paper.

Funding

Conflict of interest

The authors declare that they have no conflict of interest.

Availability of data and materials

Berkeley Drosophila Genome Project, RRID:SCR_013094

GenomeCompiler, RRID:SCR_013988

ImageJ, RRID:SCR_003070

NCBI, RRID:SCR_006472

SoftBerry, RRID:SCR_000902

Acknowledgement

References

1. Dabhole MP (2014) Recent Innovations in Therapeutic Recombinant Protein. Innov Recent Protein, Ther Recomb Pharma Bio World 10−18

2. Walsh G (2018) Biopharmaceutical benchmarks 2018. Nat Biotechnol 36:1136−1145. https://doi.org/10.1038/nbt.4305

3. Swartz JR (2001) Advances in *Escherichia coli* production of therapeutic proteins. Curr Opin Biotechnol 12:195−201. https://doi.org/10.1016/S0958-1669(00)00199-3

4. Huang CJ, Lin H, Yang X (2012) Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. J Ind Microbiol Biotechnol 39:383−399. https://doi.org/10.1007/s10295-011-1082-9

5. Retnoningrum DS, Santika IWM, Kesuma S, Ekowati SA, Riani C (2019) Construction and Characterization of a Medium Copy Number Expression Vector Carrying Auto-Inducible *dps* Promoter to Overproduce a Bacterial Superoxide Dismutase in *Escherichia coli*. Mol Biotechnol 61:231−240. https://doi.org/10.1007/s12033-018-00151-5

6. Briand L, Marcion G, Kriznik A, Heydel JM, Artur Y, Garrido C, Seigneuric R, Neiers F (2016) A self-inducible heterologous protein expression system in *Escherichia coli*. Sci Rep 6:1–11. https://doi.org/10.1038/srep33037

7. Studier FW (2005) Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 41:207–234. https://doi.org/10.1016/j.pep.2005.01.016

8. Lacour S, Landini P (2004) S-Dependent Gene Expression at the Onset of Stationary Phase in *Escherichia coli*: Function of S-Dependent Genes and Identification of Their Promoter Sequences. Society 186:7186–7195. https://doi.org/10.1128/JB.186.21.7186

9. Shimada T, Makinoshima H, Ogawa Y, Miki T, Maeda M, Ishihama A (2004) Classification and strength measurement of stationary-phase promoters by use of a newly developed promoter cloning vector. J Bacteriol 186:7112–7122. https://doi.org/10.1128/JB.186.21.7112-7122.2004

10. Loewen PC, Hu B, Strutinsky J, Sparling R (1998) Regulation in the *rpoS* regulon of *Escherichia coli*. Can. J. Microbiol. 44:707–717

11. Battesti A, Majdalani N, Gottesman S (2011) The RpoS-mediated general stress response in *Escherichia coli*. Annu. Rev. Microbiol. 65:189–213

12. Schellhorn HE (2014) Elucidating the function of the RpoS regulon. Future Microbiol 9:497–507

13. Puigbò P, Guzmán E, Romeu A, Garcia-Vallvé S (2007) OPTIMIZER: A web server for optimizing the codon usage of DNA sequences. Nucleic Acids Res 35:. https://doi.org/10.1093/nar/gkm219

14. Naville M, Ghuillot-Gaudeffroy A, Marchais A, Gautheret D (2011) ARNold: A web tool for the prediction of Rho-independent transcription terminators. RNA Biol 8:11–13. https://doi.org/10.4161/rna.8.1.13346

15. Reese MG (2001) Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. Comput Chem 26:51–56. https://doi.org/10.1016/S0097-8485(01)00099-7

16. Xayaphoummine A, Bucher T, Isambert H (2005) Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. Nucleic Acids Res 33:605–610. https://doi.org/10.1093/nar/gki447

17. Jajesniak P, Wong TS (2015) QuickStep-Cloning: A sequence-independent, ligation-free method for rapid construction of recombinant plasmids. J Biol Eng 9:1 https://doi.org/10.1186/s13036-015-0010-3

