Genome-Scale Metabolic Model Based Engineering of *Escherichia Coli* Enhances Recombinant Single Chain Antibody Fragment Production

Aidin Behravan
Shahid Beheshti University of Medical Sciences School of Pharmacy
Atieh Hashemi (at_hashemi@sbmu.ac.ir)
Shahid Beheshti University of Medical Sciences School of Pharmacy https://orcid.org/0000-0001-7121-5306
Sayed-Amir Marashi
University of Tehran College of Science
Hamideh Fouladiha
University of Tehran College of Science

Research Article

Keywords: AntiEpEX-scFv, Escherichia coli, FVSEOF, Genome-scale metabolic model, glk, Glucokinase.

Posted Date: April 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-412220/v1

Abstract

Escherichia coli is an attractive and cost-effective cell factory for producing recombinant proteins such as scFvs. AntiEpEX-scFv is a small antibody fragment receiving considerable attention for the epithelial cell adhesion molecule (EpCAM) targeting. EpCAM is one of the first discovered-cancer-associated biomarkers highly expressed on various types of solid tumors. Hereby, a genome-scale metabolic model guided engineering strategy was proposed to recognize gene targets for improved antiEpEX-scFv production. Flux balance analysis and FVSEOF algorithm identified several potential genetic targets localized in the glucose import system and pentose phosphate pathway that probably guaranteed an improved yield of scFv. Among the targets predicted by the model, *glk* gene encoding glucokinase was selected to be overexpressed in the parent strain *Escherichia coli* BW25113 (DE3). Due to metabolic burden, scFv recombinant expression caused a remarkable decrease in the maximum specific growth rate of the transformed strain. By means of overexpressing *glk*, presumably increasing carbon flux through the PP pathway, the growth capacity of the *E. coli* recombinant strain was recovered. Moreover, the engineered strain with *glk* overexpression successfully increased scfv production. The titer of antiEpEX-scFv reached 235.41 ± 9.53 μg/mL (0.428 g/g DCW) in the engineered strain compared with the parent strain (110.236 ± 7.68 μg/mL; 0.202 g/g DCW). So, model-based prediction was experimentally validated. This approach can be considered for the improvement of other recombinant proteins production.

Introduction

Nowadays, majority of clinically valuable proteins are recombinant ones whose production is directly influenced by the metabolism of the producing cell factory. Bacteria and yeasts, are two of the most well established microbial cell factories characterized by high-yield production of recombinant proteins (Ferrer-Miralles and Villaverde 2013). Despite continuous efforts, there are still limiting barriers in overexpression of recombinant proteins that need to be overcome. Metabolic burden observed in producing cells leads to biomass yield reduction, which is part of a stress reaction triggered by protein overexpression and results in reduction of recombinant protein productivity. Circumventing these challenges, metabolic engineering is being successfully employed in developing industrial cell factories (Fernández-Cabezón and Nikel 2020). Using this approach, specific biochemical reactions can be modified by deletion and/or amplification of specific genes leading to regulatory circuits rearrangement (Herrgård et al. 2006). These modifications lead to better productivity of the strains by changing the flux of some metabolic reactions. Based on this strategy, a significant improve in production rate of numerous heterologous products has been reported. For example, heterologous lycopene was produced up to 102 mg/L in an engineered *Escherichia coli* strain. However, the systematical understanding of metabolic pathways as well as regulation mechanisms is highly needed for rational metabolic engineering.

System metabolic engineering strategies that use omics data, have been expanded to guide metabolic engineering. These novel approaches such as genome-scale metabolic model (GEMM)- guided engineering have been evolved building on enormous advances in system biology. Genome-scale modeling can quantitatively predict the cellular behavior at system level and try to recommend genetic

manipulations to regulate the relationship between target protein overexpression and biomass production so that strains with high growth and protein productivity can be attained. High accuracy in prediction of cell phenotype and minimization of consumed laboratory resources and time for developing productive strains are the most momentous advantages of the model-guided metabolic engineering (Orth et al. 2011). The flux variability analysis (FVA) or the Flux balance analysis (FBA) can be utilized to compute the *E. coli* system model by considering the maximum amount of biomass as an objective target to achieve potential genes for overexpression or down-regulation. FVA or FBA result a flux variability or a flux distribution value respectively. Various algorithms such as FSEOF (flux scanning based on enforced objective flux) (Choi et al. 2010) as a FBA-based method and FVSEOF (flux variability scanning based on enforced objective flux) (Park et al. 2012) and OptForce (Ranganathan et al. 2010) as two FVA-based methods have been employed in the metabolic engineering to determine appropriate genes to be amplified in order to improve industrial strains. In this study, scFv (single chain variable fragment) antibody against EpEX (EpCAM extracellular domain) which has drawn great attention in biomedicine for dual therapeutic and diagnostic applications (Eyvazi et al. 2018), was considered as a model protein to be overproduced in *E. coli* BW25113 (DE3). The FVSEOF method was used to determine appropriate modifications of the genes in order to increase flux towards antiEpEX-scFv overproduction. Based on the results obtained from *in silico* FVSEOF analysis, *glk* gene was selected to be overexpressed. Experimental tests were then performed to evaluate the effect of *glk* overexpression on antiEpEX-scFv overproduction.

Materials And Methods

Metabolic modelling and target gene prediction

The *i*JO1366 metabolic model of *E. coli* (Orth et al. 2011) was used in the COBRA Toolbox v2.0 (Schellenberger et al. 2011) under MATLAB 2014b (Mathworks, USA) with glpk as the solver. The metabolic reaction of antiEpEX-scFv production was added to *i*JO1366 according to the previously described method (under review). FBA was used to determine maximum theoretical amount of antiEpEX-scFv (0.034 $mg.gDCW^{-1}h^{-1}$) but this value can be achieved only when biomass leads to zero and cannot be used as a realistic constraint for further analysis, therefore FVSEOF method was used by considering a range of 0 to 0.017 $mg.gDCW^{-1}h^{-1}$ of antiEpEX-scFv as constraint range in order to achieve maximum biomass. FVSEOF discretizes the range (0 to 0.017 $mg.gDCW^{-1}h^{-1}$) into five consecutive FVAs and computes the reaction fluxes in the model to maximize biomass production as an objective function. Comparing the result of FVAs declares the reaction fluxes variation, most of the reactions in the model remain unchanged or with less than 0.1 change in flux rate which was ignored. Reactions with significant ascending fluxes are favorable for experimental investigations due to their positive effect on antiEpEX-scFv production.

