A fusion protein strategy for soluble expression of *Stevia* glycosyltransferase UGT76G1 in *Escherichia coli*

Liangliang Chen¹ · Ping Sun¹ · Yan Li^{1,2} · Ming Yan¹ · Lin Xu¹ · Kequan Chen¹ · Pingkai Ouyang¹

Received: 16 May 2017 / Accepted: 5 September 2017 / Published online: 3 October 2017
© Springer-Verlag GmbH Germany 2017

Abstract The UDP-glucosyltransferase UGT76G1 from *Stevia rebaudiana* converts stevioside to rebaudioside A via a one-step glycosylation reaction, which increases the amount of sweet-tasting rebaudioside A and decreases the amount of stevioside that has a bitter aftertaste. This enzyme could, therefore, conceivably be used to improve the organoleptic properties of steviol glycosides and offer a cost-effective preparation of high-purity rebaudioside A. Producing soluble enzymes by overexpression is a prerequisite for large-scale biocatalysis. However, most of the UGT76G1 overexpressed in *Escherichia coli* is in inclusion bodies. In this study, three N-terminal fusion partners, 3'-phosphoadenosine-5'-phosphatase (CysQ), 2-keto-3-deoxy-6-phosphogluconate aldolase (EDA) and N-utilisation substance A (NusA), were tested to improve UGT76G1 expression and solubility in *E. coli*. Compared with the fusion-free protein, the solubility of UGT76G1 was increased 40% by fusion with CysQ, and the glucosyltransferase activity of the crude extract was increased 82%. This successful CysQ fusion strategy could be applied to enhance the expression and solubility of other plant-derived glucosyltransferases and presumably other unrelated proteins in the popular, convenient and cost-effective *E. coli* host.

Keywords Fusion partner · Inclusion body · Rebaudioside A · Stevioside · UGT76G1 · Glucosyltransferase

Introduction

Steviol glycosides give the leaves of *Stevia rebaudiana* their characteristic sweet taste and are used in *stevia*-based sweeteners (Urban et al. 2015). In addition to their intense sweetness, these compounds are low in calories and possess antihyperglycemic, antihypertensive, anti-inflammatory, antitumor and immunomodulatory properties (Wolwer-Rieck 2012). Rebaudioside A and stevioside are the predominant sweetening agents in steviol glycosides, but rebaudioside A possesses higher sweetness and better taste than stevioside (Tada et al. 2013). The United States Food and Drug Administration has recognised rebaudioside A to be 'generally safe' since 2008 (Rumelhard et al. 2016). As a consequence, the physical separation of rebaudioside A from the *stevia* extracts has become a fast-growing area of research (Chen et al. 1999; Li et al. 2012; Lui et al. 2011).

UDP-glucosyltransferase UGT76G1 (GenBank Accession no. AAR06912) from *S. rebaudiana*, which belongs to the class I family of glycosyltransferase multigenes (Paquette et al. 2003), catalyses the conversion of stevioside to rebaudioside A via a one-step glycosylation reaction (Brandle and Telmer 2007). By increasing the amount of the sweet-tasting rebaudioside A relative to steviol glycoside products that have a residual bitter aftertaste, enzymatic synthesis of rebaudioside A from stevioside by UGT76G1 can be used for the preparative separation and purification of rebaudioside A (Adari et al. 2016; Wang et al. 2016). We established a cell-free system for conversion of stevioside to rebaudioside A by coupling the activities of recombinant UGT76G1 from *S. rebaudiana* and sucrose synthase AtSUS1 from *Arabidopsis thaliana* (Wang et al. 2016). However, many plant-derived enzymes form insoluble inclusion bodies when overexpressed in

✉ Yan Li
liyan@njtech.edu.cn

¹ College of Biotechnology and Pharmaceutical Engineering, Nanjing Tech University, Nanjing 211800, China

² Yichang Key Laboratory of Biocatalysis, China Three Gorges University, Yichang 443002, China

Escherichia coli, resulting in low enzyme activities that impede effective scale-up (Desmet et al. 2012; Georgiou and Valax 1996; Xu et al. 1998). In our previous work (Wang et al. 2016), recombinant UGT76G1 was mainly expressed in inclusion bodies in *E. coli*. Production of glycosyltransferases can be improved in a stepwise manner by manipulating the promoter strength (strong T7 promoter vs. intermediate P34 promoter), testing inducible vs. constitutive expression systems (pET21 vs. pCXP34h vectors) and changing the *E. coli* expression strain (BL21 vs. Origami2) (Dewitte et al. 2016). In the present study, we decided to try another method to improve protein solubility for efficient enzyme production.

