ELSEVIER

International Journal for Parasitology 30 (2000) 113–118

INTERNATIONAL Journal for PARASITOLOGY

www.elsevier.nl/locate/ijpara

Rapid communication

Overcoming codon bias: A method for high-level overexpression of *Plasmodium* and other AT-rich parasite genes in *Escherichia coli*

Arthur M. Baca[a], Wim G.J. Hol[a,b,*]

[a]*Department of Bioengineering, Howard Hughes Medical Institute, Biomolecular Structure Center, Box 357742, University of Washington, WA 98195 Seattle USA*

[b]*Department of Biological Structure and Biochemistry, Howard Hughes Medical Institute, Biomolecular Structure Center, Box 357742, University of Washington, Seattle, WA 98195, USA*

Received 9 December 1999; received in revised form 18 January 2000; accepted 18 January 2000

Abstract

Parasite genes often use codons which are rarely used in the highly expressed genes of *Escherichia coli*, possibly resulting in translational stalling and lower yields of recombinant protein. We have constructed the "RIG" plasmid to overcome the potential codon-bias problem seen in *Plasmodium* genes. RIG contains the genes that encode three tRNAs (Arg, Ile, Gly), which recognise rare codons found in parasite genes. When co-transformed into *E. coli* along with expression plasmids containing parasite genes, RIG can greatly increase levels of overexpressed protein. Codon frequency analysis suggests that RIG may be applied to a variety of protozoan and helminth genes.

Keywords: *argU*; *ileX*; *glyT*; *Plasmodium falciparum*; Rare codon; RIG-plasmid

Overexpression of parasite genes is important for providing sufficient amounts of recombinant protein for biophysical studies, such as X-ray crystallography and NMR, as well as for immunological applications, such as vaccine production and the production of recombinant antigens for the generation of research antibodies. The overexpression of parasite genes in *Escherichia coli*, in particular those of *Plasmodium falciparum*, has often been a challenge because of the codon bias of these organisms. *Plasmodium falciparum* has an extremely AT-rich genome of about 80%, which has in many instances made heterologous expression of *Plasmodium* genes in *E. coli* very difficult [1,2]. Certain codons that are preferentially used by *P. falciparum* are rarely used by *E. coli* in highly expressed genes [2]. Rare codons have been shown to greatly diminish expression levels of recombinant protein in *E. coli* because of translational stalling [3,4]. One method by which this codon bias has been overcome is to re-engineer the *Plasmodium* gene to be overexpressed so that it uses the preferred codons of *E. coli* [5,6]. This method is both costly and time consuming. Alternatively, co-transformation of *E. coli* with a plasmid carrying the gene for the rare codon's cognate tRNA can increase levels of heterologous expression [3,4,7–10].

We have constructed a plasmid to overcome some of the potential codon-bias problems seen in *P. falciparum* genes. This "RIG" plasmid contains the genes that encode three tRNAs (Arg, Ile, Gly) which recognise codons that are both frequently found in *Plasmodium* and other parasite genes [11], but which are among the most rarely used by *E. coli* in highly

* Corresponding author. Tel.: +1-(206)-685-7044, fax: +1-(206)-685-7002.

E-mail address: hol@gouda.bmsc.washington.edu (W.G.J. Hol).

0020-7519/00/$20.00
PII: S0020-7519(00)00019-9

expressed genes [3]. Here we report the use of the RIG-plasmid for the overexpression of a variety of parasite proteins, including two *P. falciparum* proteins: dihydropteroate synthase (DHPS) and aldolase. The RIG-plasmid also proved to be beneficial for overexpressing *Plasmodium berghei* aldolase and *Cryptosporidium parvum* dihydrofolate reductase (DHFR). Our analysis of the rare codon frequency in parasite genes suggests that the RIG-plasmid may be broadly applicable for the overexpression of a variety of AT-rich protozoan and helminth genes. The time and expense of redesigning AT-rich parasite genes into the optimal codon bias of *E. coli* may now be circumvented by use of the RIG-plasmid in certain instances.

