Electronic Journal of Biotechnology 71 (2024) 19–27

Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

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

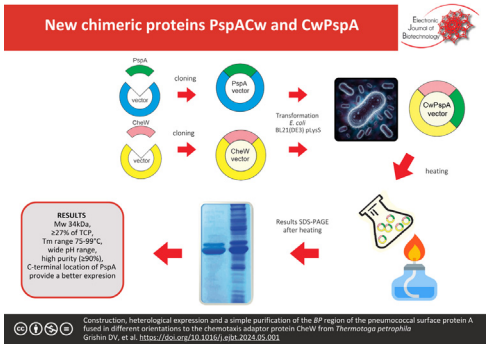
Research article

Construction, heterological expression and a simple purification of the *BP* region of the pneumococcal surface protein A fused in different orientations to the chemotaxis adaptor protein CheW from *Thermotoga petrophila* ☆

Dmitry V. Grishin [a], Nikita G. Sidorov [a,b], Olga K. Parfenova [a], Roman V. Kurkin [a], Ekaterina Y. Kasap [a,*]

[a] *Institute of Biomedical Chemistry (IBMC), 10, Pogodinskaya Street, 119121 Moscow, Russia*
[b] *Federal State Budgetary Scientific Institution "I. Mechnikov Research Institute of Vaccines and Sera", 5A, Maly Kazenny Lane, 105064 Moscow, Russia*

GRAPHICAL ABSTRACT



New chimeric proteins PspACw and CwPspA

ARTICLE INFO

Article history:
Received 27 March 2024
Accepted 10 May 2024
Available online 19 June 2024

Keywords:
Chimeric protein
Escherichia coli
Fusion technology
Heterologous expression
Hybrid proteins
Novel chimeric proteins
Pneumococcal surface protein
Streptococcus pneumoniae
Thermostable protein

ABSTRACT

Background: The important challenge to the biotechnology is to find new effective fusion partners that enable to improve solubility, expression, and optimize the subsequent fine purification of the target protein.
Results: The most invariant part of the most immunogenic region of the surface virulence factor A of *Streptococcus pneumoniae* was selected as a model target protein, while the thermostable chemotaxis polypeptide of W-type from *Thermotoga petrophila* was used as a fusion partner. The genes encoding fusion variants of these proteins were constructed and cloned into a plasmid vector under the control of the strong bacteriophage T7 transcription regulatory system. Effective *Escherichia coli* producer strains were obtained, and optimal conditions were chosen for the production of resulting constructs. The optimal pH and temperature ranges of recombinant proteins were determined, and three-dimensional shapes of their molecules were also predicted. Methods of low-stage protein purification were improved. Some of the isolated proteins demonstrated a high level of expression, solubility and purity.
Conclusions: Novel chimeric proteins were obtained which had not previously been observed in nature in such domain combinations. It was shown that the biotechnologically valuable characteristics of the

☆ Audio abstract available in Supplementary material.
Peer review under responsibility of Pontificia Universidad Católica de Valparaíso
* Corresponding author.
E-mail address: E.Kasap321@rambler.ru (E.Y. Kasap).

https://doi.org/10.1016/j.ejbt.2024.05.001

Thermotoga petrophila
Virulence factor A

hybrid proteins were more marked when the thermal-resistant partner was combined with the N-terminus of pneumococcal protein. The principles of their low-stage purification were performed which does not require any special equipment. That is a basis for significant reduction of prices for diagnostic test-systems components and subunit vaccine production.

How to cite: Grishin DV, Sidorov NG, Parfenova OK, et al. Construction, heterological expression and a simple purification of the *BP* region of the pneumococcal surface protein a fused in different orientations to the chemotaxis adaptor protein CheW from *Thermotoga petrophila*. Electron J Biotechnol 2024;71. https://doi.org/10.1016/j.ejbt.2024.05.001.

1. Introduction

Fundamental and applied interest in chimeric recombinant proteins is associated with the possibility of their use to obtain new medicinal and food polypeptides with new and interesting properties, as well as to simplify the technological process of isolating proteins of biotechnological value [1,2].

Currently, cells from various biological sources are widely utilized to obtain such proteins. However, *E. coli* is still the most frequently used microorganism for the high-level production of a wide variety of recombinant proteins [3,4,5]. Meanwhile, this expression system has a number of disadvantages, among which it is possible to distinguish the frequent aggregation of foreign proteins into insoluble inclusion bodies, which leads to a significant complication of the purification procedure of the target product from cell lysate. However, the advantages of *E. coli* such as easy manipulation and scaling determine the relevance of developing new approaches aimed at optimizing the production of recombinant proteins in such cells. Various areas of genetic and protein engineering provide an opportunity to implement new individual strategies in accordance with the requirements of a particular technological process. One of these areas is the protein hybridization, otherwise called fusion technology. The basis of this technology is a gene fusion or new genes formed by joining parts of two or more previously independent protein-coding genes. With this methodology, different fusion systems (carbohydrate-binding modules, proteases, immunoglobulins, stress proteins, etc.) are used for various technological purposes and tasks: for improving the heterologous expression and solubility of recombinant protein, optimizing its purification from cellular debris, for increasing a protein's immunogenicity, etc. [6,7,8].

A number of studies demonstrate that fusion to thermotolerant carrier proteins can affect the thermal stability of mesophilic proteins hybridized with it, which can be useful from a biotechnological point of view [9]. Thus, in Huang's scientific work, thermally stable maltodextrin-binding proteins of *Pyrococcus furiosus* (pfMBP) significantly increased the thermal stability of the target green fluorescent protein [6,9]. In addition, it has been shown that some thermoresistant proteins can also improve the solubilization parameters of fused partners [9]. It was noted that in most cases, in order to impart such useful properties, the thermostable domain:-target protein size ratio should be approximately 1.6:1 or 2:1 [9].

Such a strategy has become the subject of interest in applied biotechnology not so long ago. Therefore, it is essential to supplement the traditional solutions in this branch of science by expanding the libraries of effective fusion partners.

In the present work, special attention was paid to adaptive chemotaxis polypeptide of W-type (CheW, Cw) from hyperthermophilic bacteria of the genus *T. petrophila* [10] as a potential fusion system that could be used to obtain target polypeptides in soluble form and optimize their subsequent fine purification. The choice was due to the fact that some chemotaxis proteins, in addition to their main transport functions, can be involved in protein folding and their protection from stress. In other words, proteins such as CheW may have both chaperone-like properties [11] and are able to participate in interaction with chaperones, thereby affecting protein biosynthesis, which actualizes the problem that is being investigated in the study.

