Biochemical and Biophysical Research Communications 419 (2012) 472–476

Contents lists available at SciVerse ScienceDirect

Biochemical and Biophysical Research Communications

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

Systematic metabolic engineering for improvement of glycosylation efficiency in *Escherichia coli*

Jagroop Pandhal, Pratik Desai, Caroline Walpole [1], Leyla Doroudi [1], Dmitry Malyshev, Phillip C. Wright *

ChELSI Institute, Biological and Environmental Systems Group, Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, UK

ARTICLE INFO

Article history:
Received 19 January 2012
Available online 10 February 2012

Keywords:
Glycosylation
Escherichia coli
Efficiency
Proteomics
Selective reaction monitoring

ABSTRACT

Recently, efforts to increase the toolkit which *Escherichia coli* cells possess for recombinant protein production in industrial applications, has led to steady progress towards making glycosylated therapeutic proteins. Although the desire to make therapeutically relevant complex proteins with elaborate human-type glycans is a major goal, the relatively poor efficiency of the N-glycosylation process of foreign proteins in *E. coli* remains a hindrance for industry take-up. In this study, a systematic approach was used to increase glycoprotein production titres of an exemplar protein, AcrA, and the resulting glycosylation efficiency was quantified using a combination of Western blots and pseudo Selective Reaction Monitoring (pSRM). Western blot and pSRM results demonstrate that codon optimising the oligosaccharyltransferase, PglB, for *E. coli* expression, increases efficiency by 77% and 101%, respectively. Furthermore, increasing expression of glycosyltransferase, WecA, in *E. coli* improves efficiency by 43% and 27%, respectively. However, increasing the amount of donor lipid used in the glycosylation process did not impact on the glycosylation efficiency in this system, with this specific protein.

1. Introduction

The ambition of using bacteria to make complex post-translationally modified human protein therapeutics remains a major challenge for the bioprocessing community. The motivation is to use production systems that are less costly than mammalian expression systems with a higher level of final product control, i.e. little or no heterogeneity, is leading to exciting progress. Although expression of large functional complex proteins in *Escherichia coli*, present problems of their own, steady progress has been made [1]. The production of smaller therapeutic fragment proteins, working alone or intended for therapeutic fusion proteins, such as antibody fragments, has been demonstrated as an attractive alternative [2–4]. Advantages of small therapeutic proteins include flexibility in structure which increases binding possibilities and enhanced penetration of tissues [2]. A significant advantage is the reduced costs of production in *E. coli*. The ability to perform bacterial N-glycosylation of target proteins in *E. coli* has been demonstrated using the well characterised N-glycosylation pathway from *Campylobacter jejuni* [5–7]. Although performed in an exemplar protein, AcrA, sourced from *C. jejuni*, the process has been shown to be functional although very inefficient [6,8,9]. In previous work, the glycosylation efficiency, i.e. the percentage of glycosylated target protein compared to total target protein, of AcrA was reported to be just 13% [8]. It is thought that the occupancy rate of a protein with a specific sugar structure could be dependent on factors such as the site of the glycosylation consensus sequence and export pathways, but there is not a general rule [10–12]. The Pgl pathway is not native to *E. coli* and therefore metabolic constraints may also play a significant role in glycosylation efficiency. For this reason, an iterative metabolic engineering strategy was developed to identify bottlenecks and highlight potential targets to improve glycosylation efficiency [8]. A discovery-driven proteomics workflow, using chemical tagging was employed with a mixture model on graphs (MMG) approach [13] to suggest pathways that could be altered. An increase in expression of isocitrate lyase in the glycoxylate shunt resulted in an increase in glycosylation efficiency of AcrA by almost 3-fold to 48% [8].

