Molecular & Biochemical Parasitology 176 (2011) 69–79

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

Artificial linear episome-based protein expression system for protozoon *Leishmania tarentolae*

Susanna Kushnir[a], Ion Cristian Cirstea[a,d,1], Lyudmyla Basiliya[d], Nataliya Lupilova[a], Reinhard Breitling[b], Kirill Alexandrov[c,*]

[a] Department of Chemical Biology Max-Planck-Institute for Molecular Physiology, Otto-Hahn-Strasse 11, 44227 Dortmund, Germany
[b] Jena Bioscience GmbH, Löbstedter Str. 80, 07749 Jena, Germany
[c] Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia
[d] Department of Physical Biochemistry Max-Planck-Institute for Molecular Physiology, Otto-Hahn-Strasse 11, 44227 Dortmund, Germany

ARTICLE INFO

Article history:
Received 12 November 2010
Received in revised form 1 December 2010
Accepted 3 December 2010
Available online 15 December 2010

Keywords:
Leishmania
Kinetoplastida
Episomal elements
Gene expression
Recombinant protein expression

ABSTRACT

The trypanosomatid protozoon *Leishmania tarentolae* is a well-established model organism for studying causative agents of several tropical diseases that was more recently developed as a host for recombinant protein production. Although several expression architectures based on foreign RNA polymerases have been established for this organism, all of them rely on integration of the expression cassette into the genome. Here, we exploit a new type of expression architecture based on linear elements. These expression vectors were propagated in *Escherichia coli* as circular plasmids and converted into linear episomes with telomere-like structures prior to transfection of *L. tarentolae*. Overexpression of recombinant proteins in transgenic organisms exceeding 10% of total cellular protein, one of the highest overexpression levels obtained in a eukaryotic organism for a cytosolic protein. We show that the linear elements are stably propagated in *L. tarentolae* cells over long periods of time (>90 generations) without major changes in structure or expression yields. Overexpressing cultures can be obtained without clonal selection of the transfected cells. To establish the utility of the developed system for protein production in a parallelized format, we expressed 37 cytosolic, peripheral, and membrane proteins as fusions with EGFP in *L. tarentolae* using linear vectors. We detected the expression of 30 of these targets and describe the preparative purification of two arbitrarily selected proteins.

1. Introduction

The family *Trypanosomatidae* (*Euglenozoa*, *Kinetoplastida*), which includes the genera *Leishmania* and *Trypanosoma*, is one of the oldest groups of eukaryotes and has attracted researchers' attention for two interconnected reasons. Firstly, they are notorious for finding unique solutions to general processes of the eukaryotic cell. Among them 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 (for review, see [1]). Secondly, a number of *Leishmania* and *Trypanosoma* species are extra- and intracellular parasites of humans and livestock and cause a range of debilitating or fatal diseases. Due to its public health significance, *Trypanosomatidae* is one of the best-studied eukaryotic groups after yeast. The development of methods of trypanosomatid cultivation and genetic manipulation has allowed the dissection of the mechanisms of non-conventional gene expression [2]. Trypanosomatids can be efficiently transfected by various types of genetic vectors [2]. Expression cassettes can be maintained episomaly or stably integrated into different parts of the genome, and their expression can be driven by read-through of endogenous RNA polymerase I or II. Alternatively, expression of genes can be driven by co-integrated promoters for RNA polymerase I or heterologous RNA polymerases, such as T3 and T7 [3–6]. We and others exploited these expression architectures to produce recombinant proteins in *Leishmania* and *Trypanosoma* species [5,7–9]. Most of the efforts in developing trypanosomatid organisms as protein expression hosts were focused on *Leishmania tarentolae* due to its rapid growth on simple media and the availability of tools for genetic manipulation [10]. Two main expression architectures have been employed to achieve protein expression in *L. tarentolae*. The constitutive expression system utilizes RNA polymerase I-mediated transcription in the ribosomal RNA locus into which an expression cassette is inserted [11]. This system is simple but not suitable for the expression of proteins that are toxic to the host organism. The alternative expression

* Corresponding author. Tel.: +61 7 334 62017; fax: +61 7 334 62101.
E-mail address: k.alexandrov@uq.edu.au (K. Alexandrov).
[1] Present address: Institute of Biochemistry and Molecular Biology II, Universitaetstr. 1, 40225 Dusseldorf, Germany.

0166-6851/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2010.12.002

architecture relies on the use of a transgenic strain of *L. tarentolae* that expresses T7 RNA polymerase and the TET repressor. The tetracycline-regulated expression cassette is integrated into the genome, and expression of the target gene is induced by the addition of tetracycline to the culture medium. In addition to regulated expression, this system demonstrated significantly higher expression yields than the constitutive system [6]. A shared limitation of both expression architectures is the limited number of template gene copies per *Leishmania* genome. Although the template copy number can be increased by integration of several cassettes using alternative antibiotic resistance selection markers, the largest number of integrated gene copies for the constitutive expression system has been 4 [11] and 2 for the inducible system [6]. Recently developed strategies for the recycling of selection markers has enabled further increases in template copy number, but these procedures are slow and laborious [12,13]. In the present work, we resolved the limitations of the inducible *L. tarentolae* system by constructing linear and circular extrachromasomal multicopy expression vectors. However, the mechanisms of episomal maintenance and propagation appear to be unique in each case. With both architectures, expression levels up to 300 mg/l of culture were obtained. We demonstrate that both architectures can be used for preparative production of recombinant proteins.

2. Materials and methods

2.1. L. tarentolae culture and transfection

L. tarentolae laboratory strain T7-TR (Jena Bioscience) was cultivated as a static suspension culture in LEXSY BHI Medium (Jena Bioscience) at 26 °C according to the instructions of the manufacturer. For transfection, cultures were grown by re-inoculation of the strain at OD_{578} of 2–3 into fresh medium at 1:10 dilution. After two to three successive passages, 0.4 ml of a suspension culture of 10^8 cells/ml was transferred to a 2 mm gap electroporation cuvette (BioRad) and kept on ice for 10 min. Subsequently, ca. 5 μg DNA in 10 mM Tris, pH 8.0 were added into the cuvette and mixed well. Electroporation was performed using a BioRad Gene Pulser Xcell™ set to 450 V and 3.5 ms (time constant protocol). After electroporation, the cuvettes were placed back on ice for 10 min and subsequently the cells were transferred into ventilated flasks with pre-warmed BHI medium at 26 °C and cultivated overnight without selection. Some transfections were performed using Amaxa electroporator (Lonza) with programs X024 and U033 in BHI and programs X001 and U033 in buffers P1 or P2.

2.2. Vector construction

Construction of the artificial episome pLAC-EGFP based on the plasmid pUC18 F4egfp1.4neo (Kushnir, unpublished).

First, the plasmids containing eight telomeric repeats CCCTAA on each side of the spacer pUC18 Tel-left F4egfp1.4neo and pUC18 Tel-right F4egfp1.4neo were constructed.

pUC Tel-left F4egfp1.4neo was obtained by re-insertion of a 2.4 kb fragment amplified from pUC18 F4egfp1.4neo with primers 5′-AA TTA TTA CAT GTA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA CCC TAT GTG AGC AAA AGG CCA GCA AAA G-3′ and 5′-TGA ACT TGT GGC CGT TTA C-3′ into the template plasmid following restriction of insert and plasmid with PciI (NEB) and HindIII (MBI Fermentas). pUC Tel-right F4egfp1.4neo was constructed by re-insertion of a 1.3 kb fragment amplified from plasmid pUC18 F4egfp1.4neo with primers 5′-AATTATTACATGTA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA CCC-TAA CCC TAT CCC CTG ATT CTG TGG ATA AC-3′ and 5′-TAA TAA GGA TCC AAT ATG GGA TCG GCC ATT G-3′ into the template plasmid following restriction of insert and plasmid with PciI and SpeI (NEB) into plasmid pUC18 F4egfp1.4neo digested with the same enzymes.

The second step was the construction of pUC TeLeft VP1 F4egfp1.4neo. This plasmid was created by insertion of a 0.8 kb PciI and NcoI trimmed PCR product amplified from plasmid pF4 VP1 1.4sat with primers 5′-TAA TAA CCA TGG CTC GAG AAG AAA GTG GCA AAT GGT GG-3′ and 5′-TAA TAA ACA TGT CTC GAG TGG CAT TGT GGC AGC TAC TA-3′ into the plasmid pUC TeLeft F4egfp1.4neo pre-cut with the same enzymes.

In the last step, pLAC-EGFP was obtained by insertion of a 1.322 kb fragment from plasmid pUC Tel-right F4 egfp1.4neo digested with PciI × SpeI into the plasmid pUC Tel-left VP1 F4egfp1.4neo cut with PciI × SpeI.

