CrossMark

Molecular Cloning, Structural Modeling and the Production of Soluble Triple-Mutated Diphtheria Toxoid (K51E/G52E/E148K) Co-expressed with Molecular Chaperones in Recombinant *Escherichia coli*

Naphatsamon Uthailak[1,2] · Pornpimol Mahamad[1,2] · Pamorn Chittavanich[3] ·
Somchai Yanarojana[3] · Wassana Wijagkanalan[4] · Jean Petre[4] ·
Watanalai Panbangred[1,2]

Abstract CRM197 is a diphtheria toxin (DT) mutant (G52E) which has been used as a carrier protein for conjugate vaccines. However, it still possesses cytotoxicity toward mammalian cells. The goal of this project was to produce a non-toxic and soluble CRM197EK through introduction of triple amino acid substitutions (K51E/G52E/E148K) in *Escherichia coli*. The expression of CRM197EKTrxHis was optimized and co-expressed with different molecular chaperones. The soluble CRM197EKTrxHis was produced at a high concentration (97.33 ± 17.47 μg/ml) under the optimal condition (induction with 0.1 mM IPTG at 20 °C for 24 h). Cells containing pG-Tf2, expressing trigger factor and GroEL-GroES, accumulated the highest amount of soluble CRM197EKTrxHis at 111.24 ± 10.40 μg/ml after induction for 24 h at 20 °C. The soluble CRM197EKTrxHis still possesses nuclease activity and completely digest λDNA at 25 and 37 °C with 8- and 4-h incubation, respectively. Molecular modeling of diphtheria toxin, CRM197 and CRM197EK indicated that substitutions of two amino acids (K51E/E148K) may cause poor NAD binding, consistent with the lack of toxicity. Therefore, CRM197EK might be used as a new potential carrier protein. However, further in vivo study is required to confirm its roles as functional carrier protein in conjugate vaccines.

Introduction

Diphtheria toxin (DT) is a single polypeptide chain, with the molecular mass approximately 62 kDa, produced from toxigenic strains of *Corynebacterium diphtheriae* [1]. DT is comprised of two fragments bound together by a disulfide bridge between fragment A at its *N* terminus and fragment B at its *C* terminus. Fragment A, containing the catalytic site of diphtheria toxin, catalyzes ADP-ribosylation and inactivates elongation factor 2 (EF-2). Fragment A binds to nicotinamide adenine dinucleotide (NAD), a substrate of DT, and transfers the ADP-ribosyl group of NAD to the modified histidine 715 of EF-2, resulting in the inhibition of protein synthesis in eukaryotic cells. In contrast, fragment B does not exhibit catalytic activity and instead plays a role in translocation of fragment A into cells [2, 3]. Many natural variants of DT have been isolated with reduced toxicity but capable of stimulating the immunogenicity and are designated as cross-reacting materials (CRMs). The cross-reacting material 197 (CRM197) contains a single amino acid substitution at position 52 from glycine to glutamic acid (G52E) and its toxicity is substantially reduced relative to unaltered DT. Therefore, CRM197 has been utilized as a suitable mutant for vaccine

✉ Watanalai Panbangred
Watanalai.pan@mahidol.ac.th

1 Department of Biotechnology, Faculty of Science, Mahidol University, 272 Rama VI Road, Ratchathewi, Bangkok 10400, Thailand

2 Mahidol University - Osaka University Collaborative Research Center for Bioscience and Biotechnology (MU-OU: CRC), Faculty of Science, Mahidol University, 272 Rama VI Road, Ratchathewi, Bangkok 10400, Thailand

3 Department of Pharmacology, Faculty of Science, Mahidol University, 272 Rama VI Road, Ratchathewi, Bangkok 10400, Thailand

4 BioNet-Asia Co., Ltd., 19 Udomsuk 37, Sukhumvit 103, Prakanong, Bangkok 10260, Thailand

Springer

development [1, 4]. In medical applications, CRM197 has been used as a carrier protein for conjugate vaccines such as *Haemophilus influenzae* type b (Hib) vaccine, a quadrivalent meningococcal conjugate vaccine against meningococcal, and 13-valent pneumococcal vaccines [5]. While there have been several reports on CRM197 production, these studies have only obtained low levels of expression. The production of fragments A and B of CRM197 in *E. coli* resulted in yields ranging from 0.4 to 10 mg/L [6] and secretion of CRM197 from *Bacillus subtilis* gave 7.1 mg/L [7, 32]. Recently, our group successfully produced soluble CRM197 in *E. coli* Origami B at high concentrations by cultivation at low temperature and with co-expression of molecular chaperones [8]. Although CRM197 contains single amino acid mutation, there are no major differences between DT and CRM197 conformations [9]. CRM197 shows cytotoxicity on yeasts and mammalian cells [10], due to its weak ADP-ribosylation activity, and its overexpression leads to the inhibition of protein synthesis [11, 12]. In previous studies, another mutated DT, containing two amino acid substitutions at position 51 and 148 (K51E/E148K), was found to be substantially less toxic to yeast cells compared to wild-type DT and CRM197 [10]. DT residue E148 plays an important role in the active site of diphtheria toxin, whereas K51 is positioned in the active site loop (CL2 loop) located over NAD binding site. The active site loop is important for the binding between NAD and EF-2. Mutation at position 52 causes flexibility in the CL2 loop resulting in reduced NAD binding [9].

This study aimed to produce soluble CRM197EK, a new non-toxic mutant of DT having triple amino acid substitutions at position 51, 52 and 148 (K51E/G52E/E148K) in recombinant *E. coli*. CRM197EK was expressed as a fusion protein with the full-length thioredoxin (Trx) protein (109 amino acid residues) and contained a His-tag to increase the production of soluble protein and allow purification, respectively. The effect of induction temperature and molecular chaperones was also investigated. Nuclease activity of CRM197TrxHis was detected and similar to that found in both DT and CRM197. Molecular modeling of DT and its derivatives was also performed in order to examine the effects of amino acid substitutions on protein conformation and NAD binding, which is essential for DT toxicity. Due to its lack of toxicity, this protein might be used as the vaccine carrier protein in the future.

Materials and Methods

Chemical Reagents and Enzymes

All chemicals and reagents used in this study were purchased from either Merck (Darmstadt, Germany), Sigma-Aldrich (MO, USA) or Thermo Scientific (IL, USA). Restriction and modifying enzymes were purchased from Biolab (GA, USA).

Bacterial Strains and Culture Conditions

Escherichia coli Origami B (DE3) and *E. coli* Origami 2 (DE3) (Novagen; CA, USA) were used as expression hosts for CRM197EKTrxHis production. Both strains contain mutations in thioredoxin reductase (*trx*B) and glutathione reductase (*gor*) genes to facilitate cytoplasmic disulfide bond formation in *E. coli*. *E. coli* strains and their transformants were cultured on LB agar and LB medium supplemented with appropriated antibiotics. All strains were cultured at 37 °C with shaking at 200 rpm (New Brunswick INNOVA™ 4300; MN, USA).

