

Genetic and Process Engineering Strategies for Enhanced Recombinant N-glycoprotein Production in Bacteria

Fenryco Pratama

The University of Manchester Manchester Institute of Biotechnology

Dennis Linton

The University of Manchester

Neil Dixon (✉ neil.dixon@manchester.ac.uk)

Manchester Institute of Biotechnology (MIB) <https://orcid.org/0000-0001-9065-6764>

Research

Keywords: Glycoengineering, N-glycosylation, Bacterial host engineering, Protein folding, Process optimisation

Posted Date: December 7th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-119485/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Microbial Cell Factories on October 14th, 2021. See the published version at <https://doi.org/10.1186/s12934-021-01689-x>.

Abstract

Background

The production of N-linked glycoproteins in genetically amenable bacterial hosts offers great potential for reduced cost, faster/simpler bioprocesses, greater customisation and utility for distributed manufacturing of glycoconjugate vaccines and glycoprotein therapeutics. Efforts to optimize production hosts have included heterologous expression of glycosylation enzymes, metabolic engineering, use of alternative secretion pathways, and attenuation of gene expression. However, a major bottleneck to enhance glycosylation efficiency, which limits the utility of the other improvements is the impact of target protein sequon accessibility during glycosylation.

Results

Here, we explore a series genetic and process engineering strategies to increase recombinant N-linked glycosylation mediated by the *Campylobacter*-derived PglB oligosaccharyltransferase in *Escherichia coli*. Strategies include increasing membrane residency time of the target protein by modifying the cleavage site of its secretion signal, and modulating protein folding in the periplasm by use of oxygen limitation or strains with compromised oxidoreductases or disulphide-bond isomerase activity. These approaches could achieve up to 90% improvement in glycosylation efficiency. Furthermore, we also demonstrated that supplementation with the chemical oxidant cystine enhanced glycoprotein production and improved cell fitness in the oxidoreductase knock out strain.

Conclusions

In this study, we demonstrated that improved glycosylation in the heterologous host could be achieved by mimicking the coordination between protein translocation, folding and glycosylation observed in native such as *Campylobacter jejuni* and mammalian hosts. Furthermore, it provides insight into strain engineering and bioprocess strategy, to improve glycoprotein yield and to avoid physiological burden of unfolded protein stress to cell growth. The process and genetic strategies identified herein will inform further optimisation and scale-up of heterologous recombinant N-glycoprotein production

Full Text

This preprint is available for [download as a PDF](#).