The partial surface adhesin from *S. pneumoniae* (PspA), which is of interest in the field of pharmacology and immunobiotechnology, was chosen as the target protein for the association [12,13]. The PspA was not full-sized, but represents a BP interdomain region including part of the most immunogenic B-region (CDR) and the part of the proline-rich domain (PRD) involving the non-proline block (NPB) (Fig. 1). These regions represent valuable vaccine targets due to their structure and potential cross-reactivity [12,13].

It should be underlined that these pneumococcal proteins belong to the category of non-thermophilic macromolecules, since its maximum thermal stability does not exceed 55°C [14].

The obtained knowledge will improve understanding of the functioning of such systems and will serve as an impetus for the development of new specialized tools for the production of biomedical valuable proteins in bacterial-producing systems.

2. Materials and Methods

2.1. Bacterial strains

The *E. coli* strain *XL1-Blue [recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacIqZdeltaM15 Tn10(TetR)]]* Stratagene (USA) and strain *BL21(DE3) pLysS E. coli str. B F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB +]K-12(λS) pLysS[T7p20 orip15A](CmR)* were used for cloning and expression procedures.

2.2. DNA and oligonucleotides

The commercial bacterial T7 promoter-based plasmid vector pET21a (Novagen, USA) contains the *bla* gene for ampicillin resistance and C-terminal histidine *tag* encoding sequence. This expression system is induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) and supports high level of transcription in *E. coli* cells that lysogenised with phage λ DE3 fragment encoding the T7 RNA polymerase. Short, synthetic single-stranded oligonucleotides known as primers were used in PCR, while long overlapping oligos were used to assemble intermediate fragments of target genes six oligos per stage. The length of each oligonucleotide for assembly did not exceed 98 bases. DNA oligonucleotides were commercially synthesized by the solid-state amidophosphite method by Evrogen (Moscow, Russia).

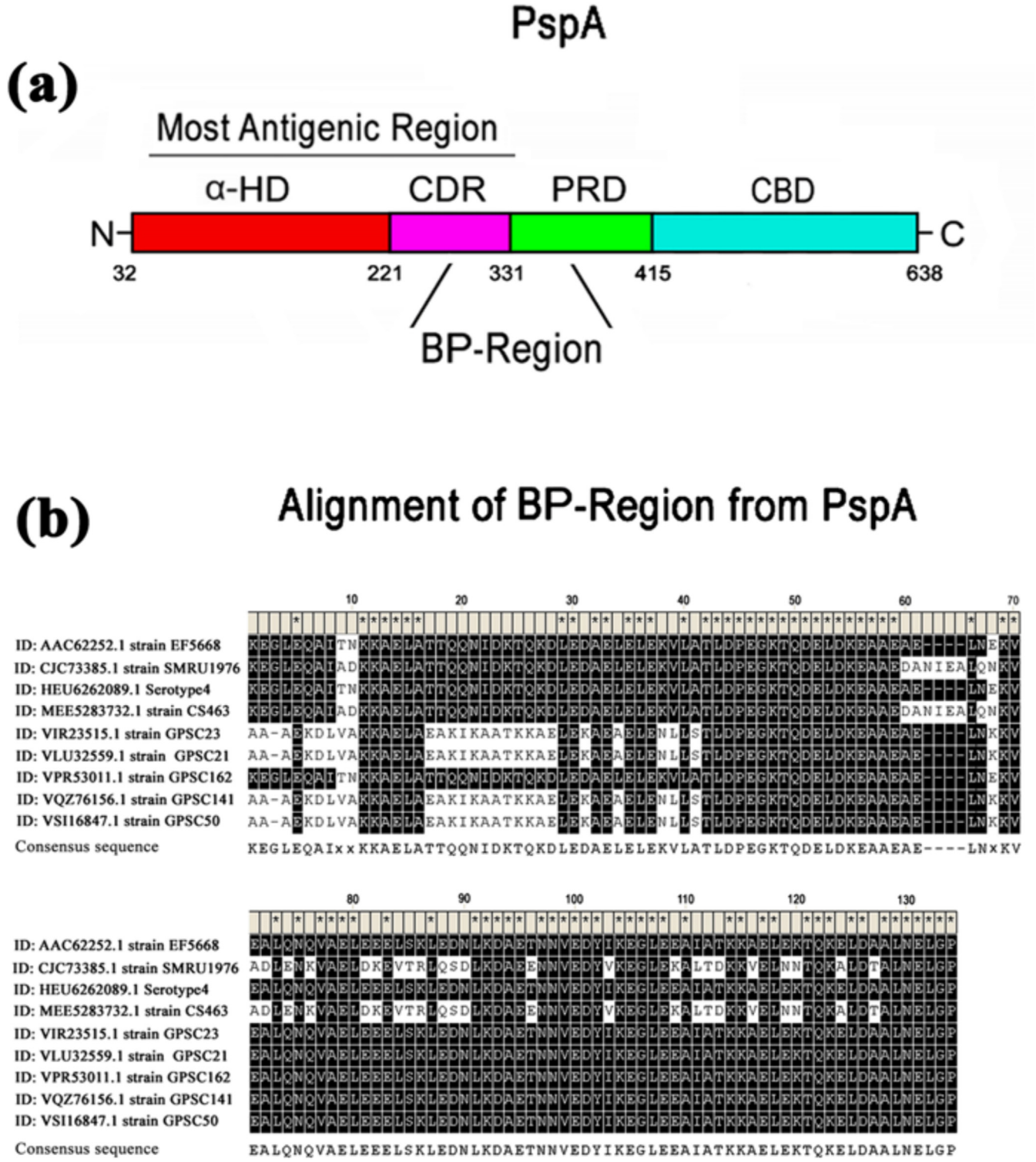


Fig. 1. Schematic diagram and amino acid sequence alignment. (a). Schematic diagram of primary structure of the pneumococcal surface protein A. (b). Multiple alignment of amino acid sequence for a BP interdomain region of PspA from different pneumococcal strains. The last row shows the consensus sequence.

2.3. Data analysis

Homologous Sequences related to target genes and proteins annotated in the GenBank international database were analyzed using the Basic Local Alignment Search Tool (BLAST, USA) [15] provided by the National Center for Biotechnology Information (NCBI, USA). DNASIS v 2.5 (Hitachi Software Engineering Co., Ltd.) and Clustal W [16] programs were used to align two or more amino acid or nucleotide sequences, as well as for subsequent visualization of sites with different homology. Cloning simulation and mapping of recombinant plasmids were performed using the Clone Manager 4.0 program (Scientific & Educational Software, USA). Polyacrylamide gel electrophoresis analysis was performed using ImageJ software [17,18,19].