Codon frequency analysis shows that for some amino acids such as Arg, Ile, and Gly, *P. falciparum* preferentially uses codons [11] which are among the most rarely used codons by *E. coli* in highly expressed genes [3]. Table 1 lists the codon usage of Arg AG(A/G), Ile AUA, and Gly GGA for genes that are highly and continuously expressed in *E. coli* during exponential growth [12], as well as for a variety of protozoan and helminth parasites [11]. It is noteworthy that the codon usage of highly expressed genes is distinctly different from the overall codon usage of all *E. coli* genes [3,11–13]. Heterologous genes which contain the codon usage of *E. coli*'s highly expressed genes are generally overexpressed in *E. coli* at better yields [3,4]. Note that there is almost a hundred fold difference between *E. coli*'s preferred codon usage for highly expressed genes of AG(A/G) and AUA and the

Table 1
Usage of rare codons in *Escherichia coli* and selected parasites. All organisms use certain of the selected rare codons more than 10% of the time. Genes overexpressed from these organisms in *E.coli* may benefit from the RIG-plasmid

	AGA/AGG (Arg)		AUA (Ile)		GGA (Gly)		
Organism	/1000[a]	Fraction[b]	/1000[a]	Fraction[b]	/1000[a]	Fraction[b]	Reference
*Escherichia coli*Highly expressed genes	nr[d]	0.91	nr[d]	0.57	nr[d]	1.97	[12]
Plasmodium Parasites							
Plasmodium falciparum	20.6	76.3[c]	44.9	53.3[c]	16.6	43.5[c]	[11]
Plasmodium vivax	23.5	62.7[c]	10.8	27.3	32.3	43.4[c]	[11]
Plasmodium malariae	15.7	87.8[c]	15.2	52.8[c]	98.5	73.0[c]	[11]
Plasmodium berghei	22.7	72.3[c]	30.8	42.6[c]	25.0	44.9[c]	[11]
Plasmodium chabaudi	17.1	77.6[c]	28.2	46.0[c]	20.7	42.2[c]	[11]
Plasmodium knowlesi	29.9	70.0[c]	17.5	34.6[c]	32.9	44.4[c]	[11]
Plasmodium yoelii	15.9	71.3[c]	41.7	49.6[c]	22.5	50.2[c]	[11]
Protozoan Parasites							
Toxoplasma gondii	17.4	30.4[c]	5.2	11.0	20.9	27.2	[11]
Theileria parva	29.9	79.6[c]	15.1	29.7	35.3	51.1[c]	[11]
Babesia bovis	19.2	42.4[c]	13.2	26.3	10.7	21.4	[11]
Eimeria tenella	18.1	38.2	8.3	18.3	26.4	32.1	[11]
Naegleria fowleri	27.6	58.3[c]	1.7	2.4	26.6	36.6	[11]
Cryptosporidium parvum	30.0	79.0[c]	19.5	24.7	25.3	38.2	[11]
Entamoeba histolytica	31.0	83.2[c]	10.3	13.6	50.4	71.6[c]	[11]
Giardia lamblia	15.2	33.3	10.6	20.4	17.4	23.5	[11]
Trichomonas vaginalis	13.3	30.6	1.8	2.6	8.4	11.2	[11]
Trypanosoma brucei	14.8	24.8	16.5	23.0	10.6	23.2	[11]
Trypanosoma cruzi	11.4	18.8	5.0	12.5	12.2	16.5	[11]
Leishmania spp.	6.6	10.1	2.3	6.6	5.3	7.4	[11]
Helminth Parasites							
Echinococcus spp.	14.4	23.4	8.3	16.1	7.5	12.8	[11]
Taenia spp.	14.0	24.9	10.7	17.9	16.1	27.2	[11]
Schistosoma mansoni	13.9	29.2	19.7	30.5	19.3	29.3	[11]
Schistosoma japonicum	17.1	32.1	16.5	28.4	21.0	28.9	[11]
Onchocerca volvulus	11.6	20.8	14.2	21.6	20.9	34.2	[11]
Brugia malayi	14.7	25.9	13.3	23.8	29.2	37.1	[11]
Brugia pahangi	12.7	24.9	13.3	22.6	17.6	29.6	[11]
Opportunistic Yeasts							
Pneumocystis carinii	20.5	54.1[c]	18.3	32.6	28.3	49.2[c]	[11]
Candida albicans	26.8	70.9[c]	11.9	18.1	13.8	22.8	[11]