Figures

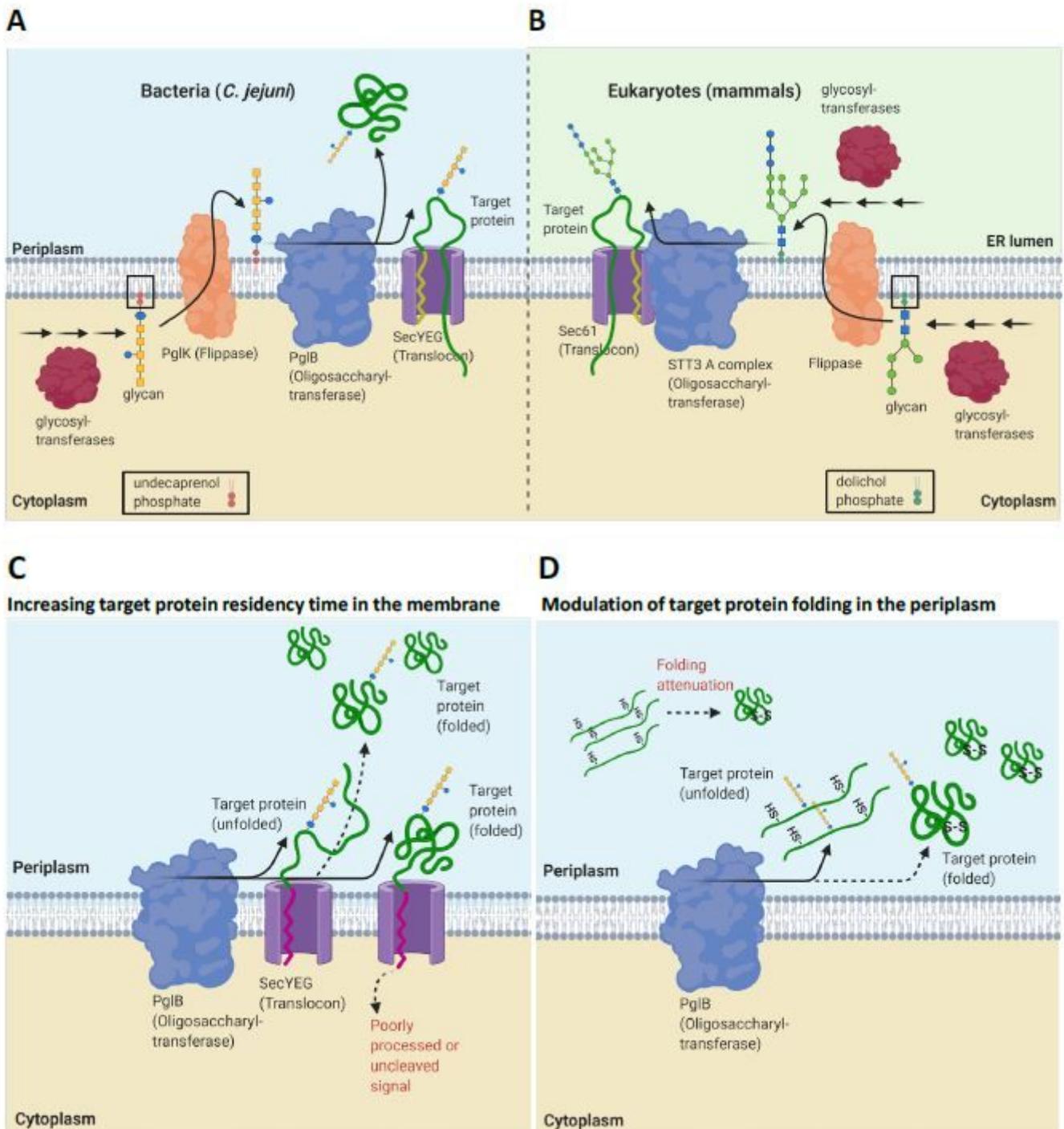


Figure 1

Schematic overview of native N-linked protein glycosylation pathway and proposed strategies in this study to improve sequon accessibility of recombinant target protein to PglB during glycosylation in glycoengineered *E. coli*. N-linked glycosylation in bacteria/*C. jejuni* (A) compared to early-stage of eukaryotic/mammalian N-linked glycosylation pathway (B). (A) In *C. jejuni*, undecaprenol-linked glycan is synthesised by glycosyltransferases in the cytoplasm, flipped to the periplasm by flippase (PglK), and then transferred co or post-translocationally to a target protein by PglB. (B) In mammalian N-linked

glycosylation, dolichol-linked glycan is synthesised both in the cytosol and endoplasmic reticulum (ER) lumen, and glycan is transferred co- or post-translocationally (later not shown) by different STT3 isoforms. (C-D) Two different strategies are proposed to enhance heterologous protein glycosylation in *E. coli*. (C) Approach based on AB CIncreasing target protein residency time in the membraneModulation of target protein folding in the periplasmD increasing PglB interaction with target protein in the membrane. Increased membrane residence of target protein was achieved by introducing signal peptide mutations with poor processivity or uncleaved signal. (D) Approach based on increasing sequon accessibility of target proteins to PglB via modulation of target protein folding state during or after translocation. Using a disulphide bonded model protein, protein folding and maturation in the periplasm are delayed by expressing the protein under suboptimal conditions for disulphide formation such as under oxygen-depleted conditions, in the absence of oxidoreductases, or under chemical redox treatment. Solid arrow = increased reaction, dashed arrow = reduced reaction.