The Phyre[2]software was used for the procedure of 3D modeling and verifying proteins based on their primary structure and the PDB framework of the closest homologs.

Hydropathy values (hydrophobicity) were calculated for proteins using the DNASIS v 2.5 (Hitachi Software Engineering Co., Ltd.).

A preliminary theoretical assessment of the solubility and thermal stability of target proteins by their primary amino acid sequence was conducted using the analytical programs "Recombinant Protein Solubility Prediction" and "Tm Predictor (TI)" according to the principle described in the article [20].

2.4. E. coli competent cell preparation and transformation

To prepare competent cells, the overnight culture was diluted 100 times in SOC medium (Dia-M, Russia) and grown at 37°C for 2–3 h in a bacterial shaker incubator ES-20, Biosan (Latvia). The bacterial culture was centrifuged in a Rotina 380R centrifuge (Hettich Zentrifugen, Tuttlingen, Germany) at 4,000 rpm for 5 min at 4°C. Plastic tubes with the remaining cell pellets were kept on ice and after that, they were carefully put back into suspension in 1 ml of a cold 50 mM solution of $CaCl_2$. This suspension was incubated at 0°C for 30 min and centrifuged again under similar conditions. The precipitate was resuspended in 100 µl of cooled 50 mM $CaCl_2$ solution and carefully flicked to mix cells and 5–10 µL of the diluted plasmid DNA or ligation mixture for transformation. After incubation for 30 min at 0°C, the mixture of competent cells and DNA was exposed to 42°C heat shock for 40–50 s and cooled to 0°C on ice for 1 min.One ml of preheated to 37°C SOC medium was added and the bacteria incubated with shaking at 37°C for 1 h. Bacterial cells were plated out onto LB agar with ampicillin at 100 µg/ml. The next day, discrete bacterial colonies were inoculated into 5 ml of LB medium with appropriate antibiotics for subsequent screening of plasmid DNA.

2.5. DNA isolation

Plasmid DNA was extracted from bacterial cells by the standard alkaline lysis procedure [21,22]. Purification of the DNA fragments of interest from TAE agarose gel was carried out using the "Cleanup S-Cap" kit (Evrogen, Russia) in accordance with the manufacturer's recommendations.

2.6. Construction of recombinant DNA molecules

Artificial DNA constructs were assembled from synthetic oligonucleotides (Evrogen, Russia) according to the Gibson method with minor modifications [23,24]. During subcloning, the ligation of restriction DNA fragments was carried out in a reaction volume of 10 µl, using the DNA ligase of phage T4 (50-100U/µl) (Thermo Fisher Scientific, USA) with 10X Overnight ligation buffer (Evrogen, Russia).To successfully junction DNA fragments, the resulting mixtures were incubated for 14–16 h at 14°C, under the condition that a recommended molar ratio of insert:vector DNA is 3:1–7:1. All assembled constructs encoding chimeric proteins were verified by PCR, restriction mapping and DNA sequencing.

2.7. Polymerase chain reaction

PCR was performed in the automatic mode in 0.5 mL thin-walled plastic tubes (Axygen, USA) on a thermal cycler "Tertsik" (DNA-technology, Russia). Reactions were set up as follows: 25 µl reactions with 2.5 µl of 1X Phusion buffer ("NEB", USA), 400 µM dNTP mix, 500 nM of each primer, 2 – 5 U of Phusion DNA polymerase ("NEB", USA) per sample and 1 µl of template DNA. Amplification parameters: 98°C – 30–60 s; (98°C – 10 s, 59–65°C – 30 s, 72°C – variable) x30; 72°C – 5 min; 10°C – storage.

2.8. Confirmation of recombinant DNA sequence

DNA samples were quantified using the *NanoDrop* One Spectrophotometer (Thermo Fisher Scientific, USA). DNA sequencing using fluorescent dye-labeled terminators was performed with an automatic sequencer ABI PRISM 3730xl ("Applied Biosystems", USA) in the company "Evrogen" (Russia). Nucleotide base signal ratios in complex sequencing data were analysed using Chromas 2.6.6 software (Technelysium Pty Ltd, Australia).

2.9. Expression and purification of fusion proteins

Recombinant *E. coli* cells were cultivated in 100 ml Erlenmeyer flasks containing 30 mL of LB medium. Before induction, the cells were intensively stirred in an orbital shaker ES-20 (Biosan, Latvia) at 37°C and 250 rpm rotational speed. A mixture of isopropyl β-D-1-thiogalactopyranoside (IPTG) was used as a chemical inducer of the expression of target proteins in cells of *E. coli* BL21(DE3) pLysS transformed with a plasmid containing a T7 promoter. When the cells reached an OD_{600} of 0.7–0.9, the inducer was added in the recommended amount of 0.5 mM IPTG. The strain was then moved to a 27°C shaker incubator and induced overnight. Uninduced samples were also run in parallel as negative controls.

Heat treatment (thermolysis procedure) was designed for simplify the purification protocol of thermotolerant proteins as well as remove any soluble aggregates. The biomass with the accumulated recombinant protein was suspended in 5 mL of thermolysis buffer (20 mMTris-HCl, pH (5.7–8.5), 50 mMNaCl, 1 mM EDTA, 0.1% Triton X-100). Ultrasonic cavitation at a frequency of 22 kHz using the sonicator equipment Scientz JY96-IIN (Drawell, Shanghai) was used to disintegrate bacterial cells. Samples were sonicated 2–3 times for 50 s with an interval of 20 s on ice. Cell lysates obtained after the disintegration of a cell population were incubated for 40 min at 75–99°C, and then, the cellular debris along with aggregates and denatured proteins could have been removed by centrifugation at 8500 rpm for 5 min.

2.10. Determination of the molecular weight of a protein by electrophoresis

All prepared samples have been studied by gel electrophoresis under denaturing conditions in the presence of sodium dodecyl sulfate (SDS) using 12% acrylamide gels (SDS-PAGE) and stained with Coomassie Brilliant Blue R-250 according to the procedure of Laemmli [25]. 2-mercaptoethanol was a reducing agent in the loading buffer. Maximum sample loading volume per well was 9 µL for samples without heat treatment, and 16 µl for samples with heat treatment. Prestained molecular weight marker PageRuler Low Range Protein Ladder ("Thermo Scientific") was used to monitor the molecular weight of studied and control proteins on SDS-PAGE.