Construction of pET48b-crm197ektrxhis Plasmid

A synthetic gene (CRM197EK) with codon optimization for *E. coli* expression and triple mutations (K51E/G52E/E148K) cloned into pUC57 (GenScript; NJ, USA) was used as a template for PCR amplification using specific primers: forward (F-CRM197EK-*Bam*HI): CgggATCCCgATATgggCgCAgACgATgTTgT and reverse (R-CRM 197EK-*Not*I): ATAAgAATgCggCCgCCCgTTACgATTTgATTTCgAAgAACAgg. The CRM197EK gene was subcloned into plasmid pET-48b (Novagen; CA, USA) at *Bam*HI and *Not*I restriction recognition sites, in-frame with Trx and His tags at the *N* terminus, to generate plasmid pET48-crm197ektrxhis (Fig. 1). The pET48-crm197ektrxhis plasmid was transformed into both *E. coli* strains and subsequently each of the chaperone expression plasmids (pKJE7, pGro7, pG-KJE8, pG-Tf2 and pTf16; Takara, Tokyo, Japan) was transformed into pET48-crm197ektrxhis containing *E. coli* for co-expression of molecular chaperones and CRM197EKTrxHis.

Optimization of CRM197EKTrxHis Expression in Recombinant *E. coli*

Comparison of CRM197EKTrxHis Expression in E. coli Origami B (DE3) and Origami 2 (DE3)

Escherichia coli transformants harboring the pET48b-crm197ektrxhis plasmid were cultured in LB medium supplemented with kanamycin (15 µg/ml) and tetracycline (12.5 µg/ml) for *E. coli* Origami B (DE3), whereas kanamycin (15 µg/ml), streptomycin (50 µg/ml) and tetracycline (12.5 µg/ml) were used for *E. coli* Origami 2 (DE3). Transformants were cultivated at 37 °C with shaking (200 rpm) until O.D.$_{600}$ equal to 0.8 was reached. The inducer (IPTG) was added into culture broth to obtain a
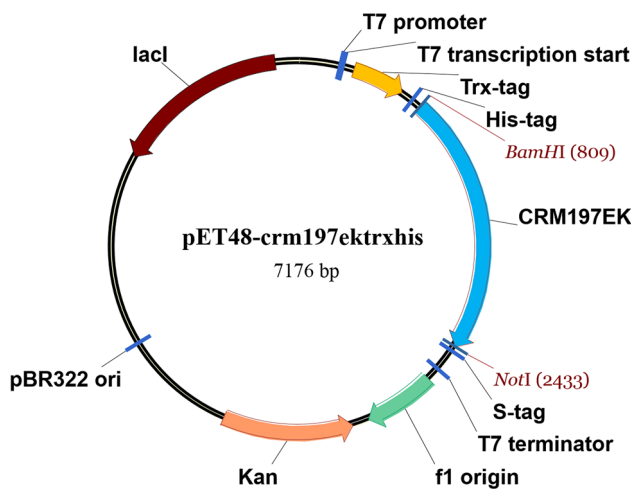
Fig. 1 pET48b-crm197ektrxhis (7176 bp). The gene encoding CRM197EK harbors mutations for three amino acid residues (K51E/G52E/E148 K). *Arrows* indicate direction of transcription

final concentration of 0.4 mM. Cultures were further incubated at 25 °C with shaking; cells were collected by centrifugation (8000 $\times g$ for 10 min) 20 h after induction.

Optimization of CRM197EKTrxHis Expression Condition in E. coli Origami B (DE3)

Escherichia coli Origami B (DE3) harboring the pET48b-crm197ektrxhis plasmid was cultured as described above. For optimization of the IPTG concentration, the inducer (IPTG) was added to obtain a final concentration of 0.1, 0.4 and 1 mM. Culture temperature was shifted to 25 °C, and cells were collected at 16 and 24 h after induction. Temperature optimization was performed by induction with 0.1 mM IPTG and followed by incubation at either 15, 20 or 25 °C. Cells were collected at 16 and 24 h after induction. The concentration of CRM197EKTrsHis was determined by ELISA.

Effect of Co-expression of Molecular Chaperones on CRM197EKTrxHis Production

Escherichia coli Origami B (DE3) harboring pET48b-crm197ektrxhis in combination with each of the chaperone plasmids (pKJE7, pGro7, pG-KJE8, pG-Tf2 and pTf16) were pre-cultured in LB medium supplemented with kanamycin (15 μg/ml), tetracycline (12.5 μg/ml) and chloramphenicol (20 μg/ml) at 37 °C for 16 h. Cultures were inoculated into 100 ml of fresh LB medium supplemented with same antibiotics and chaperone expression was immediately induced by addition of 0.5 mg/ml L-arabinose for plasmids pKJE7, pGro7, pG-KJE8 and pTf16, whereas 5 ng/ml tetracycline was added to induce chaperone protein production in pG-KJE8 and pG-Tf2.

Cultures were further incubated at 37 °C with shaking (200 rpm) until reaching an O.D.$_{600}$ equal to 0.8. Induction of CRM197EKTrxHis was initiated by addition of IPTG to obtain a final concentration of 0.1 mM. Cultures were further incubated at 20 °C, and induced cells were collected at 24 h following IPTG addition by centrifugation (8000 $\times$ g for 10 min). Proteins were analyzed by SDS-PAGE and ELISA.

Analysis of CRM197EKTrxHis

Preparation of Soluble and Insoluble CRM197EKTrxHis Fractions

Cell pellets from 1 ml of culture were collected and suspended in 200 μl of lysis buffer [50 mM Tris–HCl pH 8.0 and 1 mM phenylmethane sulfonyl fluoride (PMSF)]. Cells were disrupted using a bead beater with 0.1 g of zirconia–silicate beads (diameter 0.1 mm; Biospec Products, OK, USA) two times (1 min at each cycle) at a speed of 6.5 m/s. The supernatant containing the soluble fraction of CRM197EKTrxHis was collected after centrifugation at 12,000 rpm for 15 min. The pellet portion was resuspended in 200 μl of solubilization buffer (50 mM Tris–HCl pH 8.0, 1 mM PMSF, 500 mM NaCl, 1%Triton X-100 and 6 M Urea), incubated at room temperature for 1 h and centrifuged at 12,000 rpm, and the supernatant was collected as the insoluble fraction.

CRM197EKTrxHis Purification

CRM197EKTrxHis was purified using Ni–NTA affinity chromatography (Qiagen; Hilden, Germany). The soluble fraction of the 6xHis-tagged protein was loaded into a pre-equilibrated Ni–NTA column (600 μl/column). After washing columns three times with 600 μl of buffer NPI-20 (50 mM Tris–HCl pH 8.0, 300 mM NaCl and 20 mM imidazole), proteins were eluted twice with 200 μl of NPI-500 buffer (50 mM Tris–HCl pH 8.0, 300 mM NaCl and 500 mM imidazole). The protein in supernatant was dialyzed to remove NaCl and imidazole using dialysis buffer (20 mM Tris–HCl pH 8.0) and concentrated with an Amicon Ultra-4 50 K centrifugal filter device (Millipore, Cork Ireland).

