Chapter 15

Recombinant Protein Production in the Eukaryotic Protozoan Parasite *Leishmania tarentolae*: A Review

Tomoaki Niimi

Abstract

Leishmania tarentolae is a trypanosomatid protozoan parasite of the gecko, and has been established as a new eukaryotic expression system for the production of recombinant proteins. It seems that a protozoan parasite is a curious choice as the expression host; however, Trypanosomatidae are rich in glycoproteins with a pattern of glycosylation closely related to those in mammals and higher vertebrates. Thus, one of the main advantages of a *L. tarentolae* expression system is the mammalian-type posttranslational modification of target proteins. Although there are few examples of recombinant protein expression using this system, it can be an attractive alternative to using mammalian cells. This chapter presents an overview of the newly developed protein expression system based on *L. tarentolae*.

Key words: Trypanosomatidae, Protozoan parasite, Mammalian-type posttranslational modification, Glycosylation

1. Introduction

The protozoan parasites of the genera *Leishmania* and *Trypanosoma* are members of the family Trypanosomatidae (*Eugelenozoa*, *Kinetoplastida*), which comprises unicellular organisms characterized by the presence of a single flagellum and of a DNA-rich, mitochondria-like organelle, the kinetoplast (1–3). They are causative agents of a wide spectrum of tropical diseases responsible for substantial human and livestock morbidity and mortality. Pathogenic species of *Leishmania* cause a diverse group of diseases, collectively called leishmaniases (4). According to the World Health Organization, more than 12 million people in 88 countries are affected by this disease and 350 million people are at risk of infection (http://www.who.int/leishmaniasis/burden/en/). In addition to their importance

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_15, © Springer Science+Business Media, LLC 2012

as agents of diseases, Trypanosomatidae have also been considered as attractive model organisms for studying some intracellular processes because they have unusual mechanisms of gene expression, such as polycistronic transcription and RNA editing. In these organisms, mRNAs are transcribed as polycistronic precursors that are posttranscriptionally processed into individual mRNAs by *trans*-splicing and polyadenylation (5–7). Thus, regulation of protein expression in these organisms occurs predominantly posttranscriptionally through the structure of the intergenic untranslated regions (UTRs). Furthermore, they have been valuable for studying molecular and cellular phenomena, such as glycosylphosphatidylinositol (GPI) anchoring of proteins, antigenic variation, and telomere organization (8–10). Due to their importance for public health, and as model organisms, methods of genetic manipulation for Trypanosomatidae are well-established (6, 11), which has led to several attempts of heterologous gene expression for studying the immunogenicity and biological activity of proteins from parasites (12). It is also known that Trypanosomatidae are rich in glycoproteins, with a pattern of glycosylation closely related to those of mammals and higher vertebrates (13). Possibly due to their parasitic lifestyle, the oligosaccharide structures of their glycoproteins may be similar to those in mammals; therefore, Trypanosomatidae have emerged as potential hosts for heterologous protein production.

Leishmania tarentolae is a nonpathogenic parasite of the white-spotted wall gecko *Tarentola annularis* (14), and has been established as a new eukaryotic expression system for the production of recombinant proteins with a mammalian-type N-glycosylation pattern (15, 16). Using this system, Breitling et al. reported that the yield of green fluorescent protein was up to 30 mg/L of culture. They also succeeded in obtaining biologically active human erythropoietin (EPO); a molecule which requires glycosylation and appropriate modification for activity (17), showing that the *L. tarentolae* expression system can be considered as an alternative to mammalian cells in cultures. This chapter discusses the advantages and disadvantages of this newly developed expression system based on *L. tarentolae*.

2. Genetic Manipulation

2.1. Media and Growth Conditions

L. tarentolae cells can be cultivated in complex media (brain–heart infusion based or yeast–soybean based) or chemically defined media, both supplemented with hemin, which is essential for growth of *Leishmania* (18–20). Hemin is the chlorinated derivative of the heme molecule and traditionally produced from bovine or porcine blood. Therefore, a complete defined medium without animal-derived products is not yet available. The cost of culture medium is not inexpensive.

