Expression of Bacterial GshF in *Pichia pastoris* for Glutathione Production

Shaolin Ge,a,b Taicheng Zhu,a and Yin Lia

Institute of Microbiology, Chinese Academy of Sciences, Beijing, People’s Republic of China,a and School of Life Sciences, University of Science and Technology of China, Hefei, People’s Republic of Chinaa

Conventionally, two consecutive enzymatic reactions catalyzed by γ-glutamylcysteine synthetase and glutathione synthetase are most commonly used for glutathione production. Here we demonstrate that bacterial bifunctional GshF can be used for glutathione production in a eukaryotic system without accumulation of the intermediate γ-glutamylcysteine.

Glutathione (γ-glutamylcysteinylglycine; GSH) is the predominant nonprotein thiol present in living organisms (16). It plays an important role in many physiological functions and has a number of diverse applications in the food, pharmaceutical, and cosmetic industries (12, 18). GSH is industrially produced by fermentation or enzymatic synthesis. By either process, GSH is synthesized through two consecutive enzymatic reactions catalyzed by γ-glutamylcysteine synthetase (γ-GCS, encoded by *gshA* in *Escherichia coli*, or GSHI in *Saccharomyces cerevisiae*) and GSH synthetase (GS, encoded by *gshB* in *E. coli* or GSHII in *S. cerevisiae*), which are widely present in eukaryotes and Gram-negative bacteria. These two enzymes, especially those derived from *E. coli* and *S. cerevisiae*, have been overexpressed in various model microorganisms to generate genetically engineered strains with higher GSH production capabilities. Previous studies have reported the overexpression of the *E. coli* *gshA* and *gshB* genes in *Lactococcus lactis* (13), *Clostridium acetobutylicum* (23), *S. cerevisiae* (17), and *E. coli* (8, 14). Overexpression of the *S. cerevisiae* GSHI and GSHII genes in *S. cerevisiae* (5, 20) and *Pichia pastoris* (6) has also been reported.

While researchers have focused mainly on improving GSH productivity, one fact that is often overlooked is accumulation of the intermediate metabolite, i.e., the dipeptide γ-glutamylcysteine (γ-GC), during these processes. For instance, when the γ-GC and γ-GS modules in GshF is expected to increase the effective concentration of γ-GC, thereby increasing the catalytic rate of the second enzymatic reaction. GshF also has other potential advantages over the conventional γ-GCS/GS systems when used for GSH production. First, the feedback inhibition of γ-GCS activity by GSH is considered a physiological mechanism to prevent the overaccumulation of GSH (1, 2). GshF from *S. agalactiae* and *Streptococcus thermophilus* was shown to be insensitive to an increased concentration of GSH (9, 11). Therefore, the use of GshF might be able to maintain the GSH biosynthesis rate even in the presence of a high concentration of GSH. Second, the manipulation and engineering of a single gene is supposed to be easier than the manipulation of two genes. To date, the value of GshF in the biotechnological production of GSH has not been fully exploited. Only recently, Li et al. (11) have successfully expressed GshF from *S. thermophilus* in *E. coli*. A recombinant *E. coli* strain with GshF activity was used for the enzymatic production of GSH using three precursor amino acids. However, due to the low ATP regeneration efficiency of *E. coli*, this process was coupled with an extraneous ATP regeneration system of *S. cerevisiae*, making further process optimization complex.

The objective of this study was to introduce bacterial GshF into eukaryotic cells so that the strong ATP-generating ability of the eukaryotic system can be used for one-step GSH biosynthesis. The methylotrophic yeast *P. pastoris* was chosen as the host because it had several attractive features beneficial for high-level GSH production, such as a low nutrient requirement and ease of growth to a high cell density (2a, 3, 4). The GshF proteins, from *L. monocytogenes*, *S. agalactiae*, and *Lactobacillus plantarum*, were each cloned and expressed in *P. pastoris* GS115. The GSH production capabilities and other characteristics of the resulting recombinant strains were then determined. The data demonstrated that bacterial bifunctional GshF can be used for GSH production in a eukaryotic system.

**Cloning of gshf genes derived from bacteria.** Lm-gshF, Sa-gshF, and Lp-gshF, derived from *L. monocytogenes*, *S. agalactiae*, and *L. plantarum*, were cloned and expressed in *P. pastoris* GS115. The GSH production capabilities and other characteristics of the resulting recombinant strains were then determined. The data demonstrated that bacterial bifunctional GshF can be used for GSH production in a eukaryotic system.
and L. plantarum, respectively, were amplified by PCR from the respective genomic DNA. The sequences of the primers used and other PCR-related information are summarized in Table 1. The Lm-gshF gene was obtained from L. monocytogenes strain ATCC 19114-3 genomic DNA using LmgshF-F and LmgshF-R, creating ClaI and NotI sites (underlined in Table 1), respectively. Similarly, Sa-gshF and Lp-gshF were amplified using SagshF-F/SagshF-R and LpgshF-F/LpgshF-R, respectively. Three fragments of the expected size (around 2.3 kb) were thus obtained (Fig. 1A). The nucleotide sequence of Lm-gshF showed 99% identity with those of L. monocytogenes M7, L99, and HCC23. The nucleotide sequence of Sa-gshF had 100% identity with those of S. agalactiae NEM316, S. agalactiae 2603V/R and A909. The nucleotide sequence of Lp-gshF had 100% identity with that of L. plantarum WCFS1.