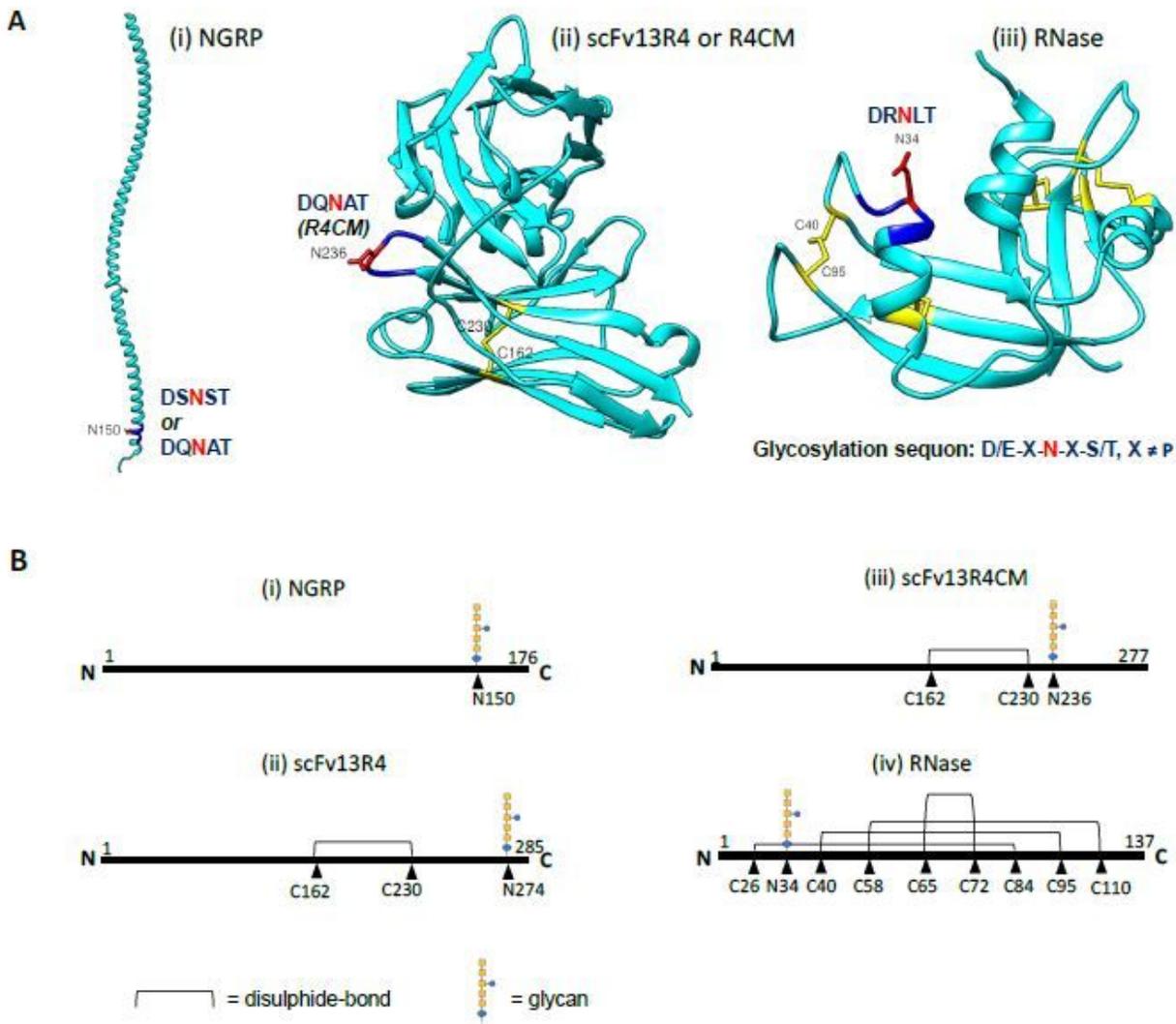


Figure 2

Structural variation of model glycoproteins. (A) Structural models of (i) NGRP, (ii) scFv13R4 and scFv13R4CM were generated by Phyre2 based on protein homology prediction ($\geq 99\%$ confidence) [60]. (iii) X-ray crystallographic structure of RNase A (PDB code 3WMR). Ribbon model of the proteins was drawn by UCSF Chimera [61]. Position of the sequon variants (D/E-X-N-X-S/T, $X \neq P$) within the protein is indicated. Disulphide-bonds are highlighted in yellow. C-terminal sequon (DQ_NAT) of scFv13R4 is not displayed in the protein model. (B) Linear representation of proteins; (i) NGRP, (ii) scFv13R4, (iii) scFv13R4CM, and (iv) RNase A. Position of disulphide-bonds (C-C) and glycosylation sites (N) are indicated with amino acid positions.

