High level expression of *Thermococcus litoralis* 4-α-glucanotransferase in a soluble form in *Escherichia coli* with a novel expression system involving minor arginine tRNAs and GroELS

Hiromi Imamura, Beong-Sam Jeon, Takayoshi Wakagi*, Hiroshi Matsuzawa[1]

Department of Biotechnology, The University of Tokyo, Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

Received 9 July 1999; received in revised form 3 August 1999

Abstract The *Thermococcus litoralis* 4-α-glucanotransferase (GTase) gene has a high content of AGA and AGG codons for arginine, which are extremely rare in *Escherichia coli*. Expression of the GTase gene in *E. coli* resulted in low protein production and the accumulation of inclusion bodies. However, simultaneous expression of GTase with $tRNA_{AGA}$, $tRNA_{AGG}$ and GroELS affected both the production and solubility of GTase, and production of soluble GTase increasing about 5-fold. This new *E. coli* expression system should be applicable to the expression of not only archaeal but also eukaryotic genes, which usually contain a large number of AGA and AGG codons.

Key words: Rare codon; Transfer RNA; GroELS; Recombinant protein; 4-α-Glucanotransferase; *Thermococcus litoralis*

1. Introduction

The expression of heterologous genes in *Escherichia coli* is one of the most frequently used techniques in the laboratory and industry. When a strong promoter is used to express recombinant proteins, a yield of more than 10% of the total *E. coli* proteins can often be obtained. But the problem sometimes occurs that recombinant proteins are expressed at seriously low levels. One of the factors causing low level expression of recombinant proteins in *E. coli* is the unfavorable codon usage of heterologous genes [1]. It is well known that the codon usage of *E. coli* is highly biased. In particular, arginine AGA and AGG codons are extremely rare, and tRNAs specific for these codons are poorly produced in *E. coli* [2,3]. On the other hand, AGA and AGG codons are major arginine codons in eukarya and archaea (Table 1). Thus, heterologous expression of eukaryotic and archaeal genes in *E. coli* could lead to disadvantages including low expression levels [1], frame shifts [4], truncation [5] or amino acid misincorporation [6,7] during protein translation. Codon optimization through gene synthesis or site-directed mutagenesis of a gene has sometimes been performed to overcome these problems. Alternatively, coexpression of tRNA specific for AGA ($tRNA_{AGA}$) with heterologous genes that contain AGA codons also suppress these negative effects due to AGA codons [1,4,7]. Even if a gene contains only few AGA codons and a sufficient expression level is achieved without $tRNA_{AGA}$, coexpression of $tRNA_{AGA}$ can further improve the expression level [8]. $tRNA_{AGA}$ is encoded by the *argU* gene (formerly called *dnaY*) in *E. coli* [9]. This tRNA had previously been thought to recognize both AGA and AGG codons through wobbling [1,9], but was found not to recognize the AGG codon because of modification of the first anticodon [4,10]. The tRNA for AGG ($tRNA_{AGG}$) is encoded by the *argW* gene in *E. coli* [3]. The arginine AGG codon is extremely rare in *E. coli* like the AGA codon. However, coexpression of $tRNA_{AGG}$ with a heterologous gene has not yet been reported.

It is also a serious problem that recombinant proteins expressed in *E. coli* often form insoluble inclusion bodies. Especially for X-ray crystallography, it is important to obtain large amounts of recombinant proteins in soluble forms because recombinant proteins refolded from inclusion bodies can possibly lose their intact structures. Several techniques have been developed to obtain recombinant proteins in soluble forms. coexpression of GroELS, molecular chaperones in *E. coli* involved in the correct folding of premature proteins in an ATP-dependent manner, is one of the effective techniques. Many recombinant proteins expressed in *E. coli* have been obtained in soluble forms through coexpression of GroELS (review in [11]).