[a] Number of times this codon occurs per 1000 codons.
[b] Percentage usage of synonymous codons (i.e. usage of this codon relative to all other codons which code for the same amino acid).
[c] Preferred codon.
[d] nr: Not Reported.

usage by *P. falciparum*, and that there is a twenty-fold difference in the usage of the GGA codon.

We have tested the effects of the addition of cognate tRNAs for the rare codons AGA, AGG, AUA, and GGA on the overexpression of four different apicomplexan proteins in different *E. coli* expression systems. (Table 2). These codons were chosen because when the codon usage of highly expressed genes in *E. coli* was compared with the codon usage of *P. falciparum*, these four codons showed the greatest difference. Each of the *E. coli* expression systems we chose to investigate results in very poor yields of the target parasite protein. (See below.) We hypothesised that the poor expression is due to the presence of a large number of evenly distributed rare codons found in these genes, as shown in Table 2, and that coexpression of arginyl, isoleucyl, and glycyl tRNAs, which recognise the codons AG(A/G), AUA, and GGA, respectively, would relieve the translational stress present in *E. coli* during heterologous gene expression. In fact, the genes shown in Table 2 also contain rare codon doublets and triplets, which can result in translational stalling [3]. We thus designed the RIG-plasmid to carry the following tRNA genes under control of their natural promoters: the minor $tRNA_{AG(A/G)}$ gene (*argU*), the minor $tRNA_{AUA}$ gene (*ileX*), and the minor $tRNA_{GG(A/G)}$ gene (*glyT*).

pRI952, a pBR322-compatible vector, is a pACYC184 derivative which carries the *argU* and *ileX* genes and served as the parent plasmid from which the RIG-plasmid was constructed [10]. The oligonucleotide primers 5′ CAGCTGCtcatgaggtacctaggCTGTTTCTT-CATCGTGTCGCATAAAATG 3′ (forward) and 5′ GCAGCTGtcatgatctagacgtcATCGATTGTCCCTC-TAAGACACGGATAAATC 3′ (reverse) were used in a PCR to amplify the *glyT* gene, which encodes the tRNA recognising the codon GGA, from genomic *E. coli* DNA. The first primer contains *Bsp*HI, *Avr*II, and *Kpn*I sites and the second primer contains *Aat*II, *Xba*I, and *Bsp*HI sites. The PCR reaction contained 10 mM Tris-HCl (pH 8.8), 1.5 mM $MgCl_2$, 75 mM KCl, 5 nM EDTA, 50 ng of genomic DNA, 75 pmol of each of the two primers, 10 nmol of the four dNTPs, and 1.2 units of Taq 2000 polymerase (Stratagene) in a total volume of 50 μl. The cycling conditions were 94°C for 30 s, 62°C for 30 s, and 72°C for 30 s for a total of 25 cycles with a final extension time of 7 min at 72°C. A PCR fragment of expected size of 707 bp was gel-purified and ligated into the pCRII-TOPO vector (Invitrogen). pCRII-TOPO-*glyT* was digested with *Bsp*HI and *Xho*I and the correct partial digest product containing

Table 2
Parasite Expression Systems Tested with the RIG-plasmid in *Escherichia coli*