The alternative variant of the episomal expression vector pLAC-Cherry was constructed with the gene coding for Cherry fluorescent protein as reporter gene and the bleomycin resistance marker gene. This was achieved by inserting a PCR amplified 0.7 kb *cherry* gene from plasmid pUC cherry ΔNcoI (Terbek, unpublished), cut with NcoI × NotI into the plasmid p LAC-EGFP digested with NcoI × NotI. In the second step, a 1.8 kb NotI × SpeI fragment carrying the *ble* resistance marker gene from plasmid pTUB AP VP1 ble (Granovsky, unpublished) was ligated into plasmid pLAC Cherry digested with NotI × SpeI.

2.3. L. tarentolae plating and clone selection

The agar plates were prepared as described in Supplementary experimental procedures. The agar plate was covered with 0.45 μm nitrocellulose membrane (Whatman) and the cells were plated directly onto the membrane. Recombinant colonies appeared after 5–7 days incubation at 26 °C. In order to induce protein expression, new plates containing solidified medium with 5 μg/ml tetracycline (TET) as inducer were prepared and the nitrocellulose membrane with grown clones was transferred onto these plates. After 24–30 h plates were illuminated with unfiltered UV light and the fluorescence was detected with a 530 nm emission filter for EGFP and a 610 nm emission filter for Cherry- and dsRED expressing clones.

2.4. Selection of clones on solid agar-free media

To plate transformants on filters we placed a ring of 1.5 mm thick Whatman filter paper (Rotilabo®, Germany) into the Petri dish and wetted it with liquid LEXSY BHI medium (Jena Bioscience) supplemented with 1% glucose, 5% fetal bovine serum, 100 mM PIPES, pH 7.4, and 1% of Pen Strep solution (Invitrogen) as shown in the Fig. S1A. Then the paper was covered with 0.45 μM nitrocellulose membrane (Whatman) and the cells were spread using a glass spreader. For induction the filter was transferred onto a new filter pad soaked in medium containing 5 μg/ml of tetracycline and incubated for 24–30 h at 27 °C.

2.5. FACS analysis

For flow cytometry analysis the cells were washed with PBS and fixed with 6% formamide in PBS for 10 min at room temperature. Subsequently cells were washed and resuspended in PBS. Analysis was performed on LRSII flowcytometr (BD) according to manufacturer's manual. The results were analyzed using DIVA or FlowJo software packages.

2.6. Pulsed-field electrophoresis

For plug preparation 2×10^7 cells/ml were resuspended in ice-cold ι buffer (0.1 M EDTA, pH 8.0, 0.01 Tris–HCl, pH 7.6, 0.02 M NaCl). The suspension was mixed with 2% of low melting agarose

(Sigma) in L-buffer in 1:1 ratio. The plugs were solidified for 15 min at room temperature and for 25 min at +4 °C. Plugs were incubated in three block volumes of L-buffer with proteinase K (0.1 mg/ml) and 1% Sarkosyl for 3 h at 50 °C. The buffer was subsequently exchanged and the samples were incubated for 16 h at 50 °C. The plugs were washed 3–5 times for 40 min in 50 volumes of TE-buffer (pH 7.6) at room temperature. The plugs were then incubated in 2 volumes of TE buffer with 40 μg/ml of PMSF for 30 min at 50 °C. Finally, the plugs were washed in 50 volumes of TE 3–5 times and stored in 10× TE at +4 °C.

For electrophoresis 1% agarose gel was prepared with 0.5× TBE buffer (45 mM Tris borate, 1 mM EDTA). The plugs were equilibrated for 15 min in 0.5× TBE buffer before loading. DNA plugs were embedded in individual wells of the gel and sealed with molten agarose. Pulsed-field electrophoresis was run in 0.5× TBE buffer at 14–16 °C.

Depending on the length of analyzed sequences the following parameters were used: 200–2000 kb – 4.6 V, ramp – 60–120 s, run length – 45 h; 500–1200 – 6 V, ramp – 5–120 s, run length – 41 h; 100–1000 kb – 6 V, ramp – 1–25 s, run length – 15–18 h.

The DNA was transferred onto nitrocellulose membranes using BioRad Vacu-blot system according to the manufacturer's instructions. The transfer was carried out in 10× SSC buffer for 90 min at 5 in. Hg negative pressure. Before blotting DNA was depurinated in 0.25 N HCl and denaturated in 0.5 N NaOH.

2.7. Southern blotting

The hybridization was carried out using DIG-labelled probes generated by a PCR-based labelling protocol. The labelling PCRs contained 10× dNTP Labelling Mixture (Roche) and BioTherm Taq DNA Polymerase (Genecraft). PCR program: 3 min – 93 °C, 30 cycles – 10 s 92 °C, 15 s 60 °C, 2 min 72 °C, 7 min – 72 °C. The sequence of the primers used for labelling egfp and Cherry was: 5′-TCTAACTCGAGGACGTAAACGGCCACAAGTTC-3′ and 5′-TCAATACTCGAGCGTCCATGCCGAGAGTGATC-3′.

For amplification of the dsRED probe the folloing primers were used: 5′-TAATAAGCTCGAGCGATGAGGTCTTCCAAGAATGTTATCAAGG-3′ and 5′-taataaGgcggccGCTTTAAAGGAACAGATGGTGGCG-3′.

50 μl of unpurifyed PCR was denaturated for 5 min at 95 °C and used for hybridization. Before hybridization the membrane was pre-hybridized for 3 h at 68 °C in Rothi-Hybri-Quick solution (Roth, Germany) with 10 μg/ml denaturated salmon sperm DNA. Hybridization with the probe was carried out for 16–20 h at 68 °C. Following hybridization the membrane was washed two times in 2× SSC solution with 0.1% SDS at room temperature and two times in 0.1× SSC + 0.1% SDS 68 °C. The hybridization products were detected using anti-DIG-AP antibody (Roche) at 1:5000 dilution. The membranes were washed 3× for 15 min with maleic buffer (0.1 M Maleic acid, 0.15 M NaCl, 0.3% Tween20) and developed with SCPD solution from Roche. Membrane was exposed to X-ray film (CL-X Posure Film from Pierce, prod.# 34093).

2.8. Purification of Rab7 and SOD expressed in L. tarentolae T7-TR strain using circular pLAC vector

The cell pellets from 2 l of culture were resuspended in Buffer A (50 mM Na phosphate pH 7.5, 100 mM NaCl, 10 μM $CuSO_4$, and 10 μM $ZnSO_4$) and desintegrated with fluidizer (Microfluidics).

The resulting homogenate was clarified by centrifugation (142,000 × *g*, 1 h at 4 °C) and filtered through a 0.45 μm nitrocellulose filter (Schleicher and Schuell). The filtrate was loaded onto a 5 ml Hi-Trap Ni Sepharose column (Pharmacia) pre-equilibrated with the lysis buffer. The column was washed extensively with lysis buffer containing 5 mM imidazol and the bound protein was eluted with a linear 5–200 mM imidazol gradient in 30 column volumes. Eluates were analyzed by SDS-PAGE followed by Coomassie blue staining. Fractions containing EGFP fusion proteins were pooled and TEV protease was added at a 1:40 molar ratio and the sample was dialyzed against 25 mM Na_2HPO_4, pH 8.0, 1 mM EDTA and 2 mM for 12 h at 4 °C. The progression of the cleavage reaction was determined by SDS-PAGE and, typically, at the end of the incubation over 90% of fusion protein was cleaved. In order to remove the 6his-EGFP assembly, uncleaved protein, TEV protease and impurities, $MgCl_2$ was added to the sample to 2 mM, imidazol to 25 mM and the protein was passed twice over a 5 ml Hi-Trap Ni–NTA column (Pharmacia). The flow-through of the column, containing target protein, was dialyzed against 25 mM HEPES, pH 7.2, 25 mM NaCl, and 5 mM DTT, concentrated using Centripreps 30 (Amicon) and stored in multiple aliquots at −80 °C. For higher purification eluates of the second Ni–NTA column were subjected to size exclusion chromatography on Superdex 75 16/60 column equilibrated with 25 mM HEPES, pH 7.2, 25 mM NaCl, and 5 mM DTT. Sample was concentrated to ca. 10 mg/ml and stored in multiple aliquots at −80 °C.