Bacterial strains, plasmids, and cultivation conditions

Escherichia coli DH5α was used for recombinant plasmid cloning and *Escherichia coli* BW25113 (DE3) (kindly provided by Prof. Dr. Silke Leimkühler, University of Potsdam, Potsdam, Germany) (Bühning et al.

2017) was utilized for expression. The plasmid pETDuet-1 (gifted from Dr. Bandehpour, Shahid Beheshti University of Medical Sciences, Tehran, Iran) was utilized as a co-expression vector. The plasmid pETDuet-antiEpEX-scFv was previously constructed in our lab (Behravan and Hashemi 2021). *Pfu* DNA polymerase was provided from mxcell. T4 DNA ligase, protein molecular weight markers and restriction enzymes were obtained from Thermo Fisher Scientific (USA). DNA fragments were purified from agarose gel using the gel extraction kit (Roche Diagnostics GmbH, Germany). M9 minimal medium contained (per liter) 0.5g of NaCl, 6g of Na_2HPO_4, 3g of KH_2PO_4, 1g of NH_4Cl supplemented with 5 g/L glucose, 2 mM MgSO4, 0.01 mM FeCl3, 0.1 mM CaCl2, and 0.1 mL 1000x trace metals element (Teknova-USA) and LB medium composed of (per liter) 10 g of tryptone, 5 g of yeast extract, and 10 g of NaCl, were used as culture media. All other chemicals were purchased from merck in analytical grade.

Construction of recombinant plasmid

The *glk* gene was amplified from the genome of *E. coli* BW25113 (DE3) employing the primers glk-F, 5-CCGGAATTCTGAAGAATGACAAAGTATGC − 3 (the *EcoR*I site is marked) and glk-R, 5-AAACTGCAGCCCGATATAAAAGGAAGGAT − 3 (the *Pst*I site is marked). Cycling condition was 94°C for 3 min followed by 30 cycles of 94°C, 30 s; 56°C, 35 s, 72°C 1min and 30 s and 1 cycle of 72°C, 10 min. To generate plasmid pETDuet-glk, the PCR product was digested with restriction enzymes *Pst*I, *EcoR*I, and then ligated to pETDuet-1 treated with the same two enzymes. pETDuet-glk-antiEpEX-scFv expression plasmid was constructed by digesting the pGH vector carrying the antiEpEX-scFv gene with *Xho*I and *Nde*I to get the gene with a hexa-histidine tag in its C-Terminal, which was then ligated with pETDuet-glk treated by *Nde*I/*Xho*I. The restriction enzyme digestion assay and sequencing were used to confirm the constructs which were transformed into the chemically competent *E. coli* BW25113 (DE3) cells for recombinant protein expression.

Expression of antiEpEX-scFv

For antiEpEX-scFv expression, a single colony of BW25113 (DE3) harboring pETDuet-glk-antiEpEX-scFv was inoculated into 5 mL of LB medium supplemented with appropriate amounts of ampicillin and incubated for 18 h at 37°C with shaking (200 rpm). After centrifugation (6000×g for 5 min at 4°C), the pellet was resuspended into 100 mL of M9 minimal medium containing ampicillin (50 mg/mL). When cell density reached an OD 600nm of 0.8, the expression of antiEpEX-scFv was induced with 0.8 mM IPTG at 37°C. Using centrifugation (10000×g, 10 min, 4°C), cells were harvested after 24h. For initial determination of protein expression, the cell pellets were suspended in 30 ml of lysis buffer containing 50 mM Tris pH 7.5, 1 mM EDTA, 1 mg/ml lysozyme, 150 mM NaCl, 1% triton X1005 and sonicated for 30 min (20s ON, 10s OFF at 400 W). After centrifugation of the cell lysate (10000×g for 30 min at 4°C), protein samples were electrophoresed on a 15% SDS-PAGE gel and visualized using Coomassie Brilliant Blue G-250 Dye. By using a wet Trasbolt (Bio-Rad, USA), the proteins were electro-transferred from gel into the polyvinylidene difluoride (PVDF) membrane to perform western blot analysis. Transferred membrane was blocked in 5% nonfat milk for 1 hour and then was washed three times with TBST and then incubated in His-tag antibody (Sigma, UK) overnight. After washing again, the membrane was incubated in anti-mouse HRP conjugated immunoglobulin (Sigma, UK) as secondary antibody for two hours and then detected by

means of a solution of 3,3'-Diaminobenzidine DAB (Sigma, UK). The recombinant antiEpEX-scFv was purified using the Ni–NTA affinity chromatography column under denaturing conditions based on the manufacturer's protocol (Qiagen, Netherlands). Utilizing the bicinchoninic acid assay (BCA assay), the concentration of the purified protein was measured (Takara BCA Protein Assay Kit, Takara, Japan).

Growth profile and glucose analysis

To investigate cell growth profile, optical density at 600 nm was determined every hour, using spectrophotometer (E-Chrome Tech, Taipei, Taiwan). Logarithmic derivation of the optical density curve was used for calculation of growth rate. In order to determine the glucose concentration, one milliliter of sample from culture broth was harvested in 1 hour intervals. The supernatant was collected following centrifugation at 10000×g (10 min). The concentration of glucose was measured using a commercial enzymatic kit (Megazyme, Wicklow, Ireland).

Real-time PCR analysis

The relative expression of *glk* as a target gene was compared between *E. coli* BW25113/Duet-glk-scFv, wild type strain *E. coli* BW25113 and *E. coli* BW25113/Duet-scFv using RT-qPCR. *E. coli* strains were cultured in 50 ml of M9 medium and induced with 0.8 mM IPTG in OD_{600} = 0.8. After 3 hours, samples were collected and diluted to OD_{600} = 0.4. Based on the manufacturer's protocol, total RNA was extracted from bacterial cells utilizing Trizol reagent (Ambion). The purity and quantity of the isolated RNA were measured by Synergy HTX multimode reader (BioTek) and was stored at -80°C for further use. cDNA synthesis kit (YT450; Yekta Tajhiz Azma) was employed to synthesize cDNA according to the instruction provided by the manufacturer. Reviewing the literature, primers were selected and assessed for GC content, specificity, secondary structures, and amplicon size. Primers sequences synthesized by Metabion are presented in Table1. StepOne Real-Time PCR System (Applied Biosystems) was employed for SYBRGreen qPCR reactions in 48 well optical reaction plates. cDNA (0.5 ng/reaction) was used as a template for qPCR reactions with 5 µl SYBR Green PCR Master Mix (2×) (YT2551; Yekta Tajhiz Azma) and primers at 10 µM final concentration. Samples were exposed to thermal plan as follows: 95°C, 30 s followed by 40 cycles of 95°C, 5 s and 60°C, 30 s. The PCR reactions were done in three technical replicates for more accuracy. $2^{-\Delta\Delta Ct}$ method was used to evaluate relative gene expression against the reference gene.

