



OPEN

Random and combinatorial mutagenesis for improved total production of secretory target protein in *Escherichia coli*

David Gonzalez-Perez^{1,2}, James Ratcliffe¹, Shu Khan Tan¹, Mary Chen May Wong¹, Yi Pei Yee¹, Natsai Nyabadza¹, Jian-He Xu³, Tuck Seng Wong^{1,4}✉ & Kang Lan Tee¹✉

Signal peptides and secretory carrier proteins are commonly used to secrete heterologous recombinant protein in Gram-negative bacteria. The *Escherichia coli* osmotically-inducible protein Y (OsmY) is a carrier protein that secretes a target protein extracellularly, and we have previously applied it in the Bacterial Extracellular Protein Secretion System (BENNY) to accelerate directed evolution. In this study, we reported the first application of random and combinatorial mutagenesis on a carrier protein to enhance total secretory target protein production. After one round of random mutagenesis followed by combining the mutations found, OsmY(M3) (L6P, V43A, S154R, V191E) was identified as the best carrier protein. OsmY(M3) produced 3.1 ± 0.3 fold and 2.9 ± 0.8 fold more secretory Tfu0937 β -glucosidase than its wildtype counterpart in *E. coli* strains BL21(DE3) and C41(DE3), respectively. OsmY(M3) also produced more secretory Tfu0937 at different cultivation temperatures (37 °C, 30 °C and 25 °C) compared to the wildtype. Subcellular fractionation of the expressed protein confirmed the essential role of OsmY in protein secretion. Up to $80.8 \pm 12.2\%$ of total soluble protein was secreted after 15 h of cultivation. When fused to a red fluorescent protein or a lipase from *Bacillus subtilis*, OsmY(M3) also produced more secretory protein compared to the wildtype. In this study, OsmY(M3) variant improved the extracellular production of three proteins originating from diverse organisms and with diverse properties, clearly demonstrating its wide-ranging applications. The use of random and combinatorial mutagenesis on the carrier protein demonstrated in this work can also be further extended to evolve other signal peptides or carrier proteins for secretory protein production in *E. coli*.

Abbreviations

dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
dNTP	Deoxynucleotide
dTTP	Deoxythymidine triphosphate
DyP4	Dye-decolorizing peroxidase 4 from <i>Pleurotus ostreatus</i> strain PC15
E_m	Emission wavelength
<i>E. coli</i>	<i>Escherichia coli</i>
E_x	Excitation wavelength
epPCR	Error-prone polymerase chain reaction
HTS	High-throughput screening
M_w	Molecular weight

¹Department of Chemical and Biological Engineering, The University of Sheffield, Sir Robert Hadfield Building, Mappin Street, Sheffield S1 3JD, UK. ²Department of Drug Discovery, Moffitt Cancer Center & Research Institute, Stabile Research Building, 12902 Magnolia Dr, Tampa, FL 33612, USA. ³Laboratory of Biocatalysis and Bioprocessing, State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, 130 Meilong Road, Shanghai 200237, People's Republic of China. ⁴National Center for Genetic Engineering and Biotechnology, 113 Thailand Science Park, Phahonyothin Road, Khlong Luang 12120, Pathum Thani, Thailand. ✉email: t.wong@sheffield.ac.uk; k.tee@sheffield.ac.uk

pNPA	<i>p</i> -Nitrophenyl acetate
pNPG	<i>p</i> -Nitrophenyl- β -D-glucopyranoside
OD ₆₀₀	Optical density at 600 nm
OsmY	Osmotically-inducible protein Y
PCR	Polymerase chain reaction
RFP	Red fluorescent protein
WT	Wildtype

Protein secretion in Gram-negative bacteria. Recombinant protein production, including technical enzymes and biopharmaceutical proteins, is a growing multi-billion-dollar market. Different eukaryotic and prokaryotic hosts are currently used for recombinant protein production, including mammalian cells, fungi, yeast and bacteria. Bacteria are especially attractive for their fast growth rates, ease in handling, broad range of carbon sources and relatively low-cost culture medium. Well-established and widely accessible genetic manipulation tools are also available, especially for the Gram-negative bacterium *Escherichia coli*¹.

Gram-positive bacteria can secrete native proteins or proteins that originate from close relatives at high yield². However, their compatibility with other proteins is low and the presence of extracellular proteolytic degradation is a challenge for product quality assurance. In contrast, Gram-negative bacteria produce a broad range of heterologous proteins. Protein secretion in Gram-negative bacteria is less common compared to other organisms like Gram-positive bacteria, yeast and fungi. However, it offers several advantages. First, Gram-negative bacteria, like *E. coli*, secrete few native proteins to provide a relatively pure background for heterologous protein production. Second, no cell disruption is required, hence reducing the complexity and the cost of downstream processing. Third, protein secretion out of the cytoplasm provides an oxidizing environment that facilitates disulphide bond formation and protein folding, for example during hormones and antibody fragments production. Fourth, toxic effect of some target proteins on the production host can be alleviated when the protein is secreted out of the cell.

Protein secretion in Gram-negative bacteria can be achieved by fusing a signal peptide or a secretory carrier to the N- or C-terminus of the target protein^{3,4}. Enhancing protein secretion is a challenge as it is currently not possible to predict the performance of a signal peptide or a secretory carrier for a specific target protein or host. Three approaches are generally used for secretory protein production in *E. coli*. The first approach is to screen a large library of signal peptides and secretory carriers^{5,6} to identify a fusion partner for optimal secretion. The second approach is to engineer the signal peptides or the transporter proteins for a better extracellular heterologous protein production in *E. coli*. Examples include: (i) inserting short random peptides libraries (1–13 amino acids) at the junction between the signal peptide and the mature protein followed by screening for the construct that secreted most protein to the *E. coli* periplasm⁷, (ii) screening error-prone mutagenesis library of an ABC transporter from *Pseudomonas fluorescens* to enhance its extracellular secretion efficiency in *E. coli*⁸, (iii) directed evolution of translation initiation region without altering the amino acid sequence of signal peptide to enhance protein secretion into the *E. coli* periplasm⁹, and (iv) using non-optimal codons in the signal sequence of TolB for a better extracellular protein secretion in *E. coli*¹⁰. The third approach involves engineering the protein production host, for instance, perturbing the cell wall peptidoglycan network via deletion of *dacA* and *dacB*¹¹ and screening *E. coli* single-gene knockout library for hypersecretory phenotypes¹².

The osmotically-inducible protein Y from *Escherichia coli*. The *E. coli* osmotically-inducible protein Y (OsmY) is expressed in response to a variety of stress conditions, especially osmotic shock^{13–15}. OsmY was selected for this work as it was identified as the most efficient fusion partner for extracellular protein secretion among the naturally excreted protein in *E. coli* BL21(DE3)¹⁶. OsmY has two BON (bacterial OsmY and nodulation) domains; one between amino acids 55–123 and the second between amino acids 134–201. The BON domain is a conserved domain that has typically about 60 amino acid residues, and has an α/β predicted fold. The exact function of the BON domain is unclear, but it is predicted to be a phospholipid binding domain that explains OsmY's secretory function¹⁷. Numerous studies used OsmY for extracellular production of recombinant proteins in *E. coli*, including endoglucanase¹⁸, β -glucosidase¹⁸, xylanase^{19,20}, xylosidase^{19,20}, single-chain antibody²¹ and various human proteins²². We also incorporated OsmY in the Bacterial Extracellular Protein Secretion System (BENNY) to accelerate the directed evolution workflow in engineering DyP4, a heme peroxidase²³. OsmY was previously reported to show chaperone activity and helped optimized unstable protein folding *in vivo*²⁴. The versatility of OsmY further prompted us to engineer this protein to expand its application.

In this study, we demonstrated the first successful example of applying random and combinatorial mutagenesis on a secretory carrier to enhance total secretory target protein production in *E. coli*. There are two mechanisms by which total secretory target protein production can be enhanced: (a) increasing the secretion capability of a carrier (i.e., altering protein partition by favouring extracellular secretion), and (b) increasing the expression level of the target protein carrying a secretion signal such as a carrier. OsmY is strategically chosen, as it has the potential to satisfy either or both requirements.