L. tarentolae cells require aerobic conditions for development. The cells can be maintained as continuous suspension culture in ventilated tissue culture flasks. Conventional static cultures are incubated in the dark at 26°C. Agitated cultures for protein expression are incubated in an orbital shaker at 140 rpm using Erlenmeyer flasks. It is possible to scale up to larger bioreactors with volumes of 1–30 L. *L. tarentolae* cells can be indefinitely grown in vitro with a doubling time of around 5 h and to high cell densities (approximately 5×10^8 cells per mL) in complex media. The cultivation of *L. tarentolae* is much easier than that of mammalian cells.

2.2. Expression Vectors

In eukaryotes, protein-coding genes are transcribed by RNA polymerase II, whereas RNA polymerase I (Pol I) transcribes the genes that encode the three ribosomal RNAs (rRNAs). In Trypanosomatidae, however, translation of Pol I-transcribed protein-coding genes is possible because of *trans*-splicing of polycistronic pre-mRNAs (21). *Trans*-splicing adds a capped 39 nt spliced leader (or mini-exon) to the 5′ end of the mRNA, which is necessary for RNA transport, stability, and translation efficiency. Thus, in these organisms, the higher transcription rate of Pol I is used to ensure high levels of gene expression. In a *L. tarentolae* expression system, integration of the expression cassette into the chromosomal small subunit rRNA locus (ssu) enables the generation of large numbers of transcripts for constitutive expression. The expression cassette of pLEXSY-2 expression vectors for *L. tarentolae* is flanked by two fragments of the small subunit rRNA locus (5′ and 3′ ssu) for homologous recombination.

The minimal vector for gene expression in *Leishmania* consists of a circular plasmid containing a signal for *trans*-splicing of mRNA, followed by a target gene (22). Better expression can often be obtained if the target gene is followed by the 3′-UTR from a highly expressed gene. Since gene regulation in *Leishmania* occurs mainly posttranscriptionally through intergenic UTRs, the choice of suitable UTRs is important for construction of an efficient expression vector. The pLEXSY-2 vectors contain three optimized UTRs: *utr1* derived from the 0.4 kb intergenic region of *L. tarentolae* adenine phosphoribosyl transferase gene; *utr2* from the 1.4 kb intergenic region of *L. tarentolae* calmodulin cluster containing three tandemly arranged calmodulin genes (*camCB*); and *utr3* from the 1.7 kb intergenic region of the *L. major* dihydrofolate reductase-thymidylate synthase gene, flanking the target and marker gene insertion sites, which provide the *trans*-splicing signal for posttranscriptional mRNA processing.

In a *L. tarentolae* expression system, heterologous proteins can either be expressed cytosolically or secreted into the medium. Alternative cloning strategies result in cytosolic or secretory expression of target genes in pLEXSY-2 vectors. For secretion, a signal sequence derived from secreted acid phosphatase of *L. mexicana*

(lmsap) is utilized in pLEXSY-2 vectors (23); however, the native signal sequence has been used successfully for secretion of human EPO and rat proprotein convertase 4 (15, 24). Although the expression efficiency of recombinant EPO with lmsap was significantly higher than that with the native signal sequence, the usage of the native signal sequence enables the production of natively processed proteins at the N-terminus.

In conjunction with the unique organization of the transcription/translation machinery of Trypanosomatidae, the inducible protein expression system based on *L. tarentolae* can also be used (25). In Trypanosomatidae, expression of genes can be driven by not only endogenous RNA polymerase I or II, but also cointegrated promoters for heterologous RNA polymerase (26). Therefore, a combination of bacteriophage T7 RNA polymerase with its promoter controlled by a repressor element responsive to tetracycline (Tet) is used for the inducible expression system. Target genes are inserted into the expression cassette under the control of a T7 promoter with the bacterial Tet responsive element (TRE) and integrated into the chromosomal ornithine decarboxylase (*odc*) locus of the host *L. tarentolae* strain, which is constitutively expressing T7 RNA polymerase and the Tet repressor under the control of endogenous Pol I. When overexpression of heterologous proteins affects the physiology of the host, the inducible expression system may be valuable.

