Cellular Microbiology (2015) doi:10.1111/cmi.12467

Breaking report

Different proteomic strategies to identify genuine Small Ubiquitin-like MOdifier targets and their modification sites in *Trypanosoma brucei* procyclic forms

P. A. Iribarren,[1] M. A. Berazategui,[1] J. C. Bayona,[1] I. C. Almeida,[2] J. J. Cazzulo[1] and V. E. Alvarez[1]*

[1]*Instituto de Investigaciones Biotecnológicas Dr. Rodolfo A. Ugalde-Instituto Tecnológico de Chascomús (IIB-INTECH), Universidad Nacional de San Martín (UNSAM)–Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Campus Miguelete, Av. 25 de Mayo y Francia, 1650 San Martín, Buenos Aires, Argentina.*

[2]*The Border Biomedical Research Center, Department of Biological Sciences, University of Texas at El Paso, El Paso, TX 79968, USA.*

Summary

SUMOylation is an important post-translational modification conserved in eukaryotic organisms. In *Trypanosoma brucei*, SUMO (Small Ubiquitin-like MOdifier) is essential in procyclic and bloodstream forms. Furthermore, SUMO has been linked to the antigenic variation process, as a highly SUMOylated focus was recently identified within chromatin-associated proteins of the active variant surface glycoprotein expression site. We aimed to establish a reliable strategy to identify SUMO conjugates in *T. brucei*. We expressed various tagged variants of SUMO from the endogenous locus. His-HA-*Tb*SUMO was useful to validate the tag functionality but SUMO conjugates were not enriched enough over contaminants after affinity purification. A Lys-deficient SUMO version, created to reduce contaminants by Lys-C digestion, was able to overcome this issue but did not allow mapping many SUMOylation sites. This cell line was in turn useful to demonstrate that polySUMO chains are not essential for parasite viability. Finally, a His-HA-*Tb*SUMO^{T106K} version allowed the purification of SUMO conjugates and, after digestion with Lys-C, the enrichment for diGly-Lys peptides using specific antibodies. This site-specific proteomic strategy led us to identify 45 SUMOylated proteins and 53 acceptor sites unambiguously. SUMOylated proteins belong mainly to nuclear processes, such as DNA replication and repair, transcription, rRNA biogenesis and chromatin remodelling, among others.

Introduction

SUMOylation is a reversible post-translational modification that involves the covalent attachment of a Small Ubiquitin-like MOdifier (SUMO) to certain lysine residues within a diverse range of target proteins. The SUMOylation cycle begins with the proteolytic processing of a SUMO precursor by specific peptidases. The cleavage of a short C-terminal fragment after a conserved diglycine (diGly) motif enables SUMO to be first activated by the E1-activating enzyme and next conjugated to its substrate proteins by the E2-conjugating enzyme, in occasions with the assistance of a variety of E3-ligases. Finally, SUMO is linked to its targets by the formation of an isopeptide bond between the carboxyl group of the C-terminal glycine residue in SUMO and the ε-amino group of the lysine residue in the target proteins. The modified lysine residue is usually embedded in a consensus motif, being preceded by a large hydrophobic residue and followed by any amino acid and then a glutamic or aspartic acid residue. Protein SUMOylation can be reverted by the action of specific peptidases in a highly dynamic process. The downstream effects of SUMO conjugation can be attributed to the modification of the surface interaction properties of SUMOylated proteins, and possible outcomes include changes in biological activity, subcellular localization or protein stability (Flotho and Melchior, 2013).

Received 1 April, 2015; revised 1 June, 2015; accepted 5 June, 2015.
*For correspondence. E-mail valvarez@iib.unsam.edu.ar; Tel. (+54) 11 4006 1500 ext 2123; Fax (+54) 11 4006 1559.

© 2015 John Wiley & Sons Ltd

The SUMOylation pathway is strictly conserved in eukaryotic organisms and has been proven to be essential in many of them. Trypanosomes belong to an early branch of the eukaryotic lineage and previous works from our group and others have demonstrated that they possess a functional SUMOylation system (Liao *et al*., 2010; Bayona *et al*., 2011; Obado *et al*., 2011; Annoura *et al*., 2012; Klein *et al*., 2013; Lopez-Farfan *et al*., 2014). These protozoan parasites have a major impact on human and animal health. *Trypanosoma cruzi* is the aetiological agent of Chagas disease, a chronic human disabling illness endemic in Latin America, whereas different *T. brucei* subspecies cause sleeping sickness in humans and nagana in cattle, both prevalent in sub-Saharan African countries (Barrett *et al*., 2003). *T. brucei* is a model organism amenable for genetic manipulation and by performing knock-down experiments SUMO was shown to be essential for both procyclic (PCF) and bloodstream (BSF) forms of the parasite, which are the replicative forms present in the tsetse fly vector and the infected mammal respectively (Liao *et al*., 2010; Obado *et al*., 2011). Furthermore, SUMO has been recently linked to *T. brucei* antigenic variation, a process where the major surface antigenic protein is replaced by a different variant with certain frequency as a strategy to elude the specific immune response of the host. SUMO was found to be enriched in a particular region of the nucleus of BSF parasites together with the E3 ligase Siz1 and the RNApolI, specifically at the chromatin region that is actively transcribing the variant surface glycoprotein (VSG), suggesting that SUMOylated unknown factors are involved in VSG monoallelic active expression (Lopez-Farfan *et al*., 2014).

