Available online at www.sciencedirect.com
SCIENCE DIRECT®
Protein Expression and Purification 42 (2005) 37–46
Protein Expression & Purification
www.elsevier.com/locate/yprep

Development of an inducible protein expression system based on the protozoan host *Leishmania tarentolae*

Susanna Kushnir [a], Klaus Gase [b,1], Reinhard Breitling [b,*], Kirill Alexandrov [a]

[a] *Department of Physical Biochemistry, Max-Planck-Institute for Molecular Physiology, Otto-Hahn-Strasse 11, 44227 Dortmund, Germany*
[b] *Jena Bioscience GmbH, Löbstedter Str. 78, 07749 Jena, Germany*

Received 30 November 2004, and in revised form 27 February 2005
Available online 25 March 2005

Abstract

Production of functional eukaryotic proteins in recombinant form is a bottle-neck in various post-genomic applications and in life science in general. At least partially this is due to the problems associated with the use of endogenous RNA polymerase II for high-level transcription of heterologous genes in eukaryotic expression systems. To circumvent these problems we developed a new inducible protein expression system based on the protozoan host *Leishmania tarentolae* (Trypanosomatidae). We have created a strain of *L. tarentolae* constitutively co-expressing T7 RNA polymerase and tetracycline repressor. This strain could be stably transformed with the heterologous target gene under control of the T7 promoter/TET operator assembly, which can initiate transcription upon addition of tetracycline to the culture medium. Using this system, we demonstrated that enhanced green fluorescent protein (EGFP) could be overexpressed to a level of ca. 1% of total cellular protein. The developed system was tested for its ability to inducibly co-express multiple genes. Using two copies of the *egfp* gene integrated at two different genomic sites, we could obtain expression levels reaching 4% of total cellular protein. Further possible improvements and applications of the developed system are discussed.

Keywords: Eukaryotic expression systems; T7 polymerase

With the completion of several genome projects, attention has turned to the elucidation of structure, properties, and functional activities of the encoded proteins. Due to the enormous number of newly discovered open reading frames (ORFs), progress in the analysis of the corresponding proteins depends on the ability to perform characterization in a parallel and high-throughput format. The success of such approaches keenly depends on our ability to produce the proteins in a pure form rapidly and at a reasonable cost. Yet, despite a 30-year-old history, no universally applicable protein production system suitable for such task has been developed [1]. Due to technical restrictions imposed by the high-throughput format of the trials, investigators were forced to resort to the most efficient expression system based on *Escherichia coli*. This system, however, can produce only ca. 15% of eukaryotic proteins in their active form [2]. The only eukaryotic expression system considered for high-throughput application purposes is based on *Saccharomyces cerevisiae*, but this suffers from low yields and lack of proper post-translational modifications.

The fundamental problem in the production of recombinant proteins in eukaryotic expression systems is rooted in the biology of the organisms chosen as expression hosts and has not been properly recognized. All available eukaryotic protein expression systems are based on free-living organisms and utilize the endogenous RNA polymerase II (Pol II)[2] for transcription. Due to constant

* Corresponding author. Fax: +49 3641 464919.
E-mail address: reinhard.breitling@jenabioscience.com (R. Breitling).
[1] Present address: Department of Molecular Ecology, Max-Planck-Institute for Chemical Ecology, Hans-Knöll-Straße 8, D-07745 Jena, Germany.

1046-5928/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2005.03.004

changes in their environment, free-living eukaryotes are forced to control their gene expression very tightly, primarily at the transcriptional level. High-level overexpression of heterologous proteins affects the physiology of the host, leading to down-regulation of protein expression via activation of transcriptional control mechanisms. The complexity of the transcription control machinery precludes engineering eukaryotes that reproducibly produce large amounts of recombinant polypeptides. An exception is the baculovirus expression system, where the Pol II based transcriptional machinery is hijacked by a virus and driven to extreme behaviour. However, effective control of the transcriptional machinery requires viral integrity, which sets limits to its engineering.

The most successful example of manipulation of the transcription/translation machinery is implemented in T7-based *E. coli* expression systems. In this case, the highly processive phage polymerase that cannot be controlled by the transcriptional machinery of the host swamps the cell with heterologous RNA. Unfortunately, this approach cannot be transplanted into standard eukaryotic expression platforms. The main problem is rooted in the coupling of DNA transcription with RNA processing. Transcripts generated by foreign RNA polymerases remain uncapped in eukaryotic cells and yield no protein product [3]. There are two possible solutions to this problem: one is the development of a system where the requirement for RNA capping can be bypassed, the second is employing an organism where transcription and translation are naturally uncoupled. The first approach is difficult since it requires identification of internal ribosomal entry sites (IRES), which mediate translational initiation nearly as efficiently as capped RNA. Such sequences have not been found so far in any existing system [4]. Therefore, in our efforts to develop an efficient eukaryotic protein expression system we chose to pursue the second approach and take advantage of the unique organization of the transcription/translation machinery of Trypanosomatidae protozoa. The absolute majority of these organisms are parasites that colonize hosts ranging from plants to higher mammals. Amongst the unique features of trypanosomatids are RNA editing, arrangement of genes in tandem arrays, polycistronic transcription followed by trans-splicing, and regulation of gene expression almost exclusively at the post-transcriptional level [5]. It was demonstrated that *Leishmania* and *Trypanosoma* species can efficiently translate mRNA generated by RNA polymerase I or a foreign polymerase such as T7 or T3 RNA polymerases [6–9]. Several efforts were made to express heterologous proteins in Trypanosomatidae using episomal vectors or genomic integrations transcribed by Pol II [10–13]. Earlier, we demonstrated that a non-pathogenic kinetoplastida, *Leishmania tarentolae*, can be used for heterologous expression of recombinant proteins [14]. The availability of methods of genetic and microbiological manipulation in combination with short doubling times of ca. 5 h leading to densities of $>5 \times 10^8$ cells/ml in suspension culture make this organism an attractive host for production of recombinant proteins. Moreover, we reported previously that expression of mammalian glycoproteins in this organism results in their modification with mammalian-like oligosaccharides [14]. The originally reported expression system based on *L. tarentolae* exploited RNA polymerase I mediated constitutive transcription of heterologous genes integrated into the 18S rDNA gene locus (Fig. 1A). In this work, we describe an inducible expression system based on transgenic *L. tarentolae* strain bearing genes for T7 RNA polymerase and tetracycline repressor.

