

Published in final edited form as:

Mol Biochem Parasitol. 2008 September ; 161(1): . doi:10.1016/j.molbiopara.2008.05.006.

Single-locus targeting constructs for reliable regulated RNAi and transgene expression in *Trypanosoma brucei*

Sam Alford and David Horn

London School of Hygiene & Tropical Medicine, Keppel Street, London, WC1E 7HT, UK

Abstract

A major obstacle to reproducible expression of recombinant transcripts lies in the epigenetic effects of the flanking chromatin following integration. We previously presented a strategy to overcome this problem in bloodstream form *Trypanosoma brucei*, using a reporter to identify a ribosomal-spacer locus that supports optimal expression and then marking that locus for subsequent targeting. Advantages include elimination of variable-expression position-effects and the easy confirmation of correct integration. We now report a set of validated constructs that exploit this system for expression of dsRNA or recombinant protein. The current construct-set allows expression of intramolecular dsRNA for RNA interference knockdown or expression of proteins that can incorporate c-Myc epitope(s) or a fluorescent-tag for subcellular localisation, interaction and/or other functional analysis. The constructs are integrated at a single, marked locus and deliver reliable and reproducible expression.

Exploitation of trypanosomatid genome sequence data [1] (<http://www.genedb.org/>) will critically depend upon the tools employed to analyse gene function and parasite biology and to identify and validate therapeutic targets. Consequently, molecular genetic tools for *T. brucei* studies are constantly being improved and valuable plasmid constructs are available for modifying genes *in situ* [2-4], for ectopic expression of recombinant proteins and for RNA interference (RNAi) [5-7]. As predicted [8], tetracycline-regulated (Tet-on) promoters have been widely used, and regulated expression, which displays rapid induction kinetics, has been critical for harnessing the power of RNAi for specific mRNA knockdown and subsequent gene-function analysis.

Unusually, certain protein-coding genes in *T. brucei* are under the control of RNA Pol I promoters [9] and many constructs available for regulated expression depend upon these 'strong' promoters or the phage T7 promoter that requires parallel expression of T7 RNA polymerase. The 'first-generation' constructs exploited the *EPI* procyclin promoter, a Pol I promoter that is highly active in insect-stage cells [8] while the T7 system [10] has been more widely used in the bloodstream-form, the life-cycle stage that causes disease in humans and animals and that displays developmental stage-specific features relevant to understanding and controlling disease. Tet-regulated ribosomal RNA (*RRNA*) and Variant Surface Glycoprotein expression-site promoters have been reported more recently [11, 12].

Stable genetic manipulation in *T. brucei* typically relies upon electroporation and homologous recombination-based integrative transformation, since stable episomes are not widely used in the insect-stage and have not been reported in the bloodstream-form. With few loci having been shown to be transcriptionally inactive in trypanosomatids, the non-transcribed *RRNA* spacer loci [13] are probably the most popular targets for integration of regulated transgenes. There are nine of these loci annotated in the haploid genome sequence with one each on chromosomes 1 and 7, two on chromosome 2 and five on chromosome 3 [1]. The capacity for expression differs from one of these loci to the next [11]; presumably, the local environment impacts upon the achievable expression level. Such position effects

present a challenge for transgene expression in all cell types [14]. Besides variable expression at different *RRNA* spacer loci, integration elsewhere can generate complex phenotypes [15]. Consequently, it is typically necessary to screen multiple recombinant clones to identify those that display desirable features, usually strong repression with high-level expression upon induction. To circumvent problems arising from position-effects, we pre-screened clones for insertion at a *RRNA* spacer that displayed the features outlined above and then exchanged the reporter cassette for one with an incomplete *HYG*-resistance marker [11]. The resulting bloodstream-form *T. brucei* '2T1' strain can subsequently be used with constructs containing an overlapping and complementary *HYG*, selectable-marker/targeting fragment (see Fig. 1). This latter fragment is also incomplete so integration at the marked locus is obligatory to generate an intact *HYG*-ORF and hygromycin-resistant cells. The requirement to screen multiple clones for robust regulated expression is eliminated using this system [11]. An additional, unexpected, but welcome outcome was increased homologous recombination efficiency stimulated by transcription of the target [16]. We now report a versatile set of constructs that benefit from single locus targeting (Fig. 1 and Table 1). The constructs have been functionally validated in 2T1 cells using a range of assays, including growth curve analysis and northern and western blot; they display little evidence of any leaky expression and robust induction after 24 h and are freely available upon request.