The isolation of SUMO conjugates is usually hampered by the rapid action of SUMO isopeptidases, in particular when non-denaturing conditions are used either because the purification principles to be applied require it or because co-immunoprecipitation experiments (using anti-target and anti-SUMO antibodies) to demonstrate the SUMOylation of a particular target are undertaken. Trypanosomatids are not the exception, and so far, only three SUMO targets have been identified in cell-free extracts obtained under non-denaturing conditions using N-ethylmaleimide (NEM) to inhibit SUMO-specific proteases: *Tc*MCA3 (Bayona *et al*., 2011) in *T. cruzi*, and Aurora kinase B (Hu *et al*., 2014) and RPA1 (Lopez-Farfan *et al*., 2014) in *T. brucei.* On the contrary, other groups reported the inability to obtain SUMOylated targets even when high concentrations of NEM were used in the extraction buffers (Annoura *et al*., 2012; Klein *et al*., 2013). In this work, we aimed to establish a reliable procedure to identify at a proteome-wide level the subset of SUMOylated proteins from *T. brucei* PCF. For that purpose, we evaluated the performance of transgenic cell lines expressing different SUMO variants and conclude that the use of His-HA-*Tb*SUMO^{T106K} cell line, engineered to generate a Lys-C digestion site before the diGly motif that enables the purification of SUMO modified proteins and the identification of the SUMOylated lysine residues on target proteins after the specific enrichment of SUMO modified peptides with diGly-Lys specific antibodies, allows the identification of SUMOylated proteins together with their acceptor sites in an unambiguous manner. This protocol was applied to PCF – since their culture can reach 30-fold higher cell densities than BSF – but can be certainly extrapolated to other parasite stages or growth conditions.

Results

*Generation of His-HA-*Tb*SUMO and Lys-deficient His-HA-*Tb*SUMO procyclic cell lines and purification of SUMO conjugates*

To streamline the purification of SUMO conjugates from *T. brucei*, we developed a procyclic cell line expressing an 8xHis and HA-tagged version of SUMO replacing the two endogenous genes (*T. brucei* is diploid) by successive rounds of homologous recombination (Supporting Information Fig. S1). Transgenic cell lines had a doubling time comparable to that of the parental strain, were able to form conjugates, and its subcellular distribution in the nuclear periphery was similar to that in wild-type parasites.

Small Ubiquitin-like MOdifier (SUMO) conjugates were purified from parasites resuspended in 6 M urea by nickel-affinity chromatography under denaturing conditions, which is useful not only to abrogate the action of deconjugating peptidases but also to eliminate SUMO non-covalently interacting proteins. After eluting with imidazole, the urea concentration was decreased and the SUMOylated proteins were further subjected to a second step of affinity chromatography using anti-HA agarose resin. Although SUMO conjugates were efficiently recovered, there was a considerable overlap between the experimental and control purified samples, and further experiments showed that the apparent protein targets likely represent contaminant proteins. This cell line was useful, however, since it showed that the addition of the tags does not alter the functionality of SUMO, which is essential in this organism (Liao *et al*., 2010; Obado *et al*., 2011).

Since these results confirmed that contaminant proteins are a major issue in the identification of SUMOylated proteins, we decided to implement a complementary approach (Matic *et al*., 2010) generating a new variant, named His-HA-*Tb*SUMOK9R^{T106R}, in which all lysine residues were replaced by arginine and, in addition, an arginine residue prior to the diGly motif was added in the

C-terminus of SUMO, creating a new trypsin cleavage site. Thus, SUMO becomes resistant to Lys-C digestion, a protease that cleaves proteins at the C-terminal side of lysine residues, while contaminant proteins can be degraded (Supporting Information Fig. S2). At the same time, the acceptor lysine in substrates can be mapped by tandem mass spectrometry (MS/MS) by locating the remnant diGly after a subsequent trypsin digestion. Procyclic parasites transfected with the His-HA-*Tb*SUMOK9R^{T106R} construct with a double replacement of the endogenous SUMO behaved as those expressing the 8xHis and HA-tagged version, despite the fact that this mutant has lost the ability to form polySUMO chains, thus suggesting that polySUMOylation does not play an essential role in this parasite.

However, when these samples were processed, liquid chromatography–tandem mass spectrometry (LC-MS/MS) revealed the presence of very few diGly-modified peptides. This fact suggested that the inclusion of an additional step conceived to enrich the specific diGly peptides would help to increase their identification, as described in the following section.

*Unambiguous identification of SUMOylation sites in His-HA-*Tb*SUMO^{T106K}*

To increase the number of SUMOylated lysine residues identified in the target proteins, we applied a recently developed approach (Tammsalu *et al*., 2014) that involves the enrichment of SUMOylated peptides with anti-diGly remnant (K-ε-GG) antibodies (Fig. 1A). We generated a cell line stably expressing His-HA-*Tb*SUMO with a mutation of the Thr residue at position 106 (prior to the diGly motif) to Lys and verified the performance of this variant His-HA-*Tb*SUMO^{T106K} to conjugate as WT SUMO (Fig. 1B–E) similar to what has been described for the previous cell lines. Parasites were lysed in denaturing buffer, SUMOylated proteins were affinity-purified by nickel chromatography and after digestion of the eluates with Lys-C, diGly-Lys containing peptides were immunocaptured with the specific antibodies.

This strategy led to a remarkable increase in the frequency of identification of acceptor lysine residues as reported in the original method (Tammsalu *et al*., 2014). As shown in Table 1 and Supporting Information Table S1, we were able to identify 53 sites in 45 proteins, the majority of the diGly peptides being detected for different proteins (Flotho and Melchior, 2013), whereas for two proteins (H4 and RRP6) two sites, and for another two proteins (Topo1B and H2A), four sites were detected. Out of the 53 mapped SUMO acceptor sites, 24 were in agreement with the previously established consensus for SUMOylation, ψKxE/D, 5 contained E or D at position +2 and 6 were situated within an inverted motif (Fig. 2).