Results and discussion

Random mutagenesis of OsmY. To evolve OsmY for a higher total secretory protein production, β -glucosidase Tfu0937 from *Thermobifida fusca* YX was selected as the secretion target protein for its validated application in cellulose degradation and ease of monitoring enzymatic activity. OsmY can secrete proteins in different *Escherichia coli* strains¹⁶. The levels of total secreted OsmY-Tfu0937 produced by *E. coli* C41(DE3) and BL21(DE3) were first measured using the *p*-nitrophenyl- β -D-glucopyranoside (pNPG) assay to identify the more efficient strain. C41(DE3) cultivation gave higher OsmY-Tfu0937 activity in the spent medium compared

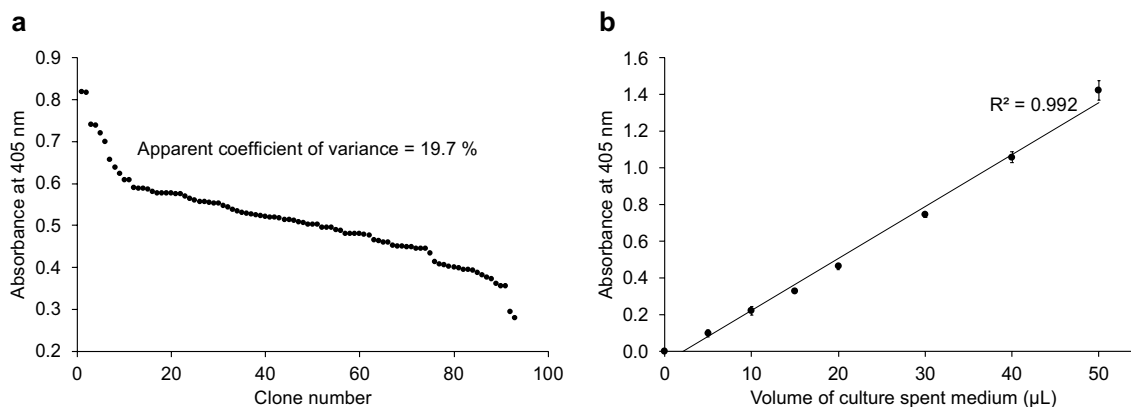


Figure 1. The OsmY(WT)-Tfu0937 activity in the spent medium was measured with the pNPG assay using an absorbance at 405 nm. **(a)** The pNPG assay was performed in a 96-well microplate for OsmY(WT)-Tfu0937-expressing C41(DE3) cells cultivated in a 96-well deep well plates. The coefficient of variance was calculated using the absolute values measured at 405 nm. **(b)** The pNPG assay was conducted using 5–50 μL of spent medium. The activity value is linear to the volume of spent medium used with an $R^2 = 0.992$.

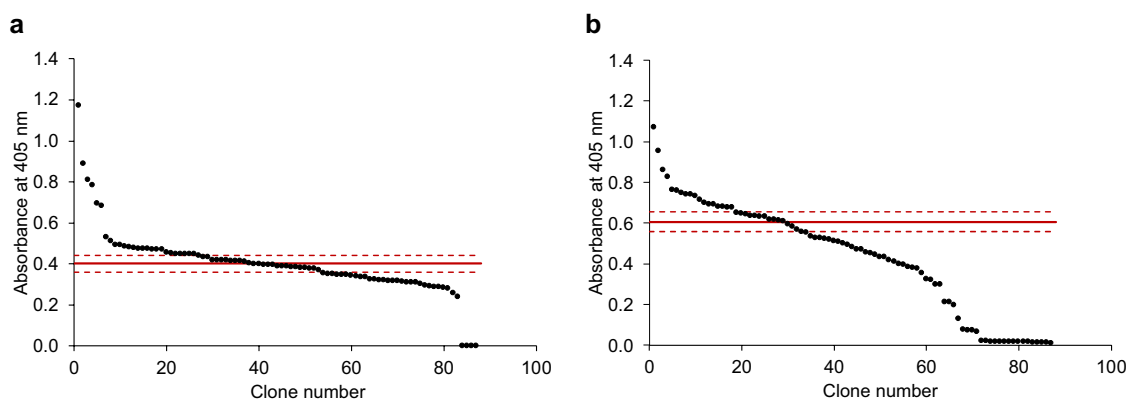


Figure 2. The OsmY-Tfu0937 activity in the spent medium of epPCR generated OsmY mutagenesis libraries was measured in a 96-well microplate. The activity was measured with the pNPG assay using an absorbance at 405 nm. The libraries were generated using **(a)** a low and **(b)** a high mutagenic rates. The activities of parental clones in column 6 were measured in the same 96-well microplate. The average of parental clones is shown as a solid red line and one standard deviation above and below the average are shown as dash red lines.

to BL21(DE3), and was therefore chosen for subsequent mutagenesis experiments. The pNPG assay was adapted to a 96-well format. The final optimised assay had an apparent coefficient of variance of 19.7% (Fig. 1a), sufficient for a high throughput screening. The enzymatic activity of secreted OsmY-Tfu0937 using pNPG as a pseudosubstrate was also directly proportional to the volume of spent medium, when 5–50 μL was used (Fig. 1b). Variable volume of spent medium was thus used to keep the assay within the linear detection range, especially for OsmY variants with higher total secretory protein production.

During random mutagenesis, mutations were only targeted to the sequence encoding for OsmY, leaving the Tfu0937 sequence unaltered. This was designed to ensure all improvements can be attributed to OsmY and were not due to changes in the target protein, Tfu0937. The libraries were created using two different conditions with high and low mutagenic rates (Fig. 2). The effects of different mutagenic rates were evident in the activity of the clones in the library. For the low mutagenic rate library (Fig. 2a), 31% of clones had the same activity as the parent (i.e., one standard deviation above and below the average of parental clones) and 30% of clones were more active than the parent (i.e., greater than one standard deviation above the parent average). For the high mutagenic rate library (Fig. 2b), 16% of clones had the same activity as the parent and 21% of clones were more active than the parent. More inactive clones were also found in the high mutagenic rate library (23%) compared to the low mutagenic rate library (5%).

Improved OsmY mutants obtained from both random mutagenesis libraries. About three hundred clones from the two random mutagenesis libraries were screened and four improved variants were identified (Table 1). These four variants were subsequently cultivated in 5-mL tube cultures and extracellular OsmY-Tfu0937 activity was measured again to confirm their improvements as shown in Table 1. Mutations in these variants were verified by DNA sequencing. The variants and wildtype can be arranged in decreasing order

Mutagenic rate	Variant name	Missense mutation	Silent mutation	Relative activity compared to OsmY wildtype ^a
n.a	OsmY(WT)	n.a	n.a	1.00 ± 0.08
Low	OsmY(S19C)	S19C (A → T)	n.a	3.40 ± 0.25
Low	OsmY(V43A)	V43A (T → C)	n.a	1.41 ± 0.13
Low	OsmY(A39)	n.a	A39 (G → T)	1.67 ± 0.16
High	OsmY(TOA4)	L6P (T → C) S154R (A → C) V191E (T → A)	K107 (A → G)	5.50 ± 0.56

Table 1. OsmY variants identified from screening of epPCR libraries. The mutations were verified by DNA sequencing. ^aVariants were cultivated in 5-mL tubes cultures and activity of extracellular OsmY-Tfu0937 was determined with the pNPG assay using an absorbance at 405 nm, and compared with the wildtype OsmY-Tfu0937.

of extracellular OsmY-Tfu0937 production as OsmY(TOA4) > OsmY(S19C) > OsmY(A39) > OsmY(V43A) > OsmY(WT). In agreement with the mutagenic conditions, the three variants from the low mutagenic rate library have a single nucleotide substitution each, while the one variant from the high mutagenic rate library has four nucleotide substitutions.

OsmY is a 201-amino acid protein with an N-terminal 28-amino acid signal peptide. Interestingly, the top two variants have mutations in their signal peptides; L6P for OsmY(TOA4) and S19C for OsmY(S19C). It was previously shown that the OsmY signal peptide secreted protein into the periplasm, but the full length OsmY protein was required for extracellular protein secretion¹⁶. Signal peptides have a conserved tripartite overall structure, which consists of a positively charged amino-terminal region, a central hydrophobic core, and a polar carboxyl-terminal region²⁵. The mutations L6P and S19C conserved the respective hydrophobic and polar nature of the amino acids. Variant OsmY(A39) has a silent mutation that changes the codon for A39 from GCG → GCT. Interestingly, codon GCG has the highest usage frequency among the four codons for alanine while GCT has the lowest usage frequency. The same was observed for the K107 silent mutation in OsmY(TOA4), where the higher usage frequency codon (AAA) was mutated to the lower usage frequency codon (AAG) for lysine. Non-optimal codon in the signal peptide had been reported to slow down the kinetics of translation to positively impact secretion efficiency¹⁰. The enhanced production of secretory protein by a switch to less optimal codons in A39 and K107 could potentially also be due to a slower translation kinetics. Important to note, codon deoptimization of signal peptides also improved the production of secretory protein in baculovirus/insect cell expression system²⁶. Our random mutagenesis data, along with the results from other signal peptide engineering studies, allow us to propose engineering strategies to guide future improvement of secretory protein production across various protein expression systems: (a) preserving the tripartite nature of the signal peptides through conservative mutations, and (b) slowing the protein translation by deoptimizing the codons of the carrier protein.