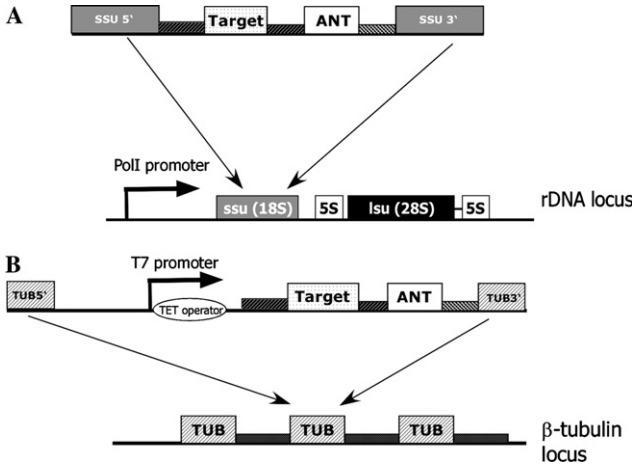

Fig. 1. Principal scheme for construction of stably transfected *L. tarentolae* strains used in this investigation. (A) Construction of constitutively expressing strains by integrating the expression cassette into the *ssu* gene cluster. 5S, small (5S) ribosomal RNA gene; *lsu*, gene for large (28S) ribosomal RNA; Target, gene intended for expression; and ANT, operably coupled antibiotic selection marker. (B) Integration of expression cassette into the tubulin gene cluster.

Experimental procedures

Plasmid construction and molecular biology

Plasmids for stable integration into rDNA (ssu) locus and RNA polymerase I mediated constitutive expression

pIRT7polNLSsat#A6 (Fig. 2A) was constructed by PCR amplification of a 2.7 kbp fragment encoding T7 RNA polymerase containing N-terminal nuclear

[2] *Abbreviations used:* APRT, adenine phosphoribosyltransferase; *ble*, bleomycin resistance gene; *cam A*, calmodulin A gene; EGFP, enhanced green fluorescent protein; *hyg*, hygromycin resistance gene; IR, intergenic region; *neo*, neomycin resistance gene; NLS, nuclear localization signal; Pol II, RNA polymerase II, rfu, relative fluorescent units; *sat*, streptothricin acetyltransferase gene; *ssu*, small subunit RNA gene in rDNA locus; T7 Pol, T7 RNA polymerase; Tet, tetracycline; *tub*, β-tubulin gene; UTR, untranslated region.

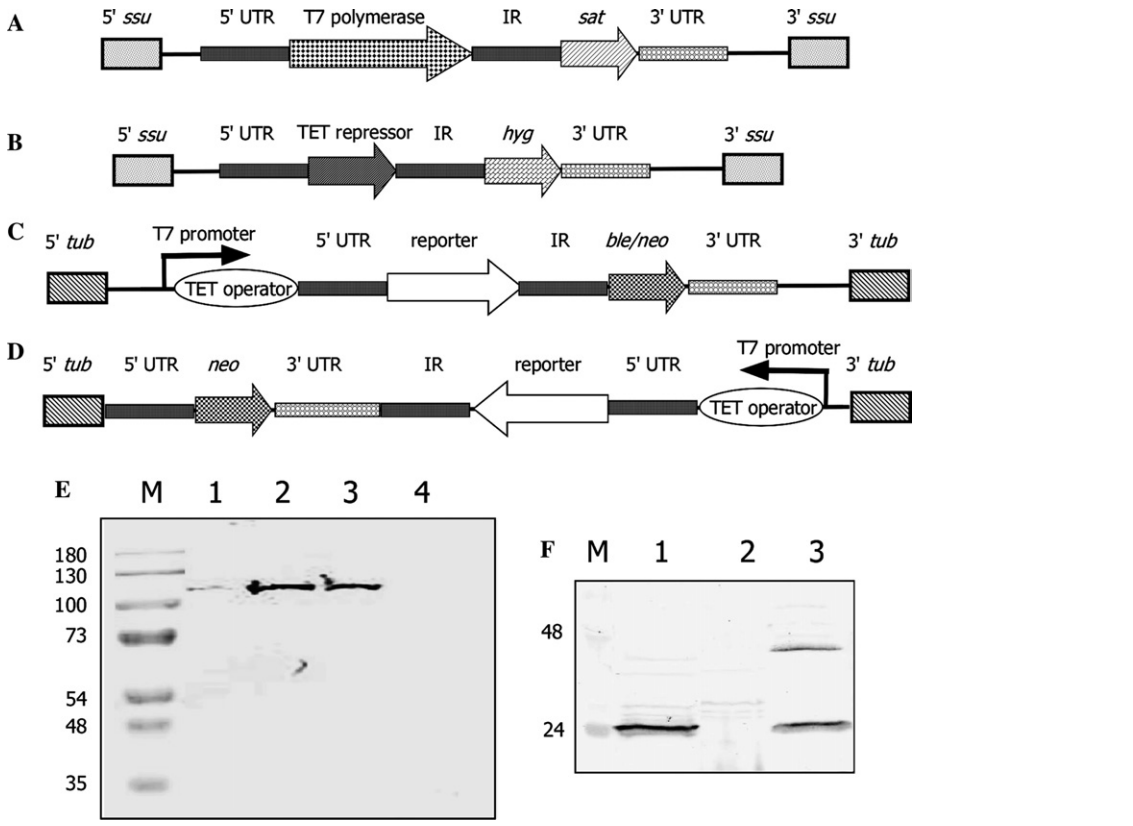

Fig. 2. Genetic architecture of *L. tarentolae* strains used for inducible expression of EGFP reporter gene. (A) Schematic representation of T7 polymerase expression cassette integrated in the *ssu* locus. (B) Schematic representation of TET repressor expression cassette integrated in the *ssu* locus. (C) Representation of a reporter gene expression cassette under control of T7 promoter/TET operator integrated in the tubulin gene cluster in sense orientation. (D) Representation of a reporter gene expression cassette under control of T7 promoter/TET operator integrated in the tubulin gene cluster in antisense orientation. (E) Western blot analysis of *L. tarentolae* strains constitutively expressing variants of T7 polymerase. M, prestained high molecular weight marker; 1, LO_NLS_T7; 2, HI_NLS_T7; 3, HI_T7; and 4, wild type strain. (F) Western blot analysis of *L. tarentolae* strains constitutively expressing TET repressor. M, prestained molecular weight marker; 1, TET repressor expressing strain; 2, wild type strain; and 3, TET repressor protein.