3. Results and discussion

3.1. Bioinformatic analysis and codon optimization

As envisioned, constructed chimeric proteins were supposed to combine the adapter protein Cw of anaerobic thermotolerant bacteria *T. petrophila* (GenBank: ABQ46258.1) sized 153 amino acid residues and the immunogenic BP region of the surface antigen PspA from mesophilic *S. pneumoniae* (GenBank: AAC62252.1) sized 131 residues in N- and C-terminal orientation relative to Cw

domain. It matters because some fusion proteins could induce different effects when located at the N- or C-terminus of the joined partner [6,26,27]. The inert poly-Glycine-Serine linker (sp) was planned to allow spatial distance between multifunctional domains in the fusion proteins. It is proved that Gly–rich linkers are useful for various types of unstable interactions, where it is necessary to create a flexible covalent bond between protein domains to form a stable protein–protein complex. Such a linker does not impose any restrictions on the conformation or interactions of joined partner proteins [28,29,30].

If in the case of the protein from *T. petrophila,* we have a high intraspecific invariance of the primary amino acid sequence, then in the case of the target region of pneumococcal antigen A, a certain intraspecific amino acid polymorphism can be observed, which requires a program alignment of homologous sequences from different biovariants within the microorganism species in order to obtain a final consensus sequence from a multiple alignment.

Before designing N- and C-analogous of the chimeric gene, it was also necessary to take into account modern principles of optimization of the structure of genes to achieve their most efficient expression in *E. coli* cells [31,32,33]. During this process, the codon composition was optimized in the nucleotide sequences encoding native Cw and consensus PspA. As a result, optimized homologs of the chimeric genes PspA-sp-Cw-6His (PspACw) and Cw-sp-PspA-6His (CwPspA) were developed *in silico*. These variants became the basis for further practical work on the creation of genetically engineered DNA cassettes (Fig. 2).

3.2. Expression vector design and molecular cloning

The assembly of planned recombinant DNA constructs was carried out using the Gibson method based on DNA recombination, which enables stepwise to assemble *in vitro* several overlapping DNA fragments into a single molecule. At the same time, the plasmid vector pET21a (Novagen, USA) was first linearized at the required sites using specific restriction endonucleases or PCR, and then used for splicing with prearranged synthetic oligonucleotides. To accomplish this, the resulting equimolar mixture of DNA was mixed with 10 µl of a 2-fold Gibson Assembly Cloning Kit master mix (ISO buffer; Phusion DNA polymerase; Taq DNA ligase; T5 exonuclease) (NEB, USA). The reaction mixture was incubated at 50°C. At the next stage, the transformation of competent *E. coli* cells of 10 µl of the ligase mixture was carried out. Verification of assembled molecules was conducted by restriction mapping, PCR with corresponding primers (Cw-F: cat atg aaa acc ctg gcg gat gcg ctg aaa ga; Cw-R: acc tag ctc att cag ggc tgc gtc taa ctc ttt ct; Pas-F: atg aaa gaa ggc ctg gaa cag gcg att a; Pas-R: gcc agt ttc acg cct tct ttc acg gta), and sequencing. Thus, recombinant plasmids pET-PspACw and pET-CwPspA were generated, in which the genes of interest were controlled by highly effective regulatory elements: the T7 promoter recognized by the phage-derived T7 RNA polymerase and the T7 transcription terminator.

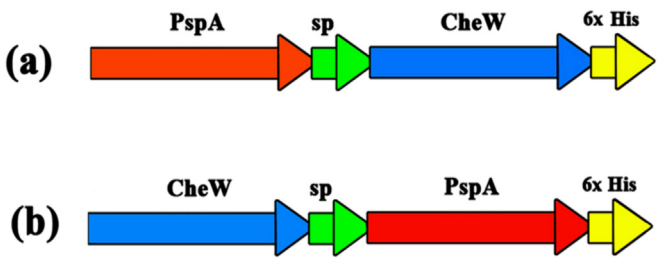


Fig. 2. Schematic diagram representing the construction of the chimeric gene cassettes of (a) PspACw and (b) CwPspA. His-tagged protein products with linkers sized 296 aa were encoded by 888 bp long gene.

3.3. Small-scale expression of recombinant PspACw and CwPspA

E. coli BL21(DE3) pLysS host cells have been made competent by using previously reported transformation solutions and then transformed with constructed uncut recombinant plasmid DNA as well as pET-TpeCheW plasmid which carried gene encoding the original thermostable Cw protein as one of the controls. A commonly used LB nutrient-rich medium was implemented for the cultivation of bacteria and higher biomass yields. The levels of inducibility and expression of target proteins in microbial cells were checked after the introduction of an inducing mixture into the medium, as described in the methodological section.

The results of SDS-PAGE revealed the appearance of a protein band of the expected molecular weight of approximately 34 kDa, corresponding to CwPspA hybrid protein (Fig. 3). In the case of the PspACw homolog, the electropherogram showed the occurrence of only minor bands at 68 – 70 kDa, which could indicate the self-association of this protein to form dimers. Perhaps, this is due to the fact that some proteins can still exist in the form of homodimeric or heteromeric complexes, even in the presence of reducing agents.

The expressed CwPspA protein represented over 27% of the total cell protein (TCP) content as determined by SDS-PAGE densitometry. At the same time, the expression level of PspACw was minor and did not exceed 6% of TCP. After purification, average protein yields for PspACw and CwPspA were in the range of 40 to 230 mg/L of culture, respectively.

It should also be noted that the growth of cells transformed with constructed recombinant plasmids before and after induction significantly differed from each other, which was clearly demonstrated in the diagram reflecting the dynamics of biomass accumulation (Fig. 4). Comparing these data with the final level of PspACw expression, it could be concluded that this protein was poorly tolerated by *E. coli*, possessing, apparently, a specific toxicity for prokaryotic cells.

3.4. Optimal pH and temperature range

A preliminary assessment of the solubility and thermal stability of the target chimeric proteins was performed using the analytical programs “Recombinant Protein Solubility Prediction” and “Tm Predictor (TI)”. For PspACw and CwPspA proteins, the calculated solubility rates exceeded the average values, while the TI values were > 1, which indicated potentially high thermal stability. To verify these data, induced *E. coli* cells were subjected to ultrasonic disintegration and thermolysis in a special buffer for 40 min at an elevated temperature of 75–99°C, followed by sedimentation of denatured proteins of host cell in the centrifugal field.