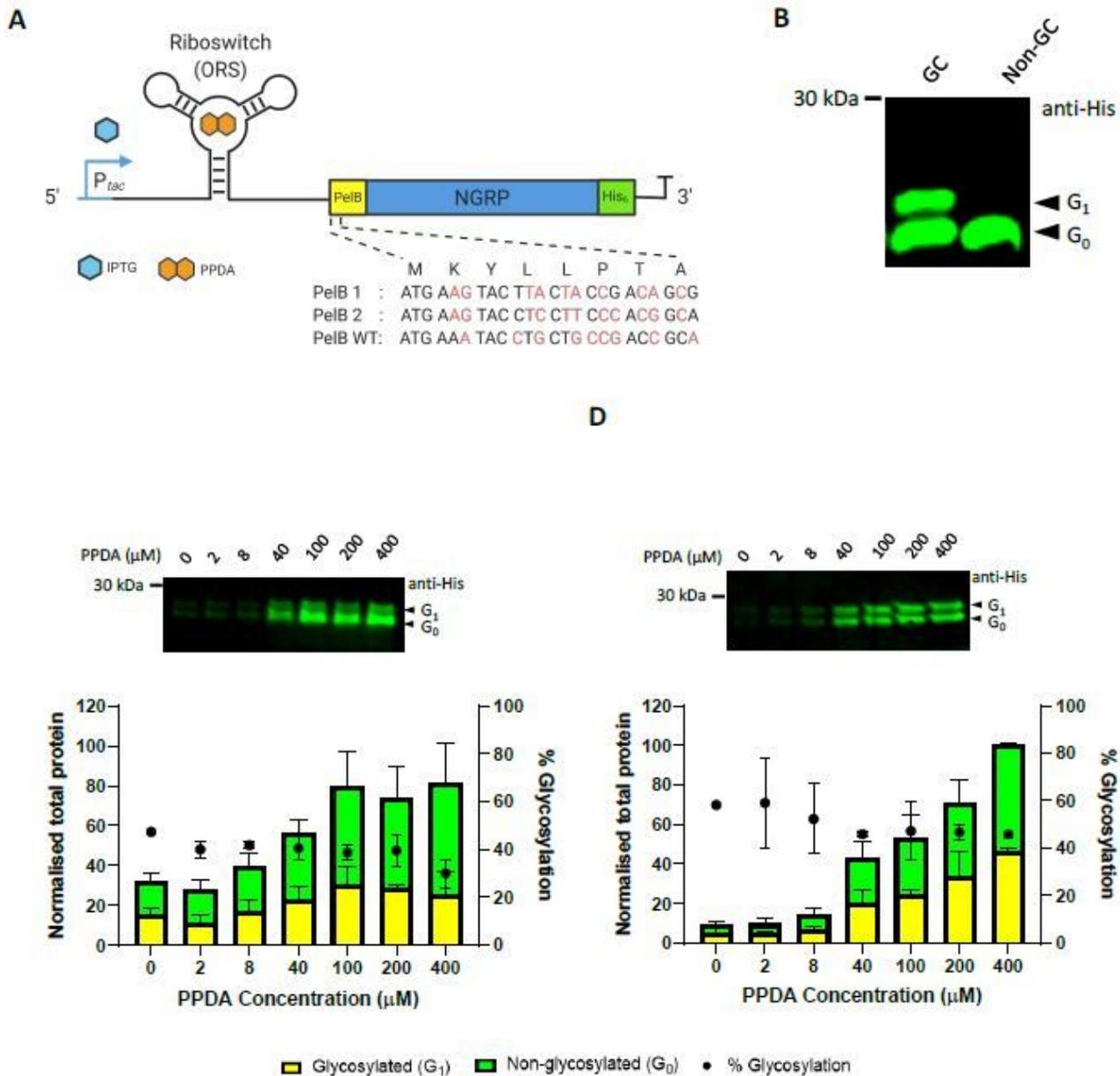


Figure 3

Glycosylation of NGRP in glyco-competent *E. coli*. (A) Organisation of pDEST-ORS expression vector used in this study. Target gene (NGRP) was fused with Sec signal peptide PelB (N-terminal) and Hexahistidine-tag (C-terminal). Expression of the target gene was regulated transcriptionally by Ptac via IPTG induction and translationally by orthogonal riboswitch (ORS) via PPDA titration. IPTG = Isopropyl β -D-1-thiogalactopyranoside, PPDA = Pyrimido [4,5-d] pyrimidine-2,4-diamine. Three synonymous nucleotide sequence variants of PelB-NGRP N-terminal codon (PelB 1, PelB 2, and PelB WT or wild-type) are shown. These variants were tested to explore the impact of the 5' codon context upon the riboswitch-dependent regulatory function. (B) Western blot analysis of periplasmic fractions of glycocompetent (GC) and non-glycocompetent (Non-GC) strains of *E. coli*. Anti-His antibody was used to detect the NGRP. Arrows indicate non-glycosylated (G0) and glycosylated (G1) NGRP. (C-D) Densitometry analysis (Western blot) of NGRP located in the periplasmic fractions from the two signal peptide variants, (C) PelB 1 and (D) PelB 2-NGRP. Proteins were expressed with increasing PPDA inducer concentrations (100 μ M IPTG, 0-400 μ M PPDA). Protein levels were normalised to the sample with highest expression level. Glycosylated (yellow bar) and non-glycosylated (green bar) protein are as shown (left y-axis). % Glycosylation (% AB30 kDaG1G0anti-His 02840100200400020406080100120020406080100PPDA Concentration (μ M)Normalised total protein% Glycosylation02840100200400020406080100120020406080100PPDA Concentration (μ M)Normalised total protein% Glycosylation30 kDaPPDA (μ M)anti-HisG1G030 kDaPPDA (μ M)anti-HisG1G0D G1/G0+G1) is indicated (black circle, right y-axis). Data were processed from three biological replicates; error bars represent standard deviation from mean values. The representative Western blots are shown as insets.