Out of the 45 proteins, 19 correspond to hypothetical proteins, whereas the remaining 26 have a functional assignment either based on some experimental evidence or predicted by sequence homology. If we do not consider the hypothetical proteins, we could say that the proteins present are in general nuclear, in good agreement with the nuclear localization of *Tb*SUMO conjugates, as judged by Indirect Immunofluorescence (IFI) analysis of the parasites. It is very interesting that many of the identified proteins are typically SUMOylated in other eukaryotic organisms or belong to cellular pathways modulated by SUMO. The results are discussed in more detail in the Discussion section.

Discussion

Proteomic identification of SUMOylated proteins is a challenging task not only because usually a minor fraction of a given target is actually modified by SUMO but also because this modification is rapidly lost if the specific isopeptidases are not inactivated. This explains why anti-SUMO polyclonal or monoclonal antibodies can only rarely be used to isolate SUMO conjugates by immuno-affinity chromatography (Becker *et al*., 2013). The most commonly employed approach involves the generation of transgenic cell lines expressing His-tagged SUMO that enables an affinity purification step using heavy denaturing conditions such as 6 M urea (Tatham *et al*., 2009; Hendriks *et al*., 2014; Tammsalu *et al*., 2014). This can be usually combined with another tag, for a second affinity purification step, and/or with strategies to increase SUMOylation, for example, heat-shock treatment of the cells (Tammsalu *et al*., 2014).

In this work, we first decided to maximize the levels of His-HA-*Tb*SUMO conjugates in *T. brucei* PCF, cultured under normal growth conditions, by employing a chromosomal tagging approach meant to avoid the competition between tagged-SUMO and the endogenous protein. Proteomic analysis of the samples obtained after the tandem-affinity purification procedure from this cell line unfortunately did not lead to the identification of SUMO targets, and all attempts to validate them produced negative results. However, since SUMO is essential in *T. brucei* PCF (Liao *et al*., 2010), and we indeed succeeded in obtaining a double replacement of the gene, this cell line was useful to demonstrate that the addition of the tags does not alter SUMO functionality.

Having shown that the proportion of contaminants masked the detection of the low abundance SUMO targets, we decided to use a different SUMO variant keeping the tags but with all lysine residues mutated to arginine and also introducing an Arg residue prior to the diGly motif in the C-terminus of SUMO. Specific digestion of cell extracts with Lys-C prior to Ni^{+2} purification was