Table 1
Sequences of RT-qPCR primers

Name	Oligonucleotide Sequences (5′3 <- ′)
16S	F: TACCGCATAACGTCGCAAGA
	R: AGTCTGGACCGTGTCTCAGT
glk	F: CTGTATTGCCATCGCTTGCC
	R: TTACCTTCGACCGGTTCTGC

Results

Prediction of overexpression targets

E. coli GEMM named *i*JO1366 was employed for prediction of metabolic engineering targets which can improve antiEpEX-scFv production. FVSEOF predicted ten metabolic reactions to enhance antiEpEX-scFv productivity via their overexpression. Two genes related to the importation of glucose to the cell (*gal*P and *glk*), four genes related to the pentose phosphate pathway (PPP) (*zwf*, *rpe*, *pgl*, *gnd*), two genes related to the folate biosynthesis pathway (*foc*A and *pur*U), and two genes related to alternative carbon metabolism sub-system (*xyl*A, *mak*) were suggested by FVSEOF analysis for overexpression. These reactions in metabolic pathway are illustrated in Fig. 1.

The transcriptome analysis showed that the *glk* gene was upregulated during the recombinant protein production in *E. coli* (Oh and Liao 2000). Therefore, in this study the *glk* gene was selected as the target gene for overexpression.

Evaluation of glk transcription level

In order to validate the overexpression of *glk*, real-time PCR experiment was employed. After 3 h of induction, the relative quantification of *glk* transcript revealed that *glk* gene was upregulated in *E. coli* BW25113/Duet-glk-scFv by 14.78-fold and 925.56-fold in comparison to the *E. coli* BW25113/Duet-scFv and the parent strain respectively.

The effect of glk overexpression on antiEpEX-scFv production

Glucokinase gene was amplified from *E. coli* BW25113 (DE3) genome by PCR reaction resulting in a 966 bp *glk* gene which can encode 321 amino acids with the molecular weight of about 35 kDa. *glk* and antiEpEX-scFv coding sequences were inserted into the first and second multiple cloning sites (MCS) of pETDuet-1 respectively to generate the plasmid pETDuet-glk-antiEpEX-scFv as shown in Fig. 2. The constructed pETDuet-glk-antiEpEX-scFv plasmid was confirmed by restriction enzyme digestion (Fig. 3) and sequencing and transformed into *E. coli* BW25113 (DE3). Recombinant *E. coli* BW25113/Duet-glk-scFv and *E. coli* BW25113/Duet-scFv were cultured in M9 minimal medium containing a suitable antibiotic to an OD_{600} of 0.8 and induced with 0.8 mM IPTG for 24 h. The SDS-PAGE analysis exhibited the presence of two separate protein bands with molecular weights of about 35 kDa (*glk*) and about 29 kDa (antiEpEX-scFv) (Fig. 4a). The expressed scFv protein was confirmed using western blot analysis and an antibody against a C-terminal histidine tag (6xHis-tag) (Fig. 4b). Recombinant antiEpEX-scFv protein was purified through affinity chromatography using Ni–NTA matrix (Fig. 4c). BCA analysis was used to calculate recombinant protein concentration. According to the concentration of the purified antiEpEX-scFv, the *glk* -overexpressed strain showed an increase in antiEpEX-scFv titer (235.41 ± 9.53 µg/mL; 0.428 g/g DCW), which was approximately 2.135 times higher than that in strain with no glk overexpression (110.236 ± 7.68 µg/mL; 0.202 g/g DCW) after 24 h post-induction cultivation (Fig. 4b). The results

suggested that the altered glucose metabolism by *glk* overexpression could improve the antiEpEX-scFv production.

Growth and glucose consumption profiles

To examine how co-expression of *glk* with antiEpEX-scFv can affect bacterial growth rate and glucose consumption rate, wild-type strain (BW25113), recombinant *E. coli* BW25113/Duet-glk-scFv and *E. coli* BW25113/Duet-scFv, were cultured in M9 minimal medium containing 5 or 10 g/L glucose at 37°C. IPTG in final concentration of 0.8 mM was added in OD_{600} = 0.8 for protein induction. In order to obtain growth profile, OD in 600 nm were measured for 24h with 1-hour intervals. Each measurement was performed in duplicate.

All strains in M9 minimal medium containing 5g/L glucose (Fig. 5a and 5b), grow logarithmically as long as glucose is available and when glucose is depleted, enter to the stationary phase (Fig. 5a). As shown in Fig. 5a, maximum specific growth rate in recombinant *E.coli* BW25113/Duet-scFv and BW25113/Duet-glk-scFv (μ_{max} = 0.462 ± 0.034 and μ_{max} = 0.552 ± 0.003 respectively) was lower than that in the parent strain (μ_{max} = 0.637 ± 0.013). Since recombinant protein production and cell growth share some common precursors, increased protein expression at the expense of decreased cell density maybe due to alteration of intracellular fluxes through the biomass precursors towards protein synthesis in recombinant strains. However, recombinant *E. coli* BW25113/Duet-glk-scFv has greater specific growth rate than the *E. coli* BW25113/Duet-scFv (Table 2) especially when more glucose is available in the media. As shown in Fig. 5C and presented in Table 2, μ_{max} is much higher for *E. coli* BW25113/Duet-glk-scFv (0.81 ± 0.043) in comparison to *E. coli* BW25113/Duet-scFv (0.592 ± 0.003) and wild type strain (0.729 ± 0.022). As expected, maximum specific growth rate of all strains in M9 minimal medium containing 10 g/L glucose, is higher than that in the medium supplemented with 5 g/L glucose.

Table 2
Specific growth rate of different strains in M9 minimal medium in different glucose concentrations

Strains	Growth rate (h^{-1}) in M9 + 5 g/L glucose	Growth rate (h^{-1}) in M9 + 10 g/L glucose
BW25113	0.637 ± 0.013	0.729 ± 0.022
BW25113-Duet-scFv	0.462 ± 0.034	0.592 ± 0.003
BW25113-Duet-glk-scFv	0.552 ± 0.003	0.81 ± 0.043

As shown in the Table 3, glucose consumption rate in *E. coli* BW25113/Duet-glk-scFv (1.145 ± 0.01) is greater than that in the *E. coli* BW25113/Duet-scFv (0.968 ± 0.02) and wild type strain (0.851 ± 0.05). Also, as illustrated in Fig. 5b, the required time for the complete consumption of glucose for parental strain and the *E. coli* BW25113/Duet-glk-scFv were 10 hours while *E. coli* BW25113/Duet-scFv needs two more hours to consume all the glucose in the medium. Interesting result from Fig. 5d is that the required time for complete consumption of glucose for parental strain and *E. coli* BW25113/Duet-glk-scFv is similar

(about 24 h) while *E. coli* BW25113/Duet-scFv doesn't consume all of the available glucose in the medium up to 24 h (Fig. 5d). It confirmed that the overexpressed *glk* gene functioned well in *E. coli* BW25113/Duet-glk-scFv.