Variant from random mutagenesis produced 1.8-times more extracellular protein. The three variants with missense mutations [OsmY(TOA4), OsmY(S19C), and OsmY(V43A)] were re-transformed into C41(DE3) to rule out any possible effects due to host strain instability. In 50-mL flask cultivations, secreted protein was detected 5.5 h after culture inoculation, however significant activity was observed after 24 h of cultivation. Results (Fig. 3a) showed that variant OsmY(TOA4) had 1.8 ± 0.5 times more OsmY-Tfu0937 activity in the spent medium compared to the OsmY(WT) after 24 h cultivation, and this improvement increased further to 2.2 ± 0.3 times after 48 h. Enhanced secretory protein production did not affect the cell viability, reflected by the similar cell growth of *E. coli* expressing OsmY(WT)-Tfu0937 and the variants.

Tfu0937 secretion by OsmY(WT) and OsmY(TOA4) were also compared at cultivation temperatures of 25 °C and 30 °C (Fig. 3b). OsmY(TOA4) was more efficient than wildtype at these lower temperatures in secretory protein production. When cultivated at 25 °C, OsmY-Tfu0937 activity in the spent medium remained low throughout the cultivation, likely due to the slower growth and protein production rates compared to cultivation at higher temperatures. However, more extracellular activity was detected for the OsmY(TOA4) variant than the OsmY(WT) after 27 h. When cultivated at 30 °C, OsmY-Tfu0937 activity in the spent medium of OsmY(TOA4) was 1.8 ± 0.2 times (23.0 h) and 1.5 ± 0.1 times (27.0 h) that of the wildtype, similar to the improvement seen at 37 °C cultivation (Fig. 3a). As in the 37 °C cultivation, there was no significant difference in cell growth between OsmY(TOA4) and OsmY(WT) at 30 °C or at 25 °C. However, higher final cell optical densities were obtained at lower temperatures, potentially due to lower protein production load.

L6P is the beneficial mutation in OsmY(TOA4) variant. OsmY(TOA4) was the best variant identified from our initial screening. It has three missense mutations (L6P, S154R, V191E), with one located in the signal peptide (L6P) and the remaining two in the second BON domain. A single mutant was created for each of these mutations using site-directed mutagenesis and their extracellular OsmY-Tfu0937 activity was measured to deconvolute their effects on secretory protein production (Fig. 4). The two single mutants, S154R and V191E, showed the same OsmY-Tfu0937 activity in the spent medium as the wildtype OsmY. However, the L6P single mutant showed 3.1 ± 1.6 times the level of secreted protein compared to the wildtype OsmY after 24.0 h cultivation. Comparing the secretory protein production levels achieved by OsmY(TOA4) and OsmY(L6P), 1.8 ± 0.5 times and 3.1 ± 1.6 times respectively, both variants secreted similar amount of proteins. There was

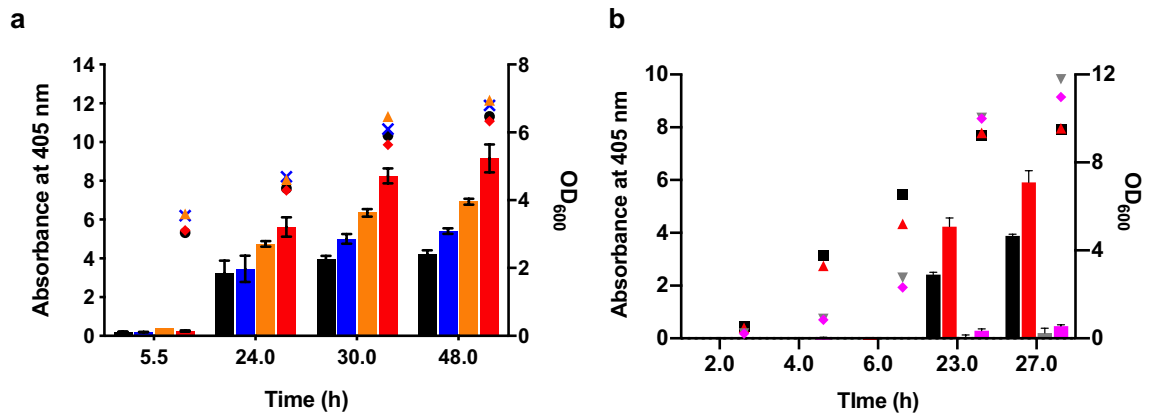


Figure 3. Characterization of improved OsmY variants in 50-mL flask cultures. The OsmY-Tfu0937 activity, measured with the pNPG assay using an absorbance at 405 nm (left axis), are plotted as bars. The cell optical densities monitored at 600 nm (right axis) are shown as symbols. Cultures were inoculated using 1:100 dilution. **(a)** OsmY(WT) and variants were expressed using C41(DE3) in $2 \times$ TY auto-induction media and at 37 °C and 200 rpm. The bars and symbols of the variants are represented using different colours; OsmY(WT) (black), OsmY(V43A) (blue), OsmY(S19C) (orange) and OsmY(TOA4) (red). Experiments were performed in triplicate. **(b)** OsmY(WT)-Tfu0937 and OsmY(TOA4)-Tfu0937 were expressed using C41(DE3) in $2 \times$ TY auto-induction media, at 30 °C or 25 °C, and 200 rpm. The bars and symbols of the sample are represented using different colours; OsmY(WT) at 30 °C (black), OsmY(TOA4) at 30 °C (red), OsmY(WT) at 25 °C (grey), OsmY(TOA4) at 25 °C (magenta).

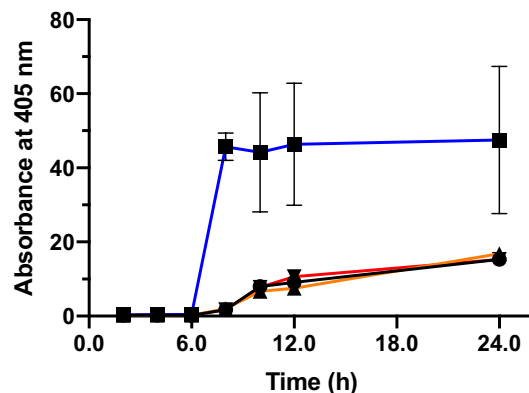


Figure 4. The OsmY-Tfu0937 activity in the spent medium of single site-directed mutants was measured with the pNPG assay using an absorbance at 405 nm. Cells were cultivated in 20-mL of $2 \times$ TY auto-induction media at 37 °C. Culture was inoculated using 1:100 dilution and 0.5 mL culture was sampled at various time points for activity assay. Experiments were performed in triplicate. The samples are represented using different colours; OsmY(WT)—black line, OsmY(L6P)—blue line, OsmY(S154R)—orange line and OsmY(V191E)—red line.

no additive interaction between S154R or V191E with L6P to further increase secretory protein production in OsmY(TOA4). The cell growth of the three single mutants were similar to the OsmY(WT), as was observed for OsmY(TOA4) (Supplementary Fig. S1). Our data indicated that the signal peptide of OsmY plays a greater role in secretory protein production compared to its BON domain.

Combinatorial mutagenesis further enhanced total secretory protein production. Site-directed mutagenesis was used to add the mutations S19C and V43A individually to the best variant OsmY(TOA4), creating mutants OsmY(M1) (i.e., TOA4 + S19C) and OsmY(M3) (i.e., TOA4 + V43A), respectively. When cultivated at the same conditions used in Fig. 3a, OsmY(M3) resulted in 2.9 ± 0.8 times more secretory protein compared to OsmY(WT) at 24 h (Supplementary Fig. S2). As with the earlier experiments, cell growth of variant OsmY(M3) was similar to that of OsmY(WT). Contrary to OsmY(M3), the spent medium of variant OsmY(M1) showed no OsmY-Tfu0937 activity, though its cell growth was similar to OsmY(WT). To determine if OsmY(M1) had lost only its protein secretion capability or also protein expression, enzymatic lysis with lysozyme was used to extract the intracellular soluble protein in OsmY(M1). The extracted soluble protein showed no OsmY-Tfu0937 activity, suggesting the loss of soluble OsmY-Tfu0937 expression in variant OsmY(M1). We subsequently also added S19C to variant OsmY(M3) and observed the same effect, where no secreted OsmY-Tfu0937 activity was