localization signal (NLS) PKKKRKV derived from SV40 T antigen with primers 5′-GTC CGG AGA TCT GTC GCC ACC ATG AAC ACG ATT AAC ATC GC-3′ and 5′-CTC ATG GAT CCG GAG TCG TAT TGA TTT G-3′ using pLew13 plasmid as a template [15]. The resulting fragment was *Bgl*II × *Bam*HI trimmed and subcloned into *Leishmania* expression vector pIRsat1 (S. Beverley, unpublished) opened with *Bgl*II. Plasmid pIRT7polNLSsat# A6 was used to generate *L. tarentolae* **LO_NLS_T7** strains overexpressing NLS tagged T7 polymerase at low levels.

pF4T7pol1.4sat#1 (Fig. 2A) was constructed by PCR amplification of a 2.7 kbp fragment encoding T7 RNA polymerase, amplified with primers 5′-GAG GCA CTC CAT GGA GAC GAT TAA CAT CGC TAA GAA CG -3′ and 5′-TAG TGA GTC GTG CGG CCG CCT ACG CGA ACG CGA AGT CCG ACT C-3′ using genomic DNA of *E. coli* BL21 (DE3) strain as a template. The product was digested with *Nco*I × *Not*I and subcloned into *Leishmania* expression vector pF4X1.4sat (Jena Bioscience) opened with the same enzymes. Plasmid pF4T7pol1.4sat# 1 was used to generate *L. tarentolae* **HI_T7** strains overexpressing T7 polymerase without NLS at high levels.

pF4T7polNLS1.4sat#7 (Fig. 2A) was created by insertion of 2.4 kbp *Bsa*BI × *Kpn*I fragment derived form pIRT7polNLSsat#A6 into pF4T7pol1.4sat#1 opened with the same enzymes, replacing the major part of T7 RNA polymerase ORF and introducing the NLS. Plasmid pF4T7polNLS1.4sat#7 was used to obtain *L. tarentolae* **HI_NLS_T7** strains overexpressing T7 RNA polymerase with NLS at high levels.

pF4TR1.4hyg#1 (Fig. 2B) was obtained by insertion of 0.6 kbp *Bgl*II × *Not*I trimmed PCR fragment harbouring the TET repressor gene, amplified with primers 5′-TAA TAA AGA TCT CCC ATG TCT AGA TTA GAT AAA AGT AAA GTG-3′ and 5′-TAA TAA TGC GGC CGC TTA AGA CCC ACT TTC ACA TTT AAG-3′ using pLew90 [15] as template into *Leishmania* expression vector pF4X1.4hyg (Jena Bioscience) opened *Bgl*II × *Not*I. Plasmid pF4TR1.4hyg#1 was used to generate *L. tarentolae* **TR** strains overexpressing TET repressor. The Tet repressor was expressed without NLS since presence of NLS is not required for its ability to

regulate T7 Pol driven transcription in the Trypanosomatidae's nucleus.

Plasmids for stable integration into host β-tubulin locus and T7 RNA polymerase mediated inducible expression of reporter genes

pTUB_F4egfp1.8ble#TB1 (Fig. 2C) was constructed by PCR amplification of the *egfp-N1* gene (Clontech) using primers 5′-TTA TTA AGA TCT CGA CAC ACG AAG CGT GTG ATG-3′ and 5′-TTA TTA GGG CCC GGA CGA ACC GGA GCT TTG TGT G-3′, and pF4egfp1.8ble#7 plasmid [14] as a template. The resulting 4 kbp PCR fragment harbouring the *egfp-N1* gene operably linked to a bleomycin selection marker (*ble*) was digested with *Bgl*II × *Apa*I enzymes and inserted into pTUB11 opened with *Bgl*II and *Apa*I. The pUC19 based vector pTUB11 contained the 5′ and 3′ parts of *L. tarentolae* β-tubulin ORF flanking the T7 promoter/TET operator assembly. The latter was introduced as *Eco*RI × *Sph*I fragment: 5′-GAA TTC **TAA TAC GAC TCA CTA TAG GG**A GAT CGt ccc tat cag tga tag aga TCT ATT TAA ATA TCG ATG GGC CCC TGC AGG CAT GC -3′ (T7 promoter bold, TET operator small caps, *Eco*RI, *Bgl*II, *Apa*I, and *Sph*I sites underlined) (K. Gase, unpublished). The 5′ β-tubulin part was PCR amplified with primers 5′-GCG CCG TTA ATT AAG TTT TGG GAG GTG ATC TCC G-3′ and 5′-GCG CCG AAT TCG GAT CCC TGC AGC GTG CGG AAG CAA ATA TCG-3′, and cloned as 0.6 kbp *Pac*I × *Eco*RI fragment, whereas the 3′ part was amplified with primers 5′-GCG CCG AAT TCG CAT GCC GAC GTT CGG TGA CCT GAA CC-3′ and 5′-GCG CCG GTT TAA ACT CGG AGA CGA GGT CGT TC-3′, and cloned as 0.6 kbp *Sph*I × *Pme*I fragment. pTUB1 plasmid was used as a template for both amplification products. This plasmid was obtained by blunt end cloning of 1.2 kbp β-tubulin gene fragment into pUC19 linearized with *Sma*I. The β-tubulin gene was amplified from genomic DNA of *L. tarentolae* Parrot laboratory strain P10 with primers 5′-CCG GCC AGT GCG GCA ACC AGA TCG G-3′ and 5′-GGT GGC GTC CTG GTA CTG CTG G-3′ (K. Gase, unpublished). The latter primers were designed on the basis of sequences conserved between *Leishmania* species.

pTUB_APegfp1.4ble#TBP11 (Fig. 2C) was obtained as described for pTUB_F4egfp1.8ble#TB1, but using the primers 5′-TTA TTA AGA TCT AGT CGC AGC CTG ACC GCA TC-3′ and 5′-TA TTA GGG CCC GGA CGA ACC GGA GCT TTG TGT G-3′, and pAPdBegfp1.4ble#1 as the template (Breitling and Kushnir, unpublished). The latter plasmid was generated by substitution of 0.4 kbp calmodulin derived 5′ UTR of *egfp-N1* gene [14] by 0.4 kbp 5′ UTR of adenine phosphoribosyltransferase (*aprt*) gene (Accession No. AF060886, [16]) amplified with primers 5′-TAA TAA GTC GAC AGT CGC AGC CTG ACC GCA TC-3′ and 5′-TAA TAA AGA TCT GGT GGC AAG GCA CAG CAA GG-3′ from genomic DNA of *L. tarentolae* Parrot laboratory strain P10.

pTUB_APegfp1.4neo#802 (Fig. 2C) was created by substituting the 1.7 kbp *ble* marker gene containing *Not*I × *Spe*I fragment of pTUB_APegfp1.4ble#TBP11 by the 2 kbp *neo* marker gene containing fragment excised with the same enzymes from *Leishmania* expression vector pF4X1.4neo (Jena Bioscience).

pTUBneoF4egfp1.8rev#934 (Fig. 2D) was constructed by insertion of 2.8 kbp *Apa*I × *Sma*I trimmed PCR fragment harbouring the *neo* gene obtained with primers 5′-TAA TCC AGF GCC CGC CCT CCT CCT CCT TTC TTG TTC CTT TC-3′ and 5′-TAA TAA CCC GGG CGC TAG CCG GAC GAA CCG GAG CTT TGT G-3′ from pF4X1.4neo (Jena Bioscience) into *Apa*I × *Sma*I digested pTUB_F4egfp1.8ble derivative, carrying the *egfp* expression cassette in antisense orientation with respect to the tubulin and neo genes.