3. Results and discussion

3.1. Construction of a shuttle Leishmania artificial chromosome (LAC)

Leishmania sp. apparently does not have a defined origin of replication and can maintain circular and linear extrachromosomal elements of both endogenous and exogenous origin [14,15]. Numerous studies demonstrated that RNA polymerase II transcription initiates and persists spontaneously on circular plasmids and can drive expression of antibiotic resistance markers and passenger genes [16,17]. The presence of continuous PolII transcription initiation on both strands may be a reason for the failure to develop a high-yield, tightly regulated expression system with the target gene placed on an episomal element(s) under the control of a promoter for a strong foreign polymerase. An alternative explanation may be the high levels of concatenation and supercoiling that disfavor productive transcription by foreign polymerases. We conjectured that one possible way to resolve both problems would be to create an artificial linear episome carrying the expression cassette. To this end, we developed a circular *E. coli* shuttle vector that could be easily converted into a linear *Leishmania* artificial chromosome (LAC).

We reasoned that the presence of telomeric repeats at the ends of the vector would protect the constructs from lateral fusion, degradation, and circularization in the *Leishmania* cells. For long-term maintenance of the LACs, we sought to test whether the randomly initiated Pol II transcription would be sufficiently strong to transcribe the selection marker gene and enable amplification and maintenance of such linear elements in *L. tarentolae*. To this end, we designed a series of shuttle vectors based on the pUC backbone, which includes a T7 promoter-TET operator-driven expression cassette composed of the gene of interest, operably coupled to a selection marker flanked by the T7 terminator (Fig. 1 A–C). Two telomere-like regions were introduced into the plasmid backbone, each containing 9 TAACCCT repeats. To improve the stability of the vector during propagation in *E. coli*, these telomere-like sequences were separated by a 1-kbp spacer flanked by XhoI sites. The resulting vector can be propagated in *E. coli* and can be converted into a linear LAC carrying telomere-like structures at either end by digestion with XhoI (Fig. 1A and B). In an alternative configuration, the marker was located upstream of the T7 promoter in order to uncouple transcription of the antibiotic resistance gene from T7 promoter activity (Fig. 1C).

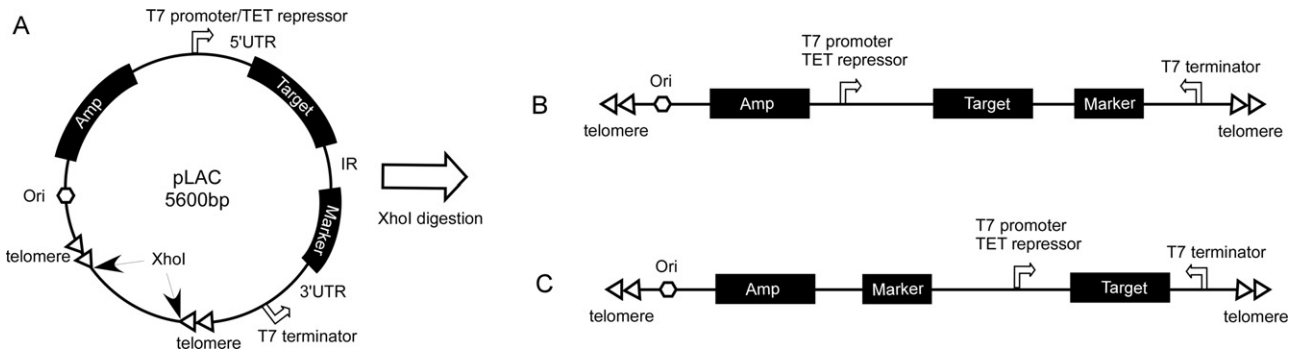

Fig. 1. Design of the *E. coli*/*Leishmania* artificial linear artificial chromosome (pLAC) shuttle vector. (A) The map of a typical pLAC shuttle vector. The triangles represent nine telomeric repeat sequences. Digestion of the vector with XhoI results in removal of the suffer sequence between telomeres and formation of (B), the linear artificial chromosome (LAC) with telomeres on both ends. (C) An alternative configuration of the LAC with the selection marker positioned outside of the T7 polymerase-transcribed cassette. Ori – *E. coli* origin of replication, IR – intergenic region, UTR – untranslated region.

3.2. Construction of the expression strains carrying LAC

Initially, we tested the ability of the designed vectors to transform *L. tarentolae*. To this end, linearized and circular constructs containing the genes coding for either EGFP, dsRed, or Cherry fluorescent reporter proteins were electroporated into *L. tarentolae* host strains constitutively expressing T7 RNA polymerase and TET repressor (*L. tarentolae ssu::T7POL SAT ssu::TR HYG*, further designated as *L. tarentolae* T7-TR) or into wild-type cells. To facilitate manipulation of the resulting clonal cell populations, we decided to depart from growing transfected cells on solid agar-based medium and to grow them on filters instead. The cells were plated on 0.45-μm nitrocellulose filters that were placed on top of a pad made of filter paper that was soaked with serum supplemented LEXSY BHI medium and appropriate antibiotics. In this case, the colonies emerged typically after 5 days when incubated at 26 °C (Fig. S1). To induce protein expression, the nitrocellulose filters with colonies were lifted and placed on top of a filter paper pad that was soaked in LEXSY BHI with tetracycline. To our knowledge, this represents the first report of cloning trypanosamatid organisms without gelated medium.

As can be seen in Fig. 2A and B, transfection with linear or circular LACs led to the emergence of inducible and highly fluorescent colonies. Interestingly, the efficiency of fluorescent clone emergence differed markedly between linear and circular constructs. While nearly 100% of all colonies that were transfected with linearized plasmid displayed strong inducible overexpression of the fluorescent reporter proteins, only ca. 10% of clones that were transfected with the circular plasmids could be induced to produce the reporter (Fig. 2B). No inducible overexpressing clones were obtained when wt or TET-repressor-expressing strains lacking T7 RNA polymerase were transfected (not shown).

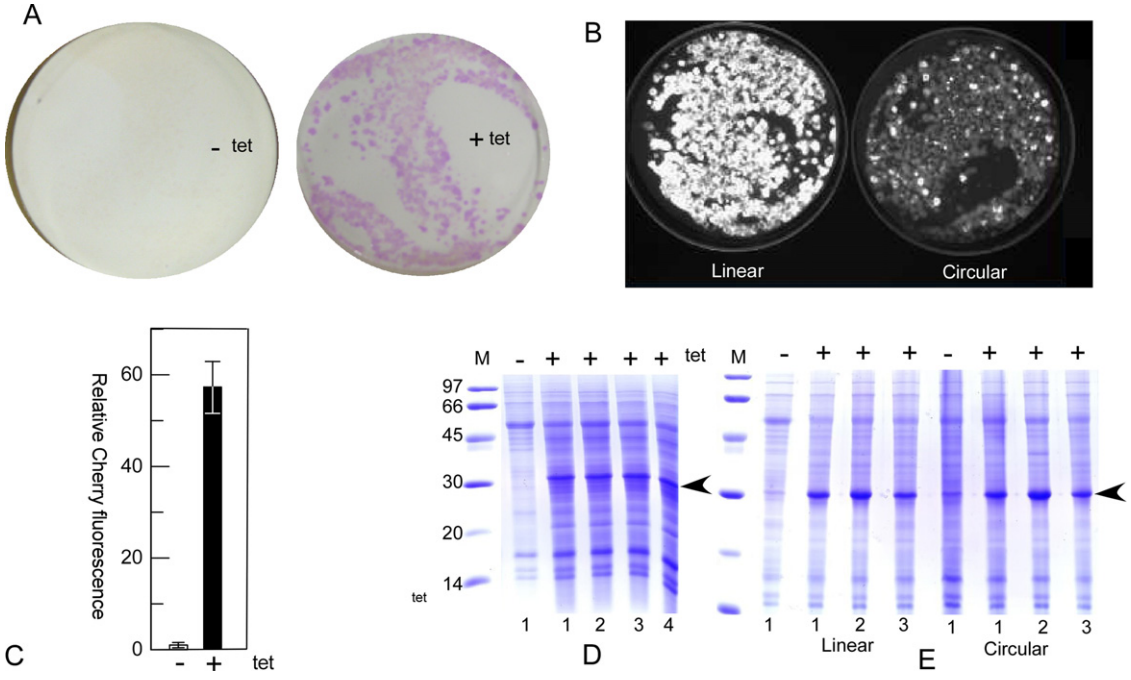

Fig. 2. Transfection of *L. tarentolae* T7-TR strain with linear and circular LACs carrying reporter genes for fluorescent proteins. (A) Colonies of *L. tarentolae* transfected with linear pLAC-Cherry vector and plated on the nitrocellulose filter. The right filter was exposed to tetracycline for 18 h. (B) Colonies of *L. tarentolae* cells transfected with linear (left panel) or circular (right panel) pLAC-dsRED vectors, plated and induced with tetracycline as in (A). The plates were photographed under unfiltered UV light. (C) Analysis of Cherry expression levels in twelve clones transfected with linear pLAC-Cherry and grown in suspension culture with or without induction. The cell suspensions were density-standardized before fluorescence measurement. (D) SDS-PAGE analysis of clones overexpressing Cherry from clones shown in (C). The arrowhead indicates the position of recombinant Cherry protein. (E) SDS-PAGE analysis of suspension cultures of clones from (B). In the case of clones transfected with circular vector only, fluorescent clones were selected for expansion, induction, and analysis.