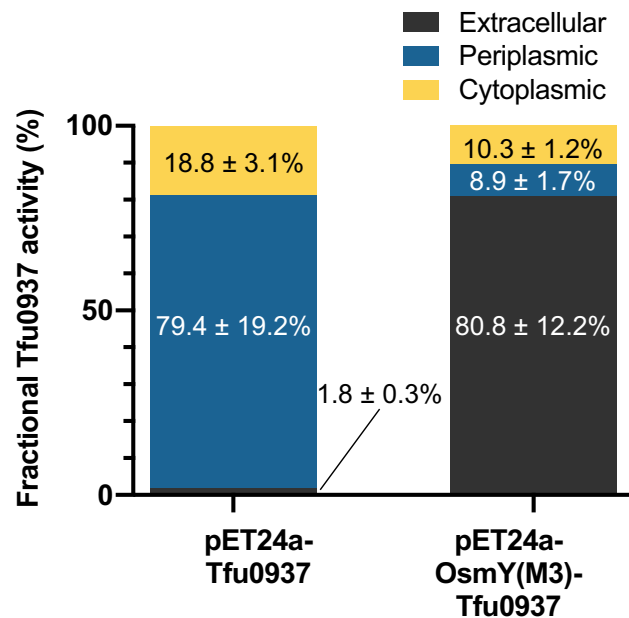


Figure 5. Subcellular localization of Tfu0937 and OsmY(M3)-Tfu0937. Proteins were expressed using pET24a-Tfu0937 and pET24a-OsmY(M3)-Tfu0937 in 2 × TY auto-induction medium at 37 °C for 15 h. The fractional Tfu0937 and OsmY(M3)-Tfu0937 activity in the extracellular, periplasmic and cytoplasmic fractions were measured with the pNPG assay using an absorbance at 405 nm. The enzyme activity in each fraction was calculated as a percentage of total activity measured in all three fractions. Experiments were performed in triplicate.

detected. We therefore concluded that combining S19C with the existing mutations in OsmY(TOA4) led to negative epistatic interaction and loss of soluble protein expression.

OsmY is required for extracellular protein secretion in *E. coli*. Tfu0937 was found in the secretome of *T. fusca*²⁷. To confirm that OsmY is necessary for the secretion of Tfu0937 in *E. coli* and to rule out the possibilities that this is solely due to cell leakage or the presence of a secretion motif in Tfu0937 recognized by *E. coli*, the plasmid pET24a-Tfu0937 was constructed. Plasmids pET24a-Tfu0937 and pET24a-OsmY(M3)-Tfu0937 were used for protein expression. When Tfu0937 was expressed without OsmY using plasmid pET24a-Tfu0937, the extracellular Tfu0937 activity constituted only 1.8 ± 0.3% of the total Tfu0937 activity in the three fractions (Fig. 5). Most of the Tfu0937 activity was found in the periplasmic fraction while the cytoplasmic fraction accounted for 18.8 ± 3.1% of total Tfu0937 activity. In contrast, 80.8 ± 12.2% of total OsmY-Tfu0937 activity was found in the spent medium when OsmY(M3) was fused to the N-terminus of Tfu0937 (Fig. 5).

OsmY(M3) variant increased the total protein production. The increased extracellular OsmY-Tfu0937 observed in OsmY(M3) compared to wildtype can be attributed to one or both of these factors: (1) higher fractional secretion of total protein, and (2) increase in total protein expression. To understand the reason behind higher secretory protein production by OsmY(M3), we performed subcellular fractionation for OsmY(WT) and OsmY(M3) at two different time points during cultivation (6 h and 24 h). After 6 h cultivation, when cells were in the exponential growth phase, both wildtype and OsmY(M3) showed the most OsmY-Tfu0937 activity in the cytoplasmic fraction, followed by the periplasmic fraction and the least activity was detected in the secreted fraction (Fig. 6). The wildtype strain has a higher fraction of OsmY-Tfu0937 found extracellularly and in the periplasm compared to OsmY(M3). However, the OsmY(M3) variant had 3.64 ± 0.13 times more total OsmY-Tfu0937 activity compared to OsmY(WT). After 24 h cultivation, both wildtype and variant showed similar protein distribution in the extracellular, periplasmic and cytoplasmic fractions (Fig. 6). However, 2.73 ± 0.24 times more total OsmY-Tfu0937 activity was measured in OsmY(M3) compared to the wildtype. A total of 4100 U/mL of extracellular OsmY(M3)-Tfu0937 was produced after 24 h of cultivation. The enhanced production of secretory protein in OsmY(M3) was thus largely attributed to higher OsmY-Tfu0937 expression in the variant, while its secretion capability is retained.

OsmY(M3) variant improves the production of secretory protein in both BL21(DE3) and C41(DE3) strains. The *E. coli* strain C41(DE3) was used for screening the mutagenesis libraries, where OsmY(TOA4) was obtained and OsmY(M3) was subsequently derived. Since OsmY(WT) can secrete OsmY-Tfu0937 in both *E. coli* BL21(DE3) and C41(DE3) strains, we wanted to determine if the enhanced secretory protein production can be transferred to BL21(DE3). All variants showed similar cell growth over time in both

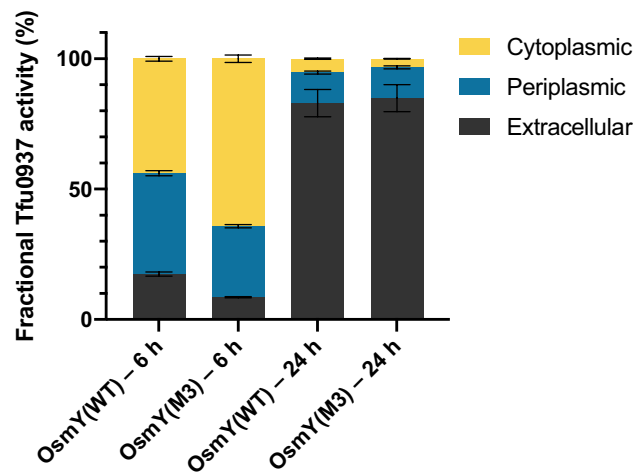


Figure 6. Fractional OsmY-Tfu0937 activity in the extracellular, periplasmic and cytoplasmic fractions of cells expressing Tfu0937 fused to either OsmY(WT) or OsmY(M3). Activity was measured with the pNPG assay using an absorbance at 405 nm. Experiments were performed in triplicate. Cells were cultivated in 2×TY auto-induction medium at 37 °C for either 6 h or 24 h. The enzyme activity in each fraction was calculated as a percentage of total activity in all three fractions.

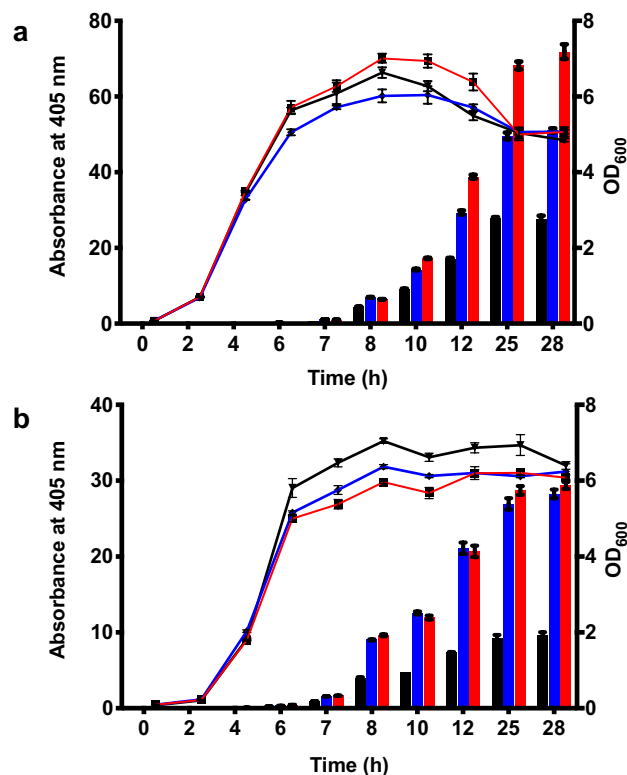


Figure 7. Protein secretion using OsmY(WT) and OsmY(M3) in (a) *E. coli* C41 (DE3) and (b) BL21(DE3). OsmY(WT)-Tfu0937 and OsmY(M3)-Tfu0937 were expressed using C41(DE3) and BL21(DE3) in 50-mL 2×TY auto-induction media and at 37 °C and 250 rpm. Cultures were inoculated using 1:200 dilution. The cell optical densities monitored at 600 nm (right axis) are shown as lines. The activity of OsmY-Tfu0937 shown as bars was measured with the pNPG assay using an absorbance at 405 nm (left axis). The lines and bars of different variants are represented using different colours; OsmY(WT)—black, OsmY(TOA4)—blue and OsmY(M3)—red.