Expression System	*P. falciparum* dihydropteroate synthase (DHPS)	*P. falciparum* aldolase	*P. berghei* aldolase	*C. parvum* dihydrofolate reductase (DHFR)
Description	Residues 373–706 of the bifunctional PPPK-DHPS enzyme fused to a C-terminal hexa-histidine tag	Residues 6–368 of aldolase fused to a N-terminal hexa-histidine tag	Residues 4–369 of aldolase fused to a N-terminal hexa-histidine tag	Residues 1–182 of the bifunctional DHFR-TS enzyme
Vector	pET-23a (Novagen)	pDS78/RBSII,6xHis[15]	pDS78/RBSII,6xHis[15]	pET-9d (Novagen)
Promoter	T7 promoter	T5 promoter	T5 promoter	T7 promoter
E. coli Expression Strain	BL21(DE3),Novagen	XL1-Blue, Stratagene	XL1-Blue, Stratagene	BL21(DE3), Novagen
Reference	[14]	[16]	[17]	[18]
Occurrences of Rare Codons in Parasite Genes Tested				
	#/(total)[a]	#/(total)[a]	#/(total)[a]	#/(total)[a]
R (AGA/AGG)	11 (13)	14 (14)	12 (14)	8 (8)
I (AUA)	18 (34)	2 (17)	5 (19)	9 (19)
G (GGA)	9 (15)	14 (31)	16 (32)	4 (7)
Rare codon doublets[b]	6	2	4	4
Rare codon triplets[c]	0	0	0	1
Total number of amino acids	334	363	365	182

[a] The first number represents the number of times the rare codon occurs in the gene. The number in parenthesis represents the total number of occurrences of the corresponding amino acid in the gene.
[b] A rare codon doublet is two tandem (both identical and different) rare codons.
[c] A rare codon triplet is three tandem (both identical and different) rare codons.

the *glyT* gene was identified and subcloned into pRI952 digested with *Bsp*HI and *Sal*I. The identity of the *glyT* gene in the RIG-plasmid (pRI952 with *glyT*) was verified by DNA sequencing analysis.

The RIG-plasmid is a pACYC184 derivative which carries a chloramphenicol resistance marker and the p15A origin of replication and is thus compatible with most commercially available expression vectors, which are of pBR322 origin and carry an ampicillin or kanamycin resistance marker and the ColEI origin of replication. The RIG-plasmid carries the *argU*, *ileX*, and *glyT* tRNA genes. A vector map as well as a complete vector sequence can be found at (http://www.bmsc.washington.edu).

Co-transformation of the RIG-plasmid along with the four expression plasmids listed in Table 2 into the different *E. coli* strains was performed using standard techniques [19]. Ten nanograms of the RIG-plasmid and 10 ng of the expression plasmid were mixed and transformed into competent *E. coli* cells. Transformants were selected on media containing chloramphenicol (25 μg ml^{-1}) and ampicillin (100 μg ml^{-1}) and were further tested for the presence of both plasmids using PCR analysis. Primers specific for pACYC184 were used to amplify regions of the RIG-plasmid, whereas primers specific for the inserted parasite gene in each expression plasmid were used to amplify regions of the expression plasmid.

Co-transformant colonies were directly picked into 30 μl of sterile water, boiled for 10 min, and 20 μl of the cleared lysate was added to the following PCR components to a final volume of 50 μl: 50 pmol of each primer, 10 nmol of the four dNTPs, 1 unit of Taq 2000 polymerase (Stratagene), 10 mM Tris-HCl (pH 8.8), 1.5 mM $MgCl_2$, 75 mM KCl, 5 nM EDTA.