ELISA and Nuclease Assay

The soluble fraction of CRM197EKTrxHis was diluted (125X) in PBS buffer (25 mM phosphate buffer pH 7.4 and 150 mM NaCl) and loaded into 96-well nickel-coated plates (100 μl/well) in triplicate for each dilution. PBS buffer alone was used as a blank. The soluble proteins were incubated in nickel-coated plates (Thermo Scientific, IL, USA) for 1 h at room temperature, and the assay was

performed using HRP-conjugated anti-diphtheria toxin polyclonal antibody (Thermo Scientific, IL, USA) as described previously [8]. For nuclease assays, the purified CRM197EKTrxHis (1 μg) was incubated with 500 ng of λDNA in reaction buffer (10 mM Tris–HCl pH 7.5, 2.5 mM CaCl$_2$ and 2.5 mM MgCl$_2$) at three different temperatures (37, 25 or 20 °C). Samples were collected (20 μl) at 0, 2, 4, 8 and 16 h of incubation. Reactions were stopped by addition of EDTA to a final concentration of 5 mM. Samples were analyzed by electrophoresis using 0.7% agarose gels. The degradation of λDNA was measured using ImageJ densitometry software (version 1.50b, MD, USA).

Molecular Modeling of Diphtheria Toxin and Its Derivatives

The ligand-free diphtheria toxin structure was modeled with diphtheria toxin structure PDB ID: 1SGK [13] as a template using the SWISS-Model ExPASy server [14]. This structure was further mutated into CRM197 and CRM197EK using PyMol software. All molecular dynamics simulations employed the AMBER12 package [15] with Amber03 force field [16]. After minimization and heating the initial structure, molecular dynamics simulations were performed for 10 nanoseconds (ns) under an isothermal–isobaric ensemble (NPT) at 310 Kelvin (K) using the Berendsen weak coupling approach [17] to maintain temperature with a time step of 2 femtosecond (fs) and cutoff radius of 12.0 Å. The coordinates were saved every 2 picosecond (ps). The root mean square deviation (RMSD) was calculated for all atoms of the selected region: fragment A (residues 1–194), CL2 loop (residues 34–52), IgG recognition site (residues 141–157) and CD4$^+$-recognition site (residues 271–290, 321–340, 331–350, 351–370, 411–430 and 431–450). Pairwise per residue and per residue energy decomposition analysis using the Molecular Mechanics–Generalized Born Surface Area approach (MM-GBSA) with parameters proposed by Onufriev et al. [18] was performed to analyze the intramolecular interactions of residues within 6 Å of CL2 loop. Average structures from 1 to 10 ns were established and aligned to fragment A (residue 1–194) to investigate structural differences between the models.

Results and Discussion

Expression of CRM197EKTrxHis in Different E. coli Hosts

The synthetic gene of CRM197EK consists of 1615 bp which encodes of 535 amino acids with three mutations (K51E/G52E/E148K). E. coli does not allow the disulfide bond formation of recombinant protein in cytoplasm if thioredoxin reductase (trxB) and glutathione reductase (gor) genes are present [19]. Fragments A and B of diphtheria toxin are bound together by a disulfide bridge between Cys186 and Cys201. Fragment A becomes enzymatically active following the reduction in this disulfide bond after translocation of diphtheria toxin into cells. Therefore, disulfide bond formation is important for the structure of diphtheria toxin and CRM197 [2]. In this study, E. coli Origami B (DE3) and E. coli Origami 2 (DE3) were used as expression hosts since both Origami strains contain mutations in trxB and gor genes that facilitate disulfide bond formation of proteins in the E. coli cytoplasm [19]. E. coli Origami strains have been used in the production of several recombinant proteins, such as human interferon, that require disulfide bonds for correct structure [20]. Soluble and insoluble fractions of CRM197EKTrxHis (~75 kDa) were produced in both Origami host strains (Fig. 2). E. coli Origami B (DE3) produced higher amounts of soluble and insoluble CRM197EKTrxHis than E. coli Origami 2 (DE3) under the same induced condition (0.4 mM IPTG, at 25 °C for 20 h). Origami B (DE3) is a BL21 derivative containing mutations in two proteases genes (lon and ompT) which decreases proteolysis in the E. coli cytoplasm. In contrast, Origami 2 (DE3) is a K-12 derivative and contains lon and ompT genes [21]. Due to the higher yield of soluble protein, E. coli Origami B (DE3) was selected as a host strain for further CRM197EKTrxHis production in this study.

Optimization of CRM197EKTrxHis Production in E. coli Origami B (DE3)

Slow expression with low levels of IPTG is required with some proteins in order to increase production of soluble protein. However, high levels of IPTG are advantageous as this can lead to production of high concentrations of total protein [22]. The highest concentration of soluble CRM197EKTrxHis was observed after induction with 0.1 mM IPTG at both 16 h (78.92 ± 3.74 μg/ml) and 24 h (71.93 ± 4.81 μg/ml) after induction (Fig. 3a). The yield of fused protein produced under both induction periods was not significantly different and was only slightly decreased upon prolonged induction. It was previously reported that production of CRM197 utilizing a pET32a expression system (ampicillin resistant) gave the highest yield when cells were induced with 0.4 mM IPTG at 15 °C [8]. However, the temperature used in this study was optimized for an IPTG concentration of 0.1 mM which was suitable for CRM197EK production from pET48b-crm197ek-trxhis (kanamycin resistant). Comparison of expression levels from transformants with 0.1 mM IPTG at 15, 20 and
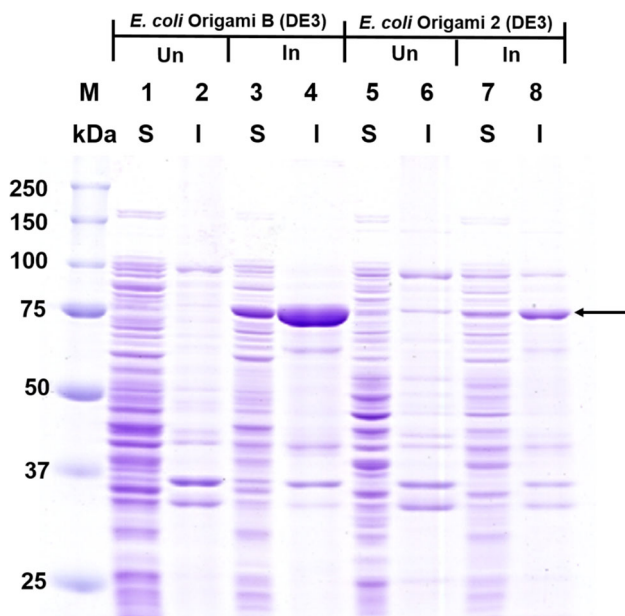
Fig. 2 SDS-PAGE of CRM197EKTrxHis expression in *E. coli* Origami B (DE3) and *E. coli* Origami 2 (DE3). Analysis was performed under un-induced (Un) and induced (In) conditions with 0.4 mM IPTG at 25 °C for 20 h; M: precision plus protein standards marker; S and I represent the soluble and insoluble fractions, respectively. *Arrow* indicates CRM197EKTrxHis

25 °C for 16 and 24 h revealed that the highest amount of soluble CRM197EKTrxHis was present after induction at 20 °C for 24 h (Fig. 3b). It was observed that cells induced at 15 and 20 °C accumulated protein at 24 h with higher yield (51.14 ± 8.83 and 97.33 ± 17.47 μg/ml, respectively) when compared to those of 16 h induction (37.25 ± 1.80 and 75.78 ± 7.31 μg/ml, respectively), whereas in cells induced at 25 °C, CRM197EKTrxHis was higher at 16 h (90.58 ± 3.53 μg/ml) compared to 24 h induction (77.49 ± 4.31 μg/ml) (Fig. 3b). Utilizing 20 °C for protein expression resulted in both a higher yield and increased biological activity of the target protein, consistent with reports that slower growing cells often exhibit enhanced production of correctly folded proteins [23]. These results indicate that induction using 0.1 mM IPTG at 20 °C for 24 h was the optimal condition for soluble CRM197EKTrxHis production. These conditions were used for production of CRM197EKTrxHis throughout the rest of this study. pET48b was used in this study because it harbors the marker gene rendering the resistance to kanamycin which is the preferred antibiotics for the production of pharmaceutical proteins [37].