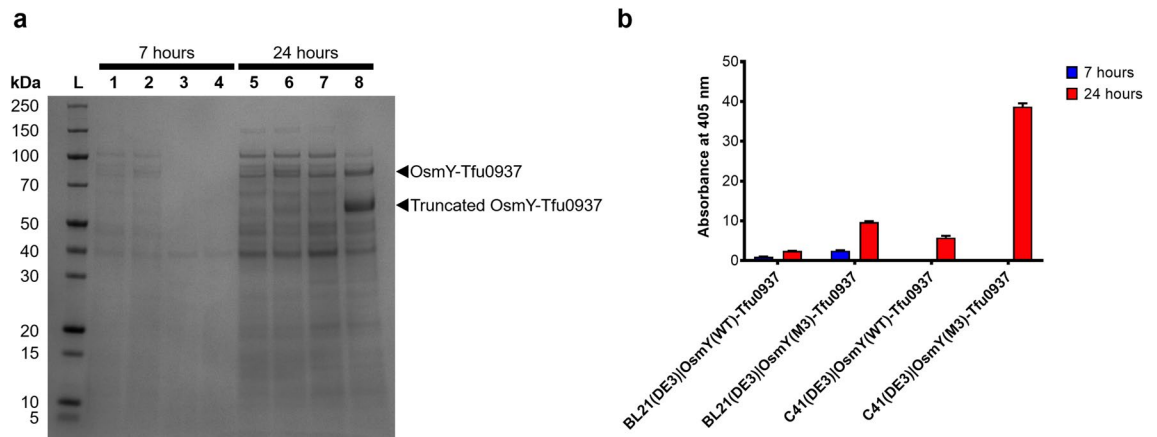


Figure 8. OsmY-Tfu0937 secretion using C41(DE3) and BL21(DE3) in 400-mL 2×TY auto-induction media, at 37 °C and 200 rpm. Cultures were inoculated using 1:100 dilution and samples were collected at 7 h and 24 h of cultivation. **(a)** SDS-PAGE of spent medium. Lane L: protein marker, Lane 1 and 5: OsmY(WT)-Tfu0937 in BL21(DE3), Lane 2 and 6: OsmY(M3)-Tfu0937 in BL21(DE3), Lane 3 and 7: OsmY(WT)-Tfu0937 in C41(DE3), Lane 4 and 8: OsmY(M3)-Tfu0937 in C41(DE3). **(b)** pNPG assay of the spent medium.

C41(DE3) and BL21(DE3) strains (Fig. 7). Protein secretion started in the late logarithmic phase of cell growth and most protein secretion occurred during the stationary phase. In agreement with initial experiments, the C41(DE3) strain produced more secreted protein compared to BL21(DE3), evident in the higher OsmY-Tfu0937 activity detected for the C41(DE3) host (Fig. 7a). Both OsmY(TOA4) and OsmY(M3) increased secretory protein level in BL21(DE3) compared to the wildtype. However, their improvements in BL21(DE3) were similar (Fig. 7b). After 25 h of cultivation, OsmY(TOA4) and OsmY(M3) in BL21(DE3) cells produced 2.9 ± 0.3 times and 3.1 ± 0.3 times more secreted protein respectively when compared to the wildtype. This was different in C41(DE3) where OsmY(M3) produced more secreted protein than OsmY(TOA4) (Fig. 7a). The improvements observed for OsmY(M3) were similar in both C41(DE3) and BL21(DE3) but OsmY(TOA4) showed greater improvement in BL21(DE3) compared to C41(DE3). We thus concluded that the improved secretory protein production by the OsmY variants in C41(DE3) can be transferred to BL21(DE3), though the mutations in OsmY(TOA4) likely affected secretion and/or expression in the two strains differently.

Secretion of OsmY(WT)-Tfu0937 and OsmY(M3)-Tfu0937 were further compared in a larger 400-mL cultivation using activity measurement and SDS-PAGE of the spent medium (Fig. 8). For all the samples collected, OsmY(M3) secreted more target protein compared to OsmY(WT) (Fig. 8b), except at 7 h for C41(DE3) where insignificant OsmY-Tfu0937 activity was detected. Consistent with the results in Fig. 7, BL21(DE3) secreted more target protein compared to C41(DE3) at 7 h (Fig. 8b), further supported by our SDS-PAGE analysis (Fig. 8a). OsmY(M3) secreted 6.7 ± 0.7 times more target protein compared to OsmY(WT) in C41(DE3) after 24 h cultivation (Fig. 8b), the highest we have observed thus far for Tfu0937. This is likely caused by a larger cultivation volume. In the 24 h sample for OsmY(M3)-Tfu0937 expressed in C41(DE3), we observed an additional truncated OsmY-Tfu0937 protein in the SDS-PAGE (Fig. 8a). This truncation was also seen in our previous work on OsmY-DyP4, where proteomic analysis suggested a potential N-terminal truncation within the first 50 amino acids of OsmY²³. The amount of OsmY(M3)-Tfu0937 in the spent medium (full length and truncated) at 24 h in C41(DE3) was estimated to be 100 mg/L of culture (Supplementary Fig. S3) using densitometry with ImageJ software²⁸.

OsmY(M3) variant improved the production of secretory RFP and Lipase 6B. Protein secretion is known to vary with the identity of the target protein. To determine if OsmY(M3) can be used to improve the production of other secretory proteins, the OsmY(WT) and OsmY(M3) were individually fused to the N-termini of the monomeric red fluorescent protein (RFP) and Lipase 6B. Secretion of OsmY-RFP in BL21(DE3) was observed after 24 h of cultivation (Fig. 9a,b), at which point OsmY(M3) showed more secretory RFP compared to the wildtype. The use of RFP allowed us to monitor intracellular protein easily (Fig. 9c). Comparing Fig. 9b,c helped shed light on the OsmY-assisted protein secretion. First, the protein production of OsmY(WT)-RFP was quicker compared to OsmY(M3)-RFP, as there was more total OsmY(WT)-RFP (intra- and extra-cellular) after 8 or 10 h of cultivation. This was expected because the codons in OsmY(M3) were deoptimized. Second, extracellular protein was accumulated after 24 h of cultivation (i.e., stationary phase), consistent with Tfu0937 discussed above. This was true for both OsmY(WT) and OsmY(M3). The enhanced production of secretory RFP is contributed by two factors: (a) enhanced secretion [41% in OsmY(M3) vs 27% in OsmY(WT)] and (b) increased in total protein expression level (23% higher).

Lipase 6B is a thermostable variant of a lipase from *Bacillus subtilis*²⁹. OsmY(WT)-Lipase 6B and OsmY(M3)-Lipase 6B were expressed in both BL21(DE3) and C41(DE3), and the secreted protein was monitored using the *p*-nitrophenyl acetate (pNPA) assay (Fig. 10). The C41(DE3) strain grew faster than BL21(DE3) with increasing extracellular Lipase 6B activity detected after 7 h of cultivation. Variant OsmY(M3) produced 2.5 ± 0.3 times more secreted protein compared to OsmY(WT) after 10 h of cultivation, as determined by the pNPA assay. On the

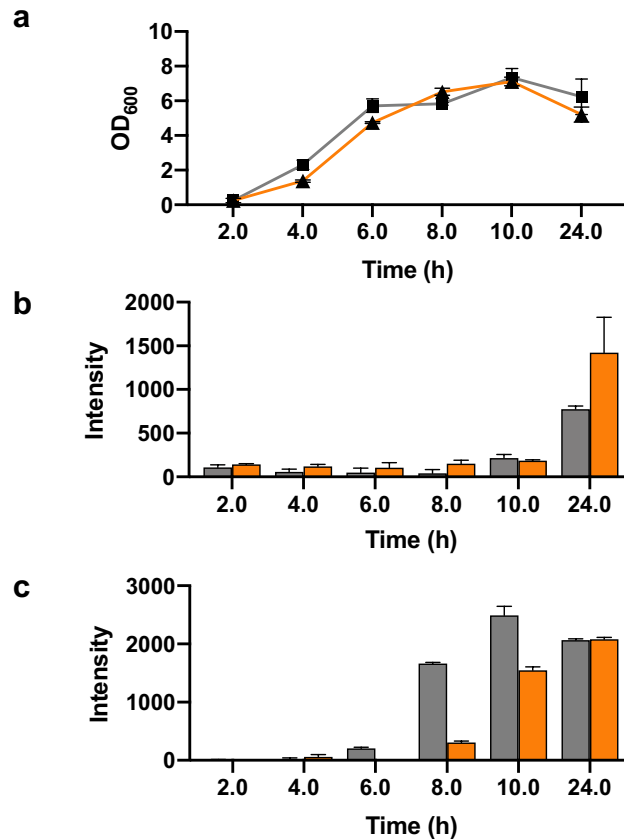


Figure 9. OsmY-RFP expression in *E. coli* BL21(DE3). Cells were cultivated in 20 mL of 2xTY auto-induction media at 37 °C using 1:100 dilution of inoculum. The supernatant and cells from 20 μ L culture were used for fluorescence measurement using $E_x = 584$ nm and $E_m = 607$ nm. (a) Growth curve of the culture. (b) Fluorescence intensity of the spent medium. (c) Fluorescence intensity of the cells. The lines and bars of different variants are represented using different colours; OsmY(WT)-RFP—grey and OsmY(M3)-RFP—orange.