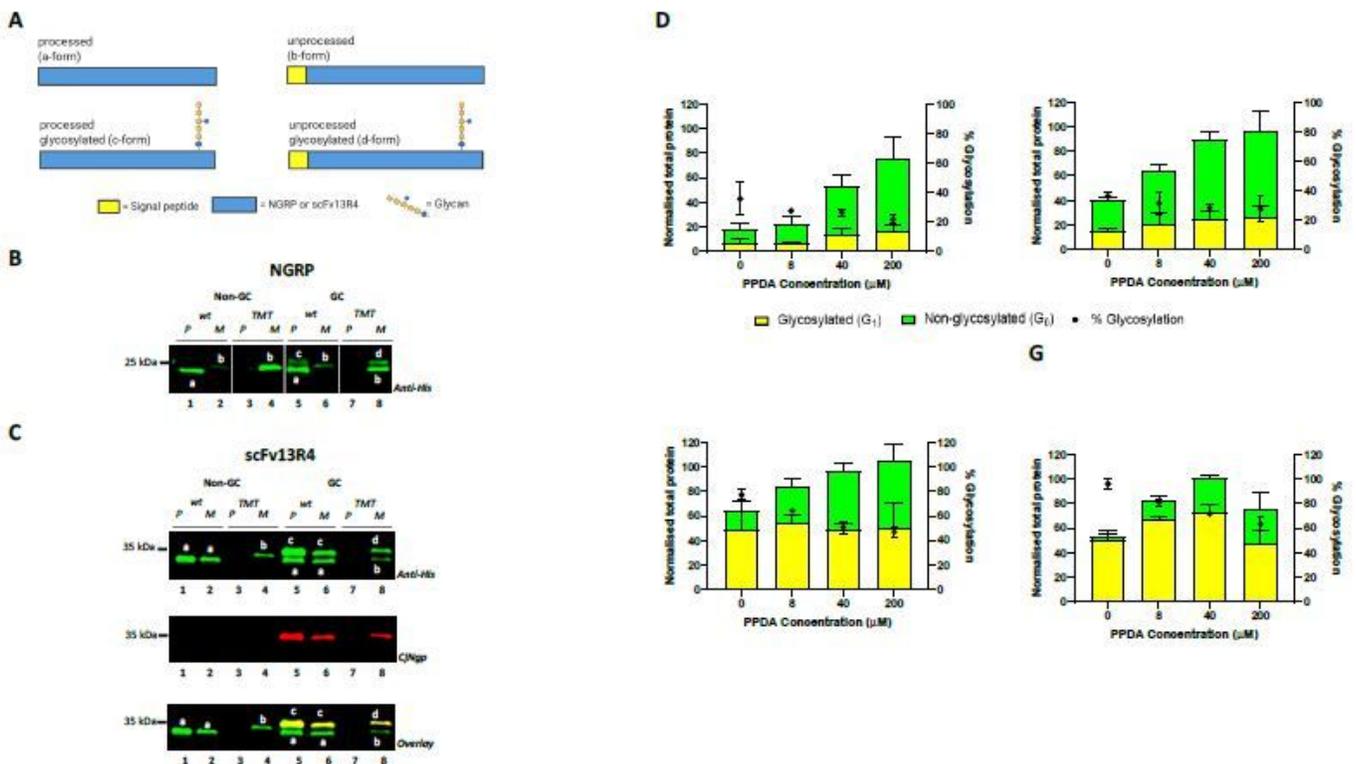


Figure 4

Production of NGRP and scFv13R4 isoforms containing signal peptide cleavage site variants. (A) Schematic representation of four predicted isoforms of NGRP or scFv13R4 (a-d-forms) based on their signal peptide processivity and protein glycosylation. (B-C) Western blot analysis of membrane (M) and periplasmic (P) expression of wild type (wt) and signal peptide cleavage mutant (TMT) NGRP (B) and scFv13R4 (C) in glyco-competent (GC) and non glyco-competent (Non-GC) *E. coli*. (B and C) All NGRP and scFv13R4 isoforms were detected by anti-His antibody. Glycosylated scFv13R4 was detected by CjNgp antibodies. Predicted a-d isoforms within the bands are indicated. (D-G) Quantitative Western blot analysis of membrane (TMT) and periplasmic (wt) localised (D-E) NGRP and (F-G) scFv13R4 produced in glyco-competent *E. coli*. Proteins were produced under different induction conditions (100 μ M IPTG, 0-200 μ M PPDA). Anti-His antibody was used to detect the proteins (glycosylated and non-glycosylated). Total proteins were determined by densitometry which signals normalised to expression at the highest induction. Glycosylated (yellow bar) and non-glycosylated (green bar) protein as shown (left y-axis). % Glycosylation (% G1/G0+G1) is indicated (black circle, right y-axis). All data (D-G) were processed from three biological replicates. Error bars indicate standard deviation from mean values.