Effect of Co-expression of CRM197EKTrxHis and Molecular Chaperones

Molecular chaperones are defined as cellular proteins which assist the folding process by preventing incorrect inter- and intracellular interactions in non-native polypeptide chains [24]. Several studies have demonstrated an increase in soluble recombinant proteins when co-expressed with molecular chaperones [25]. In an attempt to enhance the correct folding, and thus production of soluble protein, CRM197EKTrxHis was co-expressed with several chaperones including trigger factor (TF), GroEL-GroES and DnaK-DnaJ-GrpE (plasmids pG-KJE8, pGro7, pKJE7, pG-Tf2 and pTf16). The soluble and insoluble fractions of CRM197EKTrxHis were analyzed by SDS-PAGE (Fig. 4a), and only soluble material was analyzed by ELISA (Fig. 4b). The results demonstrate that the amount of soluble CRM197EKTrxHis was increased in proportion to the induction time. A high amount of soluble CRM197EKTrxHis was observed after induction for 24 h (88.80 ± 8.50 μg/ml) without co-expression of chaperones. However, co-expression with some chaperones had an impact on the yield of soluble CRM197EKTrxHis. Plasmid pG-Tf2, expressing two molecular chaperones, TF and GroEL-GroES, provided the highest amount of soluble CRM197EKTrxHis after induction at 20 °C for 24 h (111.24 ± 10.40 μg/ml). Expression of TF (pTf16) and GroEL-GroES (pGro7) alone did not enhance solubility, and similar amounts of soluble CRM197EKTrxHis were produced as transformants without chaperone co-expression (92.75 ± 11.5, 87.66 ± 18.8 and 88.80 ± 8.50 μg/ml, respectively). Expression of TF or GroEL-GroES alone can facilitate protein folding and prevent aggregation for some proteins. However, overexpression of TF together with GroEL-GroES was more effective in producing soluble human oxygen-regulated protein ORP150 and human lysozyme, likely due to their synergistic roles in vivo [26]. Surprisingly, in our previous study examining the production of CRM197 in *E. coli* the presence of TF (pTf16) provided a higher yield of soluble CRM197 compared to cells co-expressing both TF and GroEL-GroES (pG-Tf2) [8]. Different production levels for the same gene cloned gene in different expression vectors under the same induction conditions are commonly observed; however, the mechanisms behind this variability are not clear [27]. Contrasting results for the protein yield between expression systems in these studies may be due in part to the use of different incubation temperatures (15 and 20 °C) and expression vectors (pET32a and pET48b plasmids). The use of different antibiotics to maintain plasmids may have an impact on protein expression levels. Kanamycin, used to maintain pET48b-based plasmids, binds to the 30S subunit and blocks the initial formation of the 70S ribosome, reducing the rate of protein synthesis. While inclusion of ampicillin to maintain pET32a-based plasmids can reduce the activity of the enzyme transpeptidase, required for cell wall synthesis in bacteria, this antibiotic does not affect stationary phase cells [28]. In contrast to TF and GroEL-

Fig. 3 Effect of IPTG (**a**) and induction temperature (**b**) for CRM197EKTrxHis production. Cells were induced at 25 °C for 16 and 24 h using 0.1, 0.4 and 1 mM IPTG. For optimization of induction temperature; induction was performed using 0.1 mM IPTG induction at 15, 20 and 25 °C for 16 and 24 h. The concentration of the purified protein was obtained by ELISA. *Bars* indicate standard error of triplicate experiments
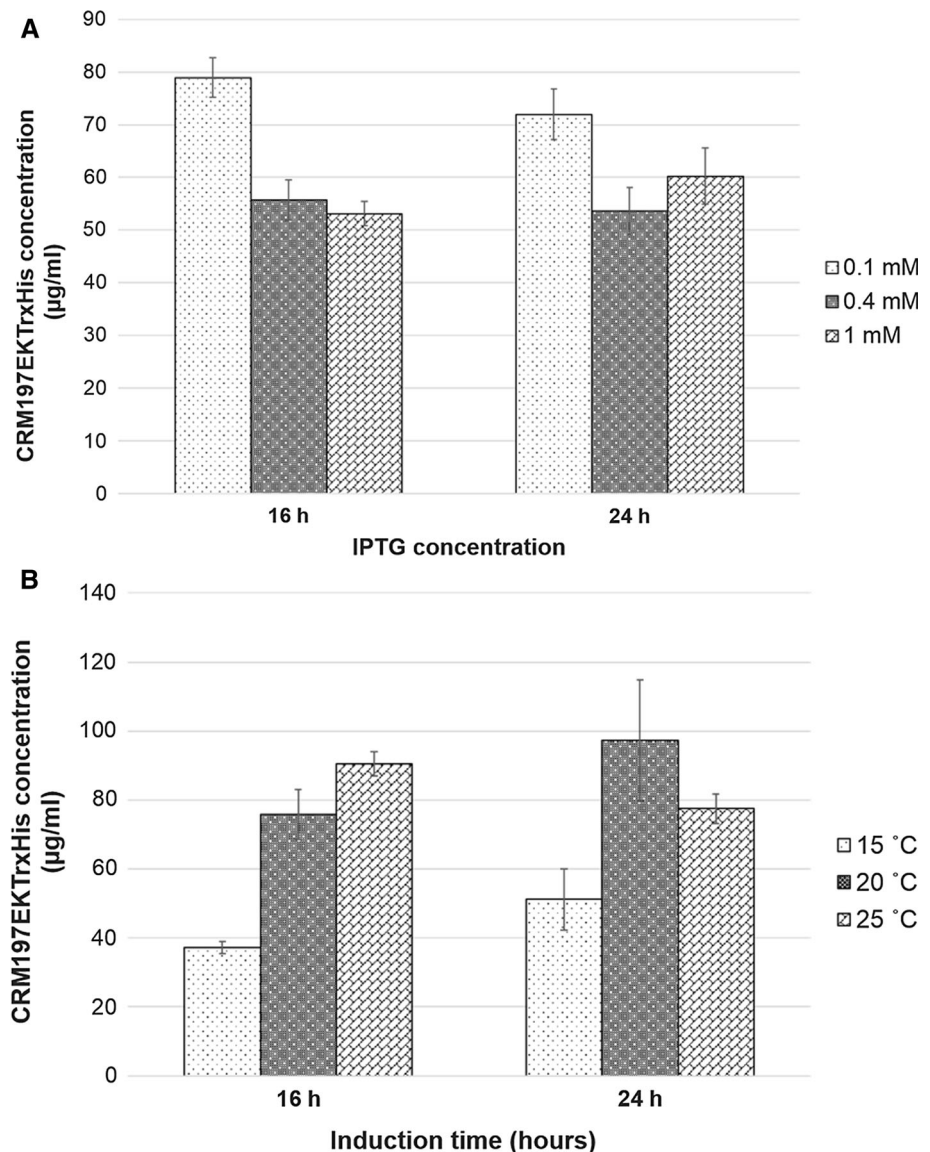

