Control of mRNA degradation in trypanosomes

Christine Clayton*‡[1], Angela Schwede*, Mhairi Stewart*, Ana Robles*, Corinna Benz*, June Po*, Martin Wurst*, Rafael Quieroz‡ and Stuart Archer*

*ZMBH (Zentrum für Molekulare Biologie der Universität Heidelberg), Im Neuenheimer Feld 282, D69120 Heidelberg, Germany, †Division of Functional Genome Analysis, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 506, 69120 Heidelberg, Germany and ‡DKFZ-ZMBH Allianz, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

Abstract

Control of gene expression in trypanosomes relies almost exclusively on post-transcriptional mechanisms. Trypanosomes have the normal enzymes for mRNA decay: both the exosome and a 5′–3′-exoribonuclease are important in the degradation of very unstable transcripts, whereas the CAF1/NOT complex plays a major role in the degradation of all mRNAs tested. Targeted RNA interference screening was used to identify RNA-binding proteins that regulate mRNA degradation, and it revealed roles for proteins with RNA recognition motifs or pumilio domains.

Introduction

The kinetoplastids are flagellated protists. The most studied members of the order are the leishmanias and trypanosomes, which cause diseases in humans and other mammals. The kinetoplastids branched very early in eukaryotic evolution from the lines leading to yeast and mammals, plants and many other unicellular organisms. As a consequence, they exhibit a variety of unique features and specializations. One of these is an almost total dependence on post-transcriptional mechanisms for the regulation of gene expression [1].

Kinetoplastid genes are strangely arranged: adjacent open reading frames are located close to each other, oriented in the same direction, and are co-transcribed, so that individual mRNAs have to be generated by processing. Results of genome sequencing [2] demonstrated that gene arrays may contain 100 genes or more. Transcription initiates at several sites within the discontinuities between the arrays ('strand-switch regions'), but there is no consensus promoter sequence there [3]. The 5′-end of each mRNA is created by a *trans* splicing reaction, in which a capped 'leader' of approx. 40 nt is added; there is so far only one documented instance of *cis* splicing. The 3′-end of each mRNA is generated, as usual, by polyadenylation, but there is no polyadenylation signal in the 3′-UTR (3′-untranslated region): instead, the polyadenylation site is determined by measuring back from the next downstream *trans* splicing site [4].

The polycistronic nature of transcription clearly precludes control of expression of individual genes at the level of transcription initiation. Correspondingly, the kinetoplastid genomes are remarkably deficient in transcription factors,

retaining only those required for the basal machinery [2]. Only in the African trypanosomes, which use RNA polymerase I to transcribe genes encoding the major surface proteins, is any mRNA-specific transcription control found. The trypanosomes express one set of surface protein genes when they are multiplying in the bloodstream of the mammalian host ('bloodstream form') and another set in the form that thrives in the tsetse fly vector ('procyclic form'). In each case, the transcription rate is controlled, almost certainly by changes in chromatin structure (see e.g. [5]).

Despite the lack of transcription control, multiple options are available for kinetoplastids to regulate gene expression: control of splicing, polyadenylation, export from the nucleus, RNA compartmentation, translation, and degradation of both protein and mRNA. Also, high levels of mRNA and protein can be obtained through increases in gene copy number. We have concentrated mainly on mRNA degradation as a means to control gene expression.

Kinetics of mRNA processing and degradation in trypanosomes

Work in many laboratories has shown that the abundances of trypanosome mRNAs are strongly influenced by the sequence of the 3′-UTR [1]. Usually, it is found that the 3′-UTR determines the mRNA degradation rate; possible contributions of processing to control and regulation have been less investigated [1]. For example, the genes encoding two developmentally regulated isoenzymes of PGK (phosphoglycerate kinase), PGKB and PGKC, are immediately adjacent in the genome and are co-transcribed, but show very different levels of regulation. In both cases, the mRNA half-life ($t_{1/2}$) is determined by the 3′-UTR [6]. To investigate the relative contributions of processing and degradation to the steady-state abundances of the *PGKB* and *PGKC* mRNA and protein, we used kinetic measurements to generate a mathematical model for *PGK* gene expression. This showed that the major determinant of *PGK* expression was indeed the mRNA $t_{1/2}$ [7].

The enzymes that degrade mRNA in trypanosomes

In yeast and mammalian cells, mRNA degradation is classically initiated by removal of the poly(A) tail. After that, the 5′-cap is removed by a decapping enzyme, then the mRNA can be degraded from both ends: in the 5′–3′-direction by the exoribonuclease XRN1 and from the 3′-end by the exosome. Trypanosomes have all three of the deadenylation systems found in mammals: the CAF1/NOT complex, PAN2/PAN3 and three possible homologues of PARN [poly(A)-specific ribonuclease]. Our recent results indicate that both CAF1 and PAN2 are involved in deadenylation ([8] and A. Schwede and C.E. Clayton, unpublished work). Decapping enzymes have, in contrast, proved elusive. Interestingly, there are three 5′–3′-exonucleases related to XRN1. XRNA is in both the nucleus and the cytoplasm, and our results showed that it plays a major role in the degradation of very unstable mRNAs. Indeed, results of RNAi (RNA interference) experiments, combined with investigations of mRNA degradation intermediates, indicated that highly unstable, developmentally regulated mRNAs may be attacked directly from the 5′-end, without prior deadenylation [9]. The functions of the other two proteins, XRNB and XRNC, both of which are in the cytoplasm, are as yet unknown.