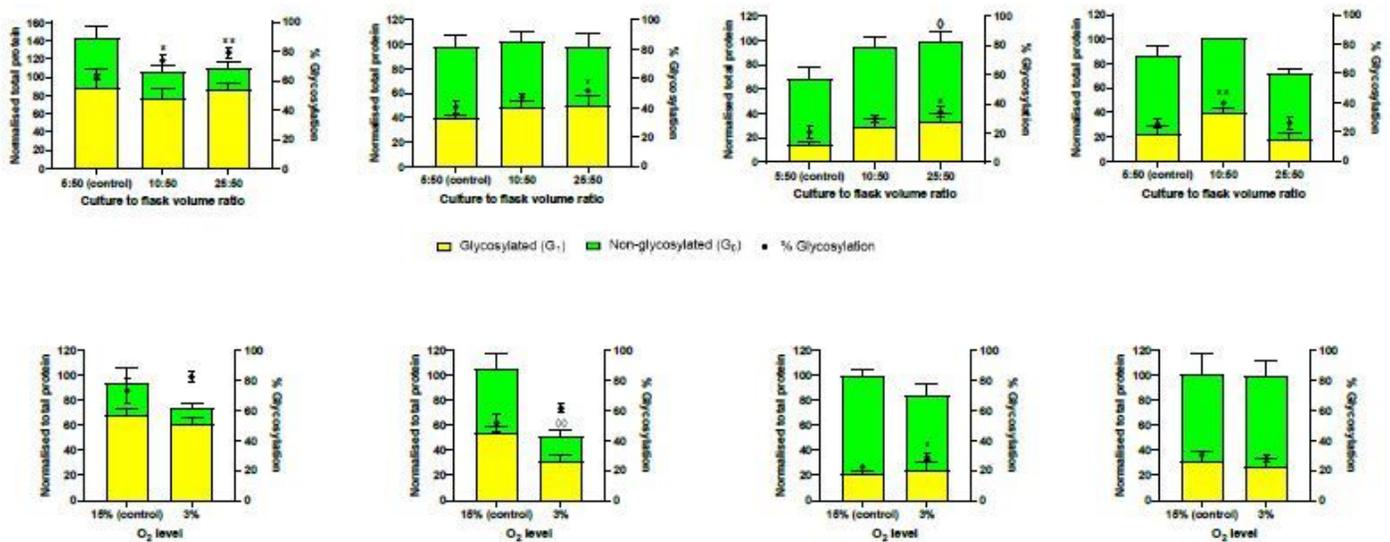


Figure 5

Effect of culture conditions and oxygen availability upon target protein glycosylation in *E. coli*. (A-D) Effect of culture to flask volume ratio (oxygen transfer efficiency) upon target protein glycosylation in *E. coli*. Western blot analysis (densitometry) of (A) scFv13R4, (B) scFv13R4CM, (C) RNase A and (D) non-disulphide control protein NGRP located in periplasm of glyco-competent *E. coli* in shake flask under three different culture to flask volume ratio (5:50, 10:50, and 25:50 mL). Total target protein (A-D) was normalised with the expression level at 10:50 culture to flask volume ratio. (E-H) Western blot analysis (densitometry) of (E) scFv13R4, (F) scFv13R4CM, (G) RNase A and (H) NGRP control non-disulphide

bond-containing protein, detected in the periplasm of glyco-competent *E. coli* under different oxygen levels culture (3% and 15% O₂). Total recombinant protein production levels (E-H) were normalised to maximal observed at 15% O₂. (A-H) Glycosylated (yellow bar) and non-glycosylated (green bar) protein as shown (left y-axis). % Glycosylation (% G₁/G₀+G₁) is indicated (black circle, right y-axis). Statistical analysis was conducted by unpaired t-test with Welch's correction to control sample at lowest culture to flask volume ratio 5:50 (A-D) or to control normoxic culture (E-H) ($P < 0.05^*$, $< 0.01^{**}$, for % glycosylation; $P < 0.05\Diamond$, $< 0.01\Diamond\Diamond$, for normalised total protein). All data were processed from three biological replicates. Error bars indicate standard deviation from mean values.