Table 3
Glucose consumption rate in different strains in M9 minimal medium in different glucose concentrations

Strains	uptake rate (g/g DCW h) in M9 + 5 g/L glucose	uptake rate (g/g DCW h) in M9 + 10 g/L glucose
BW25113	0.851 ± 0.05	0.730 ± 0.034
BW25113-Duet-scFv	0.968 ± 0.02	0.706 ± 0.039
BW25113-Duet-glk-scFv	1.145 ± 0.01	1.175 ± 0.03

As a conclusion, the enhancement of the growth rate and the glucose consumption rate by overexpressing *glk* gene is considerable in the results.

Discussion

Development of the hosts that have desirable metabolic phenotypes and ability to produce heterologous products is an important issue in microbial metabolic engineering. Utilizing various algorithms, GEMM based approaches enabled scientists to recognize gene deletion or overexpression targets for developing cell factories. For example, using MOMA simulations, L-valine biosynthesis was successfully improved in an engineered *E. coli* strain (Park et al. 2007). Also, amplification of *idi* gene selected by FSEOF together with the *dxs* gene led to lycopene overproduction (Choi et al. 2010). However, metabolic phenotypes prediction after gene deletion is much simpler than that after gene amplification. Because the corresponding metabolic fluxes of the deleted gene can be assumed as zero, while, owing to complex regulation of the metabolic network, the corresponding fluxes of the amplified genes do not certainly increase. Moreover, the amount of increase in metabolic fluxes corresponding to the gene amplification is difficult to be predicted. In this study, in order to increase the flux towards antiEpEX-scFv overproduction, the *glk* gene was selected for amplification among the several targets predicted by FVSEOF.

According to our results, recombinant expression of scFv and glk resulted in a decrease in the maximum specific growth rate of the recombinant strains compared with the parent strain. A decrease in the growth rate is normally detectable in bacteria transformed with multicopy plasmids to produce a recombinant protein. Actually, plasmid DNA replication, plasmid-encoded mRNA synthesis and translation in bacteria frequently place a metabolic burden into the engineered strains that usually results in growth retardation (Flores et al. 2004). This metabolic burden may be due to the cell inability to supply the extra demand of energy and building blocks required for plasmid replication and foreign multicopy genes expression (Li and Rinas 2020). However, a significant increase was observed in the μ_{max} of the recombinant strains from 0.592 ± 0.003 in BW25113-Duet-scFv to 0.81 ± 0.043 in BW25113-Duet-glk-scFv when the expression of the *glk* gene was increased which was comparable to the wild-type strain (0.729 ± 0.022). *glk* gene

encodes the enzyme glucokinase catalyzing the ATP-dependent phosphorylation of the glucose that was imported by GalP. *glk* overexpression probably compensates the special metabolic demands of the engineered strains via increasing the carbon flux into the PP pathway. The PP pathway which is closely interconnected with glycolysis normally provides some of the required blocks for biosynthesis of histidine, nucleotide and aromatic amino acids e.g., erythrose-4-phosphate (E4P) and ribose-5- phosphate (R5P) (Stincone et al. 2015). Also NADPH, a power of biosynthetic reactions, was reduced in its oxidative branch (Christodoulou et al. 2018). In a similar study, engineering of the pentose phosphate pathway led to reduction of the metabolic load caused by the recombinant protein production (Flores et al. 2004). Moreover, using this approach, a significant positive effect was observed on the productivity of scFv producing strains in the current study. The *glk* -overexpressed strain produced approximately 2.135 times higher titer of scFv than the strain with no *glk* overexpression. So, the metabolic engineering target predicted in our study was validated via the improvement observed in the scFv production.

Recently, the integration of the computational methods and omics at the systems level can empower metabolic engineering. Employing DNA microarray, oh et al revealed that the overproduction of recombinant non-toxic LuxA could lead to the downregulation of *ppc, fba, gnd, and atpA* genes as well as upregulation of heat shock and *glk* genes in *E. coli* strains including JM109, MC4100, and VJS676A. Based on the transcriptome profile obtained in oh et al study, glucose kinase might have the major role instead of the phosphotransferase system (PTS) to provide glucose6-phosphate in protein overproducing condition in the *E. coli* cells (Oh and Liao 2000). On the other hand, overexpression of the recombinant proteins was shown to induce heat shock genes and rapid stress response. Interestingly, *glk* has been reported to play an essential role in bacterial stress responses. Although, this gene plays a minor role in glucose metabolism, but under stress condition like heterologous protein expression or growth in acidic condition, this glycolysis enzyme is required to supply sufficient level of glucose6-phosphate (Arora and Pedersen 1995; Zhang et al. 2020). So, *glk* seems to be a suitable target gene to be overexpressed to achieve increased recombinant protein productivity, which is consistent with our results.

Here, the GEMM-guided metabolic engineering strategy was used to improve the scFv production in *Escherichia coli* BW25113 (DE3). The engineered strain with *glk* overexpression successfully increased scFv production. The titer of antiEpEX-scFv reached 235.41 ± 9.53 µg/mL (0.428 g/g DCW) in the engineered strain compared with the parent strain (110.236 ± 7.68 µg/mL; 0.202 g/g DCW). Our method for the production of scFv is a successful example which can be considered for the improvement of other recombinant proteins production.

Declarations

Funding

This work was supported by the research deputy of Shahid Beheshti University of Medical Sciences in Tehran, Iran (IR.SBMU.PHARMACY.REC.1399.372).

Conflicts of interest/Competing interests

We have no conflict of interest to declare

Availability of data and material

Not applicable

Code availability

Not applicable

Authors' contributions

AH and SAM have made substantial contributions to conception and design. AB and HF carried out the experiments and analyzed the data. AB wrote the manuscript. AH reviewed and edited the article for spelling, grammar, and intellectual content. AH and SAM organized and supervised the whole project. AH provided the facilities and materials required for the project. All authors agreed to be responsible for the content of the work.

Ethics approval

Not applicable

Consent to participate

Not applicable

Consent for publication

All authors have approved the manuscript and agree with its submission to the World Journal of Microbiology and Biotechnology.