2.3. Gene Transfer and Selection

Following construction in *Escherichia coli*, the expression plasmid is linearized and integrated into the chromosome of *L. tarentolae* by homologous recombination. *Leishmania* cells can be routinely transfected with plasmid DNA by electroporation and the transfected cells can be selected with antibiotics. Currently, four selectable marker genes are available in this system: neomycin phosphotransferase (neo), hygromycin phosphotransferase (hyg), bleomycin resistance protein (ble), and streptothricin acetyltransferase (sat) that confer resistance to G418, hygromycin, bleomycin, and nourseothricin, respectively. Therefore, up to four genes can be expressed simultaneously and it is applicable, for instance, to produce multisubunit proteins (27). Moreover, further increase of expression levels can be obtained with the additional copies of integrated genes after sequential transfection and selection (15, 28).

3. Posttranslational Modifications

The structures of N-linked oligosaccharides from various pathogenic *Leishmania* and *Trypanosoma* species have been well-investigated because they have been implicated in parasite virulence (29).

According to previous studies, the pattern of N-glycosylation is highly variable in different species and at different life cycle stages (30, 31). Most N-linked glycans in these parasites are of the high-mannose type; however, many trypanosomatids retain the enzymatic machinery for making more complex N-glycan structures. As an example, some glycoproteins of *T. brucei* have complex N-linked glycans, primarily biantennary glycans with terminal a-1, 3-linked galactose units or branched poly-*N*-acetyllactosamine sequences similar to those of higher eukaryotes (32, 33). Some proteins of *T. cruzi* are modified with a-linked galactose, fucose, and even sialic acid residues (34, 35). The N-glycosylation profile of *L. tarentolae* was investigated in recombinant human EPO (15). It was exceptionally homogenous, with higher eukaryote-like biantennary N-glycans and the $Man_3GlcNAc_2$ core structure. N-glycans were fully galactosylated and core-α-1, 6-fucosylated, whereas sialylation was missing. Higher branched, tri- and tetra-antennary glycans were not observed, probably due to a lack of N-acetylglucosaminyl transferase IV activity. Overall, the N-glycosylation pathway of *L. tarentolae* is more similar to those in mammals than those in yeast and insect cells (Fig. 1). Genetic engineering of the *L. tarentolae* host may provide a method for producing sialylated or multiantennary heterologous glycoproteins.

L. tarentolae has the potential to perform other posttranslational modifications typically associated with higher eukaryotes. These include processing of signal sequences, protein folding, and disulfide bridge formation. We successfully produced the disulfide-linked heterotrimeric glycoprotein, laminin, which contained complex structures, such as epidermal growth factor (EGF)-like repeat and coiled-coil domains, in a *L. tarentolae* expression system (27).

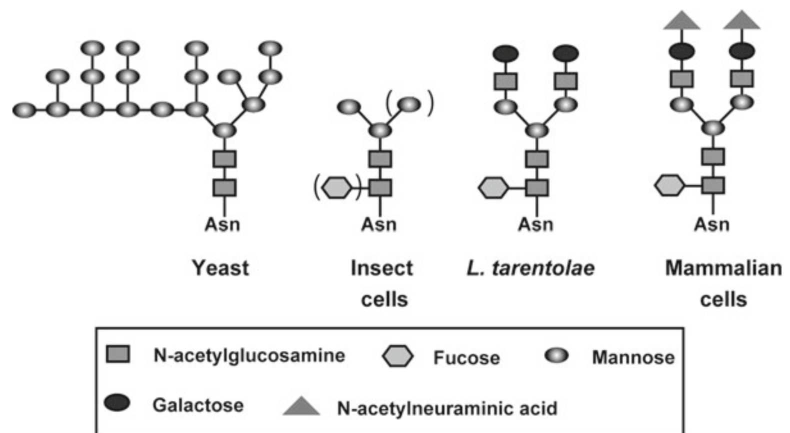

Fig. 1. N-glycosylation patterns of proteins expressed in various expression systems.