RNA interference has been widely exploited for loss-of-function experiments in *T. brucei* [17]. The most common approach is Tet-regulated expression of RNA from a transgene that is stably integrated at a *RRNA* spacer locus. One can use 'stem-loop' constructs with a single regulated promoter or '2T7' constructs with opposing, regulated T7 promoters. Stem-loop constructs require the two-step directional cloning of oppositely oriented and complementary RNAi-trigger fragments from the gene of interest while 2T7 constructs require the addition of only a single fragment cloned between the opposing promoters [6, 7]. The stem-loop constructs offer more rapid and efficient mRNA knockdown however [18], probably because intramolecular interaction, relative to intermolecular interaction, is more likely to generate a high concentration of stable dsRNA, the trigger for RNAi. Thus, requirements for efficiency of knockdown versus ease of construct assembly must be considered when selecting the favoured RNAi technology.

We previously reported a single-locus-targeting version of the 2T7 construct, pRPa^{i2T7}, used in parallel with conditional expression of T7 RNA polymerase [11]. Unfortunately, we found that conditional T7 polymerase expression, coupled with toxicity at a high expression level [11], often complicated subsequent phenotypic analysis. We tested a new construct with opposing, Tet-regulated *RRNA* promoters but this failed to trigger efficient knockdown in stably transformed cells; possibly because the Pol I transcription complex is too large to complete bi-directional transcription. We then assembled a pair of single-locus-targeting, stem-loop constructs with different multiple cloning sites (Fig. 1 and Table 1) which have been used to target several genes.

Conditional transgenic expression facilitates studies on protein subcellular localisation, protein-protein interaction, functional complementation and other functional analyses. Again, a common approach in *T. brucei* is Tet-regulated expression from a transgene that is stably integrated at a *RRNA* spacer locus. Fusion proteins incorporating eGFP, or the c-Myc epitope have been widely used for direct detection, in the case of GFP, or for immunofluorescence, immunoblotting and immunoprecipitation with few problems of cross reaction with native proteins. Neither tag displays toxicity or the presence of cryptic localisation signals.

We previously reported a single-locus-targeting construct, used for regulated expression of unmodified or *N*-terminal GFP-tagged proteins [11]. We now report a range of constructs for *N*- or *C*-terminal tagging with eGFP or c-Myc (Fig. 1 and Table 1). A second set of constructs allows expression of a second tagged protein in the same strain at either a *RRNA* spacer locus or the native locus (Table 1). These constructs facilitate dual-localisation or can be used in conjunction with RNAi to monitor knockdown (of a tagged protein) or to study the effect of knockdown on localisation (of another tagged protein), for example.

The achievable protein expression level can be predicted based on codon-usage [19] and, where codon adaptation index predicts low-level translation, strong, Pol I transcription and/or a tandem tag may be required to achieve detectable expression. Over-expression, driven by the strong *RRNA* promoter may lead to mislocalisation and/or formation of non-physiological complexes in some cases however [4]. Where such problems are suspected, the Tet concentration can be reduced [10] or, where regulated expression is unnecessary, the gene can be modified at its native RNA Pol II-transcribed locus [2-4] which should provide expression closer to the natural level.

For initial gene characterisation, we use the tools described above for a combination of RNAi knockdown and subcellular localisation of tagged protein (see legend to Fig. 1). The RNAi results often indicate whether gene knockout is feasible and knowledge regarding subcellular localisation is useful when designing phenotype assays for the knockdown or knockout strains. Thus, the combined knockdown and localisation data facilitate the design of follow-up work and the strains generated often prove useful for further, more detailed functional analyses. The system described offers savings in time and cost and provides the first opportunity to create a battery of transgenes that can be reliably expressed by integration at a single locus.