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

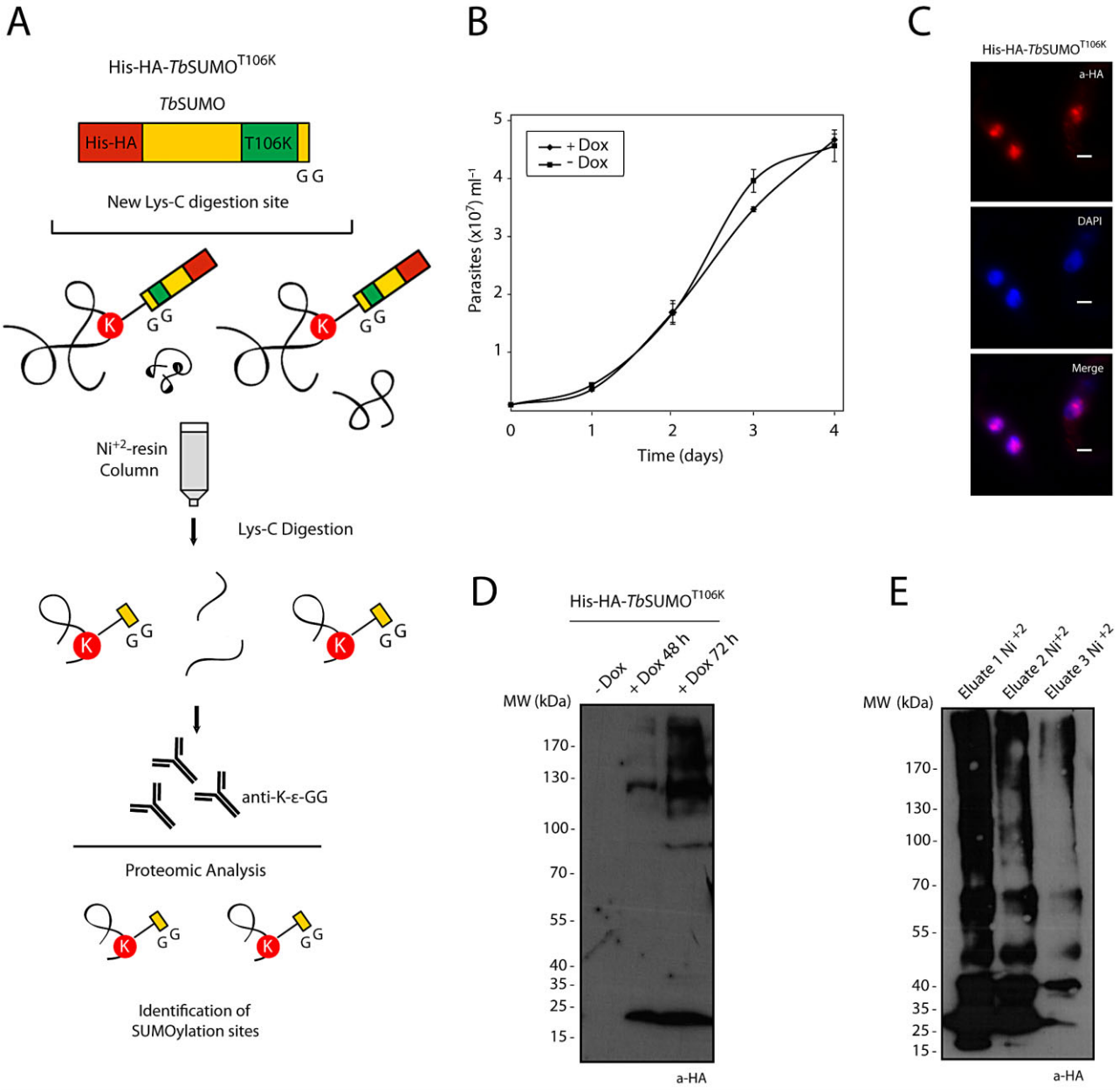

Fig. 1. Analysis of His-HA-*Tb*SUMO^{T106K} cell line.
A. Schematic representation of the purification strategy of *Tb*SUMO conjugates from His-HA-*Tb*SUMO^{T106K} parasites.
B. Growth of His-HA-*Tb*SUMO^{T106K} strain with (+ Dox) and without (– Dox) doxycycline induction.
C. Immunofluorescence analysis of His-HA-*Tb*SUMO^{T106K} parasites after 72 h of doxycycline induction. Nuclear and kinetoplast DNA were visualized by DAPI staining (blue). Representative images of anti-HA and anti-HA-DAPI merged images are shown. Scale bars: 2 μm.
D. Conjugating capacity of His-HA-*Tb*SUMO^{T106K} strain without doxycycline induction (– Dox) or after 48 h (+ Dox 48 h) or 72 h (+ Dox 72 h) of doxycycline induction. Whole-cell extracts were boiled in Laemmli's sample buffer immediately after harvesting, separated in a 7.5–12.5% discontinuous acrylamide gel (5×10^7 cells per lane), and analysed by Western blot.
E. Ni^{+2}-resin purification profile under denaturing conditions of *Tb*SUMO conjugates from His-HA-*Tb*SUMO^{T106K} strain after 72 h of doxycycline induction analysed by Western blot; Eluates 1–3 Ni^{+2}, eluates from nickel-affinity chromatography (equivalent to 5×10^8 cells per lane).

supposed to cleave all proteins but SUMO, thus reducing the presence of contaminants, while the acceptor Lys in substrates could be identified by locating the lysine residues modified by SUMO diGly-remnant after a subsequent cleavage of SUMOylated peptides with trypsin. Although this method was shown to be suitable to map SUMO-2 acceptor lysine residues in target proteins from HeLa cells (Matic *et al.*, 2010; Hendriks *et al.*, 2014), in our experience did not render a good yield of SUMO modification sites. In spite of this, we learned from this

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

Table 1. Identified SUMOylated proteins and SUMOylation sites.