Table 1
Comparison of recombinant protein expression systems

System	Growth rate	Productivity	Glycosylation	Protein folding	Availability of genetic systems	Scale-up	Media cost
Prokaryotic							
Bacteria	Very fast	Very high	No	Poor	Very good	Very good	Very low
Eukaryotic							
Yeast	Fast	High	Simple	Fair	Good	Very good	Low
L. tarentolae	Medium	Medium	Mammalian like	Good	Fair	Good	Medium
Insect cells	Slow	High	Non-mammalian like	Very good	Fair	Fair	High
Mammalian cells	Slow	Low	Yes	Very good	Fair	Fair	High

4. Conclusions

Various recombinant protein expression systems have been developed in bacteria, yeast, fungi, insect cells, mammalian cells, transgenic animals, and transgenic plants. Researchers have selected the optimal expression system based on the purpose of the desired protein product. The selection of the expression system depends on numerous factors, such as the expression levels, posttranslational modifications, production cost, availability of genetic systems, and other factors (Table 1) (36, 37). Thus, it is important to understand the advantages and disadvantages of the expression systems. At present, the *L. tarentolae* expression system provides an attractive alternative to mammalian cells for the production of recombinant mammalian proteins. When it is difficult to express a mammalian protein or domain of interest in mammalian cells, it would be advantageous to use the *L. tarentolae* expression system. Furthermore, it can be used for expression of proteins that are difficult to express in bacteria or yeast. Since the *L. tarentolae* expression system has been recently commercialized by Jena Bioscience (http://www.jenabioscience.com), there are still few examples of this recombinant protein expression (15, 24, 27, 28, 38–40). Although further investigation is required to verify a relationship between posttranslational modifications and biological activity of many other proteins, *L. tarentolae* is a promising host for production of proteins with mammalian-type posttranslational modifications. The genomes of several Trypanosomatidae have been sequenced (41–44) and are available from sources, such as TriTrypDB (http://TriTrypDB.org) (45), including the partial sequence of *L. tarentolae*. The availability of the genome sequence should facilitate future improvement of this system for large-scale production of recombinant proteins for industrial and pharmaceutical uses.

Acknowledgments

The preparation of this chapter was supported by a Grant-in-aid (no. 18108003) for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

References

1. Shapiro, T. A., and Englund, P. T. (1995) The structure and replication of kinetoplast DNA. *Annu. Rev. Microbiol.* **49**, 117–143.
2. Morris, J. C., Drew, M. E., Klingbeil, M. M., Motyka, S. A., Saxowsky, T. T., Wang, Z., et al. (2001) Replication of kinetoplast DNA: An update for the new millennium. *Int. J. Parasitol.* **31**, 453–458.
3. Liu, B., Liu, Y., Motyka, S. A., Agbo, E. E., and Englund, P. T. (2005) Fellowship of the rings: the replication of kinetoplast DNA. *Trends Parasitol.* **21**, 363–369.