The crucial indicator for the constructed proteins is thermal stability, therefore, the main phase for the study was the supernatant, which was supposed to include exactly the most thermostable and soluble protein fractions. The assay showed that only the chimeric CwPspA protein had a sufficient level of thermal stability and was distributed mainly in the soluble fraction of cellular lysate (Fig. 3 and Fig. 5). A characteristic expanded smear appears only at a temperature of 99°C, and inconspicuous stripes indicate that this chimeric protein at a temperature of ≥99°C first begins to fall apart into fragments of smaller sizes and even into individual amino acids. Along the way, it was determined how much pH affects thermal stability at different temperatures. The CwPspA separation bands appeared the same picture and intensity on the gel for all pH ranges from 5.7 to 8.5 (shown at pH 7.5). Thus, the pH did not have a special effect on the thermal stability. This is a very promising result, since it will allow further use of this protein with different buffer systems and with physiological pH values.

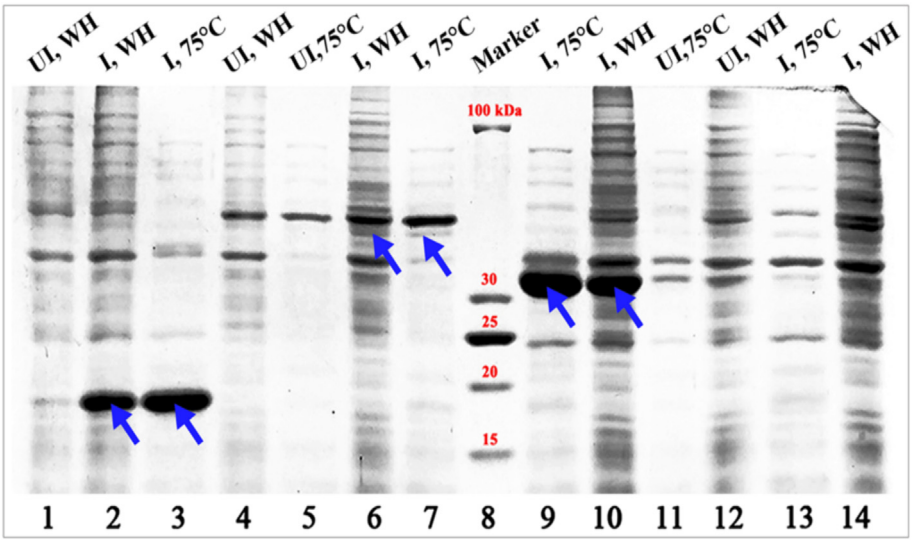


Fig. 3. SDS-PAGE analysis of soluble fractions of *E. coli* cells harboring pET-TpeCheW, pET- PspACw and pET- CwPspA. Lane 1–3: TpeCheW; Lane 4–7: PspACw; Lane 8: molecular weight marker; Lane 9–12: CwPspA; Lane 13–14: *E. coli BL21(DE3) pLysS* background control. The abbreviations above each line: I – induced, UI – un-induced, WH – without heating, 75°C – heating at 75°C. The blue arrow points to the corresponding protein band. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

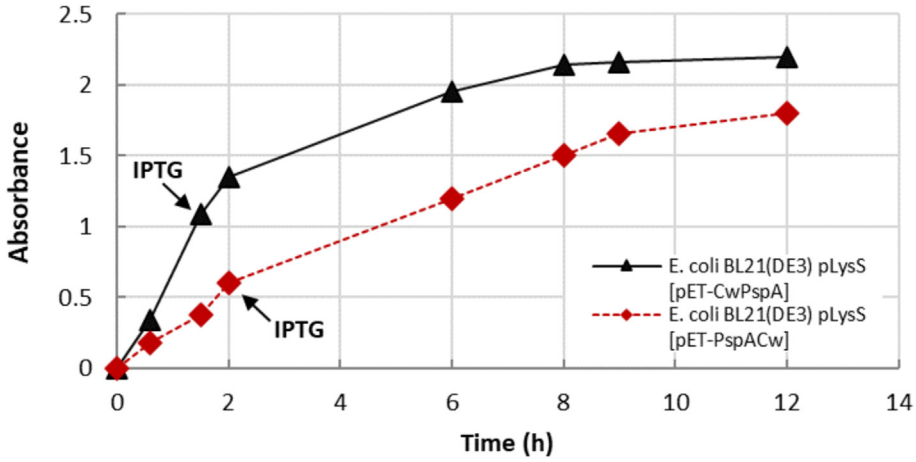


Fig. 4. Growth curves of BL21 (DE3) pLysS harboring the expression vectors of interest. Recombinant strains were cultured at 27°C; 0.5 mM IPTG was added in mid-log phase of cell growth; the OD600 values were determined at different time points.

3.5. Simple protein isolation and purification

To purify proteins, pellets of induced *E. coli* cells were suspended in a special thermolysis buffer (see **Materials and Methods**) at the rate of 10% of the culture volume. The cell suspension was treated with ultrasound in an intensive mode, on ice. The samples were heated at 80–99°C for 40–45 min, after which the cellular debris and denatured proteins were removed by centrifugation. Crystalline ammonium sulfate was added to the decanted supernatant to 80% saturation, and the mixture was left for 20 h at 4°C to concentrate recombinant proteins. The solution was purified by centrifugation (15 min, 14000 rpm); the precipitate was either lyophilized or dissolved in a storage buffer (20 mM Tris-HCl; pH 7.5; 10 mM EDTA; 10% glycerin) and stored at −20°C.

3.6. 3D structure prediction

The knowledge of proteins' tertiary structure is useful for the understanding of their functions or applications. Three-dimensional shapes for created chimeric proteins were reconstructed in a software package available on the Phyre^2 portal [34]. This server uses the alignment of hidden Markov models (HMM), simulating the operation of a process similar to the Markov process with unknown parameters, where the task is to decode unknown parameters based on the observed ones.

At the same time, to ensure maximum reliability, based on heuristic methods, templates are selected as objects for comparison, and an innovation is used – "*ab initio folding*" modeling to the reconstruction of protein regions without visible homology [34,35]. The generated 3D models were analyzed using a set of utilities of the Visual Molecular Dynamics (VMD) program for molecular visualization [36].