Acknowledgments

Our work is supported by The Wellcome Trust (079457). We thank John Kelly and Martin Taylor for comments on the draft manuscript, Taemi Kawahara for assembly of the pNAT^{XG} construct, Chris Bot and Sam Obado for revealing the location of the marked *RRNA* locus in our 2T1 cells and other members of the Horn and Kelly groups for feedback on the utility of the constructs.

Reference

- [1]. Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, Bartholomeu DC, Lennard NJ, Caler E, Hamlin NE, Haas B, et al. The genome of the African trypanosome *Trypanosoma brucei*. *Science*. 2005; 309:416–422. [PubMed: 16020726]
- [2]. Kelly S, Reed J, Kramer S, Ellis L, Webb H, Sunter J, Salje J, Marinsek N, Gull K, Wickstead B, et al. Functional genomics in *Trypanosoma brucei*: a collection of vectors for the expression of tagged proteins from endogenous and ectopic gene loci. *Mol Biochem Parasitol*. 2007; 154:103–109. [PubMed: 17512617]
- [3]. Oberholzer M, Morand S, Kunz S, Seebeck T. A vector series for rapid PCR-mediated C-terminal in situ tagging of *Trypanosoma brucei* genes. *Mol Biochem Parasitol*. 2006; 145:117–120. [PubMed: 16269191]
- [4]. Shen S, Arhin GK, Ullu E, Tschudi C. *In vivo* epitope tagging of *Trypanosoma brucei* genes using a one step PCR-based strategy. *Mol Biochem Parasitol*. 2001; 113:171–173. [PubMed: 11254965]
- [5]. Alibu VP, Storm L, Haile S, Clayton C, Horn D. A doubly inducible system for RNA interference and rapid RNAi plasmid construction in *Trypanosoma brucei*. *Mol Biochem Parasitol*. 2005; 139:75–82. [PubMed: 15610821]
- [6]. LaCount DJ, Bruse S, Hill KL, Donelson JE. Double-stranded RNA interference in *Trypanosoma brucei* using head-to-head promoters. *Mol Biochem Parasitol*. 2000; 111:67–76. [PubMed: 11087917]

- [7]. Morris JC, Wang Z, Drew ME, Paul KS, Englund PT. Inhibition of bloodstream form *Trypanosoma brucei* gene expression by RNA interference using the pZJM dual T7 vector. *Mol Biochem Parasitol.* 2001; 117:111–113. [PubMed: 11551638]
- [8]. Wirtz E, Clayton C. Inducible gene expression in trypanosomes mediated by a prokaryotic repressor. *Science.* 1995; 268:1179–1183. [PubMed: 7761835]
- [9]. Das A, Banday M, Bellofatto V. RNA polymerase transcription machinery in trypanosomes. *Eukaryot Cell.* 2008; 7:429–434. [PubMed: 17951525]
- [10]. Wirtz E, Leal S, Ochatt C, Cross GA. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol Biochem Parasitol.* 1999; 99:89–101. [PubMed: 10215027]
- [11]. Alsford S, Kawahara T, Glover L, Horn D. Tagging a *T. brucei* *RRNA* locus improves stable transfection efficiency and circumvents inducible expression position effects. *Mol Biochem Parasitol.* 2005; 144:142–148. [PubMed: 16182389]
- [12]. Glover L, Alsford S, Beattie C, Horn D. Deletion of a trypanosome telomere leads to loss of silencing and progressive loss of terminal DNA in the absence of cell cycle arrest. *Nucleic Acids Res.* 2007; 35:872–880. [PubMed: 17251198]
- [13]. White TC, Rudenko G, Borst P. Three small RNAs within the 10 kb trypanosome rRNA transcription unit are analogous to domain VII of other eukaryotic 28S rRNAs. *Nucleic Acids Res.* 1986; 14:9471–9489. [PubMed: 3797245]
- [14]. Markstein M, Pitsouli C, Villalta C, Celniker SE, Perrimon N. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nat Genet.* 2008; 40:476–483. [PubMed: 18311141]
- [15]. Motyka SA, Zhao Z, Gull K, Englund PT. Integration of pZJM library plasmids into unexpected locations in the *Trypanosoma brucei* genome. *Mol Biochem Parasitol.* 2004; 134:163–167. [PubMed: 14747155]
- [16]. Alsford S, Horn D. RNA polymerase I transcription stimulates homologous recombination in *Trypanosoma brucei*. *Mol Biochem Parasitol.* 2007; 153:77–79. [PubMed: 17316839]
- [17]. Subramaniam C, Veazey P, Redmond S, Hayes-Sinclair J, Chambers E, Carrington M, Gull K, Matthews K, Horn D, Field MC. Chromosome-wide analysis of gene function by RNA interference in the african trypanosome. *Eukaryot Cell.* 2006; 5:1539–1549. [PubMed: 16963636]
- [18]. Durand-Dubief M, Kohl L, Bastin P. Efficiency and specificity of RNA interference generated by intra- and intermolecular double stranded RNA in *Trypanosoma brucei*. *Mol Biochem Parasitol.* 2003; 129:11–21. [PubMed: 12798502]
- [19]. Horn D. Codon usage suggests that translational selection has a major impact on protein expression in trypanosomatids. *BMC Genomics.* 2008; 9:2. [PubMed: 18173843]
- [20]. Redmond S, Vadivelu J, Field MC. RNAi: an automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei*. *Mol Biochem Parasitol.* 2003; 128:115–118. [PubMed: 12706807]