4-α-Glucanotransferase (GTase) of hyperthermophilic archaeon *Thermococcus litoralis* belongs to the glycosyl hydrolase family 57 and produces cycloamylose, a novel cyclic α-1,4-glucan with a high degree (up to hundreds) of polymerization, from amylose [12]. Cycloamylose is more highly polymerized than cyclodextrins, and therefore shows improved features to form inclusion complexes, and to dissolve in water [13]. For these properties, cycloamylose is expected to be used in future industry. However, the molecular mechanism for forming such a large cyclic glucan remains unclear because structural information of the enzyme has never been reported. To elucidate the structure-function relationship of the enzyme, we have cloned the GTase gene and expressed the gene in *E. coli* [12]. But the expression level of GTase was low and over half of the expressed GTase was insoluble. The low level expression of GTase was probably due to the high contents of AGA and AGG codons in the GTase gene. In this paper, we report a powerful production system for GTase in *E. coli* involving simultaneous expression of both minor arginine tRNAs and GroELS. This system provided a sufficient amount of soluble GTase for X-ray crystallography. The ef-

*Corresponding author. Fax: (81) (3) 5841-5152.
E-mail: atwakag@hongo.ecc.u-tokyo.ac.jp

[1] Present address: Department of Bioscience and Biotechnology, Aomori University, Kohbata, Aomori 030-0943, Japan.

Abbreviations: GTase, 4-α-glucanotransferase; $tRNA_{AGA}$, transfer RNA specific for AGA codons; $tRNA_{AGG}$, transfer RNA specific for AGG codons; IPTG, isopropyl-β-D-thiogalactopyranoside

PII: S0014-5793(99)01081-9

fects of minor arginine tRNAs and GroELS on the expression level and solubility of GTase are discussed.

2. Materials and methods

2.1. Materials

Restriction endonucleases were purchased from Boehringer Mannheim (Mannheim, Germany). ExTaq DNA polymerase and the DNA ligation kit were purchased from Takara Shuzo (Kyoto, Japan). The ingredients for the *E. coli* medium were from Difco. Ampicillin was from Nacalai Tesque (Kyoto, Japan). Kanamycin was from Meiji Seika (Tokyo, Japan). Chloramphenicol was from Merck. Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from Wako Pure Chemical Industries (Osaka, Japan). Synthetic oligonucleotides were purchased from Hokkaido System Science (Sapporo, Japan).

2.2. Bacterial strains and plasmids

For DNA manipulation, *E. coli* strain MV1184 {*ara*, Δ(*lac-proAB*), *rspL, thi* (ϕ80*lacZ*ΔM15), Δ(*srl-recA*) 306::Tn*10*(*tet*^{r})/F′*traD*36, *proAB*$^+$, *lacI*^{q}, *lacZ*ΔM15} and JM109 {*recA*1, *endA*1, *gyrA*96, *thi*-1, *hsdR*17, *supE*44, *relA*1, Δ(*lac-proAB*)/F′ *traD*36, *proAB*$^+$, *lacI*^{q}, *lacZ*ΔM15} were used. For expression of the GTase gene, *E. coli* strain BL21(DE3) {*E. coli* B, F$^-$, *dcm*, *ompT*, *hsdS*($r_B^- m_B^-$) *gal* (DE3)} was used. pUBS520 [1] carrying the *argU* gene and pT-groE [14] carrying the *groESL* operon under the T7 promoter were kind gifts from Dr. P. Buckel of Boehringer Mannheim and Dr. S. Ishii of the Institute of Physical and Chemical Research (RIKEN), respectively. pJGT7 is a GTase expression vector, which has the GTase gene under the T7 promoter (Jeon et al., unpublished result).