Transfection and cultivation of L. tarentolae

Leishmania tarentolae Parrot laboratory strain P10 was cultivated as static suspension cultures in Complex LEXSY Broth—Complete (Jena Bioscience) at 26 °C. Transfections were carried out by electroporation of in vitro cultivated promastigotes essentially as described [17] except that electroporation were performed directly in growth medium at cell densities of ca. 10^8 cells/ml omitting washing steps in electroporation buffer. Recombinant clones were selected as single colonies on solidified growth medium containing 10% FCS and 40 mM Hepes, pH 7.4. The selective antibiotics used were nourseothricin (LEXSY NTC, Jena Bioscience) 100 μg/ml, hygromycin (LEXSY Hygro, Jena Bioscience) 100 μg/ml, bleomycin (LEXSY Bleo, Jena Bioscience) 100 μg/ml and neomycin (LEXSY Neo, Jena Bioscience) 50 μg/ml. Recombinant cell lines were selected by the limited dilution method in liquid static suspension cultures with the indicated drug concentrations. For electroporation, ≈5 μg of the expression plasmids was digested either with *Swa*I (for stable integration into the *ssu* locus) or with *Pac*I and *Pme*I (for stable integration into the *tub* locus). Integration of the expression constructs into the *ssu* or *tub* locus was confirmed by PCR on genomic DNA with specific primers.

Total RNA isolation and Northern blot analysis

Total RNAs were isolated from 4×10^8 cells of 24–30 h cultures with Qiagen total RNA isolation kit. Northern blot analysis was performed with the DIG System (Roche Applied Science) according to instructions of the manufacturer with minor modifications. Briefly, a 0.5 kbp *egfp-N1* probe was labelled using the "PCR DIG Probe Synthesis Kit" (Roche Applied Science). RNA was separated in denaturing 1% agarose gel, blotted onto Hybond + membrane (Amersham) with Vacublot

(Bio-Rad), and UV cross-linked to the membrane using Fluo-Link cross-linker (MWG–BIOTECH) with the intensity of 0.12 J/cm^2. Pre-hybridization and hybridization were performed in Rothi-Hybri-Quick blocking reagent (Carl Roth) supplemented with 10 μg/ml herring sperm DNA at 68 °C for 4 and 22 h, respectively. Hybridized probes were detected using "DIG Luminescent Detection Kit" (Roche Applied Science) according to the instructions of the manufacturer.

EGFP expression analysis

Fluorimetric assays of EGFP expression were performed with 0.2 ml of normalized cell lysates (10 mM Tris, pH 8.0, 1% Triton) in 96-well plates using the fluorescent plate reader Fluoroscan Ascent FL (Labsystems) with the Ascent software 2.4 and excitation and emission wavelengths set to 485 and 510 nm, respectively. The relationship between the relative fluorescence units and the total amount of EGFP in cell cultures was established by measuring fluorescence of known amounts of EGFP protein purified to homogeneity from recombinant *L. tarentolae* cultures. For Western blot detection, the cell extracts of EGFP expressing strains were resolved on 15% SDS–PAGE gels and blotted onto nitrocellulose membranes. The membranes were probed with rabbit anti-GFP antibodies (Clontech) and the signals were visualized with alkaline phosphatase-coupled goat anti-rabbit antibodies (Sigma), and colorimetric detection with BCIP/NBT substrate (Roche Applied Science) according to the instructions of the manufacturer.

Results

Construction of L. tarentolae strains co-expressing T7 RNA polymerase and TET repressor proteins

To construct a *L. tarentolae* strain capable of regulated transcription of heterologous genes, we chose to use a combination of bacteriophage T7 RNA polymerase (T7 Pol) with its promoter controlled by a repressor element responsive to a cell permeable chemical compound (Fig. 1B). Previous work on *Trypanosoma* spp. and *Leishmania* spp. demonstrated that the tetracycline repressor (TET repressor) can be used to regulate the activity of T7 and Pol I promoters in a chromosomal context [15,18,19].

As the first step in the construction of an inducible *Leishmania* expression system we generated host strains constitutively expressing T7 RNA polymerase. To determine the level of T7 Pol expression required for optimal transcription of heterologous target genes, we created three expression plasmids (see Experimental procedures). These constructs were derived from earlier described vectors and can be integrated into the *ssu* locus transcribed by RNA polymerase I (Fig. 1A) [14]. The transgenic strain designated **HI_NLS_T7** overexpressed T7 RNA polymerase N-terminally tagged with a nuclear localization sequence (NLS) derived form SV40 T antigen. By quantitative Westen blot analysis we estimated the expression level to be in the range of 6×10^4 molecules/cell for this construct. The strain designated **HI_T7** overexpressed T7 Pol without NLS at levels similar to the former strain. The third strain, designated **LO_NLS_T7**, expressed approximately 10 times less NLS-tagged T7 Pol than the HI_T7strains (Fig. 2E). Reduced levels of T7 RNA polymerase expression resulted from flanking the *T7 pol* gene with heterologous UTRs derived from the cysteine proteinase gene *cys2* of *Leishmania pifanoi* and from intergenic region of *Leishmania donovani* glycosyltransferase gene LPG1, which result in significantly lower expression in *L. tarentolae* (Breitling and Klingner, unpublished).

In the next step, the gene encoding the TET repressor was introduced into another *ssu* copy of the rDNA locus of the above described T7 pol strains. Since there are many *ssu* copies in the rDNA locus each represented by two alleles, the T7 pol gene can be integrated into one, and the TET repressor gene into another, still leaving large number of intact *ssu* genes.