Gene identification	Description	SUMOylated lysine	SUMOylated peptide
Tb927.7.2830	Histone H2A	K(5)	MATP**K**QAVKKASKGGSSRSV
		K(9)	MATPKQAV**K**KASKGGSSRSVKAGL
		K(10)	MATPKQAVK**K**ASKGGSSRSVKAGLI
		K(129)	LNKALAKKQKSGKHA**K**ATPSV
Tb927.10.10460	Histone H2B	K(5)	MATP**K**STPAKTRKEAKKTRR
Tb927.1.2430	Histone H3	K(24)	RTKKTITSKKSKKAS**K**GSDAASGVKTAQRRW
Tb927.10.15350	Histone H3V (h3vaR)	K(32)	VASRPIQAVARAPVK**K**VENTPPQKRHHRWRP
Tb927.5.4170	Histone H4	K(11)	MAKGKKSGEA**K**GSQKRQKKVLRENVR
		K(6)	MAKGK**K**SGEAKGSQKRQKKVLRENVR
Tb927.4.1330	DNA topoisomerase IB, large subunit	K(13)	MUKAQKPKSGEG**K**GKKVAVKDEEVNGKR
		K(20)	QKPKSGEGKGKKVAV**K**DEEVNGKRVVVKKED
		K(47)	KKEDMTEEKIKKVVI**K**EEENELEMVAAGMGP
		K(75)	MGPAKGAVKREGDEN**K**AAMLANEEGGGNDGE
Tb927.9.5190	Proliferative cell nuclear antigen (PCNA)	K(201)	HAPTVDPRSKGESDV**K**TEDEEADACSVRTHS
Tb927.2.4390	Endo/exonuclease Mre11 (MRE11)	K(685)	SHPNVSTVKGGGKAV**K**SETSAPKPARGRKPK
Tb927.4.1270	RuvB-like DNA helicase	K(5)	MSGI**K**IEEVISTTKKERVAA
Tb927.10.740	Structural maintenance of chromosome 4 (SMC4)	K(1302)	ERKRKRTGSTGDVQI**K**VEDEVACNNEAADIL
Tb927.8.3680	Kinetoplastid kinetochore protein 4	K(356)	ISETDEYVKKGSPVV**K**KELNSPTAEAQQTCL
Tb927.1.1170	DNA-directed RNA polymerase subunit RBP12	K(7)	MLSYTV**K**EEVKDEKLPGANNFA
Tb927.4.1630	Ribosomal RNA processing protein 6 (RRP6)	K(552)	IGSADSLKDGANGDV**K**PKPLPNVGVALRYMG
		K(691)	GEGMAGPCLQEGLEV**K**DEVEEIDVPTQKPLE
Tb927.5.1200	Exosome component CSL4 (CSL4)	K(6)	MAVIV**K**TGKCVSPGDVLYATD
Tb927.11.16600	Exosome-associated protein 2 (EAP2)	K(292)	KINGSKNSVTDPHPL**K**EG
Tb927.11.370	Repressor activator protein 1 (RAP1)	K(292)	EPKSAKVDTDSVSLV**K**NEVDESDDSIKDEFS
Tb927.9.11070	*Tb*SIZ1	K(199)	FFASRKDAGGGATQV**K**DAHSATTPPPPPPTV
Tb927.3.3590	U3 small nucleolar ribonucleoprotein protein MPP10	K(355)	WDDPSAVFSGKDAAL**K**GEGEGWYDGEDDILA
Tb927.10.14330	Hypothetical protein/U3 Small Nucleolar RNA-Associated Protein 14	K(703)	TFUKEDKEDGGKTTV**K**VMENIQLKRRRGIDA
Tb927.5.2880	Chaperone protein DNAj	K(4)	MGS**K**EDQVEEIRRLLRFSS
Tb927.11.13090	Elongation factor 1 gamma	K(352)	TFGPVKAKKAEQQAP**K**AEKAPKPAPKDDDDD
Tb927.9.7590	60S ribosomal protein L11	K(5)	MAEA**K**VSAEKKAANPMREIV
Tb927.10.13720	RNA-binding protein, putative (RBP29)	K(313)	GGEDTEPEKSATLPM**K**TESDSVSCPMEVNCN
Tb927.11.6350	AAA ATPase	K(1125)	EEKDRRKTENMEVEV**K**TMVEEEMVEVDGAPV
Tb10.v4.0062	Variant surface glycoprotein (VSG, pseudogene)	K(365)	CKQAKPTCEWKGSEA**K**EGNHFKLNTAHVEQA
Tb927.9.340	Variant surface glycoprotein (VSG, pseudogene)	K(175)	KTAEQEARSGAKEAV**K**ELEEAIYGPSKDRAD
Tb927.4.2740	Hypothetical protein/P25-alpha	K(97)	VAEILTKASPASNST**K**AEAVKFHDDKNLYTG
Tb927.3.2350	Hypothetical protein	K(35)	SRINQALELATKVTV**K**EEPDTENGGGKAAII
Tb927.11.11840	Hypothetical protein	K(69)	YCSRAPTEEKVVAQV**K**EEDGRGTGSSSRRNI
Tb927.8.6130	Hypothetical protein	K(236)	VCVKQEPSLTTAGVV**K**SEYGAHAGVNTDDED
Tb927.10.11200	Hypothetical protein	K(19)	IFSNVRTHEEVPPEV**K**DKVEPITYVPAPESA
Tb927.11.2100	Hypothetical protein	K(146)	TPPSLLTPFAEGEGV**K**ASGEFEDFAVSRSRW
Tb927.9.1410	Hypothetical protein	K(271)	APEPLQCETEAFGNI**K**EELCKLQDSTAEVRL
Tb927.9.10680	Hypothetical protein	K(171)	AEPKHEAEFSVEGDV**K**VPVAETGNDGIAFLK
Tb927.3.5370	Hypothetical protein	K(132)	RALYTSNAETSTPDQ**K**AKGGKGSLKTTPKKS
Tb927.7.2640	Hypothetical protein	K(40)	SRLVSLREGASTGVA**K**DEAESDALGGLAGGD
Tb927.3.4140	Hypothetical protein	K(455)	ASIRQQMKNKKNAEI**K**VEEIEDVAVDDGAAA
Tb927.11.4880	Hypothetical protein	K(178)	DRADFGKVPKYLKEV**K**ADIEERHALVERLKA
Tb927.11.5200	Hypothetical protein	K(163)	EDINEKDATAVEFGN**K**KKRFCVKSNAPKLPK
Tb927.6.1070	Hypothetical protein	K(107)	DDWEADTTTYEGAEV**K**AEVEAEEKLLTAQRE
Tb927.8.2830	Hypothetical protein	K(192)	YNDGRNGMIHLKDVK**K**LCESVDLDEVEFRAR
Tb927.11.5230	Hypothetical protein	K(24)	AAIYARAEEAQNAPV**K**LQPVPAFELYKAARE
Tb927.9.1410	Hypothetical protein	K(228)	APEPLQCETEAVGNI**K**EELCKLQDSTAEVRL
Tb927.9.13320	Hypothetical protein	K(498)	KKTDEEGEAEAEVAV**K**PEPVVSLDG
Tb927.10.14520	Hypothetical protein	K(371)	VPAKDPGTAPSAVVA**K**SEKEEPSAAKTPRPT

The modified Lys residues are in bold.