18. Retnoningrum DS, Rahayu AP, Mulyanti D, Dita A, Valerius O, Ismaya WT (2016) Unique Characteristics of Recombinant Hybrid Manganese Superoxide Dismutase from *Staphylococcus equorum* and *S. saprophyticus*. Protein J 35:2. https://doi.org/10.1007/s10930-016-9650-5

19. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. Methods 25:402–408. https://doi.org/10.1006/meth.2001.1262

20. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of Image Analysis. Nat Methods 9:671–675. https://doi.org/10.1007/978-1-84882-087-6_9

21. Dorak MT (2006) Real-time PCR. Taylor & Francis Group, Newcastle-upon-Tyne, UK

22. Jaishankar J, Srivastava P (2017) Molecular basis of stationary phase survival and applications. Front. Microbiol. 8:1–12

23. Shojaosadati SA, Kolaei SMV, Babaeipour V, Farnoud AM (2008) Recent advances in high cell density cultivation for production of recombinant protein. Iran J Biotechnol 6:63–84

24. Zhou Y, Gottesman S (1998) Regulation of Proteolysis of the Stationary-Phase Sigma Factor RpoS. J.bacteriol 180:1154–1158

25. Sethia PP, Rao KK, Noronha SB (2011) dps': An H_2O_2 Inducible Promoter for High Level Protein Production in *Escherichia coli*. Int Conf Environ Biosci IPCBEE 21:110–114

26. Shimada T, Yamazaki Y, Tanaka K, Ishihama A (2014) The whole set of constitutive promoters recognized by RNA polymerase RpoD holoenzyme of *Escherichia coli*. PLoS One 9:e90447. https://doi.org/10.1371/journal.pone.0090447

27. Chen G, Schellhorn HE (2003) Controlled induction of the RpoS regulon in *Escherichia coli*, using an RpoS-expressing plasmid. Can J Microbiol 49:733–740. https://doi.org/10.1139/w03-096

28. McCullen CA, Benhammou JN, Majdalani N, Gottesman S (2010) Mechanism of positive regulation by DsrA and RprA small noncoding RNAs: Pairing increases translation and protects *rpoS* mRNA from degradation. J Bacteriol 192:5559–5571. https://doi.org/10.1128/JB.00464-10

29. Battesti A, Majdalani N, Gottesman S (2015) Stress sigma factor RpoS degradation and translation are sensitive to the state of central metabolism. Proc Natl Acad Sci 112:5159–5164. https://doi.org/10.1073/pnas.1504639112

30. Sledjeski DD, Gupta A, Gottesman S (1996) The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in *Escherichia coli*. EMBO J 15:3993–4000. https://doi.org/10.1002/j.1460-2075.1996.tb00773.x

31. Soper TJ, Woodson SA (2008) The *rpoS* mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. RNA 14:1907–1917. https://doi.org/10.1261/rna.1110608

32. Bustin SA, Benes V, Garson JA, Hellemans J, Hugget J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT (2009) The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55:611–622. https://doi.org/10.1373/clinchem.2008.112797

33. Zhou K, Zhou L, Lim Q, Zou R, Stephanopoulos G, Too HP (2011) Novel reference genes for quantifying transcriptional responses of *Escherichia coli* to protein overexpression by quantitative PCR. BMC Mol Biol 12:18. https://doi.org/10.1186/1471-2199-12-18

34. Peng S, Stephan R, Hummerjohann J, Tasara T (2014) Evaluation of three reference genes of *Escherichia coli* for mRNA expression level normalization in view of salt and organic acid stress exposure in food. FEMS Microbiol Lett 355:78–82. https://doi.org/10.1111/1574-6968.12447

35. Gomes AÉI, Stuchi LP, Siqueira NMG, et al (2018) Selection and validation of reference genes for gene expression studies in *Klebsiella pneumoniae* using Reverse Transcription Quantitative real-time