Due to its relatively easy genetic manipulation, fast biomass accumulation and inexpensive manufacturing cost, *E. coli* is the preferred host for recombinant protein expression (Zerbs et al. 2014). However, unlike eukaryotic systems, translation occurs almost simultaneously with transcription, and the prokaryote host lacks posttranslational modification systems. Thus, many proteins, especially eukaryotic proteins, form undesirable aggregates (inclusion bodies) when heterologously expressed in *E. coli* (Georgiou and Valax 1999), which restricts the production of active recombinant proteins. Various methods aimed at improving protein solubility during heterologous expression in *E. coli* have been adopted, such as lowering the temperature (Jhamb and Sahoo 2012), codon optimisation (Kane 1995), coexpression with molecular chaperones (Jhamb et al. 2008; Pei et al. 2015; Schlieker et al. 2002; Tong et al. 2016) and expression with fusion partners (Ahn et al. 2007; Hellman et al. 1992; Kim and Lee 2008; Pasek et al. 2010). The most widely used fusion expression partners include maltose-binding protein (Raran-Kurussi et al. 2015), N-utilisation substance (NusA) (De Marco et al. 2004; Niiranen et al. 2007) and glutathione-S-transferase (Li et al. 2016), all of which can divert inclusion body formation and enhance the expression of functionally folded target proteins in soluble form. Additionally, several new proteins have been shown to be effective expression- and solubility-enhancers for aggregation-prone heterologous proteins in *E. coli* (Kang et al. 2015; Lee et al. 2014). To avoid lowering the yield of target proteins, aggregation-resistant fusion partners with a smaller molecular mass are deemed more appropriate. In the present study, two proteins with a molecular weight of no more than 30 kDa, 3'-phosphoadenosine-5'-phosphatase (CysQ) (Lee et al. 2014) and 2-keto-3-deoxy-6-phosphogluconate aldolase (EDA) (Kang et al. 2015) and the heavier NusA (55 kDa; Table 1) were tested as N-terminal fusion expression partners to improve the expression of recombinant UGT76G1 in *E. coli*. CysQ is believed to be required during aerobic growth in *E. coli* to help control the level of 3'-phosphoadenosine-5'-phosphosulfate in cysteine biosynthesis

(Neuwald et al. 1992). Because of the significantly increased expression of CysQ under exogenous stress conditions in *E. coli* BL21 (DE3), CysQ has an intrinsic ability to adopt its native conformation (Lee et al. 2014). As a fusion expression partner, CysQ is highly effective for enhancing the cytoplasmic solubility of various aggregation-prone heterologous proteins to ensure they retain native bioactivity and native secondary structure (Lee et al. 2014).

EDA catalyses the reversible cleavage of 2-keto-3-deoxy-6-phosphogluconate to pyruvate and glyceraldehyde-3-phosphate in the Entner–Doudoroff pathway (Han et al. 2008). Because it retains its structure even under protein denaturing conditions, EDA can be applied as an effective fusion expression partner to increase the cytoplasmic solubility of aggregation-prone heterologous proteins (Kang et al. 2015). Genes encoding the respective fusion partner and UGT76G1 were fused in plasmid pET-28a(+), transformed into *E. coli*, and protein expression and enzymatic activities of the resultant fusion proteins were measured.

Materials and methods

Cloning

The *E. coli* strain BL21 (DE3) was used as the host for protein expression. Using PCR amplification with the appropriate primers (Table 2), the codon-optimised gene (LC312448), derived from *S. rebaudiana* UGT76G1 mRNA (AY345974), was synthesized by GenScript (Nanjing, China) and then cloned into pET-28a(+) (Novagen, Darmstadt, Germany) between *Nde*I and *Xho*I restriction endonuclease sites (Fig. 1a), and genes encoding each of the three solubility-enhancer proteins CysQ, EDA and NusA were separately cloned into pET-28a(+) between *Nde*I and *Hind*III restriction endonuclease sites, with an enterokinase (EK) cleavage site (DDDDK, D₄K) at the C-terminus of the fusion partner (Fig. 1b). The resultant recombinant plasmids were named pET28a-UGT, pET28a-CysQ, pET28a-EDA and pET28a-NusA, respectively. As the heterologous gene was inserted after *Nde*I in pET-28a(+), a polyhistidine (H₆) tag was included at the N-terminus of the recombinant protein to facilitate its purification. To construct the N-terminal fusion partner of UGT76G1, the genes encoding UGT76G1 were separately incorporated between the restriction endonuclease sites *Hind*III and *Eco*RI of pET28a-CysQ, pET28a-EDA and pET28a-NusA (Fig. 1c) using the primers listed in Table 3. The resultant fusion constructs were named pET28a-CysQ-UGT, pET28a-EDA-UGT and pET28a-NusA-UGT, and all included a cleavage sequence D₄K for enzymatic digestion