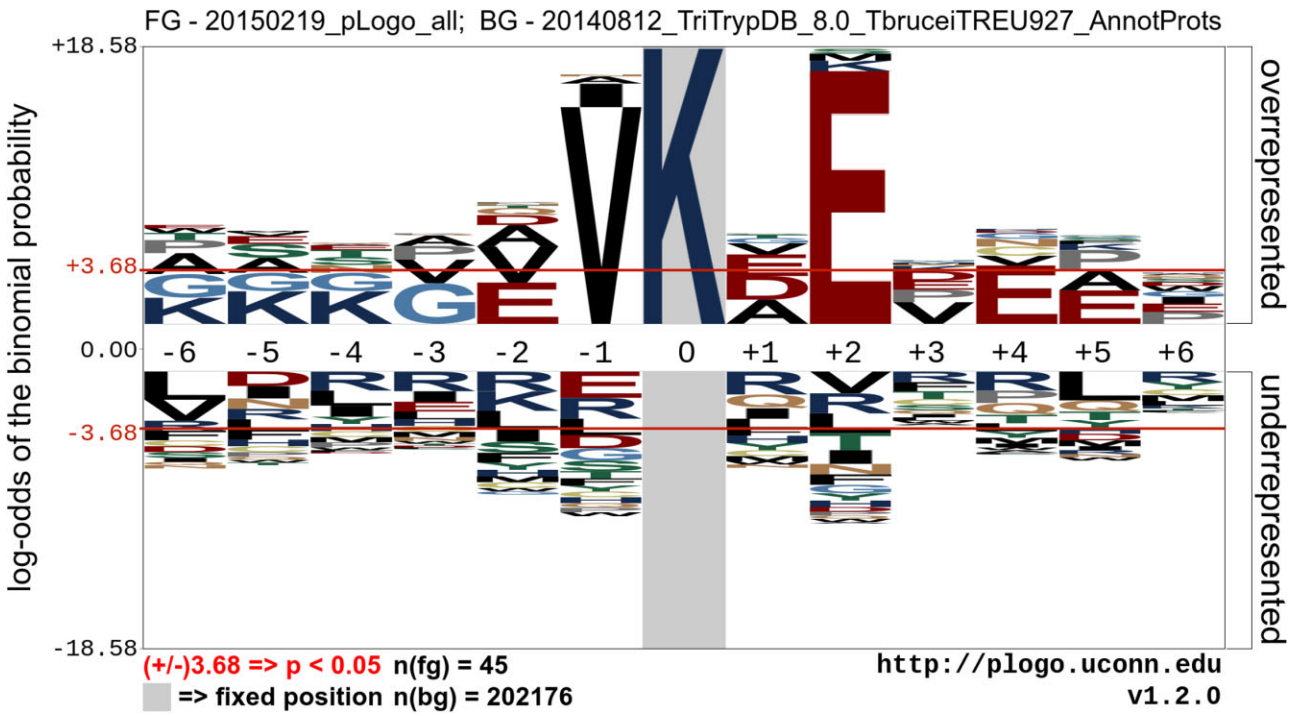

Fig. 2. Sequence analysis of His-HA-*Tb*SUMO^{T106K}-modified peptides. Sequence logo graphs of amino acid sequence conservation surrounding 53 His-HA-*Tb*SUMO^{T106K}-modified sites. The y axis corresponds to the log-odds of the binomial probability (*P*). Threshold values of 3.68 ($P < 0.05$) are marked with red horizontal lines.

double knock-in cell line that SUMO chain formation is not an essential process. Previous studies have shown that in *T. cruzi* SUMO is capable of forming chains *in vitro* and Lys23 was identified as the major branch site in these chains (Annoura *et al*., 2012). In our proteomic studies (Table 1), we were not able to detect spectra compatible with SUMO chain formation in *T. brucei* PCF, but a potential SUMOylation site is conserved in *Tb*SUMO, and we have indeed confirmed that chain formation occurs at least *in vitro* (P. A. Iribarren *et al*., in preparation).

Site-specific proteomic identification of SUMOylated proteins has been strongly upgraded recently by engineering SUMO to generate a Lys-C digestion site before the diGly motif (Tammsalu *et al*., 2014). Since trypsin digestion creates a diGly signature peptide that can also be derived from other ubiquitin-like proteins (Ubls) containing an arginine residue N-terminal to the diGly motif, such as ubiquitin or NEDD8, this approach allows the identification of exclusively SUMOylated proteins. SUMOylated peptides are then captured with specific anti-K-ε-GlyGly monoclonal antibodies and SUMO modification sites are detected in an unambiguous manner. This strategy showed a high degree of reproducibility (near 50%) between independent experiments on *T. brucei* procyclic samples, similar to what has been described for the original method using a lysate of HEK293 culture (Tammsalu *et al*., 2014).

A high proportion of the proteins identified as SUMO targets belong to functionally related nuclear processes, such as DNA replication and repair, transcription, rRNA biogenesis, RNA processing and degradation, and chromatin remodelling, among others, suggesting that SUMO signalling could be mediated by several proteins within protein groups. Some well-known SUMO targets in yeast and mammals have been identified in this study, like PCNA (Ulrich, 2009) and Topoisomerase IB (Mao *et al*., 2000; Steinacher *et al*., 2013) with four SUMOylation sites identified at the non-conserved hydrophilic N-terminal end of the large subunit (homologous to the core of the general monomeric enzyme). In kinetoplastids, Topoisomerase IB is a heterodimer and both subunits are essential for cell viability (Bakshi and Shapiro, 2004). We were also able to identify SUMOylation sites in the nuclease Mre11 that, in complex with Rad50 and Nsb1, is involved in the detection and repair of double-strand breaks (Lamarche *et al*., 2010), and the DNA helicase RuvB (Caruthers and McKay, 2002), suggesting that SUMO modulation of these processes may also be conserved in trypanosomatids.

Among the exosomal complex proteins, two structural components (EAP2 and CSL4) and one exosome-interacting protein (RRP6) (Estevez *et al*., 2001; Haile *et al*., 2007; Barbosa *et al*., 2014) were found to be SUMOylated. *Tb*RRP6 is an essential subunit in *T. brucei* PCF and the sites identified in *Tb*RRP6 (K552 and K691) are in the C-terminal end of the protein responsible for exosome interaction. Unlike yeast and humans, *Tb*RRP6

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

is found in both the nuclear and the cytoplasmic exosome complex (Haile *et al*., 2007). Interestingly, recent reports have shown a novel SUMO-dependent localization of the exosome following the induction of transcription-related DNA damage in HeLa cells (Richard *et al*., 2013).