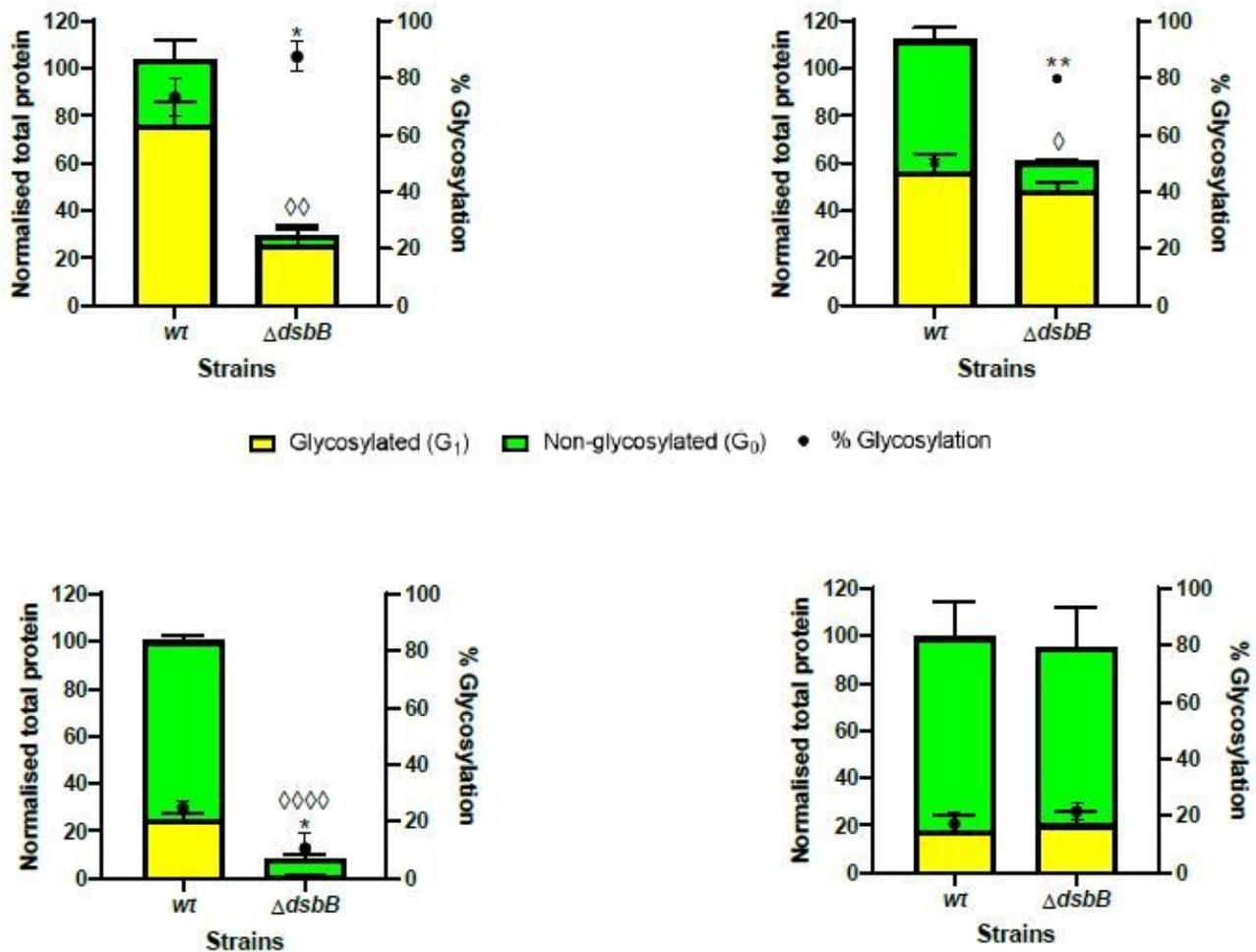


Figure 6

Glycosylation of disulphide bond-containing proteins in oxidoreductase mutant ($\Delta dsbB$) of *E. coli*. Western blot analysis (densitometry) of (A) scFv13R4, (B) scFv13R4CM, (C) RNase A and (D) control non-disulphide bond-containing protein NGRP expressed in periplasmic of glyco-competent *E. coli* wild-type (wt) or $\Delta dsbB$ strain. Total proteins (A-D) were normalised with the expression level of the wt strain. Glycosylated (yellow bar) and non-glycosylated (green bar) protein as shown (left y-axis). % Glycosylation (% G₁/G₀+G₁) is indicated (black circle, right y-axis). Statistical analysis was conducted by unpaired t-

test with Welch's correction to control sample expressed in wt strain ($P < 0.05^*$, $< 0.01^{**}$, for % glycosylation; $P < 0.05\diamond$, $< 0.01\diamond\diamond$, $< 0.0001\diamond\diamond\diamond\diamond$, for normalised total protein). All data were processed from three biological replicates. Error bars indicate standard deviation from mean values.