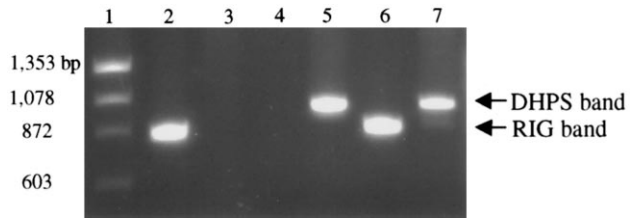

Fig. 1. PCR detection of RIG-plasmid in *E. coli* Co-transformants: Lane 1 is the marker ΦX174*HaeIII* digest. Lanes 2 and 3 are PCR amplification products of lysates of *E. coli* transformed only with the RIG-plasmid using RIG and DHPS primer pairs, respectively. A 837-bp band corresponding to a RIG-plasmid fragment is only seen in Lane 2. Lanes 4 and 5 are PCR amplification products of lysates of *E. coli* transformed only with the DHPS expression plasmid using RIG and DHPS primer pairs, respectively. A 1000-bp band corresponding to a DHPS fragment is only seen in Lane 5. Lanes 6 and 7 are PCR amplification products of lysates of *E. coli* transformed with both the RIG-plasmid and the DHPS expression plasmid using RIG and DHPS primer pairs, respectively. A 837-bp band corresponding to a RIG-plasmid fragment is seen in Lane 6 and a 1000-bp band corresponding to a DHPS fragment is seen in Lane 7.

The cycling conditions were 94°C for 30 s, 62°C for 30 s, and 72°C for 30 s for a total of 25 cycles with a final extension time of 7 min at 72°C. Primers specific to the pACYC184 backbone were used to detect the RIG-plasmid: P-RIG1 5′ CGCCGACATCACCGATGGGGAAG 3′ and P-RIG2 5′ CGCGCCCACCGGAAGGAGCTGACTG 3′, and primers specific to *P. falciparum* DHPS were used to detect the expression plasmid: P-DHPS1 5′ AAAAGGATATATGTATTAAAAGATAG 3′ and P-DHPS2 5′ CACTTGGTCTATTTTTGTTAAAAC 3′ (Fig. 1).

Escherichia coli co-transformed with the expression plasmid and RIG were compared with *E. coli* co-transformed with the expression plasmid and pACYC184 (control). Overnight cultures were prepared by picking single colonies of co-transformed *E. coli* and inoculating 3 ml Luria–Bertani (LB) media supplemented with ampicillin (100 μg ml^{-1}) and chloramphenicol (25 μg ml^{-1}). The following morning, 100 μl of the overnight culture were used to inoculate flasks containing 25 ml of LB media supplemented with ampicillin (100 μg ml^{-1}) and chloramphenicol (25 μg ml^{-1}). These cultures were incubated for 3–4 h at 37°C until cell density OD_{600} reached 0.5–0.7. The cultures were induced by the addition of IPTG to a final concentration of 0.5 mM and then harvested 3 h after induction. Expression of protein was monitored by SDS-PAGE of pre-induction and harvest whole cell lysates of cells with and without the RIG-plasmid. Western Analysis was performed on *P. falciparum* DHPS to confirm protein identity; enzyme assays were performed on *P. falciparum* and *P. berghei* aldolases, and *C. parvum* DHFR. The level of expression of all recombinant proteins was enhanced by the presence of the RIG-plasmid.

Plasmodium falciparum DHPS could only be overexpressed in the presence of the RIG-plasmid. In the absence of the RIG-plasmid, no induced band could be seen. With the RIG-plasmid, an induced band of 40.2 kDa corresponding to *P. falciparum* DHPS constituted up to 30% of *E. coli* protein at harvest. (Fig. 2). The protein was purified from inclusion bodies and Western Analysis using an antibody directed against residues *P. falciparum* PPPK-DHPS 631–644 [14] confirmed the identity of the protein as *P. falciparum* DHPS.

The expression of both *P. falciparum* and *P. berghei* aldolases was greatly enhanced by the RIG-plasmid. Without the RIG-plasmid, no induced band could be seen by SDS-PAGE, and expression of both aldolases was only 2–3 mg protein l^{-1} cell culture after a three-step purification protocol [16,17]. With the RIG-plasmid, an induced band of 41 kDa corresponding to the *Plasmodium* aldolases could be seen by SDS-PAGE and the aldolases were produced at a yield of 12 mg protein l^{-1} cell culture (L.Y. Fuchs, personal communi-

cation). (Fig. 2). With these improved yields, only one Ni-affinity chromatography step was needed to purify the protein. Enzyme activity assays of the purified protein [16,17] confirmed the identity as aldolase.