One main finding of our proteomic study was the identification of proteins involved in the epigenetic regulation of gene expression. This is particularly interesting for trypanosomatids, organisms that display unusual transcriptional features such as the lack of polII promotors for protein-coding genes. We found the telomeric protein TbRAP1 to be SUMOylated. This protein has been shown to generate a silencing gradient from the telomere ends, both in PCF (Pandya *et al*., 2013) and in BSF (Yang *et al*., 2009). It is intriguing, nevertheless, how SUMO correlates with this effect. All core histones (H2A, H2B, H3 and H4) were found to be modified by SUMO in K residues that do not match the consensus, similar to what has been reported for H2A and H2B SUMOylation sites in yeast and humans (Nathan *et al*., 2006; Hendriks *et al*., 2014; Tammsalu *et al*., 2014). Since histones are very abundant proteins, they are expected to be easily detected when general SUMOylation patterns are analysed by Western blot; however, it is difficult to predict the apparent molecular weight of SUMOylated histones mainly because they can be SUMOylated at multiple K residues. It is noteworthy that five out of the seven Lys residues identified as SUMOylated had been previously shown to be acetylated in *T. brucei*, suggesting that SUMO and acetylation could compete and produce different consequences. The possible impact of histone SUMOylation in trypanosomatids certainly deserves future studies.

Finally, we find very interesting the fact that the SUMO E3 ligase Siz1 is SUMOylated. It is tempting to speculate that SUMOylation of this protein could influence either its subnuclear localization at the active VSG expression site or could provide an interaction platform to recruit RNA polymerase I subunits (RPA1, RBP7 or other subunits).

Experimental procedures

Trypanosome culture

Procyclic form (PCF) *T. brucei brucei* Lister 29–13 cell line (T7RNAPol NEO TetR HYG) was a gift from G.A.M. Cross (Rockefeller University). This cell line was maintained axenically at 28°C in SDM-79 medium (Life Technologies, Carlsbad, CA, USA) supplemented with 10% (v/v) heat-inactivated foetal calf serum (Natocor, Córdoba, Argentina) and 7.5 mg l^{-1} hemin.

Plasmid constructions

To generate His-HA-*Tb*SUMO^{T106K} variant, we used a synthesized plasmid construction with the complete open reading frame of His-HA-*Tb*SUMO (GenScript, Piscataway, NJ, USA) (see Supplementary Experimental Procedures) as template for polymerase chain reaction (PCR) amplification (sense primer CCATGG ACGAACACCACCAC and antisense primer GCTAGCTCACC CGCCCTTCTGCTC). The amplification product, flanked by NcoI and NheI restriction sites, was cloned into pGEM-T Easy vector (Promega, Madison, WI, USA) and completely sequenced (Macrogen, Seoul, Korea). Inserts were liberated with EcoRI and NheI restriction enzymes (New England Biolabs, Ipswich, MA) and cloned into the same sites of the tetracycline inducible pLEW100v5 (Addgene, Cambridge, MA) vector with phleomycin resistance marker cassette.

Generation of Tb*SUMO-transfectant cell lines*

The above described vector was linearized and electroporated as described by the Cross laboratory (http://tryps.rockefeller.edu/) into PCF 29-13. Briefly, log phase cells (6×10^6 ml^{-1}) were collected by centrifugation, washed with Cytomix and resuspended in this solution at a concentration of 40×10^6 ml^{-1}. Then, 0.50 ml of cells was mixed with 10 μl of linearized DNA (15 μg) in a 0.2 cm electroporation cuvette (BTX, Harvard Apparatus, Holliston, MA, USA) and subjected to three pulses from a BTX Electro Square Porator ECM 830 set at 1700 V and 25 μF. Transfected cells were selected with 1 μg ml^{-1} of phleomycin (InvivoGen, San Diego, CA, USA) and protein expression was induced with 5 μg ml^{-1} of doxycycline (Sigma, Saint Louis, MO, USA) for 72 h.

Growth curves

Growth curves were obtained by counting cell number daily by quadruplicate using a Neubauer chamber under the light microscope (×400).

Indirect immunofluorescence

Parasites collected by centrifugation were washed twice with phosphate-buffered saline (PBS), fixed with 4% paraformaldehyde in PBS and then attached to poly-lysine coated glass coverslips for 30 min. Parasites were incubated with 25 mM NH_4Cl (15 min) and permeabilized and blocked with 2% bovine serum albumin, 0.1% saponin and 2% normal goat serum in PBS for 30 min. This was followed by 1 h incubation with high-affinity rat monoclonal antibodies anti-HA (Roche, Basel, Switzerland) in blocking solution. Coverslips were washed three times with PBS and then incubated for 1 h with secondary antibodies diluted 1:1000 in blocking solution (polyclonal goat anti-rat Alexa Fluor® 546). After extensive washing with PBS, coverslips were mounted using FluorSave reagent (Merck, Darmstadt, Germany) containing 5 mg ml^{-1} DAPI (4,6-diamidino-2-phenylindole) (Life Technologies) to stain nuclear and kinetoplast DNA. Images were analysed with a fluorescence microscope Nikon Eclipse E600 and captured with a digital camera Sport RT Slider (model 2.3.1, Diagnostic Instruments, Sterling Heights, MI, USA).