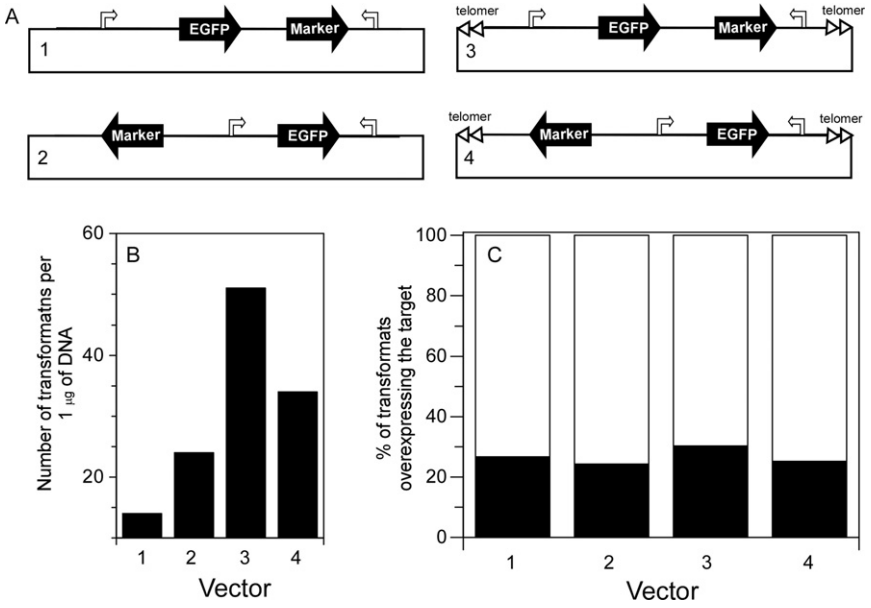

Fig. 3. Correlation of plasmid structure, transfection efficiency, and emergence of highly overexpressing clones from the transformed *L. tarentolae* T7-TR strain. (A) Structure of EGFP expression vectors introduced in circular form into the *L. tarentolae* T7-TR strain. (B) Analysis of the transfection efficiency of the vectors depicted in (A). (C) Frequency of overexpressing clone emergence in cells transfected with circular vectors depicted in (A). The shaded area reflects the fraction of overexpressing clones. The bars represent individual constructs, numbered as in (A).

In order to estimate the level of protein overexpression, we grew the transfomants in suspension culture and induced protein expression by the addition of tetracycline to the medium. As can be seen in Fig. 2C–E, recombinant fluorescent proteins were overproduced to more than 10% of total cellular protein (ca. 300 mg of recombinant proteins per liter of *L. tarentolae* suspension culture as determined by scanning of Coomassie stained gels). To our knowledge, this represents the highest overexpression level achieved for a cytosolic protein in *Leishmania* to date. The expression level exceeds those obtained in yeast, mammalian cells, and plants and rival those obtained in the baculovirus expression system.

3.3. Analysis of clonal heterogeneity of L. tarentoalae transfected with circular LACs and co-expression of multiple plasmids

The high heterogeneity of clones that were transfected with circular LAC prompted us to perform additional analyses. Firstly, we compared the correlation of the occurrence of high-overexpressing clones with plasmid organization. We varied two parameters: the presence of telomeric repeats and the orientation of transcription of the antibiotic resistance marker in relation to the T7 promoter. As shown in Fig. 3A–C, neither of the parameters significantly influenced the frequency of emergence of overexpressing clones, while the presence of telomeres significantly improved the transformation efficiency.

We then tested the strain's ability to maintain two types of episomes. In this experiment, the T7-TR *L. tarentoalae* strain was co-transfected with pLAC-EGFP-Neo and pLAC-Cherry-Ble. The clones that inducibly overexpressed both reporters emerged with ca. 10% frequency (not shown). SDS-PAGE analysis of the expression levels showed that in the co-transfected strains, individual proteins were expressed at somewhat lower levels than in strains expressing only one protein (Fig. 4). This may indicate competition between mRNAs coding for fluorescent proteins.

3.4. Analysis of structure of linear and circular LACs

In order to gain insight into the molecular events that underlie the formation of overexpressing clones, we performed Southern blot analysis of genomic DNA from T7-TR *L. tarentolae* strains that were transfected with circular or linear LACs and cultivated in suspension under drug pressure for up to 90 generations (approximately 2 months of continuous cultivation). As shown in Fig. 5A, the linear LACs reached a maximum of abundance after 10 generations and migrated predominantly at the expected position of 6 kb. Surprisingly, digestion of total DNA with NcoI resulted in the formation of probe-reactive bands of 6.5 and 2 kb instead of a single band of 3.6 kb. However, an additional 3.5-kb band was identified in clone nr. 3 (Fig. 5B), potentially indicating rearrangement of the construct. An alternative explanation could be the partial digestion of the full-length construct and aberrant migration of the resulting fragment. However, the digestion of the control plasmid does not support this conclusion (not shown). Interestingly, the clones that were transfected with circular episomes displayed large heterogeneity in both the amounts and sizes of the episomes. They also exhibited large variations in the structure and abundance of the episomes in time (Fig. 5C). Remarkably, there was no clear relationship between the size or abundance of the episome or expression levels

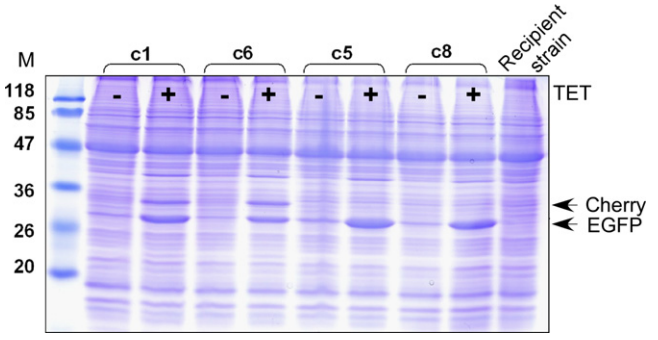

Fig. 4. Co-transfection of *L. tarentolae* T7-TR strain with circular pLAC-EGFP-Neo and pLAC-Cherry-Ble plasmids and selection on solid medium containing both antibiotics. Several clones displaying EGFP fluorescence were picked and expanded in suspension cultures. The aliquots of induced and uninduced cultures were resolved by SDS-PAGE and stained with Coomassie. The positions of recombinant EGFP and Cherry are indicated by the arrows.