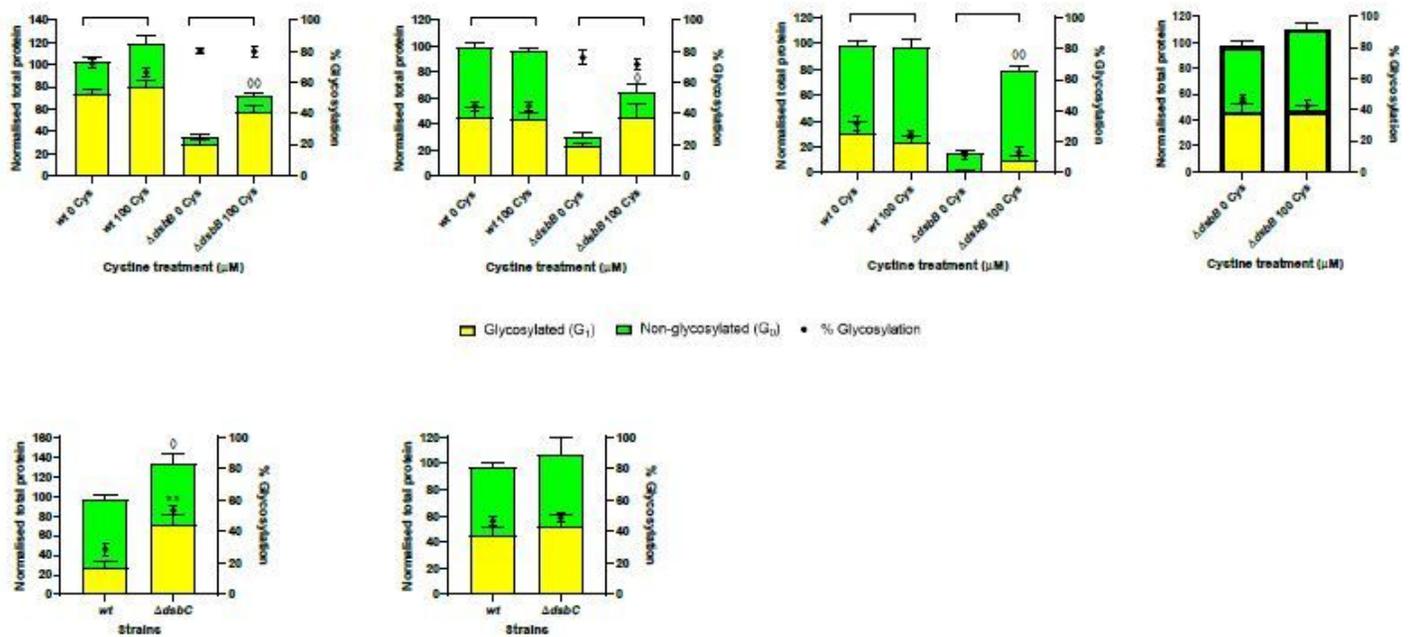


Figure 7

Impact of cystine supplementation upon glycosylation of recombinant proteins in $\Delta dsbB$ strain and glycosylation of RNase A in disulphide-bond isomerase mutant ($\Delta dsbC$). (A-D) Western blot analysis (densitometry) of (A) scFv13R4, (B) scFv13R4CM, (C) RNase A and (D) NGRP non-disulphide control protein produced in the periplasm of glyco-competent *E. coli* wild-type (wt) or $\Delta dsbB$ strain supplemented with 100 μM cystine during protein expression. Total recombinant protein production was normalised with the expression level in the wt (A-C) or $\Delta dsbB$ (D) strain without cystine treatment. (E and F) Western blot analysis (densitometry) of (E) RNase A and (F) control non-disulphide bond-containing protein NGRP expressed in periplasmic of glyco-competent *E. coli* $\Delta dsbC$. Total proteins (E and F) were normalised with the expression level of the wild-type (wt) strain. (A-F) Glycosylated (yellow bar) and non-glycosylated (green bar) protein as shown (left y-axis). % Glycosylation (% G1/G0+G1) is indicated (black circle, right y-axis). Statistical analysis was conducted by unpaired t-test with Welch's correction to control sample expressed in wt or $\Delta dsbB$ strain without cystine treatment (A-D) or to control sample expressed in wt strain (E and F) ($P < 0.01^{**}$, for % glycosylation; $P < 0.05\diamond$, $< 0.01\diamond\diamond$, for normalised total protein). All data were processed from three biological replicates. Error bars indicate standard deviation from mean values.

Supplementary Files

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

- [FigS1.JPG](#)
- [FigS2.JPG](#)
- [FigS3.JPG](#)
- [FigS4.JPG](#)
- [FigS5.JPG](#)
- [FigS6.JPG](#)
- [FigS7.JPG](#)
- [FigS8.jpg](#)