Western blot analysis of the resulting cell lines demonstrated that the TET repressor was expressed at levels close to 10^4 molecules/cell (Fig. 2F). The genomic co-integrations of TET repressor and T7 polymerase expression constructs were stably maintained over 100 generations (data not shown) and were used as recipients for target gene expression constructs.

Inducible overexpression of EGFP in L. tarentolae

As a test protein for our overexpression studies, we chose the EGFP, which can be readily detected in vivo and in vitro. We constructed two expression plasmids in which *egfp* and selection marker genes are operably coupled and flanked by segments of the β-tubulin gene (*tub*). The TET controlled T7 promoter was inserted at the 5′ end of the *egfp* and selection marker assembly (Fig. 2C). The EGFP expression cassettes were integrated into the tubulin gene cluster of *L. tarentolae* by homologous recombination via the *tub* segments on both sides of the constructs. Selection of clones bearing the reporter construct could be performed in the absence of tetracycline. This is likely to reflect a sufficient level of read-through transcription of the endogenous RNA polymerase II (see Discussion). Although, as a consequence, in this configuration some background expression of EGFP could not be avoided (Table 1, L1), it was reduced in the presence of TET repressor (Table 1, L2), indicating on one hand functional interaction of TET operator with TET repressor and on the other negative effect of TET repressor on blocking Pol II read-through (see Discussion).

Table 1
Summary of generated *L. tarentolae* strains and their fluorescence in suspension cultures in the presence or absence of tetracycline

Lane	*L. tarentolae* strain	T7 Pol expression	T7 Pol with NLS	TET repressor	EGFP (rfu)		EGFP expression cassettes
					−Tet	+Tet	
1	Wild type	No	—	No	13	12	F4egfp1.8ble
2	TR	No	—	Yes	2	2	F4egfp1.8ble
3	HI_NLS_T7	High level	Yes	No	22	24	F4egfp1.8ble
4	LOW_NLS_T7	Low level	Yes	No	61	71	F4egfp1.8ble
5	HI_T7-TR	High level	No	Yes	12	20	F4egfp1.8ble
6	LOW_NLS_T7-TR	Low level	Yes	Yes	4	80	F4egfp1.8ble
7	HI_NLS_T7-TR	High level	Yes	Yes	4	100	F4egfp1.8ble
8	HI_NLS_T7-TR	High level	Yes	Yes	5	150	APegfp1.4ble
9	HI_NLS_T7-TR	High level	Yes	Yes	8	200	APegfp1.4neo
10	HI_NLS_T7-TR	High level	Yes	Yes	21	400	F4egfp1.8ble, APegfp1.4neo
11	HI_NLS_T7-TR	High level	Yes	Yes	2	150	neoF4egfp1.8rev

The fluorescence of EGFP is expressed in relative fluorescent units (rfu) of normalized cell extracts.

Initially, we used all three cell lines with different localization signals and levels of expression of T7 RNA polymerase. The established suspension cultures were divided and 5 μg/ml tetracycline was added to the medium at OD_{600} 0.2–0.3. The results shown in Table 1, L6 and L7, and Fig. 3B indicate that incubation with tetracycline resulted in ca. 20-fold increase in EGFP fluorescence compared to the control cultures without Tet. Interestingly, the levels of EGFP overexpression were only marginally higher in the EGFP-HI_NLS_T7-TR cell line as compared to the EGFP-LO_NLS_T7-TR cell line, indicating that the T7 polymerase was not a limiting factor. The strain expressing T7 polymerase without NLS displayed almost no tetracycline-dependent EGFP expression (Table 1, L5, Fig. 3B). This was not surprising owing to the nuclear localization of the target gene. On the other hand, this finding confirmed the functionality of the SV40 derived NLS in *L. tarentolae*.

To ascertain that the observed induction effect was indeed due to the enhanced T7 mediated transcription of the target gene, we performed a Northern blot analysis of total RNA samples isolated from induced and uninduced strains. As can be seen in Fig. 3C, tetracycline addition led to a ca. 20-fold increase in the amount of EGFP mRNA, thus further confirming that T7 polymerase was active in the genomic context of *L. tarentolae.*

Subsequently, we determined the optimal concentration of tetracycline required for maximal expression of EGFP. All three cell lines with different localization signals and expression levels of T7 polymerase were included in this study. Concentrations of Tet were varied from 0 to 20 μg/ml and measurements were performed 18, 42, and 64 h post-induction (data not shown). We found that 1 μg/ml tetracycline was sufficient for maximal EGFP induction. Further increase did not enhance the expression of EGFP. Instead, we observed that concentrations of tetracycline above 20 μg/ml inhibited cell growth. Therefore, in our further experiments we used the working concentration of 10 μg/ml Tet.

Time resolved measurements of Tet-regulated EGFP expression demonstrated that after induction the EGFP fluorescence increased rapidly (Fig. 3B). Withdrawal of Tet from this culture after 30 h resulted in decrease of fluorescence as expected. However, even after 60 h it was only three times lower than in the presence of Tet. Among several possible explanations of this phenomenon the high stability of the EGFP protein appears most likely.

Optimization of pTUB vectors for tight expression control

Since regulation of gene expression in Trypanosomatidae occurs almost exclusively post-transcriptionally and to a major extent mediated by 3′ and 5′ UTRs that regulate RNA stability, we decided to test whether the developed expression vector could be further improved by UTR engineering. To this end, we compared the influence of two combinations of 5′ and 3′ UTRs flanking the *egpf* reporter gene on EGFP expression. The first combination used in plasmid F4egfp1.8 encompassed segments of the 1.8 kb intergenic region (IR) of the calmodulin (*cam*) gene cluster of *L. tarentolae* described in [14] and in the previous section. The second construct designated APegfp1.4 consists of 5′ UTR of adenine phosphoribosyl transferase gene (*aprt*) and 1.4 kb IR of *cam* operon of *L. tarentolae*. When both assemblies were integrated into the Pol I transcribed *ssu* locus of wild type *L. tarentolae* it resulted in comparable EGFP expression levels (Breitling and Klingner, unpublished). Interestingly, when expression of these gene assemblies was mediated by T7 polymerase a 2-fold effect was observed (Fig. 4): First, EGFP expression by Pol II readthrough in the transfected wild type strain was reduced by ca. 80% when using the APegfp1.4 construct. Second, upon induction EGFP expression was 50% higher as compared to the strain with F4egfp1.8 resulting in a 30-fold induction rate (Table 1, L7 and L8). The interpretation of these observations is provided in the Discussion.