GroES alone or in combination, co-expression with chaperones DnaK-DnaJ-GrpE (pKJE7) or both DnaK-DnaJ-GrpE and GroEL-GroES complexes (pG-KJE8) decreased the production of CRM197EKTrxHis at all induction times, even though chaperon proteins were produced (Fig. 4a lanes 3 and 7). Although DnaK can function as a folding modulator, it has also been reported to cause inhibition of cell growth and enhance proteolysis of recombinant proteins. Therefore, the amount of recombinant proteins could be reduced due to DnaK induced proteolysis [29]. DnaK has also been identified as a negative regulator of heat shock proteins, including GroEL [29]. Reduced effectiveness from co-expression of both DnaK-DnaJ-GrpE and GroEL-GroES complexes (pG-KJE8) may be due to decreased expression of GroEL when DnaK is present. In addition, the overproduction of several molecular chaperones simultaneously can cause decreased growth rates and reduced biomass production, resulting in a lower levels of recombinant protein production [29]. Our results are similar to those seen for the production of soluble human collagen in *E. coli* co-expressed with DnaK-DnaJ-GrpE and GroEL-GroES complexes (pG-KJE8 and pKJE7) [30]. In this case, it was proposed that energy used by host cells for chaperone production limited the rate of expression of the target protein [30].

Nuclease Assay

Both wild-type diphtheria toxin and CRM197 possess nuclease activity [31]. Since CRM197EKTrxHis is a fusion protein as well as containing triple amino acid substitutions, we examined the impact of these features on its
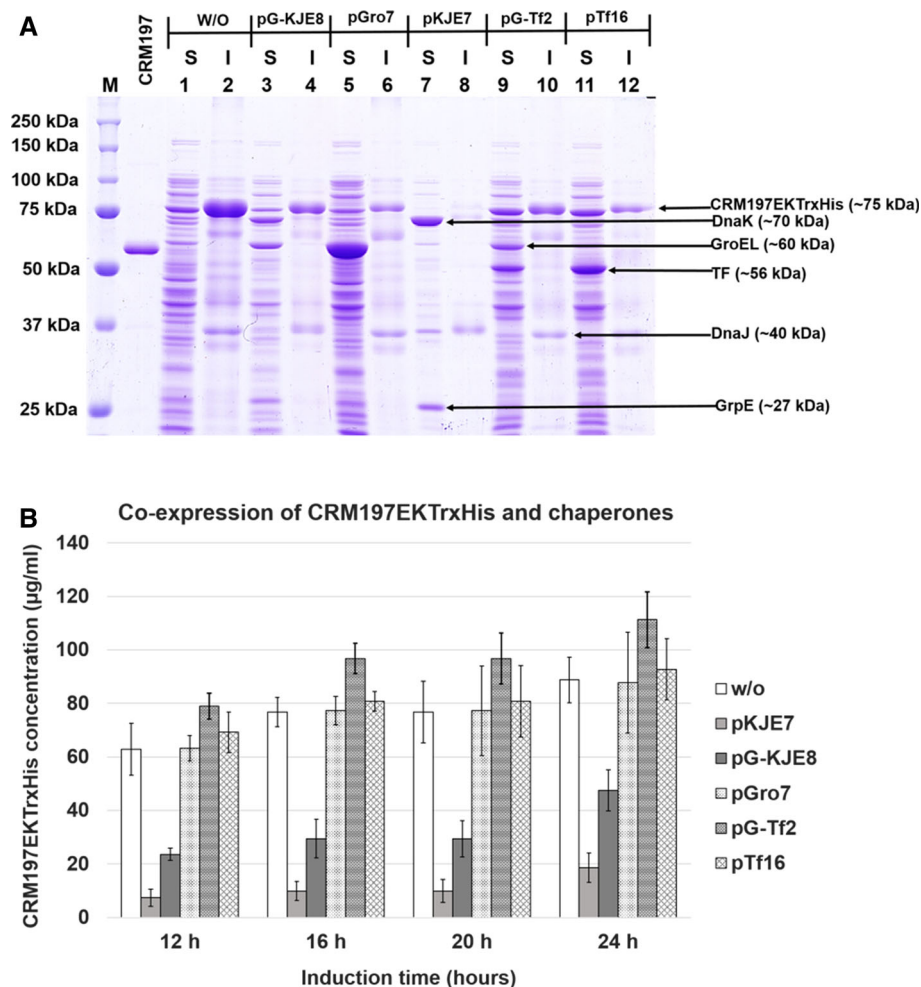
Fig. 4 a SDS-PAGE of CRM197EKTrxHis with and without chaperone plasmid. *E. coli* Origami B (DE3) harboring pET48b-crm197ektrxhis and each of the five different chaperone expression plasmids were induced with 0.1 mM IPTG at 20 °C for 24 h; M: precision plus protein standards marker; *S* and *I* represent the soluble and insoluble fractions, respectively; purified CRM197 is used as a positive control; *Lanes* 1–2: without (w/o) chaperone plasmid; *Lanes* 3–4, 5–6, 7–8, 9–10 and 11–12: CRM197EKTrxHis after co-expression with chaperone genes from pG-KJE8, pGro7, pKJE7, pG-Tf2 and pTf16 plasmid, respectively. *Arrows* indicate CRM197EKTrxHis and different chaperones proteins. **b** Effect of chaperones on soluble CRM197EKTrxHis production in *E. coli* Origami B (DE3) (pET48b-crm197ektrxhis). Recombinant protein was induced with 0.1 mM IPTG at 20 °C for 12, 16, 20 and 24 h. The amount of soluble CRM197EKTrxHis was determined by ELISA; w/o represents without chaperone plasmid. *Bars* indicate standard error of triplicate experiments

nuclease activity at three different temperatures. CRM197EKTrxHis exhibited nuclease activity and could degrade DNA at 20, 25 and 37 °C (Fig. 5). At 20 °C, λDNA was incompletely degraded, even after incubation for 16 h only a low amount of smear bands indicating lower molecular weight DNA fragments were observed. λDNA was degraded completely within 8 h at 25 °C, and the strongest nuclease activity on λDNA was observed at 37 °C, with λDNA being completely digested after incubation for 4 h. The presence of nuclease activity in CRM197EKTrxHis is indicative that the protein is correctly folded, consistent with the previous report that a fusion tag did not abolish the biochemical properties of CRM197 [32]. Cells overexpressed CRM197EKTrxHis

rendered higher nuclease activity and caused cell lysis (data not shown) similar to that observed in cells producing high amount of CRM197 [8]. This toxin-dependent DNA cleavage and cell lysis might therefore limit the yield of soluble toxin in overexpressed cells.