TABLE 1 PCR-related information for cloning of gshF genes

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Source</th>
<th>Primer pair used for amplification</th>
<th>Gene size (bp)</th>
<th>Closest homologue(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lm-gshF</td>
<td>L. monocytogenes</td>
<td>LmgshF-F, 5'-AGATCGATACCATGATAAACTTGATATGAC-3'; LmgshF-R, 5'-CAGCGGCGCTTTAGGAAATGACTCTCTAATTTTAATTTGAAT-3'</td>
<td>2,346</td>
<td>L. monocytogenes M7, L99, HCC23 (99)</td>
</tr>
<tr>
<td>Sa-gshF</td>
<td>S. agalactiae</td>
<td>SagshF-F, 5'-AGATCGATACCATGATTATAGATCGACTGTTACTAAGGAAAGCCACTTCA TCTATGTT-3'; SagshF-R, 5'-CAGCGGCGCTTTAGGAAATGACTCTCTAATTTTAATTTGAAT-3'</td>
<td>2,276</td>
<td>S. agalactiae NEM316, 2603V/R, A909 (100)</td>
</tr>
<tr>
<td>Lp-gshF</td>
<td>L. plantarum</td>
<td>LpgshF-F, 5'-AGATCGATACCATGGAATTAGATCGACTGTTACTAAGGAAATGACTCTCTAATTTTAATTTGAAT-3'</td>
<td>2,256</td>
<td>L. plantarum WCFS1 (100)</td>
</tr>
</tbody>
</table>

*a* ClaI and NotI sites are underlined.

and L. plantarum, respectively, were amplified by PCR from the respective genomic DNA. The sequences of the primers used and other PCR-related information are summarized in Table 1. The Lm-gshF gene was obtained from L. monocytogenes strain ATCC 19114-3 genomic DNA using LmgshF-F and LmgshF-R, creating ClaI and NotI sites (underlined in Table 1), respectively. Similarly, Sa-gshF and Lp-gshF were amplified using SagshF-F/SagshF-R and LpgshF-F/LpgshF-R, respectively. Three fragments of the expected size (around 2.3 kb) were thus obtained (Fig. 1A). The nucleotide sequence of Lm-gshF showed 99% identity with those of L. monocytogenes M7, L99, and HCC23. The nucleotide sequence of Sa-gshF had 100% identity with those of S. agalactiae NEM316, S. agalactiae 2603V/R and A909. The nucleotide sequence of Lp-gshF had 100% identity with that of L. plantarum WCFS1.

TABLE 2 Strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant characteristics</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli DH5α</td>
<td>his4&lt;sup&gt;+&lt;/sup&gt; Mut&lt;sup&gt;4&lt;/sup&gt;</td>
<td>TaKaRa</td>
</tr>
<tr>
<td>P. pastoris GS115</td>
<td></td>
<td>Invitrogen</td>
</tr>
<tr>
<td>S. cerevisiae S288c</td>
<td></td>
<td>Our lab</td>
</tr>
<tr>
<td>L. monocytogenes</td>
<td></td>
<td>Our lab</td>
</tr>
<tr>
<td>S. agalactiae</td>
<td></td>
<td>Our lab</td>
</tr>
<tr>
<td>L. plantarum</td>
<td></td>
<td>Our lab</td>
</tr>
<tr>
<td>GS115/LmgshF</td>
<td>GS115 integrated with recombinant plasmid pGAPZH-LmgshF</td>
<td>This study</td>
</tr>
<tr>
<td>GS115/SagshF</td>
<td>GS115 integrated with recombinant plasmid pGAPZH-SagshF</td>
<td>This study</td>
</tr>
<tr>
<td>GS115/LpgshF</td>
<td>GS115 integrated with recombinant plasmid pGAPZH-LpgshF</td>
<td>This study</td>
</tr>
<tr>
<td>GS115/Sc288c</td>
<td>GS115 integrated with recombinant plasmid pGAPZH-ScGSH</td>
<td>This study</td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGAPZαA</td>
<td>Vector for constitutive secreted protein expression; Zeo&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pPIC9k</td>
<td>Vector for inducible secreted protein expression; HIS4; Kan&lt;sup&gt;+&lt;/sup&gt;; Amp&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pGAPZαAH</td>
<td>pGAPZαA-based expression vector carrying HIS4 fragment from pPIC9k; Zeo&lt;sup&gt;+&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pGAPZH-Scgsh1</td>
<td>pGAPZαA containing gsh1 from S. cerevisiae S288c; Zeo&lt;sup&gt;+&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pGAPZH-Scgsh2</td>
<td>pGAPZαA containing gsh2 from S. cerevisiae S288c; Zeo&lt;sup&gt;+&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pGAPZH-ScGSH</td>
<td>pGAPZαA containing gsh1 and gsh2 from S. cerevisiae S288c; Zeo&lt;sup&gt;+&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pGAPZH-LmgshF</td>
<td>pGAPZαA containing gshF from L. monocytogenes; Zeo&lt;sup&gt;+&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pGAPZH-LpgshF</td>
<td>pGAPZαA containing gshF from L. plantarum; Zeo&lt;sup&gt;+&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pGAPZH-SagshF</td>
<td>pGAPZαA containing gshF from S. agalactiae; Zeo&lt;sup&gt;+&lt;/sup&gt;</td>
<td>This study</td>
</tr>
</tbody>
</table>