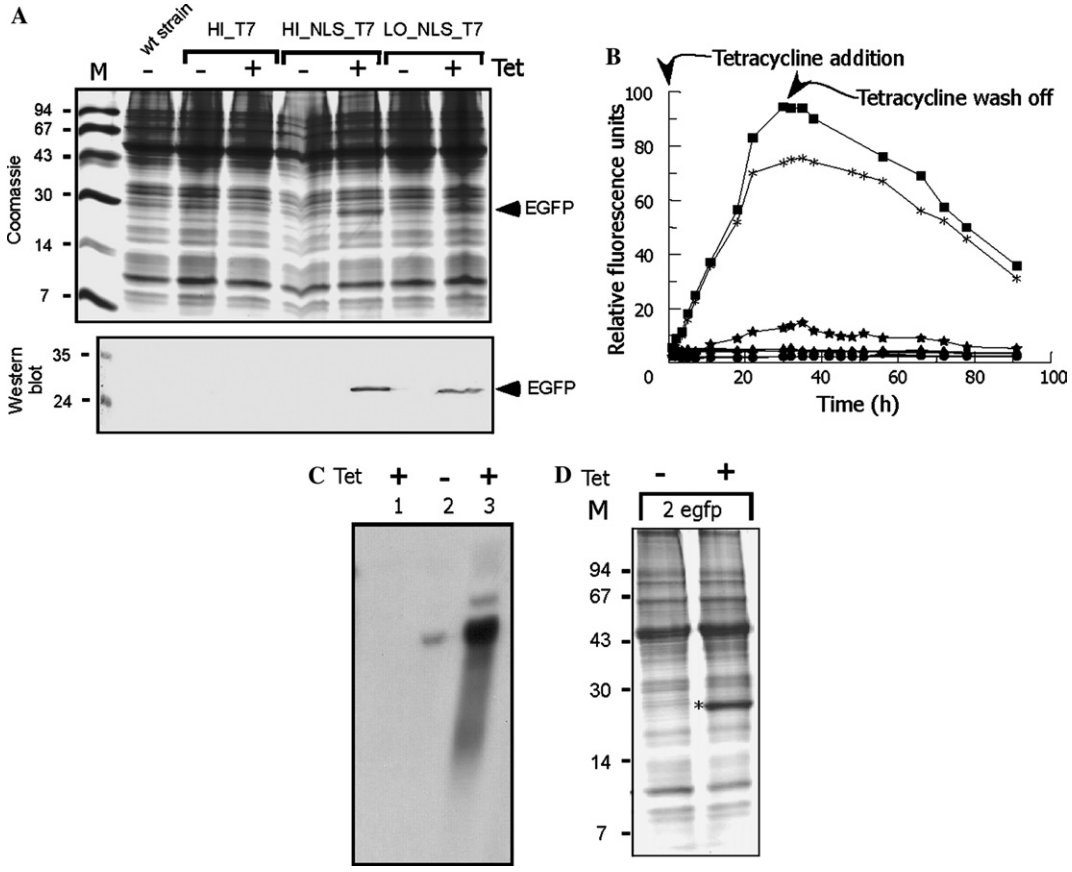

Fig. 3. T7 polymerase mediated expression of EGFP reporter protein in *L. tarentolae*. (A) Coomassie stained SDS–PAGE and Western blot analysis of induced and uninduced cultures of wild type and transgenic *L. tarentolae* strain bearing 1 copy of *egfp* gene. M, molecular weight marker. The position of EGFP protein on the gel and blot is indicated by an arrow. (B) Time course of EGFP production by *L. tarentolae* bearing different variants of T7 polymerase. EGFP expression was induced by addition of 10 μg/ml of tetracycline to cultures at OD_{600} 0.1 and cells were cultivated for the indicated period of time. The fluorescence was recorded as described in Experimental procedures. After 30 h, cells were washed and transferred into medium without tetracycline. (■) induced EGFP-HI_NLS_T7-TR; (▲) uninduced EGFP-HI_NLS_T7-TR; (∗) induced EGFP-LOW_NLS_T7-TR; (◆) uninduced EGFP-LOW_NLS_T7-TR; (★) induced EGFP-HI_T7_TR; and (●) uninduced EGFP-HI_T7_TR. (C) Northern blot analysis of total RNA extracted from induced and uninduced cultures of HI_NLS_T7-TR strains 30 h after induction with tetracycline. 1, total RNA of HI_NLS_T7-TR line without the reporter cassette in the presence of tetracycline; 2, EGFP-HI_NLS_T7-TR strain in the absence of tetracycline; and 3, induced EGFP-HI_NLS_T7-TR strain. (D) SDS–PAGE analysis of induced and uninduced cultures of *L. tarentolae* 2EGFP-HI_NLS_T7-TR strain bearing two copies of *egfp* gene. M, molecular weight marker.

A second strategy to achieve tight repression of the reporter gene was to change the gene organization of the expression constructs. We initially attempted to reduce the background expression by inserting the T7 promoter/operator-EGFP-marker cassette into the *tub* locus in antisense orientation relative to *tub* gene. This, however, led to the failure to obtain transformants unless tetracycline was present in the medium during clone selection, indicating that Pol II read-through and T7 leakage (if any) were insufficient to produce adequate amounts of antibiotic resistance protein. To circumvent this problem we created an expression vector with antibiotic resistance marker positioned in sense and the T7 controlled transcription unit in antisense orientation relative to *tub* gene (Fig. 2D). Integration of this construct into the HI_NLS_T7-TR strain yielded comparable expression levels as obtained in sense orientation and improved 75-fold induction/repression level ratio (Table 1, L11).

Construction of L. tarentolae strains inducibly co-expressing two EGFP genes

Recombinant overexpression of protein complexes is one of the particularly difficult issues in protein production. Since the number of possible genomic integrations in *Leishmania* species is limited only by the number of available selection markers, we attempted co-expression of two independently transcribed genes. To this end, we substituted the *ble* selection marker of EGFP expression plasmid pTUB_F4egfp1.8ble (Fig. 2C) with the *neo* resistance gene. The resulting construct pTUB_APegfp1.4neo was transfected into HI_NLS_T7-TR cells containing integrated *ble*-linked copy of the EGFP expression cassette. The selected cell line 2EGFP-HI_NLS_T7-TR displayed growth rates and viability similar to the parental strain. In parallel, the *neo*-linked EGFP expression cassette was introduced into the