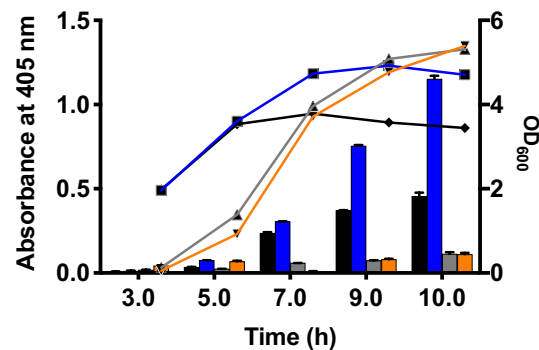


Figure 10. OsmY-Lipase 6B expressed in BL21(DE3) and C41(DE3). Cells were cultivated in 50 mL of 2xTY auto-induction media at 37 °C using 1:100 dilution of inoculum. The cell optical densities monitored at 600 nm are shown as lines. The activity of OsmY-Lipase6B in the spent medium measured with the pNPA assay using an absorbance at 405 nm are shown as bars. The lines and bars of different samples are represented using different colours; OsmY(WT)-Lipase 6B in C41(DE3)—black, OsmY(M3)-Lipase 6B in C41(DE3)—blue, OsmY(WT)-Lipase 6B in BL21(DE3)—grey and OsmY(M3)-Lipase 6B in BL21(DE3)—orange.

contrary, protein secretion was minimal in BL21(DE3) throughout the 10 h of cultivation monitored. Overexpression of recombinant lipase in *E. coli* can be toxic to the host^{29,30}. This can lead to the loss of the plasmid containing the toxic protein gene³¹, potentially explaining the poor protein expression in BL21(DE3) (Fig. 10). The strain C41(DE3) was derived from BL21(DE3) and is commonly used for the expression of membrane proteins and toxic recombinant proteins³², likely explaining its better performance for the production of secretory Lipase 6B.

Target protein	M _w (kDa)	Theoretical pI	Protein source organism	Total secretory protein production by OsmY(M3) relative to OsmY(WT)/cultivation time (h)	Recombinant protein expression host
Tfu0937	53.4	4.81	<i>Thermobifida fusca</i> YX	3.1 ± 0.3/25 2.9 ± 0.8/24	BL21(DE3) C41(DE3)
RFP	25.4	5.65	Synthetic construct	1.8 ± 0.6/24	BL21(DE3)
Lipase 6B	19.7	6.96	<i>Bacillus subtilis</i> 168	2.5 ± 0.3/10	C41(DE3)

Table 2. Production of secretory protein using OsmY(WT) and OsmY(M3) for three different target proteins.

Conclusions

In this study, we used random and combinatorial mutagenesis to improve the production of total secretory protein by OsmY. The best variant, OsmY(M3), enhanced the production of secreted β -glucosidase, OsmY(M3)-Tfu0937, in two different *E. coli* strains by approximately 3 folds (Table 2). While directed evolution had been used to randomise the junction between the signal peptide and target protein⁷ and to improve transporter proteins (e.g., ABC transporter)⁸, this is the first report that random and combinatorial mutagenesis of a secretory carrier was used to enhance secretory target protein production in *E. coli*. The critical role of the signal peptide was reflected by the two important mutations (L6P ad S19C) found in the signal peptide. Importantly, this methodology can be further extended to evolve other signal peptide or carrier proteins for secretory protein production in *E. coli*.

Subcellular fractionation of Tfu0937 with and without OsmY fusion showed that OsmY was essential for the extracellular secretion and that > 80% of OsmY-Tfu0937 protein was found extracellularly after 15 h of cultivation (Fig. 4). The transferability of improved secretory protein production by OsmY(M3) to a different *E. coli* strain and for other proteins, RFP and Lipase 6B (Table 2), showed the versatility of the OsmY(M3) variant identified. It is worth noting that these are three proteins originating from diverse organisms, having diverse molecular weights (19.7–53.4 kDa) and protein functions (Table 2). One potential limitation of a fused carrier protein is its effect on the characteristics and activities of the target protein. Introduction of a protease cleavage site for carrier protein removal can however be used to overcome this limitation¹⁶. Together with previous studies on OsmY as a secretory carrier^{18–22}, this work demonstrated the potential application of OsmY and its variants for secretion of a broad range of proteins.

Methods

Materials. Chemicals were purchased from Sigma-Aldrich (Dorset, UK) and ForMedium (Norfolk, UK). DNA modifying enzymes, deoxyribonucleotides, and DNA ladders were purchased from New England Biolabs (Hitchin, UK), Thermo Fisher Scientific (Loughborough, UK), and Agilent Technologies (Cheadle, UK). Nucleic acid purification kits were purchased from Machery-Nagel (Düren, Germany) and Omega Bio-tek (Norcross, USA). All oligonucleotides were synthesized by Eurofins Genomics (Ebersberg, Germany), and summarized in Supplementary Table S1.

Strains. *Escherichia coli* DH5 α was used for all molecular cloning, plasmid propagation and maintenance. *E. coli* BL21 (DE3) (Merck; Darmstadt, Germany) and C41(DE3) (Lucigen; Wisconsin, USA) were used for protein expression.

Molecular cloning of Tfu0937, OsmY-Tfu0937, OsmY-RFP and OsmY-Lipase 6B. The DNA sequences encoding the *E. coli* osmotically-inducible protein Y (OsmY; GenBank: AUY30809.1), the *Thermobifida fusca* YX β -glucosidase (Tfu0937; GenBank: AAZ54975.1), the *Bacillus subtilis* lipase mutant (Lipase 6B; Ref.²⁹), and a monomeric red fluorescent protein (RFP, GenBank: AAM54544.1) were codon-optimized for protein expression in *E. coli* and synthesized by GenScript (Piscataway, USA). OsmY, the secretory carrier, was fused to the N-terminus of Tfu0937 and cloned into a pET24a(+) vector (Merck; Darmstadt, Germany) using NdeI and EcoRI sites (Supplementary Fig. S4). The resulting plasmid (pET24a-OsmY-Tfu0937; Supplementary Fig. S4a) was used for protein engineering in this study.

To create the pET24a-Tfu0937 plasmid, the Tfu0937 gene was amplified with NdeI-Tfu0937-F and EcoRI-Tfu0937-R primers, digested with NdeI and EcoRI, and cloned into the pET-24a(+) vector using NdeI and EcoRI sites. To create pET24a-OsmY-RFP and pET24a-OsmY-Lipase 6B, the RFP gene was amplified using primers BamHI-RFP-F and EcoRI-RFP-R and the Lipase 6B gene was amplified using primers BamHI-Lipase6B-F and Lipase6B-EcoRI-R. The amplified RFP and Lipase 6B genes were digested with BamHI and EcoRI, and cloned into the BamHI/EcoRI digested pET24a-OsmY-Tfu0937 plasmid to replace the Tfu0937 gene. The sequences of all primers were summarised in Supplementary Table S1.