Molecular Modeling of DT and Its Derivatives

The average molecular structure of DT, CRM197 and CRM197EK was established, and a representative model of CRM197EK without the fusion tag protein is shown in Fig. 6a, b. The root mean square deviation (RMSD) of DT and its derivatives was measured in order to study the NAD binding ability and immunological properties. The toxicity
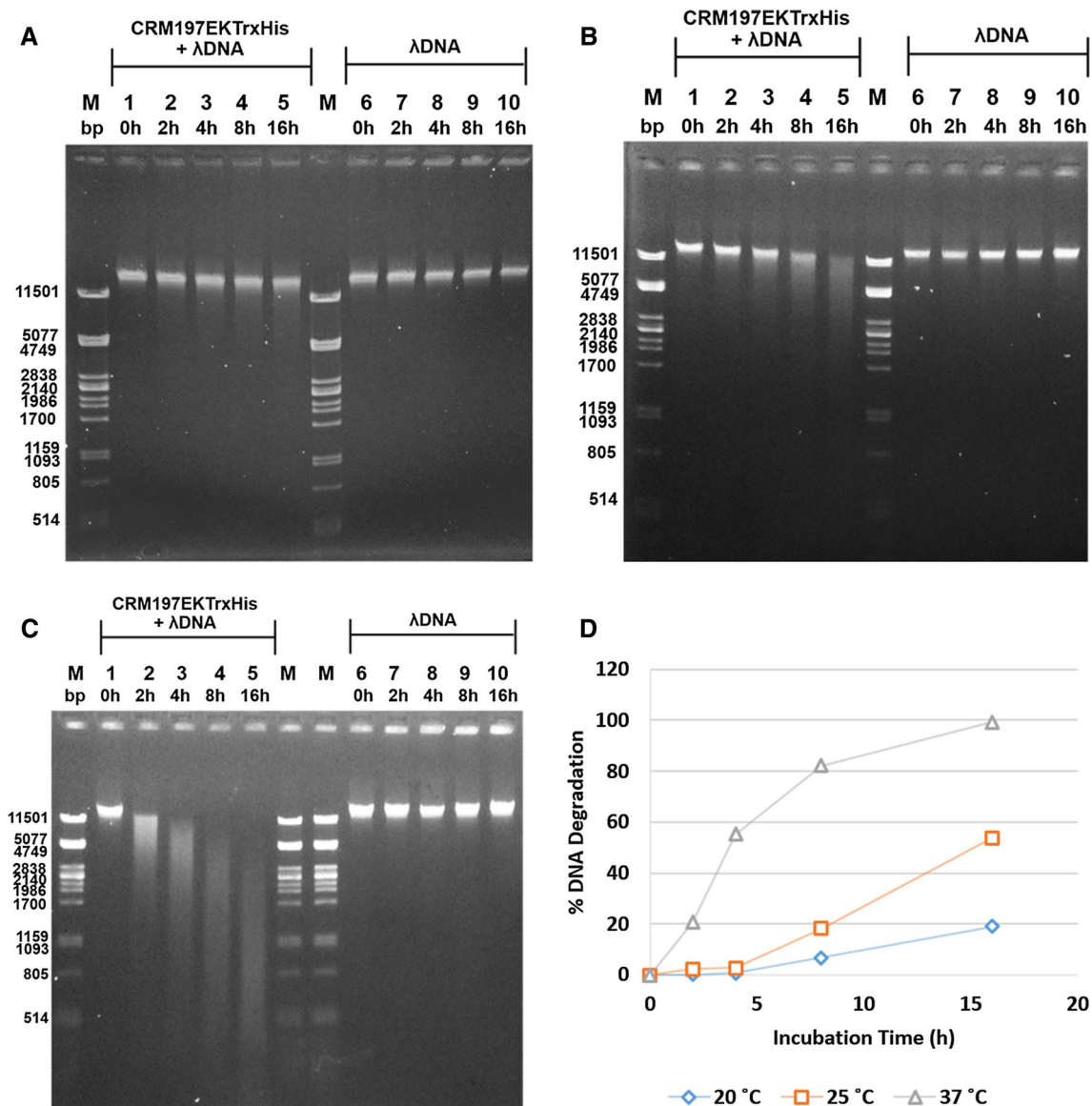
Fig. 5 Nuclease activity of CRM197EKTrxHis after incubation at **a** 20 °C **b** 25 °C **c** 37 °C. The percentage of DNA degradation after incubation for 0, 2, 4, 8 and 16 h is shown in **d**. The purified CRM197EKTrxHis was incubated with 500 ng of λDNA at three different temperatures. Samples were analyzed by agarose gel electrophoresis; M: λ*pst*I marker; λDNA without CRM197EKTrxHis was used as a negative control

of DT is related to two important steps, the ability fragment A of DT to bind NAD and ADP-ribosylation activity at the DT catalytic site (H21, Y65 and E148) [3, 9]. The active site loop (CL2 loop) located over the NAD binding site on fragment A, is important for the NAD binding activity and elongation factor 2 (EF-2) recognition [9]. The CL2 loop of CRM197 is more flexible than that of DT resulting in decreased toxicity of this variant. The structural comparison of CL2 loop of CRM197EK performed in this study indicated differences compared to that of both DT and CRM197 (Fig. 6c). These structural changes might be involved in the abrogated toxicity of this DT mutant. The structural change appears to be due primarily to the amino

acid substitution of lysine to glutamic acid at residue 51. In the DT model, the side chain of K51 was observed to form several electrostatic interaction with the oxygen backbone of CL2 loop residues, in particular T42, N45, D47 and D48, which are important for the structure of CL2 loop, as shown in Fig. 6D. The total energy stabilization of the CL2 loop gained from the presence of K51 was −12.16 kcal/mol for DT, similar to −12.19 kcal/mol of CRM197. The overall structures of the CL2 loops are similar between DT and CRM197. When K51 was changed to glutamic acid in CRM197EK, this residue (E51) is predicted to form a strong salt bridge interaction with K37 instead of stabilizing the CL2 loop. As a consequence, the stabilization