The three-dimensional model generated for the PspACw (Fig. 6a) shows that secondary structure of the protein consists of α-helices (approximately 47%) and β-strands (approximately 25%), which corresponded to the values calculated from the atomic structure of the protein. The CwPspA protein (Fig. 6b) accounts for 43% α-Helices and 21% β-sheets. Hydropathy values of peptides are

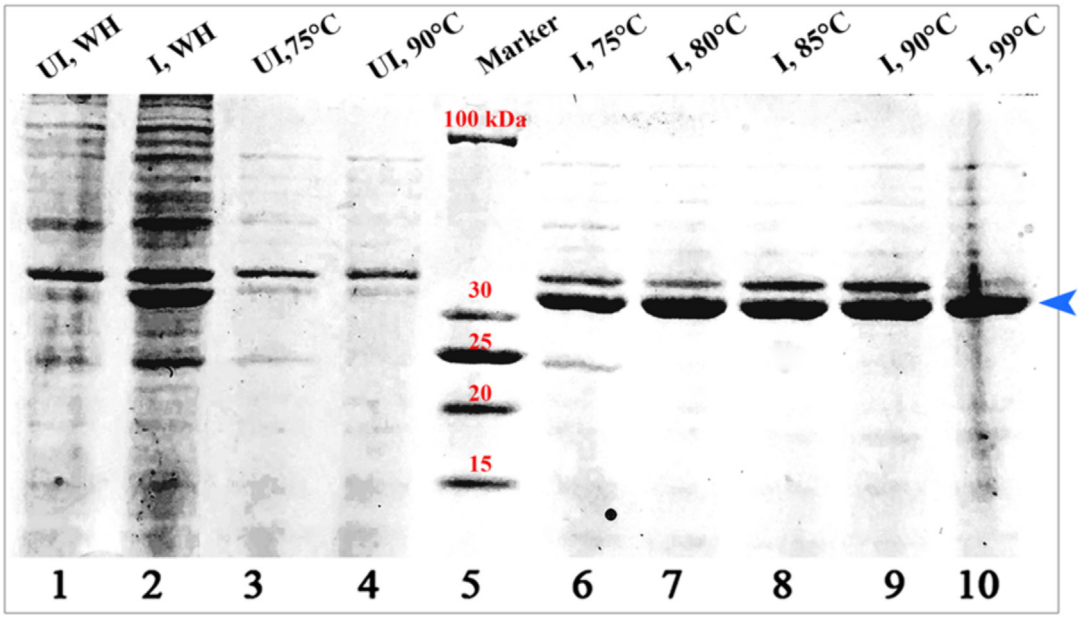


Fig. 5. SDS-PAGE of the thermolytic purified recombinant protein CwPspA. Lane 1–4: negative controls of induction or heating of CwPspA; Lane 5: molecular weight marker (M); Lane 6–10: the studied samples of induced CwPspA with thermal treatment. The abbreviations above each line: I – induced, UI – un-induced, WH – without heating, 75°C – heating at 75°C, 80°C – heating at 80°C, 85°C – heating at 85°C, 90°C – heating at 90°C, 99°C – heating at 99°C. The blue arrow points to corresponding protein bands. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

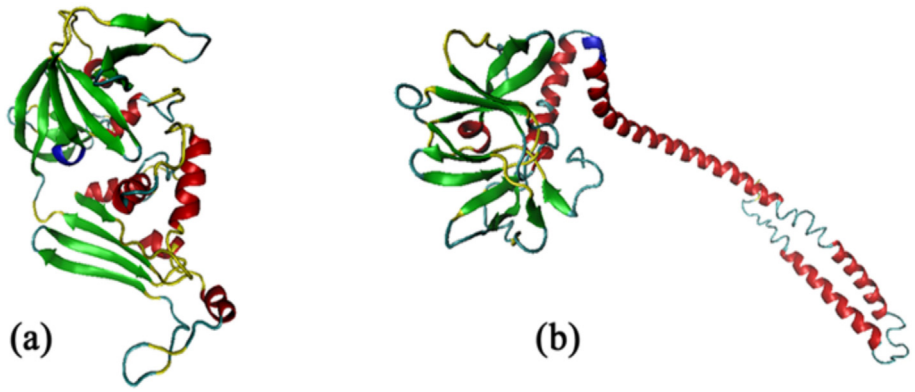


Fig. 6. Comparison of the 3D models of recombinant proteins with N- and C-terminal PspA exodomain location. (a) Predicted three-dimensional structure of PspACw chimeric protein. (b) Predicted three-dimensional structure of CwPspA chimeric protein. The predicted structures were obtained with Phyre[2] [34].

calculated based on the Hopp and Woods hydrophilicity scale [37], according to that, the average hydropathy value for the above-mentioned chimeric proteins was 0.76 and 0.28, respectively.

4. Conclusions

In the course of work, artificial genetic constructs have been developed to produce fusion proteins that did not exist in nature in such a domain combination before. These proteins contain one copy of the thermostable adaptive polypeptide CheW (Cw) of the hyperthermophilic microorganism *T. petrophila* and a copy of the immunogenic BP inter-domain of PspA antigens from *S. pneumoniae* in the N- and C-terminal mutual arrangement. The introduction of glycine-serine-rich linkers at the points of joining the CheW and PspA genes was intended to provide good flexibility and solubility of the final sequence, correct orientation and distancing of the two fused domains to allow their independent folding in the chimeric protein.

This study reveals new possibilities for using the heat-tolerant chemotaxis polypeptide Cw from *T. petrophila* as a very effective partner for fusions with some mesophilic proteins. Such an association gives the mesophilic domain additional thermal stability being as part of the chimeric protein. The principles for this phenomenon remain unclear. But thanks to the knowledge that the conformational stability of proteins may also depend on resonance energy transfer (RET) [38,39,40], it can be assumed that the phenomenon of inductive thermal stability may be based on effects similar to the RET described by Theodor Förster for a number of molecules.

The obtained C-terminal homolog of the hybrid protein (CwPspA) had molecular weight fully corresponding to the theoretically calculated one (≈ 34 kDa), overexpressed in soluble form to high level (≥27% of TCP), had the high-temperature resistance in the wide pH range (75–99°C, 40 min). Only at a temperature of 99°C, a part of the protein begins to break down after 40 min exposition. According to the results of the densitometric analysis, it has been shown that the purification of such protein can be minimized

to a low-stage process, which is based on extremely high-temperature treatment. Proteins purified in a similar manner possessed a sufficiently high purity (≥90%) and represent a well-defined protein band, which shows its homogeneity.