PCR. Sci Rep 8:1–14. https://doi.org/10.1038/s41598-018-27420-2

36. Bumgarner R (2013) DNA microarrays: Types, Applications, and their future. Curr Protoc Mol Biol 1–17. https://doi.org/10.1002/0471142727.mb2201s101

37. Hung S, Baldi P, Hatfield GW (2002) Global Gene Expression Profiling in *Escherichia coli* K12. J Biol Chem 277:40309–40323. https://doi.org/10.1074/jbc.M204044200

38. Franchini AG, Ihssen J, Egli T (2015) Effect of global regulators RpoS and cyclic-AMP/CRP on the catabolome and transcriptome of *Escherichia coli* K12 during carbon- and energy-limited growth. PLoS One 10:1–25. https://doi.org/10.1371/journal.pone.0133793

Tables

Table 1 Primers sequences used in this research

QuickStep Method and Characterization		
Primer Name	Primer Sequences (5' à 3')	References
ForQS	AGG GTT ATT GTC TCA TGA GCG GGT TAA CGC ACT GAG AGG CAG CGG CAA ATA	This research
RevQS	GGC TAT GTG TTC CGC TAT TCT GGC TGT TAA CGC ACA GAA AAG GCC AGC CT	This research
For_rpoQ	GCA CTG AGA GGC AGC GGC AAA TA	This research
Rev_rpoQ	GTT AAC GCA CAG AAA AGG CCA GCC T	This research
F$_{AanIbla}$	GGC CGC GGT TAT AAT CTT CAC CTA GAT CCT TT	[5]
R$_{NotIsod}$	GGA TTG GCG GCC GCA GGT ACG AAC TC	[5]
F$_{Konf}$	GGG CGA AAA CTC TCA AGG A	Pharmaceutical Biotechnology Lab, ITB
R$_{Konf}$	AAA CAG GCT GTG GTT CAA GTG	Pharmaceutical Biotechnology Lab, ITB
qPCR		
Primer Name	Primer Sequences (5' à 3')	Efficiencies*) References
P_F$_{SOD}$	TTG GAC AGC GTT CCG GAG AAC A	1.99 This research
P_R$_{SOD}$	CCT TTT TCT TCG CTA TTC GGG GTC A	
P_F$_{rpoS}$	CGT CAG CCG TAT GCT TCG TCT T	2.05 This research
P_R$_{rpoS}$	CTG CTT CAT ATC GTC ATC TTG CGT GG	
P_F$_{rho}$	TAA CTT CGA CAA ACC TGA AAA	1.97 This research
P_R$_{rho}$	ACC GTT ACC ACG TTC CAT AC	
P_F$_{recA}$	TTA AAC AGG CTG AAT TCC AG	1.98 This research
P_R$_{recA}$	CTG CTT TCT CGA TCA GCT TC	
P_F$_{gyrA}$	GTG ACC CGT CGT ACT ATT TT	2.72 This research
P_R$_{gyrA}$	GAT GAT CGG GTC GAT GTT CG	
P_F$_{cysG}$	TTG TCG GCG GTG GTG ATG TC	NA**) [33]
P_R$_{cysG}$	ATG CGG TGA ACT GTG GAA TAA ACG	

Table 2 *rMnSODSeq* and *rpoS* Cq and RQ values

Sample E. coli TOP10	Mid-Exponential (2 h)		Late-Exponential (5 h)		Stationary (24 h)	
	Cq*)	RQ**)	Cq*)	RQ**)	Cq*)	RQ**)
(*rMnSODSeq*)						
Non-recombinant	28.22 ± 1.16	1.0	26.82 ± 0.99	1.0	26.02 ± 1.11	1.0
pCAD$_2$_sod	19.24 ± 2.83	250.5 ± 179.4	17.56 ± 1.03	226.7 ± 86.7	19.03 ± 5.31	443.2 ± 102.3
pCAD$_2$$^+$_sod	18.01 ± 1.20	752.2 ± 364.4	16.57 ± 1.79	3206.8 ± 1870.0	17.77 ± 2.17	21334.8 ± 18981.3
(*rpoS*)						
Non-recombinant	24.61 ± 5.01	1.0	22.19 ± 3.39	1.0	23.37 ± 2.17	1.0
pCAD$_2$_sod	21.91 ± 0.62	15.6 ± 21.7	23.56 ± 1.59	0.1 ± 0.1	26.46 ± 3.05	0.4 ± 0.2
pCAD$_2$$^+$_sod	21.63 ± 3.67	7.4 ± 8.1	21.42 ± 0.27	4.2 ± 1.6	23.76 ± 2.08	43.0 ± 15.8