Random mutagenesis by epPCR. Two error-prone polymerase chain reaction (epPCR) conditions were used in this study: high and low mutagenic rates as previously reported²³. For epPCR of high mutagenic rate, the 50- μ l PCR mixture contained 1 \times standard Taq reaction buffer (Mg-free), 7 mM MgCl₂, 0.05 mM MnCl₂, 0.2 mM dATP, 0.2 mM dGTP, 1 mM dTTP, 1 mM dCTP, 20 pmol OsmY-DirEv-F primer, 20 pmol OsmY-DirEv-R primer (Supplementary Table S1), 50 ng pET-24a(+)-OsmY-Tfu0937 and 1.25 U Taq DNA polymerase (New England Biolabs). For epPCR of low error rate, the 50- μ l PCR mixture contained 1 \times standard Taq reaction buffer (Mg-free), 1.5 mM MgCl₂, 0.01 mM MnCl₂, 0.3 mM of each dNTP, 4.5 pmol OsmY-DirEv-F primer,

4.5 pmol OsmY-DirEv-R (Supplementary Table S1), 7.77 ng pET-24a(+)-OsmY-Tfu0937 and 1.25 U Taq DNA polymerase (New England Biolabs). The PCR mixtures were thermocycled using the following conditions: (i) 30 s initial denaturation at 95 °C, (ii) 25 cycles of 20 s denaturation at 95 °C, 30 s annealing at 52 °C and 1 min 30 s extension at 68 °C, and (iii) 5 min final extension at 68 °C. PCR products were purified by gel extraction. After restrictive digestion with NdeI and BamHI, the PCR products were cloned into NdeI/BamHI digested pET24a-OsmY-Tfu0937 plasmid. The recombinant plasmids were subsequently electroporated into *E. coli* C41(DE3) and plated on TYE agar (10 g/L tryptone, 5 g/L yeast extract, 8 g/L NaCl and 15 g/L agar) supplemented with 50 µg/mL kanamycin to create the OsmY mutant libraries.

Site-directed mutagenesis to create variants OsmY(L6P), OsmY(S154R) and OsmY(V191E). Site-directed mutagenesis experiments were performed using partially overlapping primers designed by the OneClick program³³. Primers OsmY(L6P)-FP and OsmY(L6P)-RP were used for OsmY(L6P), primers OsmY(V154R)-FP and OsmY(S154R)-RP were used for OsmY(S154R), and primers OsmY(V191E)-FP and OsmY(V191E)-RP were used for OsmY(V191E) (Supplementary Table S1). In the two-stage PCR designed by OneClick³³, PfuUltra High-Fidelity polymerase (Agilent Technologies) was used to introduce the L6P and S154R mutations, and PfuUltra II Fusion HS DNA Polymerase (Agilent Technologies) was used to introduce the V191E mutation.

Cultivation and protein expression in 96-well microplate. Individual colonies of the mutant library were manually transferred from the agar plates using sterile toothpicks into 96-well deep-well microplates, with each well containing 1.5 mL 2 × TY medium (16 g/L tryptone, 10 g/L yeast extract and 5 g/L NaCl) supplemented with 50 µg/mL kanamycin. Column 6 on the microplate was inoculated with the wildtype as internal control. These master plates were covered with lids, sealed and cultivated at 30 °C for 24 h.

To express protein for screening, master plates were replicated using a pin replicator into fresh 96-well deep-well plates, with each well containing 1.5 mL 2 × TY auto induction medium [16 g/L tryptone, 10 g/L yeast extract, 3.3 g/L (NH₄)₂SO₄, 6.8 g/L KH₂PO₄, 7.1 g/L Na₂HPO₄, 0.5 g/L glucose, 2.0 g/L α-lactose and 0.15 g/L MgSO₄] supplemented with 50 µg/mL kanamycin. The plates were cultivated at 30 °C for 24 h before (a) they were transferred to a fridge and kept at 4 °C overnight for the cells to settle or (b) they were plate centrifuged at 4000 rpm for 10 min (Heraeus Megafuge 8R; M10 microplate rotor). The supernatant containing the secreted OsmY-Tfu0937 was used for high-throughput screening on the next day.

Abgene 96-well polypropylene deep-well plates (Thermo Fisher Scientific; AB0661) and Abgene polypropylene plate covers (Thermo Fisher Scientific; AB0755) were used in preparing master plates and protein expression plates. All plate cultivations were conducted in Titramax 1000 plate shaker coupled to an Incubator 1000 heating module (Heidolph Instruments; Essex, UK) using a shaking speed of 1050 rpm.

High-throughput screening in 96-well microplate. Flat-bottom clear 96-well polystyrene microplates (Greiner Bio-One; 655161) were used for screening. Unless otherwise stated, 20 µL of spent medium was transferred to 96-well microplates, with each well containing 30 µL of 0.2 M sodium acetate. Fifty µL of pNPG solution (5.3 mM pNPG, 0.2 M sodium acetate, pH 4.8) was then added to each well. The microplates were incubated at 50 °C and 1050 rpm in a microplate shaker for 5 min before 100 µL of NaOH-Gly buffer (0.4 M glycine, pH 10.8) was added to stop the reaction. Absorbance at 405 nm was recorded with Multiskan FC microplate photometer (Thermo Fisher Scientific). All shaking steps were conducted in Titramax 1000 (Heidolph Instruments) using a shaking speed of 1050 rpm.

Variant characterization using tube and flask cultures. The variants were characterized using two methods, (a) tube cultures and (b) flask cultures. For tube cultures, 5 µL of cells from the master plate was transferred to culture tubes containing 5 mL of 2 × TY auto induction medium supplemented with 50 µg/mL kanamycin and cultivated overnight at 37 °C. Flask cultures were used for expression studies, where 20 mL or 50 mL of 2 × TY-based auto induction medium was inoculated using 1:100 or 1:200 dilution of the pre-culture and cultivated at 37 °C and 240 rpm, unless otherwise stated. During cultivation, cell growth was monitored by absorbance reading at 600 nm (OD₆₀₀). For assay measurement, 0.5–1.0 mL of culture was sampled and secreted protein was measured using the pNPG assay (Tfu0937), the pNPA assay (Lipase 6B) or fluorescence intensity (RFP).

The pNPG assay, pNPA assay and RFP fluorescence measurement. The sampled culture (0.5–1.0 mL) was centrifuged at > 17,000 × g for 5–10 min and 50 µL of the supernatant containing the secreted protein was used for assay. The supernatant was diluted to keep within the assay linear detection range when necessary. The pNPG assay was used to measure the Tfu0937 activity. For pNPG assay, 50 µL of supernatant was transferred to a microplate and sample was assayed using the same protocol as high-throughput screening. The supernatant was diluted using either 2 × TY medium or sample buffers for the subcellular fractions when necessary. The molar extinction coefficient of 18,000/M/cm was used to calculate the *p*-nitrophenol released. The pNPA assay was used to measure the Lipase 6B activity. For pNPA assay, 20 µL of supernatant was transferred to a microplate and 180 µL of freshly prepared pNPA solution (2.0 mM pNPA, 50 mM potassium phosphate buffer, pH 7.2) was added. Reaction was incubated at room temperature for 10 min before the absorbance was recorded at 405 nm. The supernatant was diluted using 50 mM potassium phosphate buffer, pH 7.2 assay buffer when necessary. For RFP detection, 100 µL of supernatant was transferred to a black microplate (Greiner Bio-One; 655906) and fluorescence intensity was measured using the SpectraMax M2e plate reader (Molecular Device; California, USA) at excitation wavelength of 584 nm and emission wavelength of 607 nm. The RFP fluorescence intensity of the cell

was measured using cells from 100 μL of culture resuspended in 100 μL of 50 mM K_2PO_4 , pH 7.5 and at the same excitation and emission wavelengths. Assay background, for the pNPG and pNPA assays, was always removed by subtracting the assayed absorbance (at 405 nm) of the spent medium from the culture of cells without plasmids [i.e., BL21(DE3) or C41(DE3)].

Soluble protein extraction using enzymatic lysis. One mL of culture was centrifuged at $>17,000g$ for 5–10 min. The supernatant was removed, and the cell pellet was resuspended in 200 μL of lysomix (10 mg/mL lysozyme in 50 mM, pH 7.5) by vortexing. Sample was incubated on ice for 1 h followed by centrifugation at $>17,000g$ for 10 min. The supernatant, which contained the soluble protein fraction, was transferred to a fresh tube. OsmY-Tfu0937 activity in the soluble fraction was determined using the pNPG assay.

Subcellular fractionation of Tfu0937 and OsmY-Tfu0937. Cells were cultivated as described earlier using 50 mL or 100 mL of $2\times$ TY-based auto induction medium. The cells were harvested by centrifugation at 6000g, 4 $^\circ\text{C}$ for 5 min. The supernatant was transferred to a clean tube and used as the extracellular fraction. The cell pellet was resuspended in 25 mL of osmotic shock buffer (100 mM Tris-HCl, 20% sucrose, 1 mM EDTA, pH 8.0) and incubated on ice for 15 min. The cell suspension was then centrifuged at 7000g, 4 $^\circ\text{C}$ for 10–15 min and the clear supernatant was transferred to a clean tube (fraction P1). The resulting cell pellet was resuspended in 2.5 mL of ice-cold water with 1 mM MgCl_2 and incubated on ice for 15 min. The cell suspension was centrifuged at 7000g, 4 $^\circ\text{C}$ for 10 min and the clear supernatant was transferred to a clean tube (fraction P2). Fractions P1 and P2 were used to determine Tfu0937 activity in the periplasmic fraction. The cell pellet following periplasmic extraction was resuspended in 10 mL of 100 mM Tris-HCl buffer (pH 8.0) by vortexing and cells were lysed by sonication (Vibra-Cell ultrasonic liquid processors; Sonics & Materials; Newtown, USA) using 40% amplitude and 1.5 min total processing time. After sonication, 2 mL of sonicated sample was centrifuged at 20,000g, 4 $^\circ\text{C}$ for 10 min. The supernatant was collected as the cytoplasmic fraction.