The hydrophilicity profiles (not shown) for both variants of engineered proteins are equally distributed amongst the entire molecule with uniformly alternating zones of higher and lower hydrophilicity. Meanwhile, the average values of the hydropathy coefficient on the Hopp–Woods scale for the CwPspA were slightly lower, which may indicate its higher hydrophobicity.

Computer-based 3D modeling showed the presence of a very compact structure of the PspACw stabilized by tertiary interactions between secondary structure elements. In this case, the pneumococcal antigen appeared inside the protein globule. The analysis suggested that the C-terminal location of PspA region leads to their independent folding, without being enclosed inside the whole globule of the CwPspA protein. From the standpoint of immunobiotechnology (subunit vaccines and diagnostic test-systems components), this is an undoubted advantage, since this domain is well presented and remains sterically accessible.

Our results show that the fusion protein with a C-terminal location of PspA part provides an excellent expression, thermal stability and solubility benefit when compared to an analogous fusion in which PspA is the N-terminal fusion partner.

Therefore, the obtained scientific data and designed chimeric protein CwPspA have the necessary conditions that offer new opportunities for its subsequent application in various areas of biotechnology and immunology.

Financial support

This work was supported by the Ministry of Science and Higher Education of the Russian Federation within the framework of the Program for Basic Research in the Russian Federation for a long-term period (2021–2030) [No. 122022800499-5]; the work was partially supported by the equipment of the Collective Usage Center "I. I. Mechnikov NIIVS", Moscow, Russia [No. 075-15-2021-676].

Conflicts of interest

The authors declare no conflicts of interest.

CRediT authorship contribution statement

Dmitry V. Grishin: Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Nikita G. Sidorov:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Olga K. Parfenova:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Roman V. Kurkin:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Data curation. **Ekaterina Y. Kasap:** Writing – original draft, Investigation, Formal analysis, Data curation, Conceptualization.

Acknowledgments

We would like to express our additional gratitude to Olga Parfenova and Nikita Sidorov for their help with translating the materials of the article into English.

Supplementary material

https://doi.org/10.1016/j.ejbt.2024.05.001.

Data availability

Data will be made available on request.

References

[1] Gupta V, Sengupta M, Prakash J, et al. Production of recombinant pharmaceutical proteins. In: Basic and Applied Aspects of Biotechnology. Singapore: Springer; 2016. p. 77–101. https://doi.org/10.1007/978-981-10-0875-7_4.
[2] Niazi SK, Magoola M. Advances in *Escherichia coli*-based therapeutic protein expression: mammalian conversion, continuous manufacturing, and cell-free production. Biologics 2023;3(4):380–401. https://doi.org/10.3390/biologics3040021.
[3] Kleiner-Grote GRM, Risse JM, Friehs K. Secretion of recombinant proteins from *E. coli*. Eng Life Sci 2018;18(8):532–50. https://doi.org/10.1002/elsc.201700200. PMid: 32624934.
[4] Zhang ZX, Nong FT, Wang YZ, et al. Strategies for efficient production of recombinant proteins in *Escherichia coli*: Alleviating the host burden and enhancing protein activity. Microb Cell Fact 2022;21(1):191. https://doi.org/10.1186/s12934-022-01917-y. PMid: 36109777.
[5] Cai K, Tu W, Liu Y, et al. Novel fusion antigen displayed-bacterial ghosts vaccine candidate against infection of *Escherichia coli* O157:H7. Sci Rep 2015;5:17479. https://doi.org/10.1038/srep17479. PMid: 26626573.
[6] Costa S, Almeida A, Castro A, et al. Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: The novel Fh8 system. Front Microbiol 2014;5:63. https://doi.org/10.3389/fmicb.2014.00063.
[7] Yadav DK, Yadav N, Yadav S, et al. An insight into fusion technology aiding efficient recombinant protein production for functional proteomics. Arch Biochem Biophys 2016;612:57–77. https://doi.org/10.1016/j.abb.2016.10.012. PMid: 27771300.
[8] Ki MR, Pack SP. Fusion tags to enhance heterologous protein expression. Appl Microbiol Biotechnol 2020;104(6):2411–25. https://doi.org/10.1007/s00253-020-10402-8. PMid: 31993706.
[9] Luke JM, Carnes AE, Sun P, et al. Thermostable tag (TST) protein expression system: Engineering thermotolerant recombinant proteins and vaccines. J Biotechnol 2011;151(3):242–50. https://doi.org/10.1016/j.jbiotec.2010.12.011. PMid: 21168452.
[10] Grishin DV, Samoilenko VA, Gladilina YA, et al. Effect of heterologous expression of chemotaxis proteins from genus *Thermotoga* on the growth kinetics of *Escherichia coli* cells. Bull Exp Biol Med 2019;167(3):375–439. https://doi.org/10.1007/s10517-019-04530-z. PMid: 31346881.
[11] Baker MD, Wolanin PM, Stock JB. Signal transduction in bacterial chemotaxis. Bioessays 2006;28(1):9–22. https://doi.org/10.1002/bies.20343. PMid: 16369945.
[12] Akbari E, Negahdari B, Faraji F, et al. Protective responses of an engineered PspA recombinant antigen against *Streptococcus pneumoniae*. Biotechnol Rep 2019;24:. https://doi.org/10.1016/j.btre.2019.e00385. PMid: 31763198e00385.
[13] Daniels CC, Coan P, King J, et al. The proline-rich region of pneumococcal surface proteins A and C contains surface-accessible epitopes common to all pneumococci and elicits antibody-mediated protection against sepsis. Infect Immun 2010;78(5):2163–72. https://doi.org/10.1128/IAI.01199-09. PMid: 20194601.
[14] Bom APDA, Corrêa IBS, Argondizzo APC, et al. Conformational analysis of Pneumococcal Surface Antigen A (PsaA) upon zinc binding by fluorescence spectroscopy. An Acad Bras Cienc 2018;90(2):2299–310. https://doi.org/10.1590/0001-3765201820170151. PMid: 29947666.
[15] Randić M, Pisanski T. Protein alignment: Exact versus approximate. An illustration J Comput Chem 2015;36(14):1069–74. https://doi.org/10.1002/jcc.23892. PMid: 25800773.
[16] Larkin MA, Blackshields G, Brown NP, et al. Clustal W and Clustal X version 2.0. Bioinformatics 2007;23(21):2947–3298. https://doi.org/10.1093/bioinformatics/btm404. PMid: 17846036.
[17] Pillai-Kastoori L, Schutz-Geschwender AR, Harford JA. A systematic approach to quantitative Western blot analysis. Anal Biochem 2020;593:. https://doi.org/10.1016/j.ab.2020.113608. PMid: 32007473113608.
[18] Alonso Villela SM, Kraïem H, Bouhaouala-Zahar B, et al. A protocol for recombinant protein quantification by densitometry. Microbiology Open 2020;9(6):1175–82. https://doi.org/10.1002/mbo3.1027. PMid: 32255275.
[19] van Oeffelen L, Peeters E, Nguyen Le Minh P, et al. The 'Densitometric Image Analysis Software' and its application to determine stepwise equilibrium constants from electrophoretic mobility shift assays. PLoS One 2014;9(1):. https://doi.org/10.1371/journal.pone.0085146. PMid: 24465496e85146.
[20] Grishin DV, Pokrovskaya MV, Podobed OV, et al.. Prediction of protein thermostability from their primary structure: the current state and development factors. Biomed Khim 2017;63(2):124–31. https://doi.org/10.18097/PBMC20176302124. PMid: 28414283.