2.3. Construction of plasmids

The *argW* gene was amplified from chromosomal DNA of MV1184 cells by polymerase chain reaction (PCR) using the following oligonucleotide primers: 5′-GGTTTAAGCAATCGAGCGG-3′ (sense primer) and 5′-AGCGAACCATGACGAACTG-3′ (antisense primer). Amplified DNA was ligated into the T7Blue T-vector (Novagen), followed by digestion with *Sal*I and *Sma*I. A small *Sal*I–*Sma*I fragment was ligated between the *Sal*I and *Sca*I sites of pUBS520 to obtain pArgUW. pArgUW was digested with *Avi*II and then self-ligated to yield pArgW2. A small *Sal*I–*Sma*I fragment of pT-groE was ligated between the *Sal*I and *Pvu*II sites of pUBS520, pArgW2 and pArgUW to obtain pAUGE, pAWGE and pGro-ArgUW, respectively. The characteristics of these plasmids are presented in Table 2.

2.4. Coexpression of $tRNA_{AGA}$, $tRNA_{AGG}$ and GroELS with GTase

E. coli BL21(DE3)/pJGT7 was used for expressing GTase alone. For coexpression of GTase with $tRNA_{AGA}$, $tRNA_{AGG}$ and/or GroELS, *E. coli* BL21(DE3)/pJGT7, which was cotransformed with the plasmids indicated in Table 2, was used. A 30 μl aliquot of an overnight culture was added to 3 ml fresh LB medium containing appropriate antibiotics and then the cells were grown for 3 h at 37°C. The T7 promoter was subsequently induced by the addition of 0.1 mM IPTG. The cells were allowed to grow for an additional 3 h at 37°C, collected by centrifugation and then suspended in 20 mM Tris-HCl buffer (pH 8.0). The cells were disrupted by sonication, and then the resultant lysate was ultracentrifuged at 100 000 × *g* for 60 min to separate the soluble and insoluble fractions.

2.5. SDS-PAGE

SDS-PAGE was performed by the method of Laemmli [15]. The sample proteins were denatured in the sample buffer at 100°C for 5 min prior to electrophoresis. Samples derived from 0.1 OD_{660} unit cells were applied to the gels. The gels were stained with Coomassie brilliant blue.

2.6. Assaying of GTase activity

To measure GTase activity 30 μl of a sample solution, which was equivalent to 0.015 OD_{660} unit cells, was mixed with 120 μl of 1% maltotriose dissolved in 20 mM sodium phosphate buffer (pH 6.0), followed by incubation at 90°C for 15 min. The glucose liberated was determined by the glucose oxidase/peroxidase method [16]. One unit of enzyme activity was defined as the amount of enzyme that released 1 μmol glucose from maltotriose per min at 90°C.

3. Results and discussion

The GTase gene of *T. litoralis* contains 19 AGA and 10 AGG codons (2.9% and 1.5% of total codons, respectively) out of 659 total codons. When only the GTase gene was expressed in *E. coli* BL21(DE3) cells using strong T7 RNA polymerase, the expression level of GTase was low and, moreover, approximately 60% of the expressed enzyme was insoluble (Fig. 1A, lane 2).

Then the GTase gene was expressed with $tRNA_{AGA}$ and/or $tRNA_{AGG}$ and/or GroELS in BL21(DE3) cells. The cell lysates were separated into soluble and insoluble fractions and then the expression level and solubility of GTase were determined by SDS-PAGE. On the stained SDS-polyacrylamide gel, an about 85-kDa band corresponding to GTase could be clearly distinguished from *E. coli* proteins for all of the cellular fractions (upper arrow in Fig. 1A), except for the negative control, in which the cells harbored the vector without the GTase gene insert (Fig. 1A, lane 1). A band of GroEL was also observed in lanes 6–9 in Fig. 1A (lower arrow). In addition, the GTase activities of the soluble fractions were measured (Fig. 1B).