Cryptosporidium parvum DHFR could only be expressed and purified in the presence of the RIG-plasmid (V.H. Brophy and C.H. Sibley, personal communication). DHFR activity was determined spectrophotometrically by monitoring the decrease in A_{340nm} at room temperature of cofactor NADPH (100 μM) in buffer containing 50 mM TES pH 7.0, 1 mM EDTA, 75 mM β-mercaptoethanol, 1% bovine serum albumin, and 20 μM dihydrofolate. Cell lysate concentrations were adjusted to give linear rates over the 5-min duration of the assay, and reactions were initiated by adding an equal volume of 40-μM dihydrofolate in assay buffer. A unit of activity is defined as the amount of enzyme required to produce 1 μmol of product per minute at room temperature. Without the RIG-plasmid, DHFR activity in crude whole cell lysates at harvest was 0.06 U mg^{-1} protein, which was no greater than background *E. coli* DHFR activity. With the RIG-plasmid, DHFR activity in the crude whole cell lysates increased by 33% to 0.08 U mg^{-1} protein, which allowed for the purification of *C. parvum* DHFR with a methotrexate affinity column to a yield of 2.3 mg l^{-1} cell culture. Without RIG, *C. parvum* DHFR could not be purified (V.H. Brophy and C.H. Sibley, personal communication).

The RIG-plasmid appears to be a useful tool which allows for the overexpression of several AT-rich parasite genes in *E. coli*. The RIG-plasmid is compatible with most commercially available vectors and *E. coli* expression strains. When co-transformed with the RIG-plasmid, the *E. coli* expression systems tested could produce milligram amounts of recombinant parasite protein which were previously expressed in low yields.

The RIG-plasmid may help circumvent difficulties in the translation of genes containing the codons AG(A/G), AUA, and GGA. Other codons may need to be considered when expressing AT-rich genes, such as Leu (UUA) and Thr (ACA), but the RIG-plasmid encodes the cognate tRNAs for codons which showed the greatest difference in usage between *P. falciparum* and highly expressed genes in *E. coli* and thus provides a good starting point for overexpression experiments. It should be noted that the RIG-plasmid does not address difficulties encountered in the transcription of AT-rich genes or problems associated with mRNA secondary structure or stability. Additionally, the RIG-plasmid is not designed to alter the solubility, stability, folding behaviour, degradation, or toxicity to the host cell of the overexpressed recombinant protein. It is for