Acknowledgements

We thanked Prof. Dr. Silke Leimkühler and Dr. Bandehpour for providing *Escherichia coli* BW25113 (DE3) and plasmid pETDuet-1 respectively.

References

1. Arora K, Pedersen P (1995) Glucokinase of Escherichia coli: induction in response to the stress of overexpressing foreign proteins. Arch Biochem Biophys 319 2:574–578. https://doi.org/10.1006/abbi.1995.1333
2. Behravan A, Hashemi A (2021) Statistical optimization of culture conditions for expression of recombinant humanized anti-EpCAM single-chain antibody using response surface methodology. Res Pharm Sci 16:153–164. https://doi.org/10.4103/1735-5362.310522

3. Bühning M, Friemel M, Leimkühler S (2017) Functional Complementation Studies Reveal Different Interaction Partners of Escherichia coli IscS and Human NFS1. Biochemistry 56:4592–4605. https://doi.org/10.1021/acs.biochem.7b00627
4. Choi HS, Lee SY, Kim TY, Woo HM (2010) In silico identification of gene amplification targets for improvement of lycopene production. Appl Environ Microbiol 76:3097–3105. https://doi.org/10.1128/AEM.00115-10
5. Christodoulou D, Link H, Fuhrer T et al (2018) Reserve Flux Capacity in the Pentose Phosphate Pathway Enables Escherichia coli's Rapid Response to Oxidative Stress. Cell Syst 6:569–578.e7. https://doi.org/10.1016/j.cels.2018.04.009
6. Eyvazi S, Farajnia S, Dastmalchi S et al (2018) Antibody Based EpCAM Targeted Therapy of Cancer, Review and Update. Curr Cancer Drug Targets 18:857–868. https://doi.org/10.2174/1568009618666180102102311
7. Fernández-Cabezón L, Nikel PI (2020) Chap. 11 - Advanced metabolic engineering strategies for the development of sustainable microbial processes. In: New and Future Developments in Microbial Biotechnology and Bioengineering. Elsevier, pp 225–246. https://doi.org/10.1016/B978-0-444-64301-8.00011-1
8. Ferrer-Miralles N, Villaverde A (2013) Bacterial cell factories for recombinant protein production; expanding the catalogue. Microb Cell Fact 12:113. https://doi.org/10.1186/1475-2859-12-113
9. Flores S, de Anda-Herrera R, Gosset G, Bolívar FG (2004) Growth-rate recovery of Escherichia coli cultures carrying a multicopy plasmid, by engineering of the pentose-phosphate pathway. Biotechnol Bioeng 87:485–494. https://doi.org/10.1002/bit.20137
10. Herrgård MJ, Lee BS, Portnoy V, Palsson B (2006) Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in Saccharomyces cerevisiae. Genome Res 16:627–635. https://doi.org/10.1101/gr.4083206
11. Li Z, Rinas U (2020) Recombinant protein production associated growth inhibition results mainly from transcription and not from translation. Microb Cell Fact 19:83. https://doi.org/10.1186/s12934-020-01343-y
12. Oh M-K, Liao JC (2000) DNA Microarray Detection of Metabolic Responses to Protein Overproduction in Escherichia coli. Metab Eng 2:201–209. https://doi.org/10.1006/mben.2000.0149
13. Orth JD, Conrad TM, Na J et al (2011) A comprehensive genome-scale reconstruction of Escherichia coli metabolism-2011. Mol Syst Biol 7:535. https://doi.org/10.1038/msb.2011.65
14. Park JH, Lee KH, Kim TY, Lee SY (2007) Metabolic engineering of Escherichia coli for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. Proc Natl Acad Sci U S A 104:7797–7802. https://doi.org/10.1073/pnas.0702609104
15. Park JM, Park HM, Kim WJ et al (2012) Flux variability scanning based on enforced objective flux for identifying gene amplification targets. BMC Syst Biol 6:106. https://doi.org/10.1186/1752-0509-6-106

16. Ranganathan S, Suthers PF, Maranas CD (2010) OptForce: An optimization procedure for identifying all genetic manipulations leading to targeted overproductions. PLoS Comput Biol 6:. https://doi.org/10.1371/journal.pcbi.1000744
17. Schellenberger J, Que R, Fleming RMT et al (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nat Protoc 6:1290–1307. https://doi.org/10.1038/nprot.2011.308
18. Stincone A, Prigione A, Cramer T et al (2015) The return of metabolism: biochemistry and physiology of the pentose phosphate pathway. Biol Rev 90:927–963. https://doi.org/10.1111/brv.12140
19. Zhang W, Chen X, Sun W et al (2020) Escherichia coli Increases its ATP Concentration in Weakly Acidic Environments Principally through the Glycolytic Pathway. Genes (Basel) 11:991. https://doi.org/10.3390/genes11090991