FIG 1 Cloning and expression of gshF genes from three different Gram-positive bacteria in P. pastoris GS115. (A) PCR amplification of the three gshF genes. Lanes M, molecular size markers; lane 1, L. monocytogenes gshF; lane 2, S. agalactiae gshF; lane 3, L. plantarum gshF. (B) Schematic representation of the construction of a recombinant expression vector containing the gshF genes derived from Gram-positive bacteria.
Construction of recombinant strains. The nucleotide sequence of \textit{HIS4} was amplified from plasmid pPIC9k by PCR using oligonucleotide primers his4-F (5’-CTAGATCTATGACATTTCCTTGCTACCTGCA-3’ [BglII site underlined]) and his4-R (5’-CTGGATCCTTAAATAAGTCCCAGTTTCTC-3’ [BamHI site underlined]). All of the strains and plasmids used and generated in this study are summarized in Table 2. The 2,921-bp product obtained was inserted into the BamHI site of the vector pGAPZ\textsubscript{H}-9251\textsubscript{A}, yielding plasmid pGAPZ\textsubscript{H}-9251\textsubscript{AH}. The ClaI/NotI PCR fragments of \textit{Lm-gshF}, \textit{Sa-gshF}, and \textit{Lp-gshF} were cloned into the pGAPZ\textsubscript{H}-9251\textsubscript{AH} expression vector to form pGAPZH-Lmgsh, pGAPZH-Sagsh, and pGAPZH-Lpgsh, respectively. A schematic representation of the recombinant expression vectors is shown in Fig. 1B. Three recombinant \textit{P. pastoris} strains, GS115/LmgshF, GS115/SagshF, and GS115/LpgshF, were then generated by transforming linearized pGAPZH-Lmgsh, pGAPZH-Sagsh, and pGAPZH-Lpgsh, respectively, into \textit{P. pastoris} GS115. The results of \textit{gshF} gene expression are shown in Fig. 2A. The predicted molecular masses of the \textit{L. monocytogenes}, \textit{S. agalactiae}, and \textit{L. plantarum} GshF proteins were 88.8, 86.3, and 83.0 kDa, respectively. A strongly overexpressed protein band (indicated by the left arrow) was observed at the expected position on SDS-PAGE of \textit{L. monocytogenes} cell lysate, and a weak band (indicated by the right arrow) was obtained with \textit{L. plantarum} cell lysate. The bands indicated by arrows in lanes 2 and 3 of Fig. 2A were excised and subjected to matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) analysis (15). Peptide mass fingerprinting results confirmed that the proteins in the two bands (from left to right) were GshF proteins from \textit{L. monocytogenes} HCC23 and \textit{L. plantarum} WCFS1, respectively (Fig. 2B and C). For \textit{S. agalactiae} GshF, however, neither discernible protein band could be seen on SDS-PAGE and Coomassie blue staining, nor could its peptides be detected by MS analysis (data not shown). Therefore, although the three \textit{gshF} genes were expressed by exactly the same strategy in \textit{P. pastoris}. 

![Figure 2](http://aem.asm.org/)
p. pastoris, their expression levels varied from around 7% of the total soluble protein (quantified by the Gel-Pro Analyzer program [Media Cybernetics]) for GshF from L. monocytogenes to not detectable for that from S. agalactiae. We hypothesized that the codon usage biases of the three gshF genes may account for their different expression levels in P. pastoris.