[21] Delaney S, Murphy R, Walsh F. A comparison of methods for the extraction of plasmids capable of conferring antibiotic resistance in a human pathogen from complex broiler cecal samples. Front Microbiol 2018;9:1731. https://doi.org/10.3389/fmicb.2018.01731. PMid: 30150971.
[22] Figueroa-Bossi N, Balbontín R, Bossi L. Preparing plasmid DNA from bacteria. Cold Spring Harb Protoc 2022;2022:. https://doi.org/10.1101/pdb.prot107852. PMid: 35960622Pdb.prot107852.
[23] Avilan L. Assembling multiple fragments: The Gibson Assembly. In: Scarlett G, editor. DNA Manipulation and Analysis Methods in Molecular Biology. New York, NY: Humana; 2023. p. 45–53. https://doi.org/10.1007/978-1-0716-3004-4_4. PMid: 36853455.
[24] Grishin DV, Gladilina YA, Aleksandrova SS, et al. Creation of thermostable polypeptide cassettes for amino acid balancing in farm animal rations. Appl Biochem Microbiol 2017;53(6):688–98. https://doi.org/10.1134/S0003683817060072.
[25] Brunelle JL, Green R. One-dimensional SDS-polyacrylamide gel electrophoresis (1D SDS-PAGE). In: Lorsch J. Methods in Enzymology, Academic Press, 2014;541:151–9. https://doi.org/10.1016/B978-0-12-420119-4.00012-4 PMid: 24674069.
[26] Raran-Kurussi S, Keefe K, Waugh DS. Positional effects of fusion partners on the yield and solubility of MBP fusion proteins. Protein Expr Purif 2015;110:159–64. https://doi.org/10.1016/j.pep.2015.03.004. PMid: 25782741.
[27] Ko H, Kang M, Kim MJ, et al. A novel protein fusion partner, carbohydrate-binding module family 66, to enhance heterologous protein expression in *Escherichia coli*. Microb Cell Fact 2021;20(1):232. https://doi.org/10.1186/s12934-021-01725-w. PMid: 34963459.
[28] Klein JS, Jiang S, Galimidi RP, et al. Design and characterization of structured protein linkers with differing flexibilities. PEDS 2014;27(10):325–30. https://doi.org/10.1093/protein/gzu043. PMid: 25301959.
[29] Reddy Chichili VP, Kumar V, Sivaraman J. Linkers in the structural biology of protein-protein interactions. ProteinSci 2013;22(2):153–67. https://doi.org/10.1002/pro.2206. PMid: 23225024.
[30] van Rosmalen M, Krom M, Merkx M. Tuning the flexibility of glycine-serine linkers to allow rational design of multidomain proteins. Biochemistry 2017;56(50):6565–74. https://doi.org/10.1021/acs.biochem.7b00902. PMid: 29168376.
[31] Zhong C, Wei P, Zhang YP. Enhancing functional expression of codon-optimized heterologous enzymes in *Escherichia coli* BL21(DE3) by selective introduction of synonymous rare codons. Biotechnol Bioeng 2017;114(5):1054–64. https://doi.org/10.1002/bit.26238. PMid: 27943233.
[32] Elena C, Ravasi P, Castelli ME, et al. Expression of codon optimized genes in microbial systems: Current industrial applications and perspectives. Front Microbiol 2014;5:21. https://doi.org/10.3389/fmicb.2014.00021. PMid: 24550894.
[33] Nieuwkoop T, Claassens NJ, van der Oost J. Improved protein production and codon optimization analyses in *Escherichia coli* by bicistronic design. Microb Biotechnol 2019;12(1):173–9. https://doi.org/10.1111/1751-7915.13332. PMid: 30484964.
[34] Kelley LA, Mezulis S, Yates CM, et al. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 2015;10:845–58. https://doi.org/10.1038/nprot.2015.053. PMid: 25950237.
[35] Söding J. Protein homology detection by HMM-HMM comparison. Bioinformatics 2005;21(7):951–60. https://doi.org/10.1093/bioinformatics/bti125. PMid: 15531603.
[36] Lundquist K, Herndon C, Harty TH, et al. Accelerating the use of molecular modeling in the high school classroom with VMD Lite. Biochem Mol Biol Educ 2016;44(2):124–9. https://doi.org/10.1002/bmb.20940. PMid: 26751137.
[37] Barman A, Deb B, Chakraborty S. Prediction of potential epitopes for peptide vaccine formulation against *Teschovirus A* using immunoinformatics. Int J Pept Res Ther 2020;26:1137–46. https://doi.org/10.1007/s10989-019-09916-1. PMid: 32435170.
[38] Sanchez KM, Schlamadinger DE, Gable JE, et al. Förster resonance energy transfer and conformational stability of proteins: an advanced biophysical module for physical chemistry students. J Chem Educ 2008;85(9):1253–2126. https://doi.org/10.1021/ed085p1253. PMid: 19756254.
[39] Ploetz E, Lerner E, Husada F, et al. Förster resonance energy transfer and protein-induced fluorescence enhancement as synergetic multi-scale molecular rulers. Sci Rep 2016;6:33257. https://doi.org/10.1038/srep33257. PMid: 27641327.
[40] Jena S, Tulsiyan KD, Kar RK, et al. Doubling Förster Resonance energy transfer efficiency in proteins with extrinsic thioamide probes: Implications for thiomodified nucleobases. Chem Eur J 2020;27(13):4373–83. https://doi.org/10.1002/chem.202004627. PMid: 33210381.