Figures

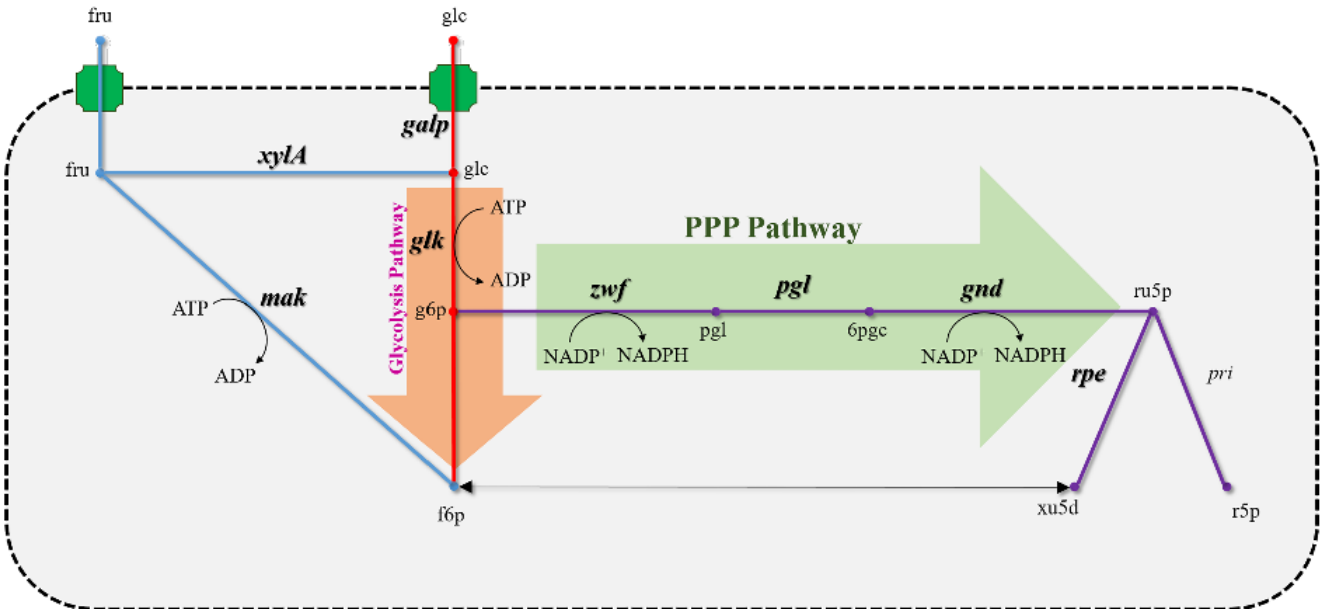

Figure 1

Overexpression targets illustrated in the metabolic map. The predicted targets for overexpression are indicated by their bold gene names.(PPP pathway: Pentose Phosphate Pathway, fru: Fructose, glc: Glucose, g6p: glucose-6-phosphate, f6p: Fructose-6-phosphate, pgl: 6-phosphogluconolacton, Ru5P: ribulose-5-phosphate, 6pgc: 6-phosphogluconate, R5P: ribose-5-phosphate, galP: galactose permease, zwf: glucose-6-phosphate dehydrogenase, xylA: xylose Isomerase, mak: fructokinase, glk: glucokinase, gnd: 6-phosphogluconate dehydrogenase, pgl: 6-phosphogluconolactonase, rpe: ribulose-5-phosphate-3-epimerase)

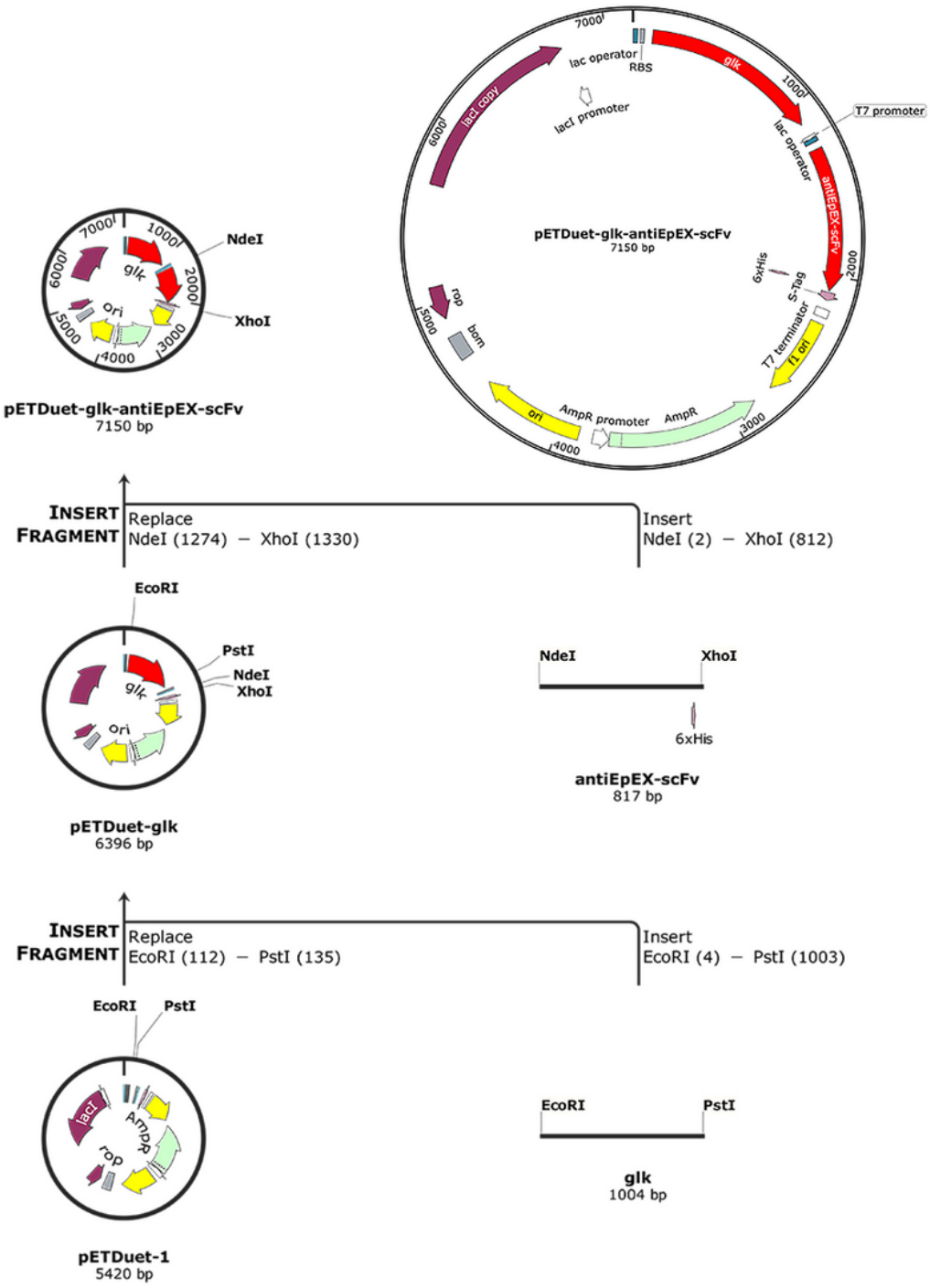

Figure 2

History and construction of the pETDuet-glk-antiEpEX-scFv expression vector: gene encoding glk was cloned between EcoRI and PstI, antiEpEX-scFv gene was cloned between NdeI and XhoI sites of pETDuet -1. Maps were designed with SnapGene Version 5.2.4