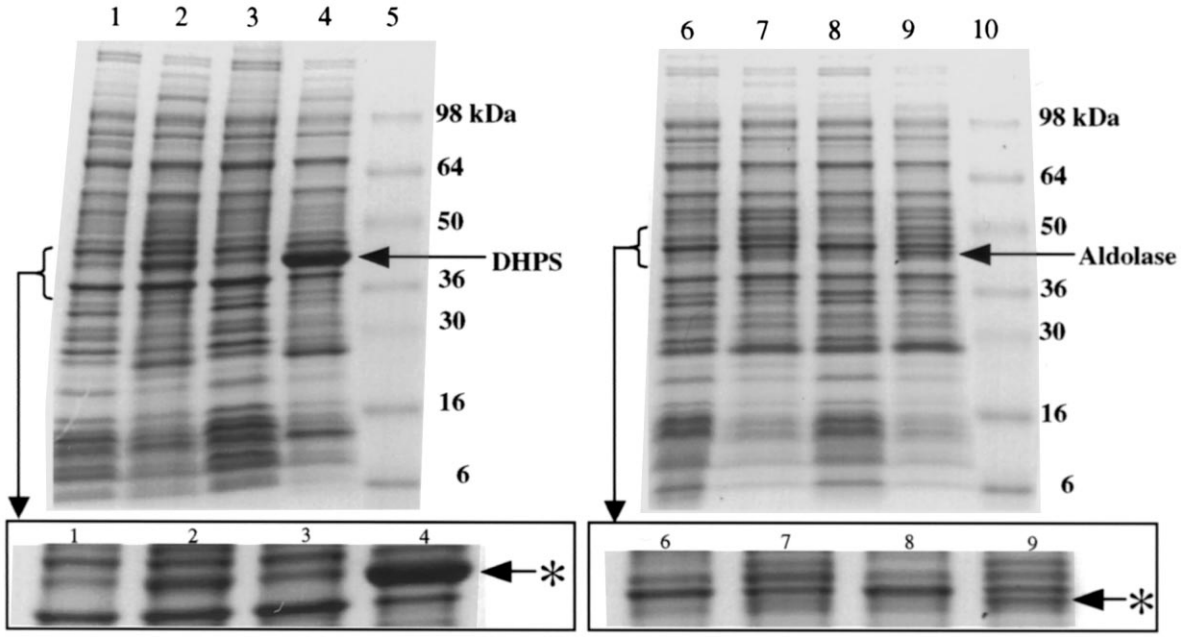

Fig. 2. SDS-PAGE Gels of Whole Cell Lysates with and without the RIG-plasmid. Lanes 1 and 2: Pre-induction and harvest samples of BL21(DE3) transformed with pET-*P. falciparum* DHPS and pACYC184. Very little induced band can be seen in the harvest sample. Lanes 3 and 4: Pre-induction and harvest samples of BL21(DE3) transformed with pET-*P. falciparum* DHPS and RIG. A band corresponding to *P. falciparum* DHPS can clearly be seen in the harvest sample. Lane 5: SeeBlue (Novex) molecular weight marker. Lanes 6 and 7: Pre-induction and harvest samples of XL1-Blue transformed with the *P. berghei* aldolase expression plasmid and pACYC184. Very little induced band can be seen in the harvest sample. Lanes 8 and 9: Pre-induction and harvest samples of XL1-Blue transformed with the *P. berghei* aldolase expression plasmid and RIG. A band corresponding to *P. berghei* aldolase can clearly be seen in the harvest sample. Lane 10: SeeBlue molecular weight marker. Results for *P. falciparum* aldolase are similar. (Not shown.)

these reasons that we believe the RIG-plasmid did not have the same magnitude of effect on each of the four expression systems tested. Each of these expression systems is different in the type of protein expressed. However, for all the expression systems tested, the RIG-plasmid had a positive effect on yields of overexpressed recombinant protein.

The RIG-plasmid may be broadly applicable for the overexpression of a variety of AT-rich protozoan, helminth, and fungal genes which contain large numbers of the codons for Arg, Ile, and Gly that are rarely used by *E. coli* in highly expressed genes, as shown in Table 1. The RIG-plasmid is a simple and convenient tool to test when performing initial overexpression experiments of AT-rich genes. In the same amount of time it takes to construct one synthetic parasite gene into the optimal codon usage of *E. coli*, various sized constructs of the same parasite gene can be tested in combination with the RIG-plasmid. In certain cases, the time and expense of redesigning parasite genes into the optimal codon bias of *E. coli* may now be circumvented by use of the RIG-plasmid. Further information on the RIG-plasmid, including information on obtaining the plasmid, can be found at (http://www.bmsc.washington.edu).

Acknowledgements

We thank James Kane, SmithKline Beecham Pharmaceuticals, for providing pRI952; Alan Cowman, Walter and Eliza Hall Institute for Medical Research (Melbourne, Australia) for providing the anti-DHPS antibody as well as pTrc-PPPK-DHPS from which the pET-DHPS expression system was derived; Ulrich Certa, Roche Pharmaceuticals, for providing the following *Plasmodium* aldolase expression vectors: p3217, p3255 and p3113; Victoria H. Brophy and Carol H. Sibley, Department of Genetics, University of Washington, for testing the RIG-plasmid with *C. parvum* DHFR; Lidia Yolanda Fuchs, Department of Biological Structure, University of Washington, for testing the RIG-plasmid with *Plasmodium* aldolases; and Carol Sibley for critically reading the manuscript. AM Baca is supported by the Whitaker Foundation.