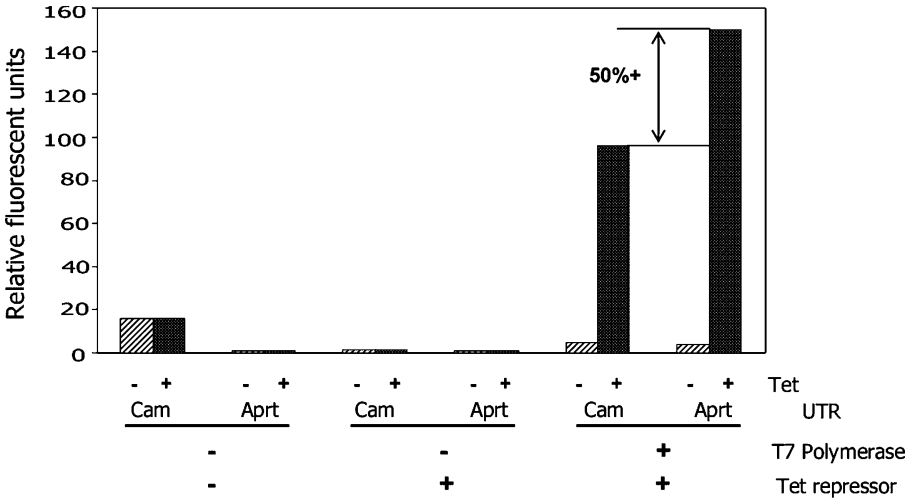

Fig. 4. Effect of 5′UTR-egfp-3′ UTR combinations on the levels of induced and background expression in the T7 based expression constructs. The F4egfp1.8 and APegfp1.4 expression cassettes bearing *L. tarentolae* UTRs of calmodulin and adenine phosphoribosyltransferase genes were transformed into *L. tarentolae* HI_NLS_T7-TR strain or control strains and analysed for expression of EGFP in the presence and absence of tetracycline. Cam, F4egfp1.8; and Aprt, APegfp1.4 expression cassette.

HI_NLS_T7-TR strain. As expected, addition of tetracycline to the medium led to a time dependent increase of EGFP fluorescence in both cell lines. The cell line containing the *neo*-coupled APegfp1.4 integration yielded a relative fluorescence intensity twice as much as the cells harbouring the *ble*-coupled F4egfp1.8 cassette (Table 1, L7 and L9). These data are consistent with our previous findings that the type of the antibiotic resistance marker used can affect the levels of the coupled reporter gene expression (Breitling and Ehrlich, unpublished).

Somewhat surprisingly, the fluorescence levels in the strain containing two copies of EGFP consistently reached 400 rfu. This is 100 rfu more than the arithmetic sum of the expression levels of the strains containing one copy of each configuration (Table 1, L7, and L9 and L10). To confirm that the increase of fluorescence did indeed reflect overexpression of the EGFP protein, we analysed the lysates of induced and uninduced cultures of this strain by SDS–PAGE and Western blotting. As shown in Fig. 3D, tetracycline induction led to emergence of a prominent band at ca. 28 kDa which was confirmed to be EGFP by Western blotting. We estimated that cell line 2EGFP-HI_NLS_T7-TR overexpressed EGFP to the level of ca. 4% of total cellular protein or ca. 40 mg/L of suspension culture.

Discussion

Here, we present a novel and biotechnologically useful inducible eukaryotic expression system based on the non-pathogenic Trypanosomatidae *L. tarentolae*. We were able to establish stable cell lines of *L. tarentolae* constitutively co-expressing the TET repressor and a nuclear-targeted T7 polymerase. When the developed strains were stably transformed with expression constructs in which the heterologous gene was placed under control of the TET controlled T7 promoter, tetracycline-dependent overexpression of heterologous target protein was observed. Using EGFP as a reporter protein we demonstrated that in this architecture, expression yields of ca. 1–2% of total cellular protein or 10–20 mg/L of suspension culture could be achieved with one copy of the expression cassette. These yields are comparable with the previously described ones obtained in a constitutive expression system based on the integration in *ssu* [14]. However, inducibility of the system provides an advantage for expression of toxic proteins.

The system demonstrated inducibility although some background expression was observed. This background expression was a direct consequence of the chosen architecture where target gene and selection marker were operably coupled and co-transcribed. It could be explained either by read-through transcription of the host RNA polymerase II or by random initiation of transcription within the 5′ UTR of the target gene by Pol II or by some leakiness of the T7 promoter/operator. There are two observations arguing against the latter two possibilities. First, positioning of additional copies of TET operator upstream of the T7 promoter or duplication of T7 promoter/operator unit in front of the 5′ UTR of the expression cassette abolished viability of the uninduced strain in the presence of antibiotic selection in contrast to single operator constructs (data not shown). Most probably this implies that in these configurations Pol II read-through was reduced to a level insufficient for adequate marker expression. If random transcription initiation within the 5′ UTR of the target gene or leakiness of the T7 promoter/operator would have accounted for the observed background these strains should have

expressed the drug resistance marker in the absence of Pol II read-through. Moreover, reporter construct could be integrated also in the host strain lacking T7 polymerase where presence of TET repressor reduced EGFP background significantly (Table 1, L1 and L2). Therefore, read-through transcription by endogenous Pol II appears to account for most of the observed background expression of the reporter gene.

We could reduce this background expression and enhance the level of EGFP expression and induction level by utilizing different combinations of 5′ and 3′ UTRs flanking the reporter gene (Fig. 4). The observed effects are likely of complex nature and may reflect on one hand reduction in Pol II read-through while on the other represent effects of the UTRs on mRNA stability and/or translation initiation. Comparing the APegfp1.4 and F4egfp1.8 constructs, we observed an opposite effect in the wild type strain vs. the T7 Pol strains. Whereas in the wild type strain the EGFP expression was reduced with the APegfp1.4 construct it was elevated in the induced T7 Pol strain. This may lead to the speculation that the RNAs transcribed by host Pol II and heterologous T7 Pol may have different stabilities or other relevant features leading to the observed effects. Alternatively, the structure of generated mRNA might also affect the transcription efficiency of T7 RNA polymerase.

In an alternative effort to reduce background expression of the target protein, we tested an antisense organization of the expression cassette. Placing of the T7 controlled transcription unit in the antisense orientation with respect to the tubulin and neo genes resulted in the lowest background and the highest induction rate compared with all other constructs. It must be noted, however, that even in this configuration some background expression of EGFP could be detected reflecting low level Pol II transcription on the antisense strand. It may be possible to create a "zero" background system by integrating the T7 controlled transcription unit into the non-transcribed ribosomal spacer [15].