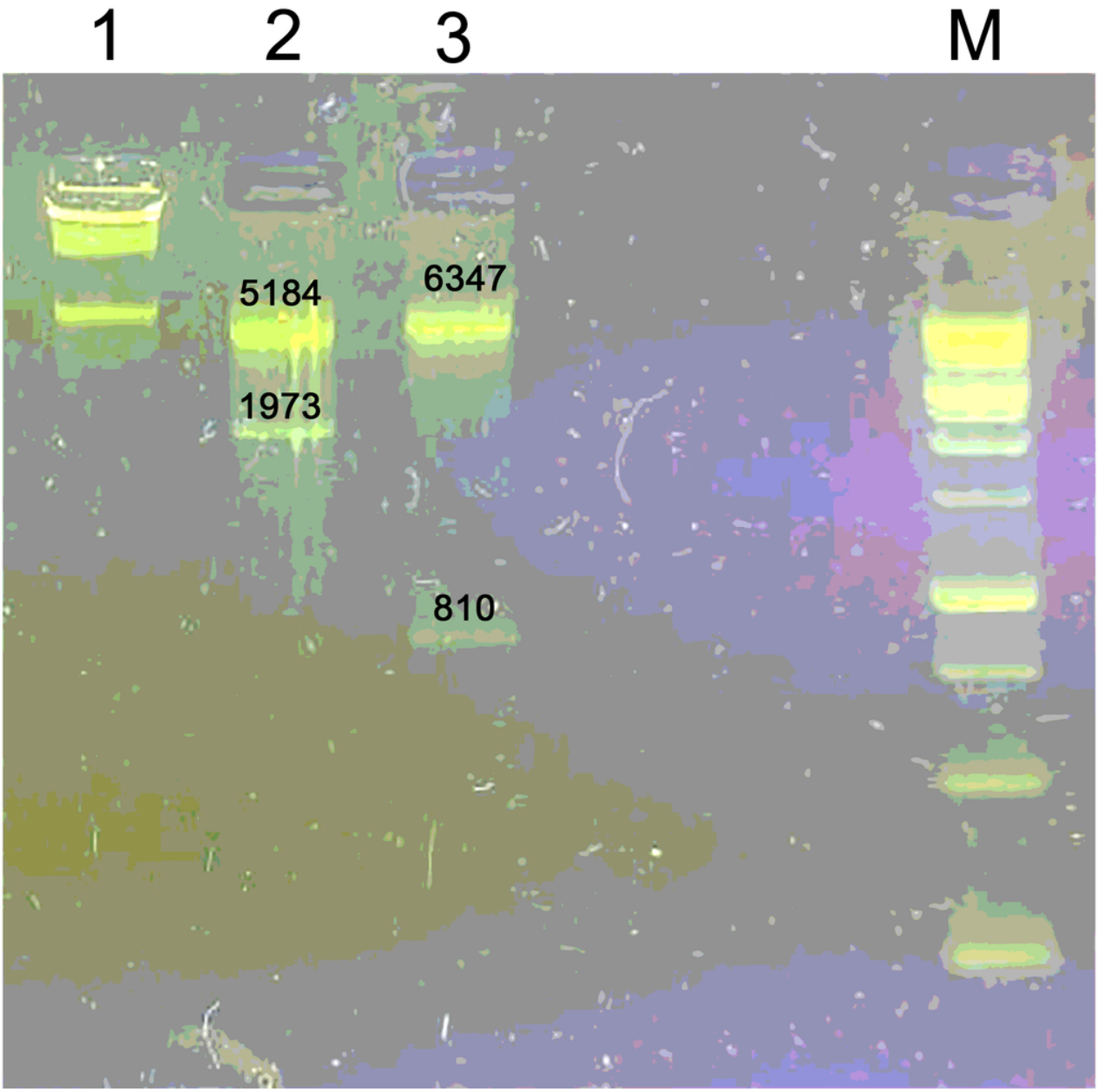

Figure 3

Confirmation of the recombinant plasmid pETDuet-glk-antiEpEX-ScFv by enzymatic digestion reaction: 1: plasmid pETDuet-glk-antiEpEX-scFv (7150bp); 2: recombinant plasmid pETDuet-glk-antiEpEX-scFv digested with EcoRI and PstI (6347bp and 1973bp); 3: recombinant plasmid pETDuet-glk-antiEpEX-scFv digested with XhoI and NdeI (5184bp and 810bp); M: DNA marker 1kb.