References

[1] Saul A, Battistutta D. Codon usage in *Plasmodium falciparum*. Mol Biochem Parasitol 1988;27:35–42.

[2] Sayers JR, Price HP, Fallon PG, Doenhoff MJ. AGA/AGG codon usage in parasites: implications for gene expression in *Escherichia coli*. Parasitol Today 1995;11:345–6.

[3] Kane JF. Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. Curr Opin Biotechnol 1995;6:494–500.

[4] Kurland C, Gallant J. Errors of heterologous protein expression. Curr Opin Biotechnol 1996;7:489–93.

[5] Prapunwattana P, Sirawaraporn W, Yuthavong Y, Santi DV. Chemical synthesis of the *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase gene. Mol Biochem Parasitol 1996;83:93–106.

[6] Pan W, Ravot E, Tolle R, et al. Vaccine candidate MSP-1 from *Plasmodium falciparum*: a redesigned 4917 bp polynucleotide enables synthesis and isolation of full-length protein from *Escherichia coli* and mammalian cells. Nucleic Acids Res 1999;27:1094–103.

[7] Brinkmann U, Mattes RE, Buckel P. High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the *dnaY* gene product. Gene 1989;85:109–14.

[8] Seidel HM, Pompliano DL, Knowes JR. Phosphonate biosynthesis: molecular cloning of the gene for phosphoenolpyruvate mutase from *Tetrahymena pyriformis* and overexpression of the gene product in *Escherichia coli*. Biochemistry 1992;31:2598–608.

[9] Rosenberg AH, Goldman E, Dunn JJ, Studier FW, Zubay G. Effects of consecutive AGG codons on translation in *Escherichia coli*, demonstrated with a versatile codon test system. J Bacteriol 1993;175:716–22.

[10] Del Tito BJ, Ward JM, Hodgson J, et al. Effects of a minor isoleucyl tRNA on heterologous protein translation in *Escherichia coli*. J Bacteriol 1995;177:7086–91.

[11] Nakamura Y, Gojobori T, Ikemura T. Codon usage tabulated from the international DNA sequence databases. Nucleic Acids Res 1998;26:334.

[12] Médigue C, Rouxel T, Vigier P, Hénaut A, Danchin A. Evidence for horizontal gene transfer in *Escherichia coli* speciation. J Mol Biol 1991;222:851–6.

[13] Karlin S, Mrázek J, Campbell AM. Codon usages in different gene classes of the *Escherichia coli* genome. Mol Microbiol 1998;29:1341–55.

[14] Triglia T, Menting JGT, Wilson C, Cowman AF. Mutations in dihydropteroate synthase are responsible for sulfone and sulfonamide resistance in *Plasmodium falciparum*. Proc Natl Acad Sci USA 1997;94:13944–9.

[15] Stüber D, Matile H, Garotta G. System for high-level production in *Escherichia coli* and rapid purification of recombinant proteins: application to epitope mapping, preparation of antibodies, and structure–function analysis. Immunol Methods 1990;4:121–52.

[16] Döbeli H, Trzeciak A, Gillessen D, et al. Expression, purification, biochemical characterization and inhibition of recombinant *Plasmodium falciparum* aldolase. Mol Biochem Parasitol 1990;41:259–68.

[17] Meier B, Döbeli H, Certa U. Stage-specific expression of aldolase isoenzymes in the rodent malaria parasite *Plasmodium berghei*. Mol Biochem Parasitol 1992;52:15–28.

[18] Vasquez JR, Gooz'e L, Kim K, Gut J, Petersen C, Nelson RG. Potential antifolate resistance determinants and genotypic variation in the bifunctional dihydrofolate reductase-thymidylate synthase gene from human and bovine isolates of *Cryptosporidium parvum*. Mol Biochem Parasitol 1996;79:153–65.

[19] Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual, 2nd edition. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1989.