The expression level of GTase greatly increased and about 40% of the expressed enzyme was soluble, when the GTase gene was coexpressed with $tRNA_{AGA}$, $tRNA_{AGG}$ and GroELS (Fig. 1A, lane 9). Also, the GTase activity of the

Table 1
AGA and AGG codon usage in *E. coli* and other organisms

	Per all codons[a] (%)		Per arginine codons[a] (%)	
	AGA	AGG	AGA	AGG
Eubacteria				
E. coli	0.25	0.15	4.5	2.7
Bacillus subtilis	1.06	0.39	25.7	9.4
Helicobacter pylori	0.94	0.85	27.2	24.6
Viruses				
HIV-1	3.88	1.62	58.4	24.4
Archaea				
Methanococcus jannaschii	2.76	0.99	71.9	25.8
Pyrococcus horikoshii	2.04	3.03	37.2	55.3
Eucarya				
Saccharomyces cerevisiae	2.13	0.93	48.0	20.9
Arabidopsis thaliana	1.82	1.11	34.7	21.2
Human	1.09	1.10	19.8	20.0

[a]Values were calculated using the codon usage database [19].

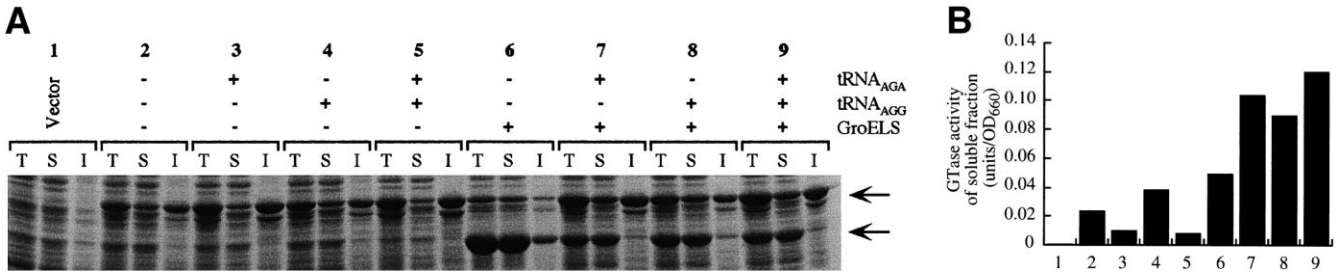

Fig. 1. Effects of coexpression of $tRNA_{AGA}$, $tRNA_{AGG}$ and GroELS on the production and solubility of GTase in *E. coli*. SDS-PAGE of the total proteins (A) and GTase activity of the soluble fraction of an *E. coli* BL21(DE3) cell lysate (B). Expression of the genes and preparation and fractionation of the cell lysates were performed as described in Section 2. A: (T) Total cell lysate, (S) soluble fraction and (I) insoluble fraction equivalent to 0.015 OD_{660} unit cells were subjected to electrophoresis. The upper arrow indicates the bands of GTase and lower arrow indicates the bands of GroELS. Lane 1: *E. coli* cells harboring a vector without the GTase gene insert; lane 2: cells harboring pJGT7 for the GTase gene only; lane 3: cells harboring pJGT7 and pUBS520 for the $tRNA_{AGA}$ gene; lane 4: cells harboring pJGT7 and pArgW2 for the $tRNA_{AGG}$ gene; lane 5: cells harboring pJGT7 and pArgUW for the $tRNA_{AGA}$ and $tRNA_{AGG}$ genes; lane 6: cells harboring pJGT7 and pT-groE for the *groESL* operon; lane 7: cells harboring pJGT7 and pAUGE for the $tRNA_{AGA}$ gene and the *groESL* operon; lane 8: cells harboring pJGT7 and pAWGE for the $tRNA_{AGG}$ gene and the *groESL* operon; lane 9: cells harboring pJGT7 and pGro-ArgUW for the $tRNA_{AGA}$ and $tRNA_{AGG}$ genes, and the *groESL* operon. B: GTase activity of the soluble fractions was measured as described in Section 2. The lanes are numbered as in A.

soluble fraction was approximately five-fold higher (Fig. 1B, lane 9) than that without coexpression (Fig. 1B, lane 2). When coexpressed with GroELS and either $tRNA_{AGA}$ or $tRNA_{AGG}$, the enzyme activity of the soluble fraction was also high (Fig. 1B, lanes 7 and 8), but a little lower than that with the complete system (Fig. 1B, lane 9).