Table 1 Description of the fusion expression partners

Gene	Gene ID	Fusion expression partner	Protein size (kDa)
<i>cysQ</i>	948728	3'-Phosphoadenosine-5'-phosphatase	27
<i>eda</i>	946367	2-Keto-3-deoxy-6-phosphogluconate aldolase	22
<i>nusA</i>	947682	N-utilizing substance A	55

Table 2 Primers used to amplify the target genes

Heterologous protein	Accession no.	Primer sequences for direct expression
UGT76G1	AAR06912	5'-gtg ccg cgc ggc agc <u>cat atg</u> gaa aat aaa acc gaa acc acc gtc c-3' 5'-gtg gtg gtg gtg gtg <u>ctc gag</u> tta cag aga gct gat gta tga aac c-3'
CysQ	BAE78215	5'-gtg ccg cgc ggc agc <u>cat atg</u> tta gat caa gta tgc cag ctt g-3' 5'-ctc gag tgc ggc cgc <u>aag ctt</u> gta aat aga cac tct gaa ccc c-3'
EDA	BAA15658	5'-gtg ccg cgc ggc agc <u>cat atg</u> aaa aac tgg aaa aca agt gca g-3' 5'-ctc gag tgc ggc cgc <u>aag ctt</u> cag ctt agc gcc ttc tac agc-3'
NusA	BAE77215	5'-gtg ccg cgc ggc agc <u>cat atg</u> aac aaa gaa att ttg gct gta g-3' 5'-ctc gag tgc ggc cgc <u>aag ctt</u> cgc ttc gtc acc gaa cca gc-3'

Fig. 1 Plasmids used for expression of **a** UGT76G1, **b** solubility enhancer proteins CysQ, EDA and NusA, and **c** fusion proteins CysQ/EDA/NusA-UGT76G1

by EK between the N-terminal fusion partner and the target protein UGT76G1.

PCRs were carried out using a Biometra Thermocycler T-Gradient Thermo Block (American Laboratory Trading, Inc., Connecticut, USA). Cycling parameters included a pre-denaturation at 94 °C for 5 min, 30 cycles of denaturation at 94 °C for 30 s, annealing at 62 °C for 30 s,

and extension at 72 °C for 100 s, and a final extension for 5 min at 72 °C. PCR products were isolated by electrophoresis on a 0.8% agarose gel and purified using an agar gel DNA Extraction Kit (Shanghai Generay Biotech Co., Ltd., Shanghai, China). Expression constructs were verified by DNA sequencing and used to transform *E. coli* BL21 (DE3). Clones were subsequently selected using Luria–Bertani (LB)-agar plates supplemented with kanamycin (50 mg/l).

Expression of recombinant proteins in *E. coli*

For expression of recombinant proteins, each *E. coli* transformant was incubated in 5 ml of LB media (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl) with kanamycin (50 mg/l) at 37 °C with shaking at 200 rpm overnight. These small cultures were used to inoculate larger 50 ml cultures that were grown in the same way. After 2 h, cultures were transferred to an incubator at 30 °C and incubated with shaking for another 0.5 h, at which point the culture turbidity (OD₆₀₀) reached 0.5–0.6, and gene expression was induced by the addition of 0.5 mM isopropyl-β-D-thiogalactoside. After a further 16 h of cultivation, cells were harvested by centrifugation (5432×g, 5 min) and cell pellets were washed twice with 5 ml 50 mM potassium phosphate buffer (pH 7.2). Cell pellets were resuspended in 1 ml of the same buffer and disrupted using a Sonifier (Ningbo Scientz Biotechnology Co. LTD., Zhejiang, China). Following centrifugation at 7000×g for 30 min, cell-free supernatants were subjected to polyacrylamide gel electrophoresis analysis on a 12.5% Tris–glycine precast gel (Bio-Rad Laboratories, Inc., CA, USA). Protein concentration was determined using the Bradford method (Bradford 1976) using bovine serum albumin as a