*) Cq mean was measured from two independent cDNAs

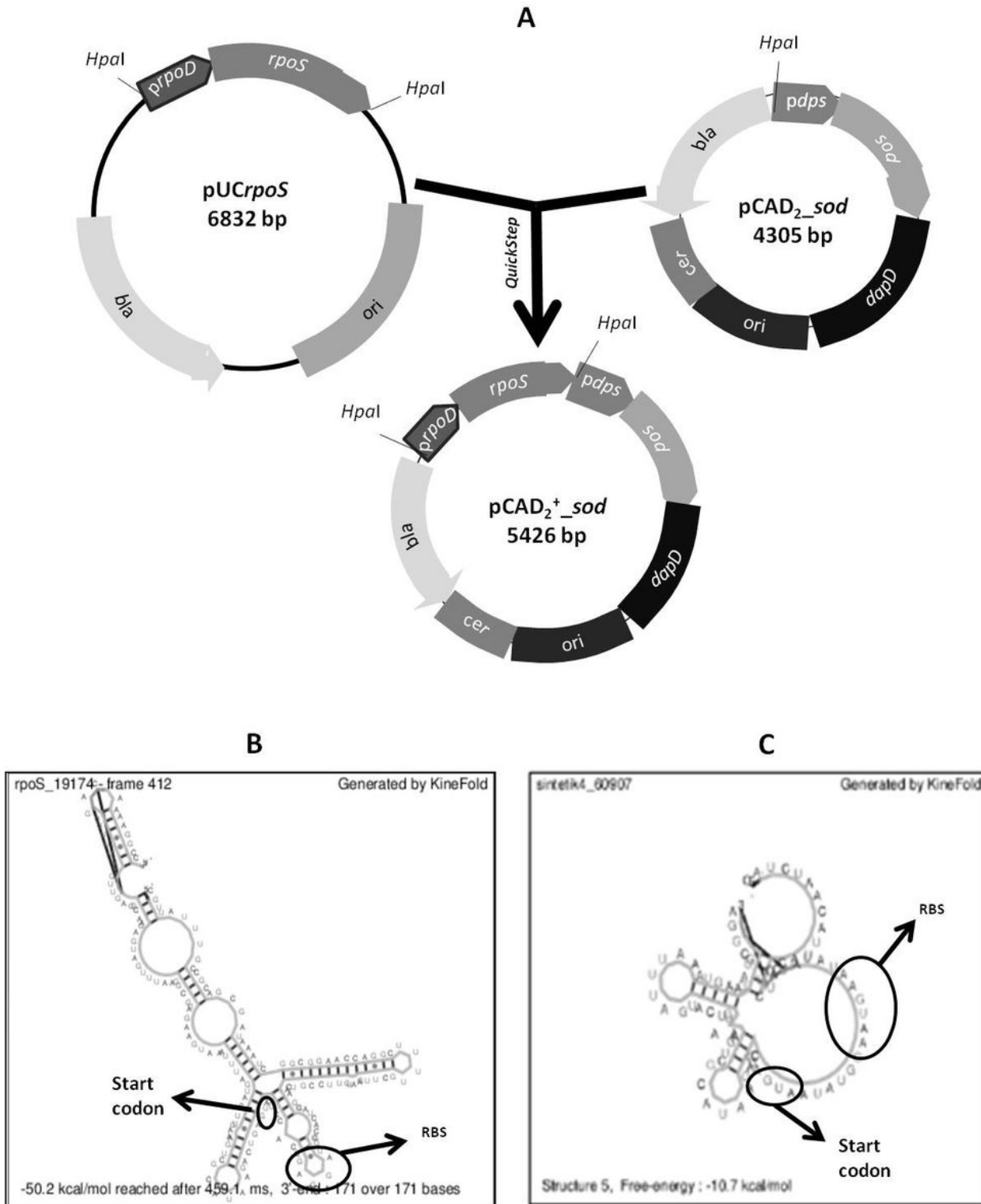**) Quantified using $2^{-\Delta\Delta Cq}$ method; relative to *rho* reference gene