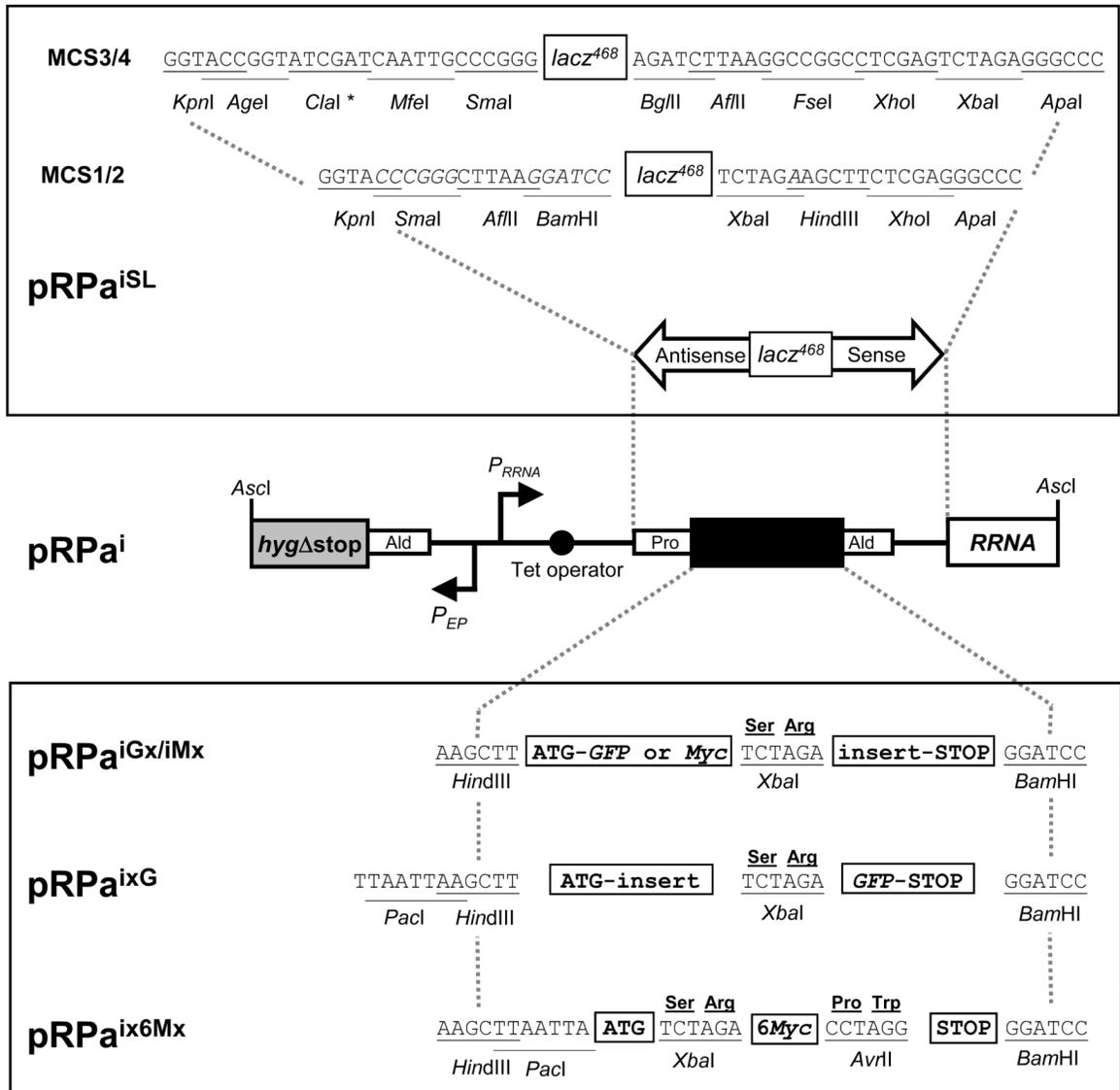


Figure 1. Schematic map of single-locus targeting constructs and available cloning sites
 Following *Ascl* digestion (cleaves GGCGCGCC), the *pRPa* constructs have terminal *hyg* stop and *RRNA* spacer sequences for targeting the *hyg* start::*PAC*::*RRNA* locus, on chromosome 2a in our 2T1 cells. Correct integration generates an intact *HYG* ORF and deletes a *PAC* ORF [11]. Selectable marker (*HYG*) transcription is then constitutive (*P_{EP}*: procylin promoter) and independent of the Tet-regulated cassette (*P_{RRNA}*: *RRNA* promoter). All cloning sites indicated are unique and several generate cohesive ends that are compatible with other restriction enzymes.

Upper box: Two stem-loop (SL) RNAi constructs are available with different multiple cloning sites (MCS1/2 and MCS3/4). In both cases the ‘loop’ is a 468-bp *lacZ* fragment. We routinely select specific RNAi gene targets (400-600 bp) and primers using the RNAit software [20]

<http://trypanofan.path.cam.ac.uk/software/RNAit.html>. A single pair of PCR primers are designed that incorporate four selected restriction sites (not present in the RNAi target fragment) such that a single PCR product can be differentially digested and sequentially cloned. For example, using MCS1/2, the following primers could be used to clone antisense

followed by sense fragments: Primer 1, *XbaI*-*Bam*HI-5 sequence; Primer 2, *ApaI*-*KpnI*-3 sequence. The antisense and sense fragments are generated by *Bam*HI-*KpnI* and *XbaI*-*ApaI* digestion respectively. We've seen no problems with SL-construct stability or yield using *E. coli* DH5 . * The *ClaI* site in MCS3 is blocked by overlapping Dam methylation in DH5 cells.

Lower box: Tagging constructs are available for expression of *N*- or *C*-terminal fusions with eGFP (G, Clontech) or c-Myc (M, EQKLISEEDL). The tandem 6Myc construct allows assembly of either *N*- or *C*-terminal fusions and improved sensitivity. x indicates location of selected ORF.