**Synthesis of GSH using recombinant P. pastoris expressing GshF.** To evaluate the GSH synthesis capabilities of GS115/LmgshF, GS115/SagshF, and GS115/LpgshF, three positive clones (confirmed by colony PCR) of each strain, together with control strain GS115, were cultured in YPD liquid medium (2% glucose, 1% yeast extract, 2% peptone) for 2 days, and the intracellular GSH content was extracted from the cells with 40% ethanol (22) and determined using the monobromobimane fluorescent labeling and high-performance liquid chromatography methods (23). Figure 3 shows that strain GS115/LmgshF had significantly improved GSH synthesis capability, compared with that of strain GS115. The intracellular GSH level was increased by almost 80%. No significant increase in intracellular GSH content was observed in strains GS115/SagshF and GS115/LpgshF. These results were not unexpected, because the GshF expression level of GS115/LmgshF was significantly higher than that of the other two strains.

**Comparison of GSH synthesis capabilities of recombinant P. pastoris strains expressing L. monocytogenes gshF and S. cerevisiae GSHI/GSHII.** In order to further evaluate the potential of bacterial GshF for GSH biosynthesis and the potential of reducing γ-GC accumulation, we compared the thiol-producing profile of strain GS115/LmgshF with that of another P. pastoris recombinant strain expressing the *S. cerevisiae* GSHI and GSHII genes (Fig. 4). The strain expressing the GSHI and GSHII genes of *S. cerevisiae* Sc288c was generated by exactly following the strategy of Fei et al. (6) and designated GS115/Sc288c in this work. Strains GS115, GS115/LmgshF, and GS115/Sc288c were cultured in YPD medium for 42 h, and their intracellular thiol concentration profiles were determined. No difference in the growth rates of these three strains was observed (data not shown). Figure 2A shows that the intracellular γ-GC content of GS115/Sc288c accounted for 11.2% of the GSH synthesized (molar ratio). This increased to 20.1% when three amino acid precursors (glycine, cysteine, and glutamic acid) were added (Fig. 2B). In contrast, as we hypothesized, no intracellular γ-GC was detected in GS115/LmgshF in the absence or presence of precursor amino acids. Although the catalytic efficiency of GshF seems slightly lower than that of GSHI/GSHII, as the GSH synthesized by GS115/LmgshF is around 81 to 83% of that synthesized by GS115/Sc288c, this problem could be solved by increasing the GshF protein expression level in *P. pastoris* or improving the GshF turnover activity by directed evolution.

**Conclusion.** In order to test the possibility of using bacterial bifunctional GshF for GSH production in a eukaryotic system, three GshF proteins from the Gram-positive species *L. monocytogenes, S. agalactiae*, and *L. plantarum* were each cloned and expressed in the methylotrophic yeast *P. pastoris* GS115. Only *L. monocytogenes* GshF displayed significant protein expression and catalytic activity in *P. pastoris*. Compared with that of *S. cerevisiae* GSHI and GSHII, the expression of GshF in *P. pastoris* could significantly reduce the accumulation of the intermediate metabolite γ-GC.

**ACKNOWLEDGMENTS**

This research was supported by the National Natural Science Foundation of China (grant 30560006) and the Knowledge Innovation Program of the Chinese Academy of Sciences (grant KSCX2-EW-G-15-03).

We thank Shutao Sun (Institute of Microbiology, Chinese Academy of Sciences) for her assistance with MALDI-TOF MS analysis.

**REFERENCES**

pastoris: a useful experimental tool in protein engineering and production.
sion of the gamma-glutamylcysteine synthetase gene in Saccharomyces
genus catalyzes the two primary activities for glutathione biosynthesis. J.
glutathione-producing strains of Escherichia coli by recombinant DNA
Streptococcus agalactiae. One protein accounts for gamma-glutamylcysteine
10. Reference deleted.
bifunctional enzyme encoded by gshF from Streptococcus thermophilus
Lactococcus lactis for glutathione overproduction. Appl. Microbiol. Bio-
technol. 67:83–90.
43:211–214.
analysis between a wild type Clostridium acetobutylicum DSM 1731 and its
mutant with enhanced butanol tolerance and butanol yield. J. Proteome
Res. 9:3046–3061.
711–760.
394.
biosynthesis in bacteria by bifunctional GshF is driven by a modular
structure featuring a novel hybrid ATP-grasp fold. J. Mol. Biol. 416:
486–494.
recombinant plasmid participating in glutathione synthesis and produc-
tion of glutathione using yeast containing said recombinant plasmid. Jap-
anean patent JP62275685.
bifunctional gamma-glutamate-cysteine ligase/glutathione synthetase
22. Xiong ZQ, et al. 2009. Efficient extraction of intracellular reduced gluta-
thione from fermentation broth of Saccharomyces cerevisiae by ethanol.
Bioreour Technol. 100:1011–1014.
Clostridium acetobutylicum by introducing glutathione biosynthetic capa-