Figures

Figure 1

Illustration of Construction of Recombinant pCAD2+_sod and mRNA secondary structure prediction of native and recombinant rpoS. a QuickStep cloning method was conducted by involving pUCrpoS and pCAD2_sod. The prpoD_rpoS fragment was amplified and inserted into pCAD2_sod to generate new plasmid, pCAD2+_sod. b In the native rpoS mRNA secondary structure, with a high free-energy (-50.2 kcal/mol), loops formation caused the RBS and start codon were closed. c Meanwhile, the recombinant

rpoS mRNA secondary structure with the highest free-energy less negative (-10.7 kcal/mol) gave an opened RBS and start codon. The in silico prediction was conducted using Kinefold [16]
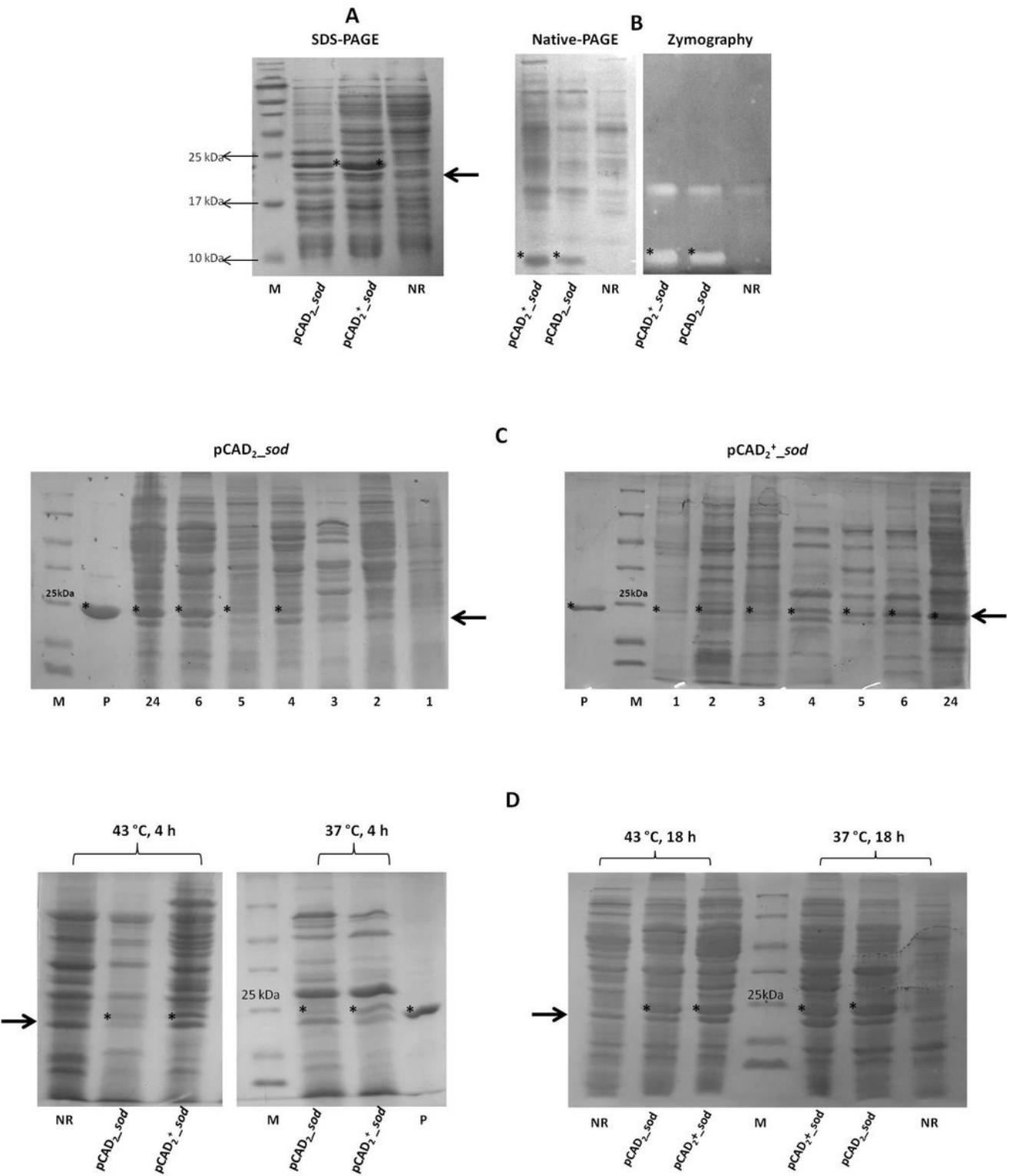