Table 3 Primers used for generating fusion partner constructs

Fusion protein	Primer sequences for fusion expression
CysQ-UGT	5'-aga gtg tct att tac <u>aag ctt</u> gat gac gat gac aag gaa aat aaa acc gaa acc-3' 5'-gtg gtg gtg gtg gtg <u>ctc gag</u> tta cag aga gct gat gta tga aac c-3'
EDA-UGT	5'-gaa ggc gct aag ctg <u>aag ctt</u> gat gac gat gac aag gaa aat aaa acc gaa acc-3' 5'-gtg gtg gtg gtg gtg <u>ctc gag</u> tta cag aga gct gat gta tga aac c-3'
NusA-UGT	5'-ttc ggt gac gaa gcg <u>aag ctt</u> gat gac gat gac aag gaa aat aaa acc gaa acc-3' 5'-gtg gtg gtg gtg gtg <u>ctc gag</u> tta cag aga gct gat gta tga aac c-3'

standard. Supernatants were used as crude enzyme extracts, and both supernatants and insoluble fractions were analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Samples containing the enzyme were separated by 12.5% SDS-PAGE and stained with Coomassie brilliant blue. Gels were analysed using a GenoSens1860 instrument (Clinx Science Instruments Co., Ltd, Shanghai, China), and UN-SCAN-IT gel analysis software (<https://www.silkscientific.com/gel-analysis.htm>) was used to calculate the relative concentration of target protein in each band based on the optical densitometry.

Purification of recombinant proteins

Purification of recombinant UGT76G1 and CysQ-UGT was carried out using metal affinity chromatography. First, 1 ml of Ni²⁺-NTA agarose resin (GenScript, Nanjing, China) was loaded into a column and prewashed twice with 10 ml of Lysis Equilibration Buffer (50 mM sodium dihydrogen phosphate, 300 mM NaCl, pH 8.0) at a flow-rate of 0.5–1 ml per min. Cell lysates (5 mg total protein) containing (H₆)-UGT76G1 or (H₆)-CysQ-UGT76G1 were loaded onto the Ni-charged resin and 6 ml of Wash Buffer I (50 mM sodium dihydrogen phosphate, 300 mM NaCl, 20 mM imidazole, pH 8.0) was passed through the column to avoid non-specific binding of untagged *E. coli* host proteins. Next, 6 ml of Wash Buffer II (50 mM sodium dihydrogen phosphate, 300 mM NaCl, 50 mM imidazole, pH 8.0) was passed through the column. Binding was carried out in batch mode at 4 °C, and resin was washed three times with 10 ml Elution Buffer (50 mM sodium dihydrogen phosphate, 300 mM NaCl, 250 mM imidazole, pH 8.0) to elute the target proteins. Resin was subsequently washed twice with 10 ml Lysis Equilibration Buffer to prepare the column for the next purification. Fractions were analysed using SDS-PAGE.

Glucosyltransferase activity assay

The substrate stevioside (1 mM) was incubated with ~0.5 mg of crude or purified UGT76G1 or CysQ-UGT for 30 min at 30 °C in a 3-ml reaction mixture containing 2 mM UDP-glucose, 50 mM potassium phosphate buffer (pH 7.2) and 3 mM MgCl₂ (Madhav et al. 2013). Before and after the incubation period, a 500-μl sample was removed, heated at 95 °C for 15 min, centrifuged at 13,523×*g* and the supernatant was collected. The control reaction contained water instead of enzyme. All experiments were performed in triplicate. Samples were extracted twice with an equal volume of water-saturated 1-butanol and dried using a QYN100-1 nitrogen purge device (Joyn Electronic, Shanghai, China). Pooled butanol fractions were dried completely and re-dissolved in 300 μl of water-saturated 1-butanol. Samples were then measured using high-performance liquid chromatography (HPLC). One unit (U) of glucosyltransferase activity was defined as the amount of enzyme that produced 1 μmol rebaudioside A per min under the described conditions. Because the proteins were of different molecular weights [UGT76G1 = 52.9 kDa; CysQ-UGT = 80.6 kDa; calculating by ExPASy Compute pI/Mw (http://web.expasy.org/compute_pi/)], we used mU/μmol as units for specific activity of the purified enzymes.

HPLC analysis

HPLC was used to measure the concentration of rebaudioside A on an Agilent Infinity LC system (Agilent Technologies, Santa Clara, CA, USA) equipped with a Galaksil BF-NH₂ column (4.6 × 250 mm, 5 μm 120 Å) maintained at 40 °C and monitored by UV detection at 210 nm (Wuxi Galak Chromatography Technology Co. Ltd., Wuxi, China) (Kolb et al. 2001; Wang et al. 2016). The flow rate was 1 ml/min using 20% H₂O in CH₃CN (pH adjusted to 4.6 with acetic acid) as the mobile phase. The injection volume was 10 μl. A rebaudioside A standard was obtained commercially (Wako Pure Chemical Industries, Ltd., Osaka, Japan).