A survey of trypanosome RNA-binding proteins

We are using conditional RNAi to screen for proteins that regulate mRNA degradation in trypanosomes, with the expectation that such proteins will be essential. In mammalian cells, proteins that modulate mRNA stability generally have one or more RNA-binding domains: RRM (RNA recognition motif); KH domain; CCCH zinc finger; and Puf (pumilio) domain. The trypanosome genome encodes 74 RRM proteins, 39 CCCH proteins and 12 Puf proteins; in most cases, their roles in RNA metabolism cannot be predicted from the sequence. Using RNAi we have found that two of nine Puf proteins ([10] and S. Archer, unpublished work), 25 of 38 RRM proteins (M. Wurst, J. Po, M. Marentije, S. Stoitsova, L. Quijoda, M. Stewart, C. Hartmann and C.E. Clayton, unpublished work) and one of four CCCH proteins (C. Benz and M. Stewart, unpublished work) tested are required for normal growth. Purification of tagged versions, followed by hybridization of the bound RNA to microarrays, has enabled us to show that some of the proteins specifically bind to a small subset of mRNAs, and control their abundance. Three very small proteins, UBP (U-rich-RNA-binding protein) 1, UBP2 and RBP3, consist mainly of the RRM and are far more abundant than mRNA on a molar basis; RNAi or overexpression nevertheless influences specific subsets of RNAs ([11] and A. Robles and R. Quieroz, unpublished work). Most interestingly, one of the Puf proteins is required to control mRNAs that accumulate in the G_2-phase of the cell cycle (S. Archer and R. Quieroz, unpublished work).

References

1 Clayton, C. and Shapira, M. (2007) Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. Mol. Biochem. Parasitol. **156**, 93–101

2 Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.-A., Adlem, E., Aert, R. et al. (2005) The genome of the kinetoplastid parasite, *Leishmania major*. Science **309**, 436–442

3 Palenchar, J.B. and Bellofatto, V. (2006) Gene transcription in trypanosomes. Mol. Biochem. Parasitol. **146**, 135–141

4 Liang, X., Haritan, A., Uliel, S. and Michaeli, S. (2003) *Trans* and *cis* splicing in trypanosomatids: mechanism, factors, and regulation. Eukaryotic Cell **2**, 830–840

5 Glover, L. and Horn, D. (2006) Repression of polymerase I-mediated gene expression at *Trypanosoma brucei* telomeres. EMBO Rep. **7**, 93–99

6 Colasante, C., Robles, A., Li, C.-H., Schwede, A., Benz, C., Voncken, F., Guilbride, D.L. and Clayton, C. (2007) Regulated expression of glycosomal phosphoglycerate kinase in *Trypanosoma brucei*. Mol. Biochem. Parasitol. **151**, 193–204

7 Haanstra, J., Stewart, M., Luu, V.-D., van Tuijl, A., Westerhoff, H., Clayton, C. and Bakker, B. (2008) Control and regulation of gene expression: quantitative analysis of the expression of phosphoglycerate kinase in bloodstream form *Trypanosoma brucei*. J. Biol. Chem. **283**, 2495–2507

8 Schwede, A., Ellis, L., Luther, J., Carrington, M., Stoecklin, G. and Clayton, C.E. (2008) A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells. Nucleic Acids Res., doi:10.1093/nar/gkn108

9 Li, C.-H., Irmer, H., Gudjonsdottir-Planck, D., Freese, S., Salm, H., Haile, S., Estévez, A.M. and Clayton, C.E. (2006) Roles of a *Trypanosoma brucei* 5′ → 3′ exoribonuclease homologue in mRNA degradation. RNA **12**, 2171–2186

10 Luu, V.D., Brems, S., Hoheisel, J., Burchmore, R., Guilbride, D. and Clayton, C. (2006) Functional analysis of *Trypanosoma brucei* PUF1. Mol. Biochem. Parasitol. **150**, 340–349

11 Hartmann, C., Benz, C., Brems, S., Ellis, L., Luu, V.-D., Stewart, M., D'Orso, I., Busold, C., Fellenberg, K., Frasch, A.C.C. et al. (2007) The small trypanosome RNA-binding proteins TbUBP1 and TbUBP2 influence expression of F box protein mRNAs in bloodstream trypanosomes. Eukaryotic Cell **6**, 1964–1978