Figure 2

rMnSODSeq expression profile in E. coli TOP10. a The expression of rMnSODSeq (*), 24.3 kDa, from pCAD2+_sod after 24 h at 37 °C showed ~1.5 times more intense protein band compared to rMnSODSeq from pCAD2_sod, while the non recombinant cells showed no expression of the recombinant protein. The

intensity of the band was compared to a reference E. coli protein band of 20 kDa (←) using ImageJ [20]. b Zymography analysis after Native-PAGE running gels showed that the dismutase activity of rMnSODSeq from both different recombinant clones was retained. c In earlier hours of incubation (1-6 h) at the exponential phase, the rMnSODSeq protein bands from pCAD2+_sod had appeared since the first hour of incubation, while from pCAD2_sod it was seen after 4 hours of incubation. d The rMnSODSeq protein bands after incubation at 43 °C from pCAD2+_sod showed a more intense band compared to that from pCAD2_sod in the mid-exponential phase (4 h). Comparison of the rMnSODSeq protein band in the stationary phase (18 h) showed that the intensity of the protein bands from both recombinant clones were quite similar. Those bands were tended to be denser compared to the bands at 37 °C when referred to another E. coli protein band of 20 kDa as reference (→).
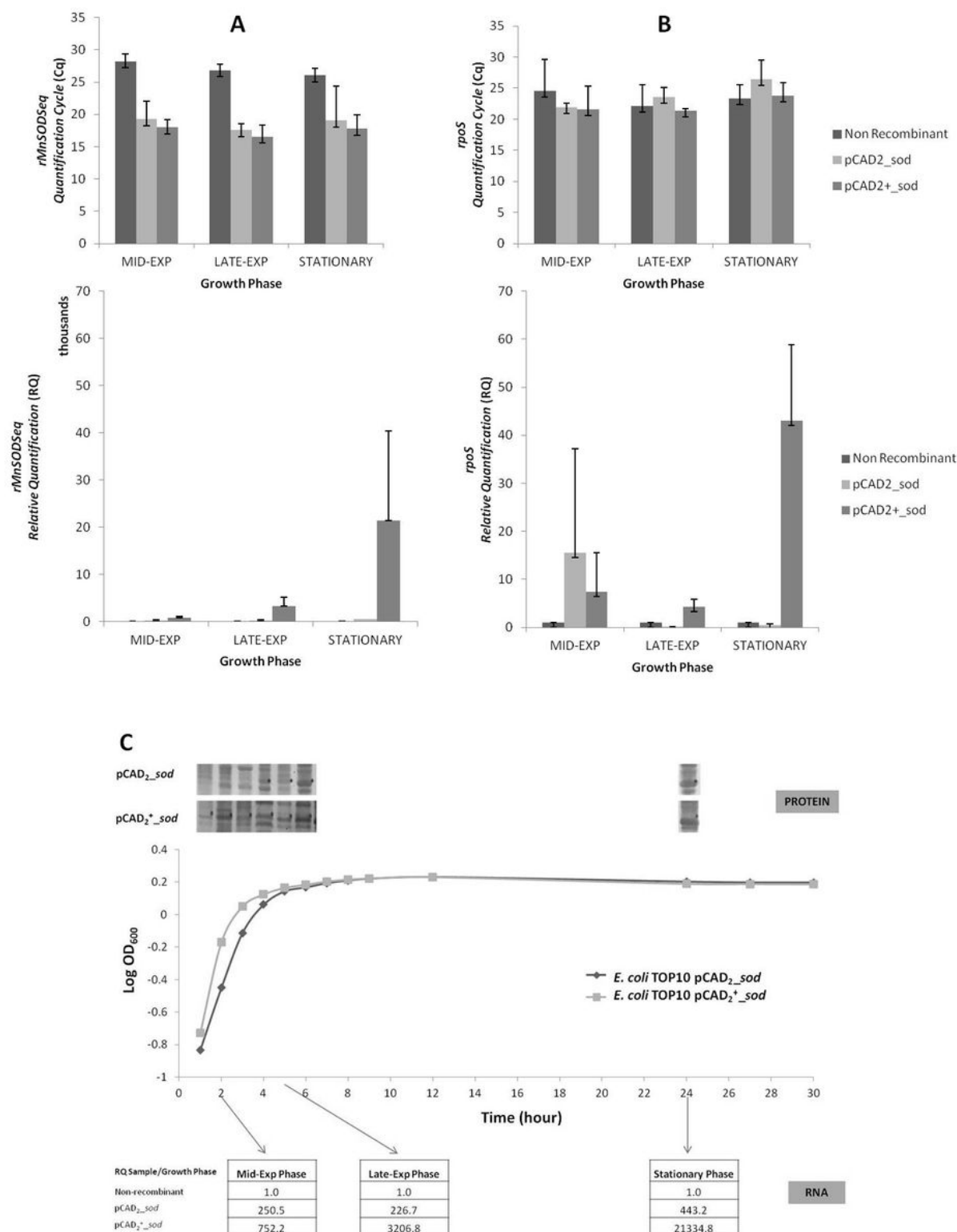
Figure 3

