

Specialised RNA decay finetunes monogenic antigen expression in African Trypanosomes

Htay Mon-Aye¹, Lianne I M Lansink¹, Leon Walther¹, Sophie Longmore¹, Madeleine Jones¹, Adam Dowle¹, Joao Cunha-Reis¹, Joana R. C. Faria¹

¹York Biomedical Research Institute, Department of Biology, University of York, Heslington YO10 5DD, York, UK

Antigenic variation is a sophisticated immune evasion strategy employed by many pathogens. *Trypanosoma brucei* expresses a single Variant-Surface-Glycoprotein (VSG) from a large genetic repertoire, which they periodically switch throughout an infection. Co-transcribed with the active-VSG within a specialised nuclear body are expression-site-associated-genes (*ESAGs*), involved in important host-parasite interactions, including protecting the parasite from human serum lytic effects, modulating the host's innate immune response and uptake of essential nutrients. Despite expression within the same polycistron, within the expression-site body (ESB), there is a significant differential expression between *ESAGs* and *VSGs* (>140-fold), however, the regulatory mechanism has remained elusive for decades.

Here, we applied TurboID-mediated proximity labelling mass spectrometry (PL-MS) to map the ESB post-transcriptional network, identifying three new components: ESB-associated protein 1 (ESAP1) and ESB-specific proteins 2 and 3 (ESB2 and 3). These are recruited in a hierarchical manner, forming discreet sub-nuclear condensates that are developmentally regulated between insect and mammalian stages of the parasite. Inducible overexpression and RNAi-mediated depletion experiments followed by transcriptomic analyses showed that ESAP1/ESB2-3 negatively regulate *ESAG* transcripts.

ESB2 is a predicted endonuclease structurally similar to SMG6, a critical component of nonsense-mediated decay in mammals. We validated this function *in vitro*, observing that the purified nuclease domain (aa 95-C) acts in a concentration- and magnesium-dependent manner. *In vivo* CRISPR/Cas9-mediated precision editing of catalytic residues impaired both ESB2 localisation and its ability to regulate *ESAG* expression, confirming that these processes are nuclease activity dependent.

To distinguish whether ESB2 targets *ESAGs* via sequence recognition or through a position-dependent mechanism mediated by spatial restriction of its activity, we integrated an *RFP::PAC* reporter flanked by tubulin UTRs (unrelated DNA elements) or *ESAG* UTRs in different locations within the active *VSG* expression site and found that ESB2's activity is spatially regulated and not driven by sequence-specificity. Moreover, we found that ESB2's recruitment to the ESB relies on a continuous flux of active transcription and RNA processing, its own catalytic activity and ESB3, as shown by co-immunoprecipitation.

Overall, we identified a novel post-transcriptional mechanism of gene expression control, underscoring how specialised RNA decay can regulate expression of specific genes.

Finally, to comprehensively profile the ESB proteome, we developed a "saturation proximity labeling" approach by integrating newly identified components into our PL-MS pipeline. Using a custom bioinformatic workflow and CRISPR/Cas9-based endogenous tagging, we screened 40 high-priority candidates for subcellular localisation. This revealed new components of the ESB and adjacent nuclear bodies. We are now characterising these factors and utilising ultrastructure expansion microscopy (UEXM) to define the architectural organisation of the ESB at high resolution.