Recently, Schwarz et al. [9] glycosylated AcrA and two human antibody fragments F8 and CH2, using a system where the native enzyme undecaprenyl-phosphate alpha-*N*-acetylglucosaminyl 1-phosphate transferase (WecA) in *E. coli* was used to add the initial sugar *N*-acetylglucosamine (GlcNAc), the same initial sugar found in human *N*-glycoproteins. They reported glycosylation efficiencies of 40% and 5% for the antibody fragments, respectively. The overall titres of recombinant protein were very low (personal communication-Flavio Schwarz) and therefore improvements are required in overall recombinant protein production as well as glycosylation efficiency. The research presented here was carried out to address the glycosylation efficiency issue.

* Corresponding author. Address: ChELSI Institute, Department of Chemical and Biological Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK. Fax: +44 (0)114 2227501.
E-mail address: p.c.wright@sheffield.ac.uk (P.C. Wright).
[1] These authors equally contributed to this work.

0006-291X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2012.02.020

In this work, the glycosylation efficiency of AcrA was tested using systematic genetic changes to the *E. coli* host cell system. A codon optimised oligosaccharyltransferase gene (*pglB*) was tested and quantified using Western blots and pseudo Selected Reaction Monitoring (pSRM) (see Pandhal et al. [8]). Western blots give an actual glycosylation efficiency percentage whereas pSRM confirms this value by providing fold difference changes in glycoprotein normalised by total protein. The effect of increasing WecA expression was also tested, as this glycosyltransferase recognises GlcNAc and attaches the sugar to asparagine residues [14]. Finally, the increase in expression of BacA was also tested. In this system, glycosylation involves the transfer of a heptasaccharide from an undecaprenyl-pyrophosphate donor to the asparagine side chain of proteins. BacA confers undecaprenyl pyrophosphate phosphatase activity [15], which potentially results in increased availability of undecaprenyl phosphate for glycoprotein synthesis. An increase in glycosylation efficiency from increased BacA presence would indicate this as a bottleneck in this system. The targets for engineering cells are shown in Fig. 1.

2. Materials and methods

All materials were purchased from Sigma–Aldrich (Dorset, U.K.) unless otherwise stated.

2.1. DNA cloning, PCR, mutagenesis and vectors

The chloramphenicol resistant vectors pACYC*pgl* [7] and pACYC*pgl2* [9] were implemented here to perform the N-glycosylation process. *PglB* is present on both vectors, but was removed in pACYC*pgl* to create pACYC*pgl* Δ*pglB*. Briefly, pACYC*pgl* was digested with BaeI (NEB, Herfordshire, UK) at 37 °C for 1 h. After confirmation by agarose gel electrophoresis, the linear DNA was digested with exonuclease BAL-31 (NEB) for 5, 10, 20 and 30 min at 30 °C. The reaction was stopped by addition of SureClean reagent (Bioline, London, UK). Overhangs were filled in using Phusion polymerase (NEB) following the manufacturer's instructions. After further clean-up using SureClean reagent, vectors from all four time points were re-ligated using T4 DNA ligase (NEB) with manufacturer's instructions for one hour at 37 °C. Plasmids were transformed into NEB5alpha competent cells (NEB) and a colony screen (forward primer GTGATAAAAATCCTATTCTC, reverse primer ACGCGATGCTTT GAAATATT) was performed. Briefly, sterile pipette tips were used to transfer colonies into 5 μL water for subsequent PCR and also to re-streak colonies on fresh LB agar (with antibiotics). The PCR program was as follows: 95 °C for 5 min, and 30 cycles of 95 °C for 30 s, annealing at 55 °C for 30 s and extension at 72 °C for times depending on the expected insert size. A final extension step of 72 °C for 10 min was included. Following agarose gel electrophoresis, PCR products that visually looked reduced in size were sent for sequencing at the Core Genome Facility (University of Sheffield). Plasmids with removed *pglB* were further tested phenotypically using Western blots to see if glycosylation was abolished (Western blot in Supplementary materials).