Results and discussion

Expression of UGT76G1 fused with CysQ, EDA and NusA

Three plasmids, pET28a-CysQ, pET28a-EDA and pET28a-NusA, each containing a different fusion partner, together with pET28a-UGT containing the gene encoding UGT76G1 without a fusion partner, were transformed into *E. coli* and expression was carried out. High levels of soluble expression were observed for all three fusion partners (Fig. 2). By

Fig. 2 SDS-PAGE analysis of soluble and insoluble fractions obtained from recombinant *E. coli* cell lysates. M, the standard molecular marker; S and I (Blank), soluble and insoluble fractions from BL21(DE3) cells transformed with pET-28a(+); S and I (UGT76G1), soluble and insoluble fractions from BL21(DE3) cells transformed with pET28a-UGT; S and I (CysQ, EDA and NusA), soluble and insoluble fractions from BL21(DE3) cells transformed with pET28a-CysQ, pET28a-EDA and pET28a-NusA. Arrows in gel images indicate recombinant protein bands

contrast, expression of the fusion partner-free UGT76G1 protein resulted in insoluble inclusion bodies.

We next introduced the gene encoding UGT76G1 into the engineered plasmids containing the fusion partners and transformed the resulting constructs into the *E. coli* host. As shown in Fig. 3, the solubility of the aggregation-prone target protein was enhanced following fusion with all three N-terminal protein partners. Fusion with CysQ resulted in the highest amount

of target protein in the soluble fraction. We speculate that CysQ might promote the binding of molecular chaperones that bind to its cluster of hydrophobic amino acids (Lee et al. 2014).

In general, a fusion partner of lower molecular weight results in a higher yield of target protein, and a larger fusion partner might be more likely to interfere with the biological activity of target proteins (Lee et al. 2014). Therefore, the CysQ fusions were chosen for subsequent experiments. The influence of temperature on the expression of the CysQ-UGT76G1 fusion was investigated, and a lower temperature enhanced the solubility (Fig. 4). At 25 °C, the percentage of soluble CysQ-UGT reached almost 21%, compared with less than 15% for the partner-free target protein. The lower temperature presumably allowed more time for correct protein folding to take place.

Glucosyltransferase activity of UGT76G1 and CysQ-UGT

It is important to check the biological activity of fusion proteins because fusion of a solubility-enhancing partner can cause unexpected structural and functional changes in target protein. We first investigated the activity of crude enzymes and found that rebaudioside A was generated in all experiments (Fig. 5), which confirmed that fusion to CysQ did not disrupt the activity of UGT76G1. Furthermore, the specific activity of UGT76G1 fused with CysQ was 82% higher than the partner-free control lysate (Table 4). Fusion of CysQ, therefore, improved the soluble

Fig. 3 SDS-PAGE analysis of soluble and insoluble fractions obtained from recombinant *E. coli* cell lysates. M, standard molecular markers; S and I (UGT76G1, CysQ-UGT, EDA-UGT and NusA-UGT), soluble and insoluble fractions from BL21(DE3) cells

transformed with pET28a-UGT, pET28a-CysQ-UGT, pET28a-EDA-UGT and pET28a-NusA-UGT. Arrows in gel images indicate recombinant protein bands

Fig. 4 SDS-PAGE analysis of soluble and insoluble fractions obtained from recombinant *E. coli* cell lysates at different temperatures. M, the standard molecular marker; S and I (20, 25 and 30 °C), soluble and insoluble fractions from BL21(DE3) cells transformed with pET28a-CysQ-UGT at 20, 25 and 30 °C. The arrow indicates recombinant protein bands

expression of UGT76G1 and enhanced the specific activity. We speculate that fusion with protein partners may improve the structural stability and/or dynamics, resulting in enhanced enzyme activity.

Recombinant fusion proteins included an N-terminal hexahistidine (H_6) tag and a D_4K EK cleavage sequence, which facilitated purification of target proteins using standard immobilised metal affinity chromatography. However, we found that EK could not cleave the fusion proteins, and we are unable to explain exactly why this was

Table 4 Enzyme activity of crude lysates

Enzyme	Specific activity (mU/mg)
UGT76G1	11.28 ± 1.68
CysQ-UGT	20.53 ± 1.02

All data were averaged from triplicate experiments

Fig. 6 SDS-PAGE analysis of the purified CysQ-UGT fusion protein. M, the standard molecular marker; S and I, soluble and insoluble fractions from BL21(DE3) cells transformed with pET28a-CysQ-UGT. P1 is the flow-through following loading of soluble fractions from BL21(DE3) cells transformed with pET28a-CysQ-UGT. P2-4 are soluble fractions from BL21(DE3) cells transformed with pET28a-CysQ-UGT eluted with Wash Buffer I, Wash Buffer II and Elution Buffer containing increasing amounts of imidazole. The arrow indicates recombinant protein bands

the case. The D_4K sequence was derived from the literature, but this may not be the optimum recognition sequence in all cases. One possible explanation for the failure of the