When the GTase gene was coexpressed with $tRNA_{AGA}$ and $tRNA_{AGG}$, without GroELS, most of the expressed enzyme accumulated as insoluble inclusion bodies (Fig. 1A, lane 5), although the total production level of GTase was increased in comparison with that without the coexpression of tRNAs (Fig. 1A, lane 2). Consequently, the GTase activity of the soluble fraction became lower with the coexpression of $tRNA_{AGA}$ and $tRNA_{AGG}$ (Fig. 1B, lane 5). Coexpression of $tRNA_{AGA}$ alone led to similar results (Fig. 1A,B, lane 3) to those with the coexpression of both (Fig. 1A,B, lane 5). Coexpression of $tRNA_{AGG}$ alone affected the production level of total GTase (Fig. 1A, lane 4) less than that of $tRNA_{AGA}$ (Fig. 1A, lane 3), which may reflect the difference between the numbers of AGA and AGG codons in the GTase gene. However, the solubility of GTase (Fig. 1A, lane 4) and GTase activity (Fig. 1B, lane 4) with $tRNA_{AGG}$ were much higher than those with $tRNA_{AGA}$ (Fig. 1A,B, lane 3), for unknown reasons.

In contrast, coexpression of GroELS alone led to a decrease in the total production level of GTase (Fig. 1A, lane 6), while the net amount of soluble GTase increased because most of the expressed GTase was soluble (Fig. 1A, lane 6). Consequently, GTase activity increased about two-fold (Fig. 1B, lane 6) in comparison with that without any coexpression (Fig. 1B, lane 2). The decrease in the total production of GTase on coexpression of GroELS may be due to the limitation of aminoacyl tRNAs and/or ribosomes. A decrease in the production level of GroELS was also observed when GTase was coexpressed with $tRNA_{AGA}$ and/or $tRNA_{AGG}$ and this phenomenon could be accounted for in the same way.

Consequently, the GTase activity of the soluble fraction was highest when the GTase gene was expressed with both GroELS and the minor arginine tRNAs. In contrast, when only the minor arginine tRNAs were coexpressed, the GTase activity became lower, probably because the folding of the GTase polypeptide failed to catch up with the synthesis.

The most effective way to overproduce recombinant proteins in *E. coli* is to exchange rare codons of the genes for favorable codons by means of site-directed mutagenesis or gene synthesis. However, the alternative method of coexpression of minor tRNAs can easily increase the expression of recombinant proteins to sufficiently high levels, suppressing translation errors in a more economic way. So far, $tRNA_{AGG}$ has never been coexpressed with a heterologous gene. The present results indicate for the first time that coexpression of $tRNA_{AGG}$ is also effective for the expression of heterologous genes in *E. coli*, like the coexpression of $tRNA_{AGA}$.

The $tRNA_{AGA}$/$tRNA_{AGG}$/GroELS expression plasmid constructed in this study has a replication origin of p15A, and is compatible with most vectors for the expression or cloning of the genes. Heterologous genes can easily be coexpressed with $tRNA_{AGA}$, $tRNA_{AGG}$ and GroELS by using *E. coli* cells harboring both an expression vector for the heterologous gene of interest and the $tRNA_{AGA}$/$tRNA_{AGG}$/GroELS expression plasmid. Coexpression of either $tRNA_{AGA}$ [1,17,18] or GroELS [11] was reported previously. Now, a novel expression system involving $tRNA_{AGA}$/$tRNA_{AGG}$/GroELS in *E. coli* has been proved to be useful for the overexpression of a gene containing rare arginine codons, and should be applicable to the expression of eukaryotic and archaeal genes with relatively high contents of $tRNA_{AGA}$ and $tRNA_{AGG}$ (Table 1) in *E. coli*.