qPCR analysis and summary of rMnSODSeq expression profile in E. coli TOP10. In all growth phases, i.e. mid-exponential, late-exponential, and stationary phase, the Cq and RQ gave the same trend value to rMnSODSeq from all samples. Cq was measured from two independent cDNAs. a The Cq values order were descending from E. coli TOP10 non-recombinant, pCAD2_sod, and pCAD2+_sod, respectively. Thus, it showed a higher rMnSODSeq mRNA expression in the host harboring pCAD2+_sod compared to others.

In the quantification of the rMnSODSeq gene expression relative to rho reference gene, the RQ values were ascending from the mid-exponential, late-exponential, and stationary phase, respectively. Thus, the expression of rMnSODSeq from E. coli TOP10 pCAD2+_sod was found switched to earlier phase of growth and higher in all phases compared to other samples. b Meanwhile, the Cq and RQ values of rpoS were found inconsistent. c In summary; at the protein level, the rMnSODSeq protein band from pCAD2+_sod was expressed earlier in the exponential growth phase (after 1 h of incubation) compared to the one from pCAD2_sod (after 4 h of incubation). The result was in line to the RNA level which showed an enhancement in the RQ value of both samples, either between samples at the same growth phase or among the growth phases.

Supplementary Files

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

- Supplementary.docx