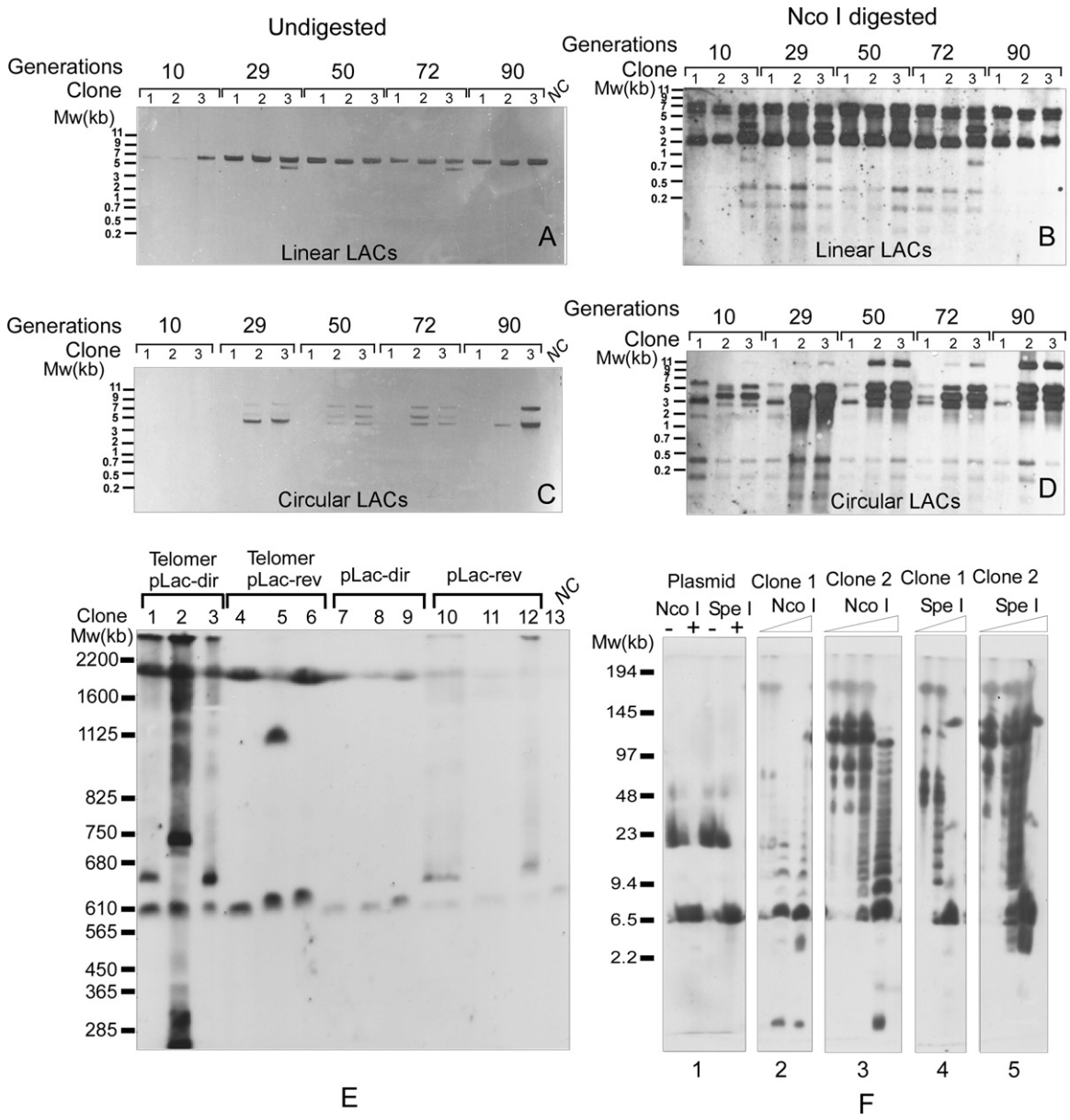

Fig. 5. Electrophoretic analysis of circular and linear LACs maintained in *L. tarentolae* T7-TR strains under antibiotic selection. (A) Sothern blot analysis of total genomic DNA (3 μg) extracted form *L. tarentolae* strains transfected with linear pLAC-dsRED constructs and resolved on a 0.7% agarose gel. Three clones per experiment (1–3) were selected based on their ability to inducibly express the target protein and were passaged for the indicated number of generations under selection pressure. The dsRed gene was used as a hybridization probe. (B) As in (A), but the total DNA extracted from the clones was digested with NcoI prior to loading. This was predicted to result in the formation of a 3.6-kb dsRed gene containing a fragment from the 6.7-kb linear LAC. Interestingly, a smaller fragment consistently emerged in all clones tested, indicating that the original construct underwent a rearrangement. (C) Same as in (A), but the cells were transfected with circular LACs. (D) As in (C), but the total DNA was digested with NcoI prior to electrophoresis. (E) Pulsed-field electrophoresis of the clones transfected with circular pLACs-EGFP and resistance markers in different orientations (see Fig. 1B and C). Telomer pLac-dir – telomere containing circular pLAC with EGFP and neomycin selection marker within the same T7 transcription unit. Telomer pLac-rev – as above, but with the neomycin selection marker in the antisense orientation. pLac-dir – circular episome identical to telomere pLac-dir but lacking telomeric repeats. pLac-rev – as above, but with the antibiotic resistance marker facing in the opposite orientation. NC – negative control, represented by *L. tarentolae* T7-TR strain. The separation of total DNA was carried out for 45 h. The DNA was transferred onto a membrane and blotted with an EGFP probe. (F) Pulsed-field gel electrophoresis of total DNA derived from clones of the *L. tarentolae* T7-TR strain transfected with telomere-containing circular pLAC-EGFP. The agar blocks with lysed cells were incubated with different concentrations of NcoI or SpeI restriction enzymes, which were expected to cut the parental construct only once, as evidenced by the digestion of the isolated plasmid (slab 1). Remarkably, in the case of clones digested with SpeI, complete digestion resulted in emergence of a high-molecular weight product, possibly representing a fragment of a chromosomally inserted cassette (slab 3 and 4).

of the fluorescent marker protein (not shown). Digestion of total DNA of the clones that were transfected with the circular vector resulted in emergence of multiple probe-reactive bands of various sizes. They included the bands exceeding the molecular weight of the original vector, indicating that the episomes underwent large rearrangements (Fig. 5D). This prompted us to analyze the series of strains transfected with various episomal constructs from the transformation experiments above (Fig. 3).

Pulsed-field analysis revealed reactive bands ranging from 0.1 to >2 Mb (Fig. 5E). In order to elucidate the structure of the detected bands, we performed partial digestion of the genomic DNA isolated from clones transfected with circular LACs using enzymes that cut the original plasmid only once. As can be seen in Fig. 5F, this resulted in the formation of typical ladders of fragments between the uncut position and the position corresponding to the monomeric plasmid. Consistent with earlier findings, this indicates that the

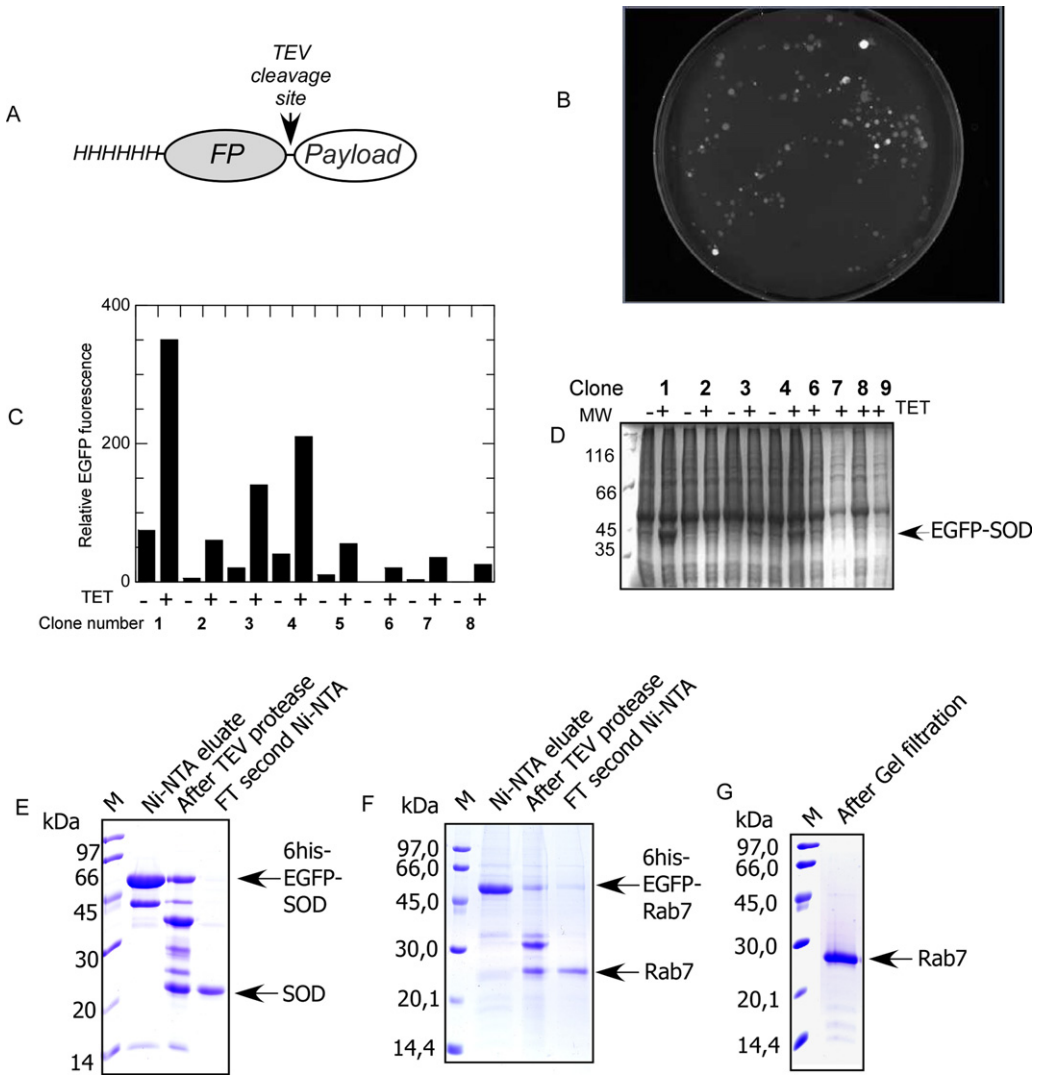

Fig. 6. Expression and purification of human superoxide dismutase (SOD) and canine Rab7 GTPase using cicular LACs. (A) Schematic representation of fusion protein used for identification of clones overproducing SOD and Rab7. (B) Photograph of the colonies of the *L. tarentolae* T7-TR strain transfected with circular pLAC-EGFP-SOD-Neo. The picture was taken under unfiltered UV light. (C) Fluorescence of tetracycline-induced and uninduced density-normalized suspension cultures of *L. tarentolae* T7-TR:pLAC EGFP SOD Neo clones. (D) As in (C), but the lysates of the clones were analyzed on a 15% SDS-PAGE gel and stained with Coomassie. (E) Coomassie-stained SDS-PAGE gel loaded with eluates from different steps of SOD purification. (F) As in (E), but loaded with eluates of Rab7 purification. (G) Pooled eluate of the gel filtration column loaded with Rab7 obtained in (F).