Consent for publication. All authors have read the manuscript and approved its submission.

Data availability

All data generated or analysed during this study are included in this published article and its supplementary information files.

Received: 3 November 2020; Accepted: 22 February 2021

Published online: 05 March 2021

References

- Makrides, S. C. Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol. Rev.* **60**(3), 512–538 (1996).
- Pohl, S. & Harwood, R. C. Heterologous protein secretion by *Bacillus* species: From the cradle to the grave. *Adv. Appl. Microbiol.* **73**, 1–25 (2010).
- Burdette, L. A., Leach, S. A., Wong, H. T. & Tullman-Ercek, D. Developing Gram-negative bacteria for the secretion of heterologous proteins. *Microb. Cell Fact.* **17**(1), 196 (2018).
- Freudl, R. Signal peptides for recombinant protein secretion in bacterial expression systems. *Microb. Cell Fact.* **17**(1), 52 (2018).
- Brockmeier, U. *et al.* Systematic screening of all signal peptides from *Bacillus subtilis*: A powerful strategy in optimizing heterologous protein secretion in Gram-positive bacteria. *J. Mol. Biol.* **362**(3), 393–402 (2006).
- Hemmerich, J. *et al.* Use of a Sec signal peptide library from *Bacillus subtilis* for the optimization of cutinase secretion in *Corynebacterium glutamicum*. *Microb. Cell Fact.* **15**(1), 208 (2016).
- Kaderbhai, M. A., Davey, H. M. & Kaderbhai, N. N. A directed evolution strategy for optimized export of recombinant proteins reveals critical determinants for preprotein discharge. *Protein Sci.* **13**(9), 2458–2469 (2004).
- Eom, G. T., Song, J. K., Ahn, J. H., Seo, Y. S. & Rhee, J. S. Enhancement of the efficiency of secretion of heterologous lipase in *Escherichia coli* by directed evolution of the ABC transporter system. *Appl. Environ. Microbiol.* **71**(7), 3468–3474 (2005).
- Mirzadeh, K. *et al.* Increased production of periplasmic proteins in *Escherichia coli* by directed evolution of the translation initiation region. *Microb. Cell Fact.* **19**(1), 85 (2020).
- Zalucki, Y. M., Shafer, W. M. & Jennings, M. P. Directed evolution of efficient secretion in the SRP-dependent export of TolB. *Biochim. Biophys. Acta.* **1808**(10), 2544–2550 (2011).
- Yang, H., Lu, X., Hu, J., Chen, Y., Shen, W., Liu, L. Boosting secretion of extracellular protein by *Escherichia coli* via cell wall perturbation. *Appl. Environ. Microbiol.* **84**(20) (2018).
- Natarajan, A., Haitjema, C. H., Lee, R., Boock, J. T. & DeLisa, M. P. An engineered survival-selection assay for extracellular protein expression uncovers hypersecretory phenotypes in *Escherichia coli*. *ACS Synth. Biol.* **6**(5), 875–883 (2017).
- Bernstein, C., Bernstein, H., Payne, C. M., Beard, S. E. & Schneider, J. Bile salt activation of stress response promoters in *Escherichia coli*. *Curr. Microbiol.* **39**(2), 68–72 (1999).
- Oh, J. T. *et al.* Cationic peptide antimicrobials induce selective transcription of micF and osmY in *Escherichia coli*. *Biochim. Biophys. Acta.* **1463**(1), 43–54 (2000).
- Yim, H. H. & Villarejo, M. osmY, a new hyperosmotically inducible gene, encodes a periplasmic protein in *Escherichia coli*. *J. Bacteriol.* **174**(11), 3637–3644 (1992).
- Qian, Z. G., Xia, X. X., Choi, J. H. & Lee, S. Y. Proteome-based identification of fusion partner for high-level extracellular production of recombinant proteins in *Escherichia coli*. *Biotechnol. Bioeng.* **101**(3), 587–601 (2008).
- Yeats, C. & Bateman, A. The BON domain: A putative membrane-binding domain. *Trends Biochem. Sci.* **28**(7), 352–355 (2003).
- Gupta, S., Adlakha, N. & Yazdani, S. S. Efficient extracellular secretion of an endoglucanase and a beta-glucosidase in *E. coli*. *Protein Expr. Purif.* **88**(1), 20–25 (2013).
- Le, Y. & Wang, H. High-level soluble expression of a thermostable xylanase from thermophilic fungus *Thermomyces lanuginosus* in *Escherichia coli* via fusion with OsmY protein. *Protein Expr. Purif.* **99**, 1–5 (2014).
- Zheng, Z., Chen, T., Zhao, M., Wang, Z. & Zhao, X. Engineering *Escherichia coli* for succinate production from hemicellulose via consolidated bioprocessing. *Microb. Cell Fact.* **11**, 37 (2012).

21. Cheng, C. M. *et al.* Functional production of a soluble and secreted single-chain antibody by a bacterial secretion system. *PLoS ONE* **9**(5), e97367 (2014).
22. Kotzsch, A. *et al.* A secretory system for bacterial production of high-profile protein targets. *Protein Sci.* **20**(3), 597–609 (2011).
23. Alessa, A. H. A. *et al.* Accelerated directed evolution of dye-decolorizing peroxidase using a bacterial extracellular protein secretion system (BENNY). *Bioresour. Bioprocess.* **6**(1), 20 (2019).
24. Lennon, C. W. *et al.* Folding optimization in vivo uncovers new chaperones. *J. Mol. Biol.* **427**(18), 2983–2994 (2015).
25. von Heijne, G. Signal sequences. The limits of variation. *J. Mol. Biol.* **184**(1), 99–105 (1985).
26. Wang, Y., Mao, Y., Xu, X., Tao, S. & Chen, H. Codon usage in signal sequences affects protein expression and secretion using baculovirus/insect cell expression system. *PLoS ONE* **10**(12), e0145887 (2015).
27. Adav, S. S., Ng, C. S., Arulmani, M. & Sze, S. K. Quantitative iTRAQ secretome analysis of cellulolytic *Thermobifida fusca*. *J. Proteome Res.* **9**(6), 3016–3024 (2010).
28. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods.* **9**(7), 671–675 (2012).
29. Kamal, M. Z. *et al.* In vitro evolved non-aggregating and thermostable lipase: Structural and thermodynamic investigation. *J. Mol. Biol.* **413**(3), 726–741 (2011).
30. Bell, P. J. L. *et al.* Prospecting for novel lipase genes using PCR. *Microbiology* **148**(Pt 8), 2283–2291 (2002).
31. Kato, Y. Extremely low leakage expression systems using dual transcriptional–translational control for toxic protein production. *Int. J. Mol. Sci.* **21**(3), 705 (2020).
32. Dumon-Seignovert, L., Cariot, G. & Vuillard, L. The toxicity of recombinant proteins in *Escherichia coli*: A comparison of over-expression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expr. Purif.* **37**(1), 203–206 (2004).
33. Warburton, M. *et al.* OneClick: A program for designing focused mutagenesis experiments. *AIMS Bioeng.* **2**(3), 126–143 (2015).

Acknowledgements

We thank EPSRC (EP/E036252/1), the Open Project Funding of the State Key Laboratory of Bioreactor Engineering (to JX and TSW), COST action (CM1303 Systems Biocatalysis; to DGP), the Royal Academy of Engineering (the Leverhulme Trust Senior Research Fellowship; to TSW; LTSRF1819\15\21), BBSRC GCRF-STARS (BB/R020183/1; to TSW), The University of Sheffield (GCRF Fellowship; to KLT), the Department of Chemical and Biological Engineering (Summer Undergraduate Fellowship; to JR, MCMW and YPY) for financial support.

Author contributions

K.L.T. and T.S.W. supervised the project and designed the experiments. K.L.T., J.R., D.G.P., S.K.T., M.C.M.W., Y.P.Y. and N.N. conducted the experiments. K.L.T., T.S.W. and J.X. analysed the data and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-84859-6>.

Correspondence and requests for materials should be addressed to T.S.W. or K.L.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021