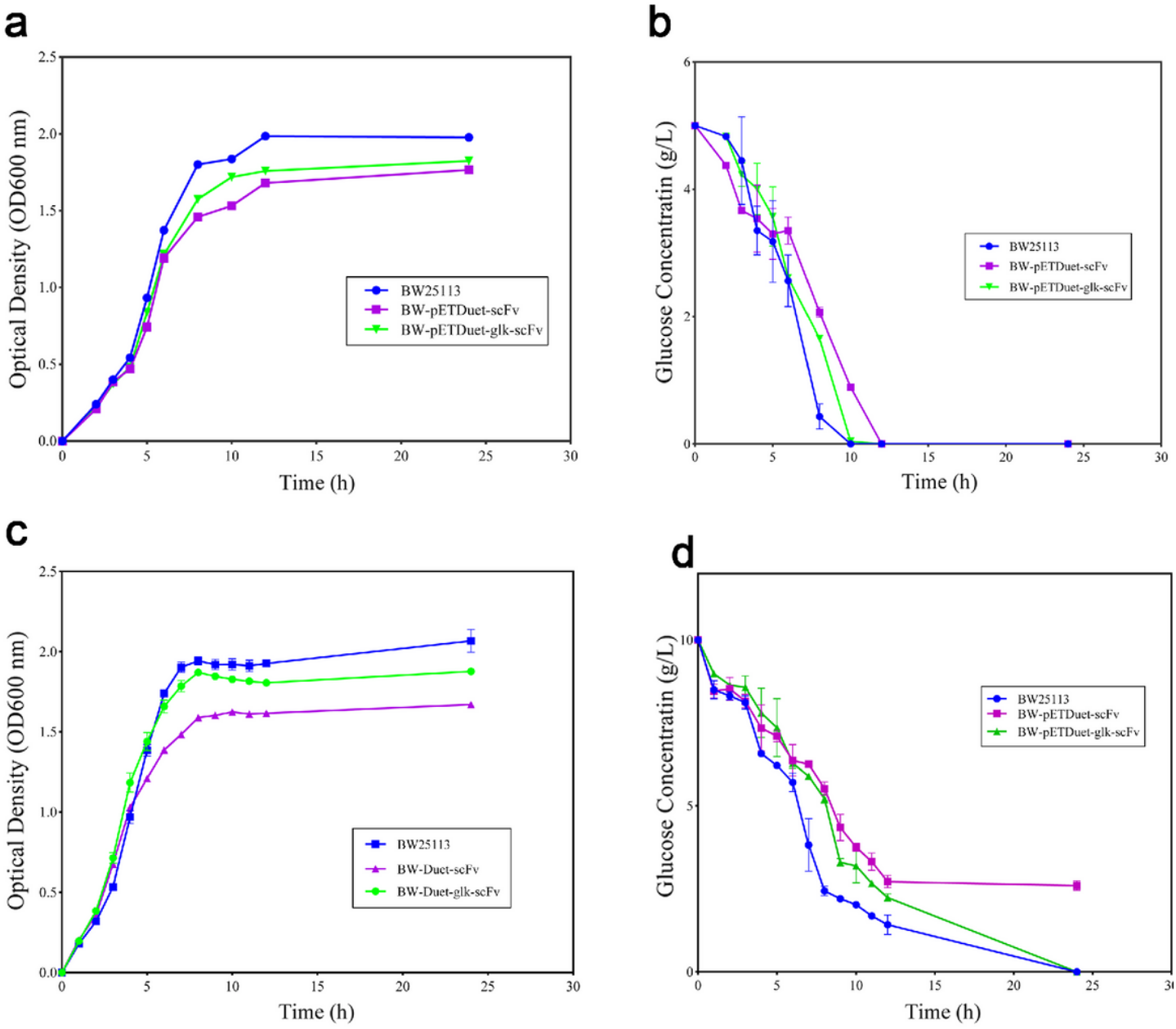

Figure 4

SDS–PAGE and western blot analyses for characterization of the antiEpEX-scFv recombinant protein in E. coli cell extracts. A: SDS–PAGE analysis of total lysate. (1,2): Induced total lysate of BW25113/Duet-scFv; (M): Molecular weight protein marker (14.4 – 116 kDa); (3,4): Induced bacterial lysate with the empty plasmid pETDuet-1, (5,6): Induced total lysate of E. coli BW25113/Duet-glk-scFv. B: Western blot analysis for the recombinant antiEpEX-scFv using anti-polyhistidine monoclonal antibody. (C-): Uninduced bacterial lysate; (M): Prestained molecular weight marker (10 – 250 kDa); (1): Induced total lysate of BW25113/Duet- scFv (2): Induced total lysate of E. coli BW25113/Duet-glk-scFv C: SDS–PAGE analysis of the purified antiEpEX-scFv. (M): Molecular weight protein marker (14.4 – 116 kDa); (C-) Uninduced bacterial lysate. (1) Eluted protein fraction from E. coli BW25113/Duet-scFv; (2) Eluted protein fraction from E. coli BW25113/Duet-glk-scFv. All experiments were done in M9 minimal medium

containing glucose as a carbon source. When OD600 reached to 0.8, cells were induced for 24 h using 0.8 mM IPTG. The experiments were performed in duplicates. The protein bands corresponded to antiEpEX-scFv and glk are shown by arrows.

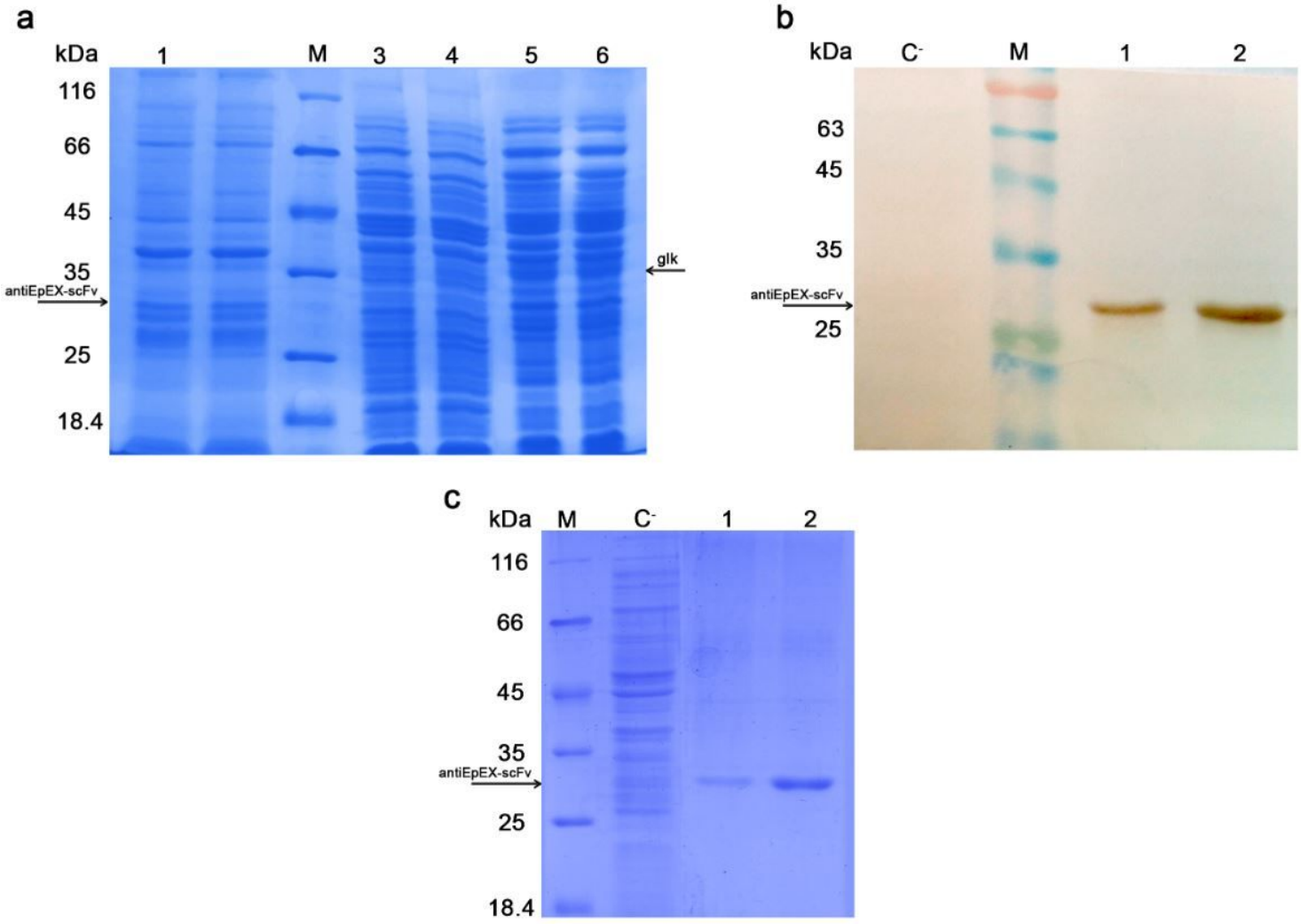

Figure 5

Growth profiles and glucose consumption rates. When OD600 reached to 0.8, cells were induced for 24 h using 0.8 mM IPTG. Growth and glucose consumption profiles of the wild type (BW25113) and the recombinant strains in M9 medium supplemented with 5 g/L glucose (A, B) and 10 g/L glucose (C, D). Error bars illustrated the standard deviation of two experimental replicates. All graphs are drawn using GraphPad Prism 8 software. Data are presented as mean ± SD, n = 2.