4. Banuls, A. L., Hide, M., and Prugnolle, F. (2007) *Leishmania* and the leishmaniases: A parasite genetic update and advances in taxonomy, epidemiology and pathogenicity in humans. *Adv. Parasitol.* **64**, 1–109.
5. Simpson, L., and Maslov, D. A. (1999) Evolution of the U-insertion/deletion RNA editing in mitochondria of kinetoplastid protozoa. *Ann. N. Y. Acad. Sci.* **870**, 190–205.
6. Clayton, C. E. (1999) Genetic manipulation of kinetoplastida. *Parasitol. Today* **15**, 372–378.
7. Clayton, C., and Shapira, M. (2007) Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol. Biochem. Parasitol.* **156**, 93–101.
8. Krakow, J. L., Hereld, D., Bangs, J. D., Hart, G. W., and Englund, P. T. (1986) Identification of a glycolipid precursor of the *Trypanosoma brucei* variant surface glycoprotein. *J. Biol. Chem.* **261**, 12147–12153.
9. Blackburn, E. H. (1991) Structure and function of telomeres. *Nature* **350**, 569–573.
10. Borst, P., and Rudenko, G. (1994) Antigenic variation in African trypanosomes. *Science* **264**, 1872–1873.
11. Beverley, S. M. (2003) Protozomics: Trypanosomatid parasite genetics comes of age. *Nat. Rev. Genet.* **4**, 11–19.
12. Fernandez-Robledo, J. A., and Vasta, G. R. (2010) Production of recombinant proteins from protozoan parasites. *Trends Parasitol.* **26**, 244–254.
13. Ferguson, M. A. (1999) The structure, biosynthesis and functions of glycosylphosphatidylinositol anchors, and the contributions of trypanosome research. *J. Cell Sci.* **112**, 2799–2809.
14. Simpson, L., Frech, G. C., and Maslov, D. A. (1996) RNA editing in trypanosomatid mitochondria. *Methods Enzymol.* **264**, 99–121.
15. Breitling, R., Klingner, S., Callewaert, N., Pietrucha, R., Geyer, A., Ehrlich, G., et al. (2002) Non-pathogenic trypanosomatid protozoa as a platform for protein research and production. *Protein Expr. Purif.* **25**, 209–218.
16. Basile, G., and Peticca, M. (2009) Recombinant protein expression in *Leishmania tarentolae*. *Mol. Biotechnol.* **43**, 273–278.
17. Wasley, L. C., Timony, G., Murtha, P., Stoudemire, J., Dorner, A. J., Caro, J., et al. (1991) The importance of N- and O-linked oligosaccharides for the biosynthesis and in vitro and in vivo biologic activities of erythropoietin. *Blood* **77**, 2624–2632.
18. Pal, J. K., and Joshi-Purandare, M. (2001) Dose-dependent differential effect of hemin on protein synthesis and cell proliferation in *Leishmania donovani* promastigotes cultured in vitro. *J. Biosci.* **26**, 225–231.
19. Fritsche, C., Sitz, M., Weiland, N., Breitling, R., and Pohl, H. D. (2007) Characterization of the growth behavior of *Leishmania tarentolae*: A new expression system for recombinant proteins. *J. Basic Microbiol.* **47**, 384–393.
20. Fritsche, C., Sitz, M., Wolf, M., and Pohl, H. D. (2008) Development of a defined medium for heterologous expression in *Leishmania tarentolae*. *J. Basic Microbiol.* **48**, 488–495.
21. Lee, M. G., and Van der Ploeg, L. H. (1997) Transcription of protein-coding genes in trypanosomes by RNA polymerase I. *Annu. Rev. Microbiol.* **51**, 463–489.
22. Curotto de Lafaille, M. A., Laban, A., and Wirth, D. F. (1992) Gene expression in *Leishmania*: Analysis of essential 5′ DNA sequences. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2703–2707.
23. Wiese, M., Ilg, T., Lottspeich, F., and Overath, P. (1995) Ser/Thr-rich repetitive motifs as targets for phosphoglycan modifications in *Leishmania mexicana* secreted acid phosphatase. *EMBO J.* **14**, 1067–1074.
24. Basak, A., Shervani, N. J., Mbikay, M., and Kolajova, M. (2008) Recombinant proprotein convertase 4 (PC4) from *Leishmania tarentolae* expression system: purification, biochemical study and inhibitor design. *Protein Expr. Purif.* **60**, 117–126.
25. Kushnir, S., Gase, K., Breitling, R., and Alexandrov, K. (2005) Development of an inducible protein expression system based on the protozoan host *Leishmania tarentolae*. *Protein Expr. Purif.* **42**, 37–46.
26. Wirtz, E., Hartmann, C., and Clayton, C. (1994) Gene expression mediated by bacteriophage T3 and T7 RNA polymerases in transgenic trypanosomes. *Nucleic Acids Res.* **22**, 3887–3894.
27. Phan, H. P., Sugino, M., and Niimi, T. (2009) The production of recombinant human laminin-332 in a *Leishmania tarentolae* expression system. *Protein Expr. Purif.* **68**, 79–84.
28. Gazdag, E. M., Cirstea, I. C., Breitling, R., Lukes, J., Blankenfeldt, W., and Alexandrov, K. (2010) Purification and crystallization of human Cu/Zn superoxide dismutase recombinantly produced in the protozoan *Leishmania tarentolae*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **66**, 871–877.
29. Descoteaux, A., and Turco, S. J. (1999) Glycoconjugates in *Leishmania* infectivity. *Biochim. Biophys. Acta* **1455**, 341–352.
30. Parodi, A. J. (1993) N-glycosylation in trypanosomatid protozoa. *Glycobiology* **3**, 193–199.