Table 2
Characteristics of minor arginine tRNAs and GroELS expression plasmids

Plasmid[a]	Product
pUBS520	$tRNA_{AGA}$
pArgW2	$tRNA_{AGG}$
pArgUW	$tRNA_{AGA}$ and $tRNA_{AGG}$
pT-groE	GroELS
pAUGE	GroELS and $tRNA_{AGA}$
pAWGE	GroELS and $tRNA_{AGG}$
pGro-ArgUW	GroELS, $tRNA_{AGA}$ and $tRNA_{AGG}$

[a]All plasmids have a kanamycin marker except for pT-groE, which has a chloramphenicol marker.

Acknowledgements: We are indebted to Dr. P. Buckel for providing the *argU* plasmid and Dr. S. Ishii for the GroELS plasmid. We wish to thank Mr. I. Yoshioka for the useful discussion regarding rare codons.

References

[1] Brinkmann, U., Mattes, R.E. and Buckel, P. (1989) Gene 85, 109–114.
[2] Ikemura, T. (1981) J. Mol. Biol. 146, 1–21.
[3] Komine, Y., Adachi, T., Inokuchi, H. and Ozeki, H. (1990) J. Mol. Biol. 212, 579–598.
[4] Spanjaard, R.A., Chen, K., Walker, J.R. and van Duin, J. (1990) Nucleic Acids Res. 18, 5031–5036.
[5] Rosenberg, A.H., Goldman, E., Dunn, J.J., Studier, F.W. and Zubay, G. (1993) J. Bacteriol. 175, 716–722.
[6] Seetharam, R., Heeren, R.A., Wong, E.Y., Braford, S.R., Klein, B.K., Aykent, S., Kotts, C.E., Mathis, K.J., Bishop, B.F., Jennings, M.J., Smith, C.E. and Siegel, N.R. (1988) Biochem. Biophys. Res. Commun. 155, 518–523.
[7] Calderone, T.L., Stevens, R.D. and Oas, T.G. (1996) J. Mol. Biol. 262, 407–412.
[8] Wakagi, T., Oshima, T., Imamura, H. and Matsuzawa, H. (1998) Biosci. Biotechnol. Biochem. 62, 2408–2414.
[9] Garcia, G.M., Mar, P.K., Mullin, D.A., Walker, J.R. and Prather, N.E. (1986) Cell 45, 453–459.
[10] Sakamoto, K., Kawai, G., Niimi, T., Satoh, T., Sekine, M., Yamaizumi, Z., Nishimura, S., Miyazawa, T. and Yokoyama, S. (1993) Eur. J. Biochem. 216, 369–375.
[11] Thomas, J.G., Ayling, A. and Baneyx, F. (1997) Appl. Biochem. Biotechnol. 66, 197–238.
[12] Jeon, B.S., Taguchi, H., Sakai, H., Ohshima, T., Wakagi, T. and Matsuzawa, H. (1997) Eur. J. Biochem. 248, 171–178.
[13] Nakamura, H., Takaha, T. and Okada, S. (1996) Shokuhin Kogyo 39, 52–59 (in Japanese).
[14] Yasukawa, T., Kanei, I.C., Maekawa, T., Fujimoto, J., Yamamoto, T. and Ishii, S. (1995) J. Biol. Chem. 270, 25328–25331.
[15] Laemmli, U.K. (1970) Nature 227, 680–685.
[16] Stanley, K. and Wellington, P. (1969) Anal. Biochem. 30, 467–471.
[17] Chesshyre, J.A. and Lowe, G. (1995) Biotechnol. Techniques 9, 237–240.
[18] Hill, A.M., Cane, D.E., Mau, C.J.D. and West, C.A. (1996) Arch. Biochem. Biophys. 336, 283–289.
[19] Namamura, Y., Gojobori, T. and Ikemura, T. (1998) Nucleic Acids Res. 334, 26.