plasmids underwent concatenation [17,18]. The fact that similar structures were observed both in Leishmania and Trypanosoma indicates that concatenation mechanism is conserved across the trypanosomatids. Remarkably, the digestion also affected the >2-Mb bands, indicating that they also contained plasmid fragments (not shown). Although the available data are not sufficient to conclusively identify the nature of these bands, we conjecture that they may represent the individual transcription units or their concatemers integrated into the chromosome(s). As all *L. tarentolae* plasmids that were employed contained UTRs that were derived from the *L. tarentolae* host genome, such integrations may be possible [6]. This may explain the remarkable stability of the overexpressing clones obtained by transfection with the circular LACs (see Sections 3.6 and 3.7).

3.5. Expression and purification of recombinant proteins using *L. tarentolae* transfected with circular LACs

The overexpression levels of recombinant proteins achieved in the *L. tarentolae* T7-TR stain rival the best available eukaryotic expression systems, such as baculovirus-infected insect cells. We next tested whether the developed methodology could be used for preparative production of recombinant proteins. We initially decided to work with cell lines transfected with circular LACs, as they were stable over long cultivation periods without antibiotic pressure (see below), making them suitable for large-scale fermentation. In order to circumvent the problem of clonal heterogeneity, we decided to express the protein of interest in an N-terminal-cleavable fusion with a fluorescent marker protein. As a test example, anti-oxidant Cu/Zn superoxide dismutase (SOD), which catalyzes the dismutation of superoxide into oxygen and hydrogen (EC 1.15.1.1), was chosen. A second target was Rab7, a small GTP-binding protein involved in the regulation of late endosomal biogenesis in eukaryotes [19]. Both proteins were expressed as fusions with 6His-tagged EGFP (Fig. 6A). To facilitate the purification of non-tagged proteins, a sequence that encoded a TEV protease cleavage site was inserted between the fluorescent protein and the target. The *L. tarentolae* T7-TR strain was transfected with constructs, as described above, plated onto nitrocellulose filters, and exposed to tetracycline to induce expression. This treat-

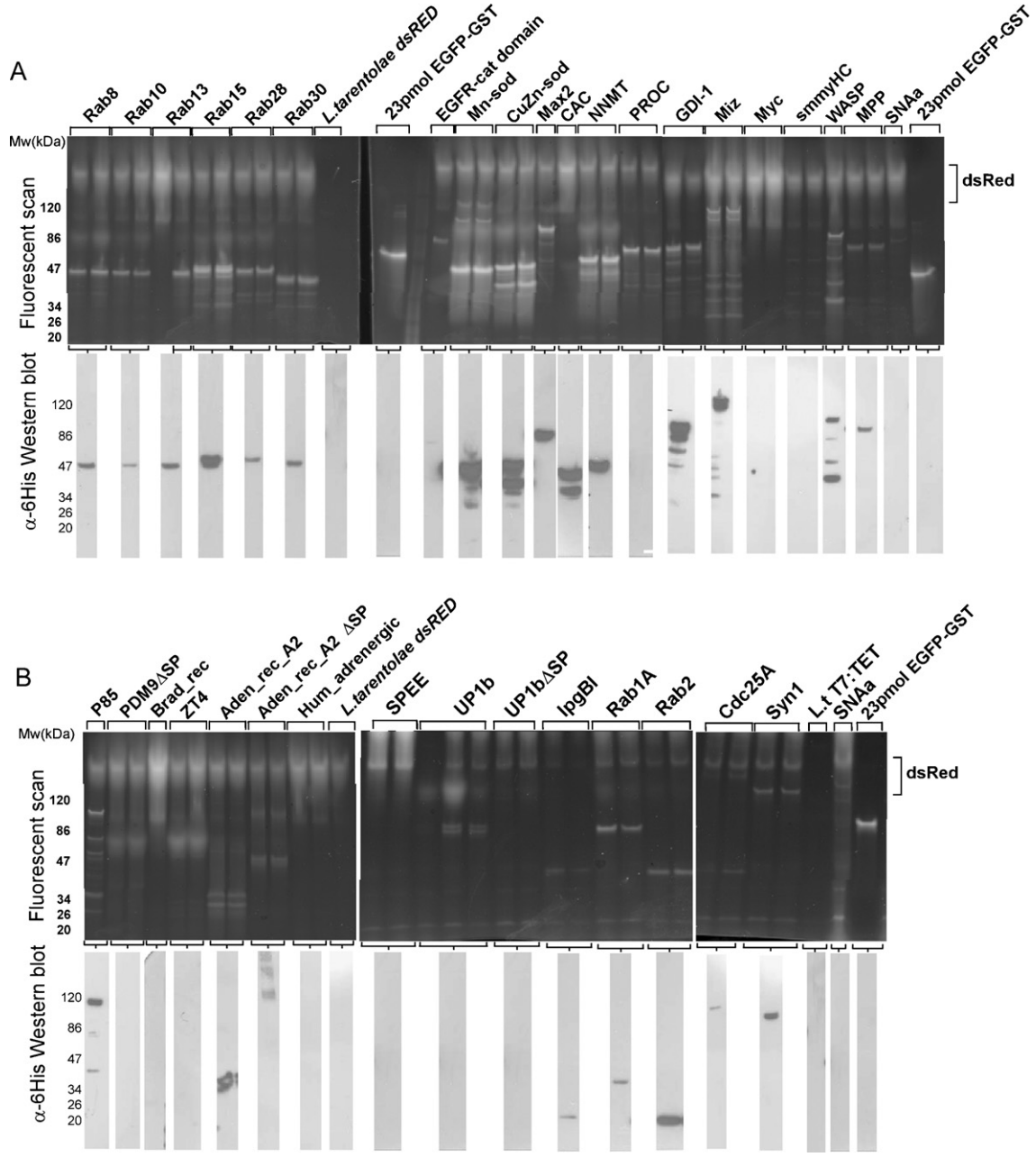

Fig. 7. SDS-PAGE and Western blot analysis of *L. tarentolae* T7-TR odc::dsRED^{Ti} Ble strains transfected with linear pLACs expressing a selection of pro- and eukaryotic genes N-terminally tagged with EGFP. The cells transfected with linear pLACs were plated on nitrocellulose filters, and two independent clones per construct were selected and expanded into 2 ml of BHI medium. Protein expression was induced when the cultures reached OD_{600} of 0.8–0.9, and 60 μl of culture was pelleted and resuspended in 40 μl of SDS-PAGE loading buffer. The unboiled samples (20 μl) were resolved on a 15% SDS-PAGE gel and scanned for EGFP fluorescence using a BioRad fluorescent scanner. The strains that were constructed inducibly co-expressed the dsRed reference protein, which migrated as a tetramer under these conditions on the gel and resulted in weak background fluorescence. The remaining samples were boiled and subjected to Western blot analysis with anti-his-tag antibody. Abbreviations of the expressed genes: Rab8 – rat Rab8 GTPase (accession AAH71176), Rab10 – mouse Rab10 GTPase (NP_057885), Rab13 – human Rab13 GTPase (accession AAH94846), Rab28 – human Rab28 GTPase (accession CAA64364), Rab30 – human Rab30 GTPase (accession NP_055303), EGFR cat. domain – catalytic domain of human EGFR receptor (accession AAH94761.1), Mn-sod – human mitochondrial superoxide dismutase 2 (accession P04179), CuZn-sod – human Cu–Zn superoxide dismutase (NP_000445.1), Max2 – human max protein isoform b (accession NP_660087.1), CIC – human calcineurin isoform CNEX9-11 (accession ABW74484.1), NNMT – human nicotinamide N-methyltransferase (accession NP_006160.1), PROC – human pyrroline-5-carboxylate reductase 1 (accession NP_008838.2), GDI-1 – bovine Rab GDP dissociation inhibitor alpha (accession NP_776489.1), Miz – human Myc interacting zinc finger and BTB domain-containing protein 17 (accession NP_003434.2), Myc – human Myc proto-oncogene protein (accession P01106.1), smmyHC – human myosin heavy chain isoform 11 (accession EAW53924.1), WASP – human Wiskott–Aldrich syndrome protein (accession AAH02961.1), MPP – human M-phase phosphoprotein 1 (accession EAW50126.1), SNAa – human serine/threonine-protein phosphatase 2B catalytic subunit alpha (accession NP_001124163.1), p85 – bovine phosphatidylinositol 3-kinase regulatory subunit alpha (accession NP_777000.1), PDM9ΔSP – human transmembrane protein with EGF-like and two follistatin-like domains 2 lacking signal peptide (accession EAX10832.1), Brad_rec – human B2 bradykinin receptor (accession NP_000614.1), ZT4 – human zinc transporter 4 (accession NP_037441.2), Aden_rec_A2 – human adenosine receptor A2 (accession NP_000666.2), Aden_rec_A2ΔSP – human adenosine receptor A2 lacking signal peptide (accession NP_000666.2), Hum_adrenergic – human beta-2 adrenergic receptor (accession NP_000015.1), SPEE – human spermidine synthase (accession NP_003123.2), UP1b – human uroplakin-1b (accession NP_008883.2), UP1bΔSP – human uroplakin-1b lacking signal peptide (accession NP_008883.2), NP_085292.1, IpgB- Shigella flexneri invasion protein [5a] (accession NP_085292.1), Rab1A – human Rab1A GTPase (accession NP_004152.1), Rab2 – human Rab2 GTPase (accession NP_002856.1), CDC25A – human cell division cycle phosphatase 25A (accession EAW64850.1), Syn1– rat synaptojanin1 (accession CAA07267.1), and SNAa – human serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform isoform 2NP (accession 001124163.1).