To test whether the developed expression system is suitable for co-expression of multiple genes, we constructed *L. tarentolae* strains containing two copies of the *egfp* gene independently integrated into the genome under the control of the T7 promoter. Surprisingly, exposure of these cells to tetracycline resulted in overexpression of EGFP to a level significantly larger than the sum of the expression values of the individual strains containing a single copy integration. Although the nature of this phenomenon is unclear at present, we speculate that this might be due to the sub-optimal TET responsive element:TET repressor ratio. Since the TET responsive element is present in the cell as a single copy, it is expected that it will be greatly overtitrated by TET repressor present at ca. 10,000 copies per cell. According to our estimates, TET repressor is present in the cells at a concentration in the range of 100 nM. Addition of tetracycline should reduce its affinity for TET operator from 100 pM to at least 100 μM [20]. However, it is possible that the local concentration of TET repressor is much higher than that calculated here. This might lead to incomplete de-repression of the T7 promoter even in the presence of tetracycline. The data summarized in Fig. 4 suggest that the latter possibility is more likely, since strains producing the TET repressor and no T7 polymerase express EGFP only to 10% of the level observed in the wild type cells both in the absence or in the presence of tetracycline. This may indicate that a fraction of TET repressor pool remains bound to its recognition sequence even at high concentrations of tetracycline.

It appears that further increase in protein overexpression yields can be achieved by increasing the copy number of the template. This can be provided by various approaches such as, for instance, by increasing the number of genomically integrated heterologous gene copies or by placing the gene of interest on a multi-copy circular or linear episome.

Acknowledgments

We thank David Campbell for advice and stimulating discussions and Steven M. Beverley and George Cross for generous gift of plasmids and advice. Vacheslav Yurchenko is acknowledged for intellectual contribution at the initial stage of this project. Paul-André Genest and Roger S. Goody are acknowledged for critically reading the manuscript. We thank Sarah Müller, Katja Merzliak, Thomas Hahn, and Ion Cirstea for help in constructing the expression plasmids. This work was in part supported by Grant B309-98020 from the Thuringian Ministry of Sciences, Research and Culture to Jena Bioscience GmbH and by DFG Grant AL 484/5-3 to K.A.

References

[1] R.F. Service, Structural genomics. Tapping DNA for structures produces a trickle, Science 298 (2002) 948–950.
[2] A.M. Edwards, C.H. Arrowsmith, D. Christendat, A. Dharamsi, J.D. Friesen, J.F. Greenblatt, M. Vedadi, Protein production: feeding the crystallographers and NMR spectroscopists, Nat. Struct. Biol. 7 (Suppl.) (2000) 970–972.
[3] K. Dower, M. Rosbash, T7 RNA polymerase-directed transcripts are processed in yeast and link 3′ end formation to mRNA nuclear export, RNA 8 (2002) 686–697.
[4] S.R. Thompson, K.D. Gulyas, P. Sarnow, Internal initiation in *Saccharomyces cerevisiae* mediated by an initiator tRNA/eIF2-independent internal ribosome entry site element, Proc. Natl. Acad. Sci. USA 98 (2001) 12972–12977.
[5] S.M. Teixeira, Control of gene expression in Trypanosomatidae, Braz. J. Med. Biol. Res. 31 (1998) 1503–1516.
[6] S. Biebinger, L.E. Wirtz, P. Lorenz, C. Clayton, Vectors for inducible expression of toxic gene products in bloodstream and procyclic *Trypanosoma brucei*, Mol. Biochem. Parasitol. 85 (1997) 99–112.

[7] C.E. Clayton, Genetic manipulation of kinetoplastida, Parasitol. Today 15 (1999) 372–378.
[8] E. Wirtz, C. Hartmann, C. Clayton, Gene expression mediated by bacteriophage T3 and T7 RNA polymerases in transgenic trypanosomes, Nucleic Acids Res. 22 (1994) 3887–3894.
[9] J.H. LeBowitz, H.Q. Smith, L. Rusche, S.M. Beverley, Coupling of poly(A) site selection and trans-splicing in *Leishmania*, Genes Dev. 7 (1993) 996–1007.
[10] C. Dumas, A. Muyombwe, G. Roy, C. Matte, M. Ouellette, M. Olivier, B. Papadopoulou, Recombinant *Leishmania major* secreting biologically active granulocyte–macrophage colony-stimulating factor survives poorly in macrophages in vitro and delays disease development in mice, Infect. Immun. 71 (2003) 6499–6509.
[11] A. Furger, T.W. Jungi, J.Y. Salomone, V. Weynants, I. Roditi, Stable expression of biologically active recombinant bovine interleukin-4 in *Trypanosoma brucei*, FEBS Lett. 508 (2001) 90–94.
[12] W.W. Zhang, H. Charest, G. Matlashewski, The expression of biologically active human p53 in *Leishmania* cells: a novel eukaryotic system to produce recombinant proteins, Nucleic Acids Res. 23 (1995) 4073–4080.
[13] A. Misslitz, J.C. Mottram, P. Overath, T. Aebischer, Targeted integration into a rRNA locus results in uniform and high level expression of transgenes in *Leishmania amastigotes*, Mol. Biochem. Parasitol. 107 (2000) 251–261.
[14] R. Breitling, S. Klingner, N. Callewaert, R. Pietrucha, A. Geyer, G. Ehrlich, R. Hartung, A. Muller, R. Contreras, S.M. Beverley, K. Alexandrov, Non-pathogenic trypanosomatid protozoa as a platform for protein research and production, Protein Expr. Purif. 25 (2002) 209–218.
[15] E. Wirtz, S. Leal, C. Ochatt, G.A. Cross, A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*, Mol. Biochem. Parasitol. 99 (1999) 89–101.
[16] O.H. Thiemann, J.D. Alfonzo, L. Simpson, Cloning and characterization of *Leishmania tarentolae* adenine phosphoribosyltransferase, Mol. Biochem. Parasitol. 95 (1998) 141–146.
[17] S.M. Beverley, C.E. Clayton, Transfection of *Leishmania* and *Trypanosoma brucei* by electroporation, Methods Mol. Biol. 21 (1993) 333–348.
[18] E. Wirtz, C. Clayton, Inducible gene expression in trypanosomes mediated by a prokaryotic repressor, Science 268 (1995) 1179–1183.
[19] S. Yan, P.J. Myler, K. Stuart, Tetracycline regulated gene expression in *Leishmania donovani*, Mol. Biochem. Parasitol. 112 (2001) 61–69.
[20] C. Berens, W. Hillen, Gene regulation by tetracyclines. Constraints of resistance regulation in bacteria shape TetR for application in eukaryotes, Eur. J. Biochem. 270 (15) (2003) 3109–3121.