Fig. 6 Molecular modeling of DT, CRM197 and CRM197EK. **a** Surface model of CRM197EK. **b** Model structure of CRM197EK. In both **a** and **b**, the IgG recognition site, CL2 loop and NAD binding site on fragment A, as well as CD4+ recognition site of R and T domains in fragment B are shown. **c** The superposition of CL2-loop structure of DT (*white*), CRM197 (*gray*) and CRM197EK (*Black*). **d** The electrostatic interaction between K51, D45 and D47. **e** The superposition of the IgG recognition site in DT, CRM197 and CRM197EK. All molecular modeling images were drawn using PyMol software version 1.3
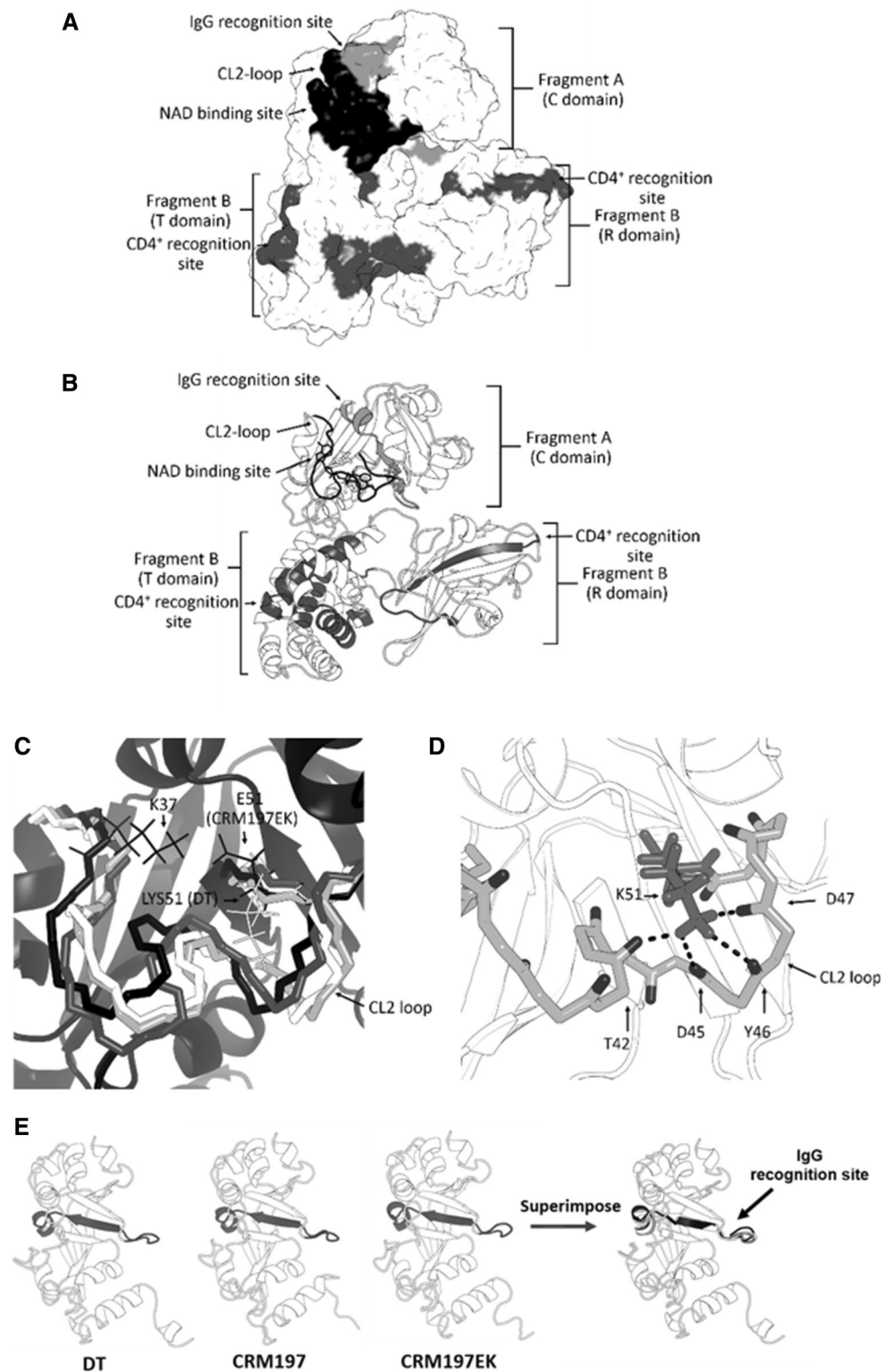
energy of CL2 loop was increased to 0.08 kcal/mol for CRM197EK. The lack of the interaction of K51 in CRM197EK could lead to a shift in position of subsequent residues. The fluctuation of all atoms in CL2 loop of CRM197EK was found to be greater than that observed in DT, and it might affect the binding stability of CRM197EK to NAD.

In addition to changes in the CL2 loop, CRM197EK contains an amino acid substitution at residue 148, which plays an important role in the active site. Therefore,

CRM197EK was expected to be a non-toxic protein due to the lack of catalytic activity and impaired NAD binding. In terms of immunogenicity, the human antibody (IgG) can react to fragment A of diphtheria toxin at amino acid residues 141–157 [33]. The CD4+ of T-cells can also react with six peptide sequences in DT including residue 271–290, 321–340, 331–350, 351–370, 411–430 and 431–450 [34]. The IgG recognition site of DT, CRM197 and CRM197EK was aligned in this analysis as is shown in Fig. 6e. Structural modeling of the IgG interacting regions found no difference in the three DT structures; therefore, CRM197EK is also expected to generate an immune response for Diphtheria. In medical applications, CRM197 can be used as a carrier protein for conjugate vaccines through the formation of cross-links between lysine residues of CRM197 and oligosaccharide antigens [35]. CRM197EK contains 40 lysine residues. In addition, it has been reported that the mutant of diphtheria toxin containing two amino acid substitutions (K51E/E148K) could induce an immunogenic response in mice and had the ability to act as a carrier protein for the MenA glycoconjugate vaccine [36]. Therefore, CRM197EK was expected to be used as a carrier protein similar to that of CRM197. In conclusion, CRM197EK with mutation at residue 52 together with 51 and 148 may be a good candidate as carrier protein because it could reduce both the stability of NAD binding and catalytic activity of the enzyme. However, in vivo study to elucidate its biological activity is required to compare its function with other carrier proteins especially CRM197.

Conclusions

This study provided useful alternative information about a way to enhance soluble production of a recombinant CRM197 derivative using a bacterial expression system. The co-expression of the target protein with chaperones examined here should have utility for production of other proteins in *E. coli*, including carrier proteins and proteins with pharmaceutical value. Our study was successful in terms of enhancement of soluble production of CRM197EKTrxHis in *E.coli* using molecular chaperones. The conditions for production were optimized, and CRM197EKTrxHis was produced as soluble protein with biochemical activity. The highest amount of soluble CRM197EKTrxHis was produced in recombinant *E. coli* co-expressed with pG-Tf2 (harboring chaperone genes for trigger factor and GroES-GroEL) after induction at 20 °C for 24 h. Molecular modeling predicted that CRM197EK will be recognized by IgG similarly to DT and CRM197. It is therefore expected to be a good candidate for vaccine development against diphtheria and/or can be used as carrier protein for conjugate vaccines.