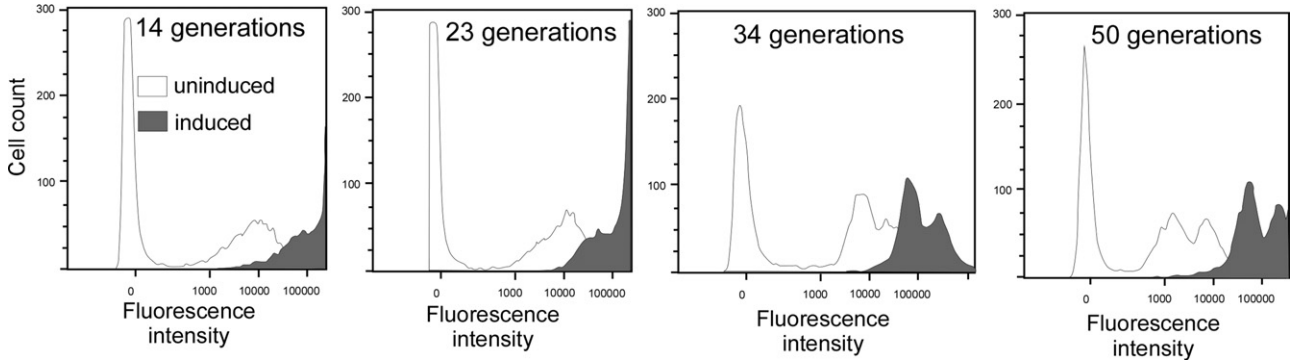

Fig. 8. Cytometric analysis of polyclonal suspension cultures transfected with linear pLAC-EGFP constructs and passaged under selection pressure for the indicated number of generations. Expression of EGFP was induced with tetracycline 12 h before measurement. Seven generations correspond to the return of the suspension culture to the original density upon 1:100 dilution.

ment resulted in the emergence of fluorescent clones of different brightnesses (Fig. 6B). In order to test the performance, a subset of clones with varying fluorescences were picked and expanded in 2 ml agitated suspension culture. Aliquots of the cultures before and after induction with tetracycline were collected and analyzed by fluorescence scanning and SDS-PAGE. These experiments confirmed that the clones that displayed high fluorescence in the filter assay also produced high expression yields in suspension (Fig. 6C and D). They also indicate that transient induction of overexpression does not appreciably compromise the stability of the selected cell lines. SDS-PAGE analysis of the induced and uninduced clones demonstrated that overexpression levels of SOD reached 10% of total cellular protein (Fig. 6D). Since the obtained yields were suitable for preparative protein production, we expanded the best-expressing clones into 2-l suspension cultures and induced protein expression at OD_{600} 0.8–0.9. The cells were allowed to grow for an additional 16 h and were collected by centrifugation. The cells were resuspended in phosphate buffer containing protease inhibitors and disrupted by fluidizing. The lysate was cleared by centrifugation and loaded onto a Ni–NTA column. The protein was eluted with a gradient of imidazol, and the fractions containing protein with more than 60% purity were collected. At this stage both EGFP-SOD and EGFP-Rab7 fusions were cleaved with TEV protease and re-chromatographed on Ni–NTA to remove the uncleaved protein and the TEV protease. This procedure resulted in >70% pure protein, as determined by SDS-PAGE analysis. In the case of Rab7, the protein was additionally chromatographed on a Superdex 75 size exclusion column. The procedures resulted in purification of 10 mg of SOD and 7 mg of Rab7 from 2 l of suspension culture. These results demonstrated that the *L. tarentolae* expression system that is based on artificial chromosomes can be used for preparative production of recombinant eukaryotic proteins.

3.6. Analysis of suitability of linear LACs for protein expression in *L. tarentolae*

The results that have been described in the previous sections indicate that the developed methodology is suitable for preparative production of recombinant proteins. While the expression cell lines obtained using circular artificial episomes have the advantage of high stability, the clonal heterogeneity complicates the identification of the productive clones. This problem is particularly pronounced when a larger number of targets are expressed. From this perspective, low clonal heterogeneity of strains generated with linear LACs offers a potential advantage. To test the suitability of this expression architecture for large-scale expression studies, we selected a set of 37 genes encoding soluble, peripheral, and membrane proteins. The genes were cloned into pLAC-6his-EGFP (Fig. 5A) using the Infusion cloning protocol [20], and an *L. tarentolae* T7-TR *odc::dsRED*^{Ti} *Ble* strain was transfected. In this recipient strain, a gene encoding red fluorescent protein is integrated into the chromosomal ODC locus under the control of the tet-R-operated T7 promoter. This provides an internal control for protein expression as the cells develop a bright red color in response to tetracycline. The transfected cells were plated for selection, and two independent clones per target were picked and expanded in 2 ml of suspension culture. Protein expression was induced by the addition of tetracycline, and cells were harvested after 16 h. In order to visualize the recombinant proteins, we took advantage of the fact that unboiled EGFP remains folded and fluorescent when resolved by SDS-PAGE (Fig. 7). Therefore, we subjected aliquots of the lysates obtained from the induced cultures to SDS-PAGE analysis, followed by fluorescence scanning. In parallel, another set of aliquots of these lysates was boiled, resolved by SDS-PAGE, and subjected to Western blotting with anti-6his antibody. The results depicted in Fig. 6 demonstrate that 30 out of the 37 tested proteins yielded fluorescent bands, indicating expression of at least partially folded fusion protein. Western blot analysis resulted in the detection of positive expression in 24 cases. Although a detailed analysis of these expression results is interesting and important, it is beyond the scope of the current study and will be performed elsewhere. These data, however, clearly indicate that the majority of heterologous genes could be successfully expressed from episomes in *L. tarentolae* and that linear expression vectors enable rapid and efficient construction of *production* clones.

3.7. Analysis of stability and inducibility of polyclonal populations of *L. tarentolae* transfected with linear pLACs

The obtained data demonstrate that *L. tarentolae* is a promising platform for expression of heterologous proteins, which combines high expression yields with the ease of construction and handling. The observed low clonal heterogeneity of *Leishmania* populations transfected with linear constructs suggested that the plating-based cloning step could be potentially omitted. This would further simplify the system and make it more compatible with high-throughput applications. To test this hypothesis, the *L. tarentolae* T7-TR host strain was transfected with a linearized telomere-bearing plasmid encoding 6his-EGFP-SOD, but instead of plating, the transfected cells were selected in suspension culture and diluted 1:100 if the cultures reached the stationary phase (7 generations per passage). After the number of generations indicated in Fig. 7, an aliquot of the cultures was withdrawn and protein expression was induced. The induced and uninduced cells were subjected to flow cytometry (Fig. 8).