Fig. 5 HPLC analysis of the enzymatic synthesis of rebaudioside A from stevioside using crude enzymes. **a** 1 g/l stevioside standard. **b** 0.8 g/l rebaudioside A standard. **c** Sample taken from reaction mixtures immediately; (d1–3) samples taken from reaction mixtures after a 0.5 h incubation with crude UGT76G1, CysQ-UGT and EDA-UGT

Table 5 Enzyme activity of purified enzymes

Enzyme	Specific activity (mU/ μ mol)
UGT76G1	0.04 $\pm$ 0.01
CysQ-UGT	0.49 $\pm$ 0.06

All data were averaged from triplicate experiments

EK cleavage could be the complex three-dimensional structure of the CysQ-UGT fusion protein, which could bury the cleavage site in the structure and prevent EK from accessing the recognition sequence.

SDS-PAGE confirmed binding of the CysQ-UGT fusion protein to the resin, as there was minimal protein with a molecular weight of $\sim$ 80 kDa in the flow-through (lane P1, Fig. 6). With Wash Buffer I containing 20 mM imidazole, contaminating *E. coli* proteins were eluted (P2, Fig. 6), and CysQ-UGT was eluted with Wash Buffer II containing 50 mM imidazole (P3, Fig. 6). The concentration of target protein was low due to the large volume of Wash Buffer used. CysQ-UGT was also eluted with Elution Buffer containing 250 mM imidazole (P4, Fig. 6). In this fraction, CysQ-UGT was at a higher concentration than in P3.

Glucosyltransferase activities of the purified enzymes are shown in Table 5. Interestingly, purified UGT76G1 lost almost all catalytic activity, whereas CysQ-UGT only lost some activity. We suspect that imidazole could affect the stability of the enzyme. However, we found that CysQ improved the specific activity of UGT76G1, and CysQ significantly increased the stability or tolerance to imidazole. Therefore, the relatively small CysQ fusion expression partner was successfully used to improve the solubility and activity of recombinant glucosyltransferase UGT76G1 in *E. coli*, thereby facilitating the enzymatic synthesis of rebaudioside A. This strategy could presumably be used to improve the expression of other plant glucosyltransferases in prokaryotic expression systems.

Acknowledgements We greatly acknowledge financial support from the NSFC (21106068), the Open fund Program of the Yichang Key Laboratory of Biocatalysis (2015NP01), Subei Science and Technology Projects (BN2015115), TAPP and Provincial Key R&D Plan of Jiangsu (BE2017703).

Author contributions LC and YL conceived and designed the study. LC and PS performed the experiments and analysed the data. LC wrote the paper. MY, LX, KC and PO reviewed and edited the manuscript. All authors read and approved the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

- Adari BR, Alavala S, George SA, Meshram HM, Tiwari AK, Sarma AVS (2016) Synthesis of rebaudioside-A by enzymatic transglycosylation of stevioside present in the leaves of *Stevia rebaudiana* Bertoni. Food Chem 200:154–158. doi:10.1016/j.foodchem.2016.01.033
- Ahn KY, Song JA, Hah KY, Park JS, Seo HS, Lee J (2007) Heterologous protein expression using a novel stress-responsive protein of *E. coli* RpoA as fusion expression partner. Enzyme Microb Technol 41:859–866. doi:10.1016/j.enzymictec.2007.07.009
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254. doi:10.1006/abio.1976.9999
- Brandle JE, Telmer PG (2007) Steviol glycoside biosynthesis. Phytochemistry 68:1855–1863. doi:10.1016/j.phytochem.2007.02.010
- Chen TH, Zhang Y, Shi ZQ, Liu XH, He BL (1999) Studies on the adsorptive selectivity of the polar resin with carbonyl group on rebaudioside A. Acta Polym Sin 50:398–403
- De Marco V, Stier G, Blandin S, de Marco A (2004) The solubility and stability of recombinant proteins are increased by their fusion to NusA. Biochem Biophys Res Commun 322:766–771. doi:10.1016/j.bbrc.2004.07.189
- Desmet T, Soetaert W, Bojarova P, Kren V, Dijkhuizen L, Eastwick-Field V, Schiller A (2012) Enzymatic glycosylation of small molecules: challenging substrates require tailored catalysts. Chemistry 18:10786–10801. doi:10.1002/chem.201103069
- Dewitte G, Walmagh M, Diricks M, Lepak A, Gutmann A, Nidetzky B, Desmet T (2016) Screening of recombinant glucosyltransferases reveals the broad acceptor specificity of *stevia* UGT-76G1. J Biotechnol 233:49–55. doi:10.1016/j.jbiotec.2016.06.034
- Georgiou G, Valax P (1996) Expression of correctly folded proteins in *Escherichia coli*. Curr Opin Biotechnol 7:190–197. doi:10.1016/S0958-1669(96)80012-7
- Georgiou G, Valax P (1999) Isolating inclusion bodies from bacteria. Methods Enzymol 309:48–58. doi:10.1016/S0076-6879(99)09005-9
- Han KY, Park JS, Seo HS, Ahn KY, Lee J (2008) Multiple stressor-induced proteome responses of *Escherichia coli* BL21(DE3). J Proteome Res 7:1891–1903. doi:10.1021/pr700631c
- Hellman J, Paavilainen S, Mantsala P (1992) Expression in *Escherichia coli* and purification of intracellular proteins by fusion to cyclomaltodextrin glucanotransferase. J Biotechnol 26:275–288. doi:10.1016/0168-1656(92)90012-X
- Jhamb K, Sahoo DK (2012) Production of soluble recombinant proteins in *Escherichia coli*: effects of process conditions and chaperone co-expression on cell growth and production of xylanase. Bioresour Technol 123:135–143. doi:10.1016/j.biortech.2012.07.011
- Jhamb K, Jawed A, Sahoo DK (2008) Immobilized chaperones: a productive alternative to refolding of bacterial inclusion body proteins. Process Biochem 43:587–597. doi:10.1016/j.procbio.2008.02.004
- Kane JF (1995) Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. Curr Opin Biotechnol 6:494–500. doi:10.1016/0958-1669(95)80082-4
- Kang YS, Song JA, Han KY, Lee J (2015) *Escherichia coli* EDA is a novel fusion expression partner to improve solubility of aggregation-prone heterologous proteins. J Biotechnol 194:39–47. doi:10.1016/j.jbiotec.2014.11.025