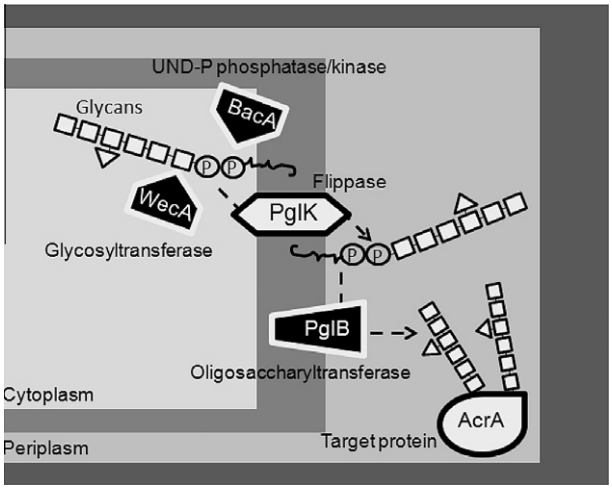

Fig. 1. A schematic representation of an *E. coli* cell harbouring N-glycosylation capability. The systematic cellular engineering targets are shown as black boxes with white borders. The order of steps involved in the N-glycosylation can be followed using the black dashed arrows. Briefly, a lipid linked oligosaccharide is built in the cytoplasm using WecA to add the initial GlcNAc sugar to the phosphorylated lipid. This is transferred using PglK flippase to the periplasm, where the oligosaccharyltransferase, PglB, recognises the structure and transfers it onto the target protein (AcrA) on the appropriate consensus sequence. P = phosphate group.

A codon optimised *pglB* gene was synthesised by DNA 2.0 (CA, USA) using the GeneDesigner software algorithm. The codon optimised sequence is given in the Supplementary materials, with screenshots of how the software is used. The gene was provided on vector pjexpress401 with kanamycin resistance and named pjexpress401*pglB**.

WecA was amplified from *E. coli* K12 DNA using the primer sequences presented in the Supplementary materials. PCR was performed as described previously [8], but with an annealing temperature of 57 °C. The amplified gene was purified using SureClean reagent and digested with restriction enzymes BamHI and XmaI for 1 h at 37 °C. The same digestion was performed on vector pjexpress401*pglB**. The vector backbone was purified using a gel extraction kit (Zymoresearch, Cambridge) and ligated to *wecA* using T4 DNA ligase. The vector was named pjexpress401*wecA*. This procedure was repeated using *bacA* amplified from *E. coli* K12 DNA using primers given in the Supplementary materials. However, BamHI and XbaI digestion was used prior to ligation. This vector was named pjexpress401*bacA*. The native (non-codon optimised) *pglB* gene was also amplified from the original pACYC*pgl* vector and inserted into pjexpress401 to make pjexpress401*pglB* using a similar procedure described above. This vector was used to re-introduce pglB activity to cells to test pACYC*pgl*Δ*pglB* mutation and serve as a control of pjexpress401*pglB**.

2.2. Cell growth, induction and harvesting

E. coli CLM24 cells [5] were used as host cells for this study. Ampicilin, kanamycin and chloramphenicol concentrations of 50 μg/ml were used where appropriate. Overnight cultures were used to seed 500 ml of LB-broth and grown at 37 °C with shaking at 180 rpm. When the optical density (O.D.) at 600 nm reached 0.5, cells were induced with 0.2% L-arabinose and 20 μM IPTG. 20 O.D. units worth of cells were harvested 3 h post induction via centrifugation at 10,000*g* for 15 min at 4 °C.

2.3. Protein purification

Proteins from the periplasm were prepped and AcrA purified using his-tag purification as described previously [8]. The resulting his-tag purification was prepared for quantitation using pSRM and Western blotting [8].

2.4. Western blots

Western blots were performed as described previously [8], except for alterations described briefly below. His-tag purified protein samples were quantified using RC/DC assay® (BioRad, UK) and 5 μg of total protein was loaded in each well (in 15 μL) combined with 4× loading buffer (1 M Tris–HCl(pH 6.8), 20% glycerol,