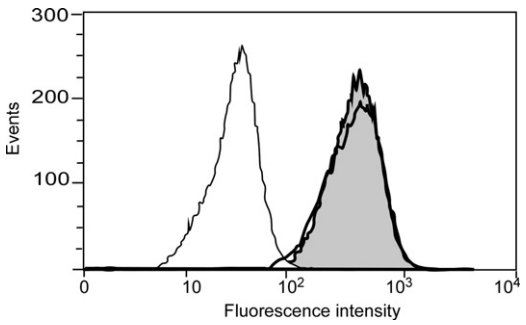

Fig. 9. FACS analysis of a clone of the *L. tarentolae* T7-TR strain transfected with a circular pLAC-EGFP construct and cultivated for 50 generations in suspension culture without a selection drug for the episome. The white peak represents the fluorescence of the uninduced culture while the grey and empty peaks represent the fluorescence of the induced starting and 3-month-grown cultures, respectively.

The results suggested that although population heterogeneity increased with time, the production of EGFP-SOD remained relatively stable after 34 generations. Interestingly, the cytometric analysis revealed the presence of a stable non-fluorescent cell subpopulation in the culture. This may reflect an early recombination event, resulting in deletion of the expression cassette. Nevertheless, the obtained non-clonal populations were sufficiently homogeneous with respect to the expressed product and demonstrated a degree of stability that is suitable for applications involving large-scale fermentation.

3.8. Stability analysis of L. tarentolae strains transfected with circular pLACs

The utility of a protein expression system largely depends on its ability to persist in functional form for a number of generations, sufficient for large-scale fermentation. To analyze the stability of the established clones transfected with circular pLACs, they were cultivated in the presence or absence of a selection drug for the episome by successive dilutions and analyzed by cytometry for the ability to inducibly express fluorescent marker proteins. It transpired from our studies that high-expressing clones that were transfected with circular expression constructs could be passaged this way for at least 50 generations without a noticeable change in expression level or inducibility (Fig. 9). Passaging under continuous induction resulted in a moderate decline of expression levels and the formation of a small pool of drug-tolerant but fluorescently silent cells (data not shown). This feature makes the system suitable for 100-l fermentation without selection, for example. Moreover, we successively cultivated one of the clones for more than 2 years (>1000 generations) under selection and found no significant changes in expression level or inducibility (not shown).

4. Conclusions

The presented data provide the first example of high-yield inducible protein expression in *Kinetoplastidae*, mediated by extrachromosomal linear elements. The linear elements bearing telomeric repeats at both ends were stably maintained under drug pressure for over >90 generations (Figs. 4 and 8). In combination with a tetracycline-regulated T7 transcription system, such linear vectors mediated the robust and inducible expression of target proteins, up to 10% of total cellular protein. This represents one of the highest yields achieved for intracellular expression in a eukaryotic organism. Remarkably high-expressing clones were also obtained when *Leishmania* was transfected with the circular version of the vectors. However, in this case, they comprised less than 20% of the transfected population. The origin of this heterogeneity is presently unclear, as the analyzed clones displayed broad diversity in copy number and size of the circular elements. Pulsed-field gel electrophoresis demonstrated that the extrachromosmal elements were concatenated, as often observed in natural or synthetic extrachromosomal elements of *Leishmania* [21]. We provide circumstantial evidence of the possible integration of the expression cassettes into chromosomes, which may also lead to the emergence of stable and highly productive clones. The fact that the resulting expression clones were stable over months of cultivation, even in the absence of antibiotic pressure, indirectly supports this idea and warrants more detailed studies of the molecular mechanisms underlying the emergence of genetically stable strains.

We positively tested the suitability of the overexpressing clones to serve as a source of recombinant proteins. Using a combination of affinity chromatography and proteolytic digestion, we were able to purify milligram amounts of two human proteins to near-homogeneity. This further demonstrates the utility of the developed system for production of recombinant eukaryotic proteins. It is noteworthy that EGFP-Rab7 fusion could not be produced in *E. coli* in soluble form due to the formation of inclusion bodies (Alexandrov, unpublished).

In order to further evaluate the developed expression system, we expressed a library of predominantly mammalian proteins in *L. tarentolae* using linear LACs. Analysis of the expression clones demonstrated that the majority of the genes were expressed at detectable levels. Remarkably, several proteins with molecular weights above 100 kDa were well expressed in this system. This further validates the use of the system for the production of large mammalian proteins. We expect that the developed expression architectures will find broad use not only among molecular parasitologists and microbiologists but also in the broader biotechnological and biochemical community.

Acknowledgments

This work was supported in part by DFG Grant AL 484/8-1, ARC DP Grant DP1094080, and ARC FF FT0991611 to K.A.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.molbiopara.2010.12.002.

References

[1] Clayton C, Shapira M. Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. Mol Biochem Parasitol 2007;156:93–101.
[2] Clayton CE. Genetic manipulation of kinetoplastida. Parasitol Today 1999;15:372–8.
[3] Wirtz E, Hartmann C, Clayton C. Gene expression mediated by bacteriophage T3 and T7 RNA polymerases in transgenic trypanosomes. Nucleic Acids Res 1994;22:3887–94.
[4] Yan S, Myler PJ, Stuart K. Tetracycline regulated gene expression in Leishmania donovani. Mol Biochem Parasitol 2001;112:61–9.
[5] Yao C, Luo J, Hsiao CH, Donelson JE, Wilson ME. Leishmania chagasi: a tetracycline-inducible cell line driven by T7 RNA polymerase. Exp Parasitol 2007;116:205–13.
[6] Kushnir S, Gase K, Breitling R, Alexandrov K. Development of an inducible protein expression system based on the protozoan host *Leishmania tarentolae*. Protein Expr Purif 2005;42:37–46.
[7] Zhang WW, Charest H, Matlashewski G. The expression of biologically active human p53 in Leishmania cells: a novel eukaryotic system to produce recombinant proteins. Nucleic Acids Res 1995;23:4073–80.
[8] LeBowitz JH, Smith HQ, Rusche L, Beverley SM. Coupling of poly(A) site selection and trans-splicing in Leishmania. Genes Dev 1993;7:996–1007.
[9] LeBowitz JH, Coburn CM, McMahon-Pratt D, Beverley SM. Development of a stable Leishmania expression vector and application to the study of parasite surface antigen genes. Proc Natl Acad Sci USA 1990;87:9736–40.
[10] Basile G, Peticca M. Recombinant protein expression in *Leishmania tarentolae*. Mol Biotechnol 2009;43:273–8.

[11] Breitling R, Klingner S, Callewaert N, et al. Non-pathogenic trypanosomatid protozoa as a platform for protein research and production. Protein Expr Purif 2002;25:209–18.
[12] Scahill MD, Pastar I, Cross GA. CRE recombinase-based positive–negative selection systems for genetic manipulation in *Trypanosoma brucei*. Mol Biochem Parasitol 2008;157:73–82.
[13] Mureev S, Kushnir S, Kolesnikov AA, Breitling R, Alexandrov K. Construction and analysis of *Leishmania tarentolae* transgenic strains free of selection markers. Mol Biochem Parasitol 2007;155:71–83.
[14] Olmo A, Arrebola R, Bernier V, Gonzalez-Pacanowska D, Ruiz-Perez LM. Co-existence of circular and multiple linear amplicons in methotrexate-resistant Leishmania. Nucleic Acids Res 1995;23:2856–64.
[15] Genest PA, ter Riet B, Dumas C, Papadopoulou B, van Luenen HG, Borst P. Formation of linear inverted repeat amplicons following targeting of an essential gene in Leishmania. Nucleic Acids Res 2005;33:1699–709.
[16] White TC, Fase-Fowler F, van Luenen H, Calafat J, Borst P. The H circles of *Leishmania tarentolae* are a unique amplifiable system of oligomeric DNAs associated with drug resistance. J Biol Chem 1988;263:16977–83.
[17] Papadopoulou B, Roy G, Ouellette M. Autonomous replication of bacterial DNA plasmid oligomers in Leishmania. Mol Biochem Parasitol 1994;65:39–49.
[18] Kelly JM, Ward HM, Miles MA, Kendall G. A shuttle vector which facilitates the expression of transfected genes in *Trypanosoma cruzi* and *Leishmania*. Nucleic Acids Res 1992;20:3963–9.
[19] Chavrier P, Parton RG, Hauri HP, Simons K, Zerial M. Localization of low molecular weight GTP binding proteins to exocytic and endocytic compartments. Cell 1990;62:317–29.
[20] Berrow NS, Alderton D, Sainsbury S, et al. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. Nucleic Acids Res 2007;35:45–9.
[21] Grondin K, Papadopoulou B, Ouellette M. Homologous recombination between direct repeat sequences yields P-glycoprotein containing amplicons in arsenite resistant *Leishmania*. Nucleic Acids Res 1993;21:1895–901.