- Kim S, Lee SB (2008) Soluble expression of archaeal proteins in *Escherichia coli* by using fusion-partners. *Protein Expr Purif* 62:116–119. doi:[10.1016/j.pep.2008.06.015](https://doi.org/10.1016/j.pep.2008.06.015)
- Kolb N, Herrera JL, Ferreyra DJ, Uliana RF (2001) Analysis of sweet diterpene glycosides from *Stevia rebaudiana*: improved HPLC method. *J Agric Food Chem* 49:4538–4541. doi:[10.1021/jf010475p](https://doi.org/10.1021/jf010475p)
- Lee JH, Lee JY, Song JA, Han KY, Lee DS, Lee J (2014) A stress-responsive *Escherichia coli* protein, CysQ is a highly effective solubility enhancer for aggregation-prone heterologous proteins. *Protein Expr Purif* 101:91–98. doi:[10.1016/j.pep.2014.06.006](https://doi.org/10.1016/j.pep.2014.06.006)
- Li J, Chen ZB, Di DL (2012) Preparative separation and purification of rebaudioside A from *Stevia rebaudiana* Bertoni crude extracts by mixed bed of macroporous adsorption resins. *Food Chem* 132:268–276. doi:[10.1016/j.foodchem.2011.10.077](https://doi.org/10.1016/j.foodchem.2011.10.077)
- Li H, Liu N, Wang WT, Wang JY, Gao WY (2016) Cloning and characterization of GST fusion tag stabilized large subunit of *Escherichia coli* acetohydroxyacid synthase I. *J Biosci Bioeng* 121:21–26. doi:[10.1016/j.jpibosc.2015.05.010](https://doi.org/10.1016/j.jpibosc.2015.05.010)
- Lui YF, Di DL, Bai QQ, Li JT, Chen ZB, Lou S, Ye HL (2011) Preparative separation and purification of rebaudioside A from steviol glycosides using mixed-mode macroporous adsorption resins. *J Agric Food Chem* 59:9629–9636. doi:[10.1021/jf2020232](https://doi.org/10.1021/jf2020232)
- Madhav H, Bhasker S, Chinnamma M (2013) Functional and structural variation of uridine diphosphate glycosyltransferase (UGT) gene of *Stevia rebaudiana*-UGTSr involved in the synthesis of rebaudioside A. *Plant Physiol Biochem* 63:245–253. doi:[10.1016/j.plaphy.2012.11.029](https://doi.org/10.1016/j.plaphy.2012.11.029)
- Neuwald AF, Krishnan BR, Brikun I, Kulakauskas S, Suziedelis K, Tomcsanyi T, Leyh TS, Berg DE (1992) CysQ, a gene needed for cysteine synthesis in *Escherichia coli* K-12 only during aerobic growth. *J Bacteriol* 174:415–425. doi:[10.1128/jb.174.2.415-425.1992](https://doi.org/10.1128/jb.174.2.415-425.1992)
- Niiranen L, Espelid S, Karlsen CR, Mustonen M, Paulsen SM, Heikinheimo P, Willassen NP (2007) Comparative expression study to increase the solubility of cold adapted Vibrio proteins in *Escherichia coli*. *Protein Expr Purif* 52:210–218. doi:[10.1016/j.pep.2006.09.005](https://doi.org/10.1016/j.pep.2006.09.005)
- Paquette S, Moller BL, Bak S (2003) On the origin of family 1 plant glycosyltransferases. *Phytochemistry* 62:399–413. doi:[10.1016/S0031-9422\(02\)00558-7](https://doi.org/10.1016/S0031-9422(02)00558-7)
- Pasek M, Boeggeman E, Ramakrishnan B, Qasba PK (2010) Galectin-1 as a fusion partner for the production of soluble and folded human beta-1,4-galactosyltransferase-T7 in *E. coli*. *Biochem Biophys Res Commun* 394:679–684. doi:[10.1016/j.bbrc.2010.03.051](https://doi.org/10.1016/j.bbrc.2010.03.051)