We maintain 2T1 *T. brucei* cells (Lister 427, MITat1.2, clone 221a), that also express the *Tn10*Tet repressor (*TetR::BLE::TUB*), in HMI-11 growth medium containing phleomycin and puromycin ($0.5 \mu\text{g ml}^{-1}$ and $0.2 \mu\text{g ml}^{-1}$ respectively). We use electroporation (Gene Pulser II, BioRad) with $\sim 2.5 \times 10^7$ cells and $\sim 3 \mu\text{g}$ (yield from one mini-prep, Qiagen) of each *AscI*-linearised construct in duplicate; each mixture is transferred to 36 ml of medium in a 25 cm² culture flask. After ~ 6 h, the appropriate drug selection is applied ($2.5 \mu\text{g ml}^{-1}$ hygromycin) and cells are distributed into 12- or 24-well plates. Each electroporation experiment typically, and conveniently, yields 2-5 transformed clones using this approach and these are visible after ~ 5 days ($\sim 100,000$ cells). To screen clones for simple, double-crossover integration and loss of the *PAC* gene, we remove $\sim 5\%$ of each day-5, hyg^R culture and grow with or without puromycin ($2 \mu\text{g ml}^{-1}$) in a 24-well plate for another 24-48 h. Typically, $\sim 50\%$ of clones are puromycin sensitive and two of these are retained for further analysis. Note that these strains are now resistant to phleomycin (*TetR::BLE::TUB*) and hygromycin (*HYG::RPa::RRNA*). Initial analysis normally consists of comparing uninduced and induced (Tet at $1 \mu\text{g ml}^{-1}$) cells. For RNAi, this is by counting cell density every 24 h for 3-4 days to generate a growth curve with parallel RNA and/or protein extraction (24 and/or 48 h) for assessment of knockdown by northern and/or western blot depending upon the availability of a suitable antibody. For recombinant protein expression, it is by western blot and/or (immuno)fluorescence analysis after 24 h induction (peak expression is normally seen at this time [10]). Construct maps and sequences, along with further details, are available at <http://homepages.lshtm.ac.uk/~ipmbdhor/dhhome.htm> (Resources). Further developments and other compatible or complementary constructs will also be documented at this site.

Table 1

Single-locus targeting (pRPa) and other complementary constructs.

| Constructs | Name ^a | Selectable Marker ^b | Expression Cassette | Tet-regulated | Linearise with |
|-------------------|----------------------|--------------------------------|-----------------------------|---------------|----------------|
| pRPa | pRPa ^{iSL} | <i>HYG</i> | Stem-loop RNAi ^c | Yes | <i>AscI</i> |
| | pRPa ^{iGx} | <i>HYG</i> | GFPx | Yes | <i>AscI</i> |
| | pRPa ^{iG} | <i>HYG</i> | xGFP | Yes | <i>AscI</i> |
| | pRPa ^{iMx} | <i>HYG</i> | Mycx | Yes | <i>AscI</i> |
| | pRPa ^{i6Mx} | <i>HYG</i> | x6Mycx | Yes | <i>AscI</i> |
| pRP ^d | pRP ^{iMx} | <i>BSD</i> | Mycx | Yes | <i>NotI</i> |
| | pRP ^{i6Mx} | <i>BSD</i> | x6Mycx | Yes | <i>NotI</i> |
| pNAT ^e | pNAT ^{xG} | <i>BSD</i> | xGFP | No | <i>f</i> |
| | pNAT ^{x12M} | <i>BSD</i> | x12Myc | No | <i>f</i> |
| | pNAT ^{Gx} | <i>BSD</i> | GFPx | No | <i>f</i> |
| | pNAT ^{6Mx} | <i>BSD</i> | 6Mycx | No | <i>f</i> |

^aRP, *RRNA* promoter controls the expression cassette; NAT, native locus controls expression; a, denotes *AscI* sites; i, inducible; SL, stem-loop; x, selected ORF; G, eGFP; M, cMyc epitope.

^b*HYG*, hygromycin-B phosphotransferase; *BSD*, blasticidin-S deaminase.

^cDifferent multiple cloning sites available: MCS1/2, MCS3/4 (see Fig. 1).

^dpRP constructs can integrate at any *RRNA* spacer locus and yield variable levels of expression upon induction in different clones.

^epNAT allows expression of a tagged ORF from its native locus; see [2-4].

^fThe restriction enzyme used depends upon the specific ORF targeted.