References

1. Uchida, T., Pappenheimer, A., & Greany, R. (1973). Diphtheria toxin and related proteins I. Isolation and properties of mutant proteins serologically related to diphtheria toxin. *Journal of Biological Chemistry, 248,* 3838–3844.
2. Deng, Q., & Barbieri, J. (2008). Molecular mechanisms of the cytotoxicity of ADP-ribosylating toxins. *Annual Review of Microbiology, 62,* 271–288.
3. Bell, C., & Eisenberg, D. (1996). Crystal structure of diphtheria toxin bound to nicotinamide adenine dinucleotide. *Biochemistry, 35,* 1137–1149.
4. Pappenheimer, J. A. (1977). Diphtheria toxin. *Annual Review of Biochemistry, 46,* 69–94.
5. Shinefield, H. (2010). Overview of the development and current use of CRM 197 conjugate vaccines for pediatric use. *Vaccine, 28,* 4335–4339.
6. Bishai, W., Rappuoli, R., & Murphy, J. (1987). High-level expression of a proteolytically sensitive diphtheria toxin fragment in *Escherichia coli. Journal of Bacteriology, 169,* 5140–5151.
7. Jianhua, Z., & Petracca, R. (1999). Secretory expression of recombinant diphtheria toxin mutants in *Bacillus subtilis. Journal of Tongji Medical University, 19,* 253–256.
8. Mahamad, P., Boonchird, C., & Panbangred, W. (2016). High level accumulation of soluble diphtheria toxin mutant (CRM197) with co-expression of chaperones in recombinant *Escherichia coli. Applied Microbiology and Biotechnology, 100,* 6319–6330.
9. Malito, E., Bursulaya, B., Chen, C., Surdo, P., Picchianti, M., Balducci, E., et al. (2012). Structural basis for lack of toxicity of the diphtheria toxin mutant CRM197. *Proceedings of the National Academy of Sciences of the United States of America, 109,* 5229–5234.
10. Kimura, Y., Saito, M., Kimata, Y., & Kohno, K. (2007). Trans-genic mice expressing a fully nontoxic diphtheria toxin mutant, not CRM197 mutant, acquire immune tolerance against diphtheria toxin. *Journal of Biochemistry, 142,* 105–112.
11. Kageyama, T., Ohishi, M., Miyamoto, S., Mizushima, H., Iwa-moto, R., & Mekada, E. (2007). Diphtheria toxin mutant CRM197 possesses weak EF2-ADP-ribosyl activity that poten-tiates its anti-tumorigenic activity. *Journal of Biochemistry, 142,* 95–104.
12. Qiao, J., Ghani, K., & Caruso, M. (2008). Diphtheria toxin mutant CRM197 is an inhibitor of protein synthesis that induces cellular toxicity. *Toxicon, 51,* 473–477.
13. Bell, C., & Eisenberg, D. (1997). Crystal structure of nucleotide-free diphtheria toxin. *Biochemistry, 36,* 481–488.
14. Schwede, T., Kopp, J., Guex, N., & Peitsch, M. (2003). SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Research, 31,* 3381–3385.
15. Case, D., Darden, T., Cheatham, T., Simmerling, C., Wang, J., Duke, R., et al. (2012). *AMBER 12*. San Francisco: University of California.
16. Duan, Y., Wu, C., Chowdhury, S., Lee, M., Xiong, G., Zhang, W., et al. (2003). A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry, 24,* 1999–2012.

17. Berendsen, H., Postma, J., Gunsteren, W., DiNola, A., & Haak, J. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics, 81,* 3684–3690.

18. Essmann, U., Perera, L., Berkowitz, M., Darden, T., Lee, H., & Pedersen, L. (1995). A smooth particle mesh Ewald method. *The Journal of Chemical Physics, 103,* 8577–8593.

19. Prinz, W., Åslund, F., Holmgren, A., & Beckwith, J. (1997). The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. *Journal of Biological Chemistry, 272,* 15661–15667.

20. Rabhi-Essafi, I., Sadok, A., Khalaf, N., & Fathallah, D. (2007). A strategy for high-level expression of soluble and functional human interferon α as a GST-fusion protein in *Escherichia coli*. *Protein Engineering, Design & Selection, 20,* 201–209.

21. Studier, F., Daegelen, P., Lenski, R., Maslov, S., & Kim, J. (2009). Understanding the differences between genome sequences of *Escherichia coli* B strains REL606 and BL21 (DE3) and comparison of the *Escherichia coli* B and K-12 genomes. *Journal of Molecular Biology, 394,* 653–680.

22. Joseph, S., & David, W. (2001). *Molecular cloning: a laboratory manual.* New York: Gold Spring Harbor.

23. Vera, A., González-Montalbán, N., Arís, A., & Villaverde, A. (2007). The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnology and Bioengineering, 96,* 1101–1106.

24. Hendrick, J., & Hartl, F. (1993). Molecular chaperone functions of heat-shock proteins. *Annual Review of Biochemistry, 62,* 349–384.

25. Weickert, M., Doherty, D., Best, E., & Olins, P. (1996). Optimization of heterologous protein production in *Escherichia coli*. *Current Opinion in Biotechnology, 7,* 494–499.

26. Nishihara, K., Kanemori, M., Yanagi, H., & Yura, T. (2000). Overexpression of trigger factor prevents aggregation of recombinant proteins in *Escherichia coli*. *Applied and Environmental Microbiology, 66,* 884–889.

27. Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Frontiers in Microbiology, 5,* 172.

28. Kohanski, M., Dwyer, D., & Collins, J. (2010). How antibiotics kill bacteria: from targets to networks. *Nature Reviews Microbiology, 8,* 423–435.

29. Martínez-Alonso, M., García-Fruitós, E., Ferrer-Miralles, N., Rinas, U., & Villaverde, A. (2010). Side effects of chaperone gene co-expression in recombinant protein production. *Microbial Cell Factories, 9,* 1.

30. Jia, Q., Fan, D., Ma, P., Ma, X., & Xue, W. (2014). The different roles of chaperone teams on over-expression of human-like collagen in recombinant *Escherichia coli*. *Journal of the Taiwan Institute of Chemical Engineers, 45,* 2843–2850.

31. Bruce, C., Baldwin, R., Lessnick, S., & Wisnieski, B. (1990). Diphtheria toxin and its ADP-ribosyltransferase-defective homologue CRM197 possess deoxyribonuclease activity. *Proceedings of the National Academy of Sciences, 87,* 2995–2998.

32. Stefan, A., Conti, M., Rubboli, D., Ravagli, L., Presta, E., & Hochkoeppler, A. (2011). Overexpression and purification of the recombinant diphtheria toxin variant CRM197 in *Escherichia coli*. *Journal of Biotechnology, 156,* 245–252.

33. Perera, V., & Corbel, M. (1990). Human antibody response to fragments A and B of diphtheria toxin and a synthetic peptide of amino acid residues 141–157 of fragment A. *Epidemiology and Infection, 105,* 457–468.

34. Raju, R., Navaneetham, D., Okita, D., Diethelm-Okita, B., McCormick, D., & Conti-Fine, B. (1995). Epitopes for human CD4 + cells on diphtheria toxin: Structural features of sequence segments forming epitopes recognized by most subjects. *European Journal of Immunology, 25,* 3207–3214.

35. Bröker, M., Costantino, P., DeTora, L., McIntosh, E., & Rappuoli, R. (2011). Biochemical and biological characteristics of cross-reacting material 197 (CRM197), a non-toxic mutant of diphtheria toxin: Use as a conjugation protein in vaccines and other potential clinical applications. *Biologicals, 39,* 195–204.

36. Pecetta, S., Vijayakrishnan, B., Romano, M., Proietti, D., Surdo, P., Balocchi, C., et al. (2016). Evaluation of the non-toxic mutant of the diphtheria toxin K51E/E148 K as carrier protein for meningococcal vaccines. *Vaccine, 34,* 1405–1411.

37. Mignon, C., Sodoyer, R., & Werle, B. (2015). Antibiotic-free selection in biotherapeutics: Now and forever. *Pathogens, 4*(2), 157–181.