31. McConville, M. J., Mullin, K. A., Ilgoutz, S. C., and Teasdale, R. D. (2002) Secretory pathway of trypanosomatid parasites. *Microbiol. Mol. Biol. Rev.* **66**, 122–154.
32. Zamze, S. (1991) Glycosylation in parasitic protozoa of the trypanosomatidae family. *Glycoconj. J.* **8**, 443–447.
33. Mehlert, A., Zitzmann, N., Richardson, J. M., Treumann, A., and Ferguson, M. A. (1998) The glycosylation of the variant surface glycoproteins and procyclic acidic repetitive proteins of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **91**, 145–152.
34. Couto, A. S., Goncalves, M. F., Colli, W., and de Lederkremer, R. M. (1990) The N-linked carbohydrate chain of the 85-kilodalton glycoprotein from *Trypanosoma cruzi* trypomastigotes contains sialyl, fucosyl and galactosyl (alpha 1–3)galactose units. *Mol. Biochem. Parasitol.* **39**, 101–107.
35. Parodi, A. J., Labriola, C., and Cazzulo, J. J. (1995) The presence of complex-type oligosaccharides at the C-terminal domain glycosylation site of some molecules of cruzipain. *Mol. Biochem. Parasitol.* **69**, 247–255.
36. Sodoyer, R. (2004) Expression systems for the production of recombinant pharmaceuticals. *BioDrugs* **18**, 51–62.
37. Cha, H. J., Dalal, N. N., and Bentley, W. E. (2005) Secretion of human interleukin-2 fused with green fluorescent protein in recombinant *Pichia pastoris*. *Appl. Biochem. Biotechnol.* **126**, 1–11.
38. Soleimani, M., Mahboudi, F., Davoudi, N., Amanzadeh, A., Azizi, M., Adeli, A., et al. (2007) Expression of human tissue plasminogen activator in the trypanosomatid protozoan *Leishmania tarentolae*. *Biotechnol. Appl. Biochem.* **48**, 55–61.
39. Dortay, H., and Mueller-Roeber, B. (2010) A highly efficient pipeline for protein expression in *Leishmania tarentolae* using infrared fluorescence protein as marker. *Microb. Cell Fact.* **9**, 29.
40. Ben-Abdallah, M., Bondet, V., Fauchereau, F., Beguin, P., Goubran-Botros, H., Pagan, C., et al. (2011) Production of soluble, active acetyl serotonin methyl transferase in *Leishmania tarentolae*. *Protein Expr. Purif.* **75**, 114–118.
41. Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., et al. (2005) The genome of the African trypanosome *Trypanosoma brucei*. *Science* **309**, 416–422.
42. El-Sayed, N. M., Myler, P. J., Bartholomeu, D. C., Nilsson, D., Aggarwal, G., Tran, A. N., et al. (2005) The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* **309**, 409–415.
43. Ivens, A. C., Peacock, C. S., Worthey, E. A., Murphy, L., Aggarwal, G., Berriman, M., et al. (2005) The genome of the kinetoplastid parasite, *Leishmania major*. *Science* **309**, 436–442.
44. Peacock, C. S., Seeger, K., Harris, D., Murphy, L., Ruiz, J. C., Quail, M. A., et al. (2007) Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nat. Genet.* **39**, 839–847.
45. Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B. P., Carrington, M., et al. (2009) TriTrypDB: A functional genomic resource for the Trypanosomatidae. *Nucleic Acids Res.* **38**, D457-462.