- Pei XL et al (2015) Chaperones-assisted soluble expression and maturation of recombinant Co-type nitrile hydratase in *Escherichia coli* to avoid the need for a low induction temperature. *J Biotechnol* 203:9–16. doi:[10.1016/j.jbiotec.2015.03.004](https://doi.org/10.1016/j.jbiotec.2015.03.004)
- Raran-Kurussi S, Keefe K, Waugh DS (2015) Positional effects of fusion partners on the yield and solubility of MBP fusion proteins. *Protein Expr Purif* 110:159–164. doi:[10.1016/j.pep.2015.03.004](https://doi.org/10.1016/j.pep.2015.03.004)
- Rumelhard M, Hosako H, Eurlings IMJ, Westerink WMA, Staska LM, van de Wiel JAG, La Marta J (2016) Safety evaluation of rebaudioside A produced by fermentation. *Food Chem Toxicol* 89:73–84. doi:[10.1016/j.fct.2016.01.005](https://doi.org/10.1016/j.fct.2016.01.005)
- Schlieker C, Bukau B, Mogk A (2002) Prevention and reversion of protein aggregation by molecular chaperones in the *E. coli* cytosol: implications for their applicability in biotechnology. *J Biotechnol* 96:13–21. doi:[10.1016/S0168-1656\(02\)00033-0](https://doi.org/10.1016/S0168-1656(02)00033-0)
- Tada A et al (2013) Absolute quantitation of stevioside and rebaudioside A in commercial standards by quantitative NMR. *Chem Pharm Bull* 61:33–38. doi:[10.1248/cpb.c12-00736](https://doi.org/10.1248/cpb.c12-00736)
- Tong YJ, Feng SS, Xin Y, Yang HL, Zhang L, Wang W, Chen W (2016) Enhancement of soluble expression of codon-optimized *Thermomicrobium roseum* sarcosine oxidase in *Escherichia coli* via chaperone co-expression. *J Biotechnol* 218:75–84. doi:[10.1016/j.jbiotec.2015.11.018](https://doi.org/10.1016/j.jbiotec.2015.11.018)
- Urban JD, Carakostas MC, Taylor SL (2015) Steviol glycoside safety: are highly purified steviol glycoside sweeteners food allergens? *Food Chem Toxicol* 75:71–78. doi:[10.1016/j.fct.2014.11.011](https://doi.org/10.1016/j.fct.2014.11.011)
- Wang Y, Chen L, Li Y, Yan M, Chen K, Hao N, Xu L (2016) Efficient enzymatic production of rebaudioside A from stevioside. *Biosci Biotechnol Biochem* 80:67–73. doi:[10.1080/09168451.2015.1072457](https://doi.org/10.1080/09168451.2015.1072457)
- Wolwer-Rieck U (2012) The leaves of *Stevia rebaudiana* (Bertoni), their constituents and the analyses thereof: a review. *J Agric Food Chem* 60:886–895. doi:[10.1021/jf2044907](https://doi.org/10.1021/jf2044907)
- Xu J, Kaloyanova D, Ivanov IG, AbouHaidar MG (1998) The low expression level of pokeweed antiviral protein (PAP) gene in *Escherichia coli* by the inducible *lac* promoter is due to inefficient transcription and translation and not to the toxicity of the PAP. *Arch Biochem Biophys* 351:82–88. doi:[10.1006/abbi.1997.0552](https://doi.org/10.1006/abbi.1997.0552)
- Zerbs S, Giuliani S, Collart F (2014) Small-scale expression of proteins in *E. coli*. *Method Enzymol* 536:117–131. doi:[10.1016/B978-0-12-420070-8.00011-8](https://doi.org/10.1016/B978-0-12-420070-8.00011-8)