Electrophoresis and immunoblotting

Parasite extracts to evaluate the conjugation ability of *Tb*SUMO^{T106K} and protein fractions from affinity purification steps

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

were resuspended in Laemmli's sample buffer [0.125 M Tris (pH 6.8), 4% (w/v) sodium dodecyl sulfate, 20% (v/v) glycerol, 100 mM dithiothreitol] and boiled for 5 min. All samples were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (7.5–12.5% discontinuous acrylamide gradient) and transferred to a nitrocellulose Hybond ECL membrane (GE Healthcare, Pittsburgh, PA, USA) for probing with high-affinity rat monoclonal anti-HA antibodies (Roche) diluted 1:500. Horseradish peroxidase-conjugated goat anti-rat secondary antibody (Sigma) diluted 1:200 was detected by chemiluminescence using SuperSignal West Pico Chemiluminescent Substrate (Pierce, Rockford, IL, USA).

Purification of Tb*SUMO conjugates*

About 5×10^{10} parasites from selected clones were collected by centrifugation and washed once with PBS supplemented with 20 mM NEM (Sigma). Cells were then resuspended in lysis buffer (6 M urea, 500 mM NaCl, 50 mM Tris–HCl, 5 mM β-mercaptoethanol, pH 7.5) at a concentration of ~ 3×10^6 parasites μl^{-1} and sonicated up to loss of viscosity. For further purification of *Tb*SUMO conjugates, lysates were cleared by centrifugation for 30 min at 23 000 × *g*. Supernatants, with the addition of imidazole at a final concentration of 20 mM, were incubated with 2.5 ml of Ni^{+2}-resin (GE Healthcare), pre-equilibrated in lysis buffer, with stirring at room temperature (RT) for 1:30 h. After centrifugation for 10 min at 4500 × *g*, the resin was washed three times with 12 ml of buffer 1 (6 M urea, 500 mM NaCl, 50 mM Tris–HCl, 20 mM imidazole, pH 7.50) and seven times with 12 ml of buffer 2 (6 M urea, 500 mM NaCl, 50 mM Tris–HCl, 30 mM imidazole, pH 7.50). For each wash step, the resin was incubated for 3 min with stirring at RT and then centrifuged for 3 min at 4500 × *g*. Proteins were eluted with buffer 3 (6 M urea, 500 mM NaCl, 50 mM Tris–HCl, 1 M imidazole, pH 7.5) and 3 fractions of 2.5 ml were collected. For each elution step, the resin was incubated for 5 min with stirring at RT and then centrifuged for 5 min at 4500 × *g*. Proteins were concentrated by acetone precipitation (80% v/v) and stored at −80°C. Protein fractions were collected for Western blot analysis.

Lys-C digestion protocol of His-HA-*Tb*SUMO^{T106K}-modified proteins and immunopurification of diGly-Lys-containing peptides was performed as described (Tammsalu *et al.*, 2014).

MS analysis

Mass spectrometry analysis, data processing and analysis, including manual validation of the results, were performed exactly as described (Tammsalu *et al.*, 2014).

Bioinformatics analysis

Sequence analysis was performed with pLogo (O'Shea *et al.*, 2013). Residues were scaled relative to their Bonferroni-corrected statistical significance using *T.brucei* TREU927 proteome as a background dataset.

Acknowledgements

The authors thank Triin Tammsalu and Ronald T. Hay from the Centre for Gene Regulation and Expression, College of Life Sciences, University of Dundee, UK, for performing mass spectrometry experiments. This work was supported by grants from the National Agency for Promotion of Scientific and Technological Research from the Argentinian Ministry of Science and Technology (ANPCyT, MinCyT). JJC and VEA are members of the Research Career, and PAI is doctoral fellow, from the Argentinian National Research Council. The authors state no conflict of interest.

References

Annoura, T., Makiuchi, T., Sariego, I., Aoki, T., and Nara, T. (2012) SUMOylation of paraflagellar rod protein, PFR1, and its stage-specific localization in *Trypanosoma cruzi*. *PLoS ONE* **7:** e37183.

Bakshi, R.P., and Shapiro, T.A. (2004) RNA interference of *Trypanosoma brucei* topoisomerase IB: both subunits are essential. *Mol Biochem Parasitol* **136:** 249–255.

Barbosa, R.L., Legrand, P., Wien, F., Pineau, B., Thompson, A., and Guimaraes, B.G. (2014) RRP6 from *Trypanosoma brucei*: crystal structure of the catalytic domain, association with EAP3 and activity towards structured and non-structured RNA substrates. *PLoS ONE* **9:** e89138.

Barrett, M.P., Burchmore, R.J., Stich, A., Lazzari, J.O., Frasch, A.C., Cazzulo, J.J., and Krishna, S. (2003) The trypanosomiases. *Lancet* **362:** 1469–1480.

Bayona, J.C., Nakayasu, E.S., Laverriere, M., Aguilar, C., Sobreira, T.J., Choi, H., *et al.* (2011) SUMOylation pathway in *Trypanosoma cruzi*: functional characterization and proteomic analysis of target proteins. *Mol Cell Proteomics* **10:** M110 007369.

Becker, J., Barysch, S.V., Karaca, S., Dittner, C., Hsiao, H.H., Berriel Diaz, M., *et al.* (2013) Detecting endogenous SUMO targets in mammalian cells and tissues. *Nat Struct Mol Biol* **20:** 525–531.

Caruthers, J.M., and McKay, D.B. (2002) Helicase structure and mechanism. *Curr Opin Struct Biol* **12:** 123–133.

Estevez, A.M., Kempf, T., and Clayton, C. (2001) The exosome of *Trypanosoma brucei*. *EMBO J* **20:** 3831–3839.

Flotho, A., and Melchior, F. (2013) Sumoylation: a regulatory protein modification in health and disease. *Annu Rev Biochem* **82:** 357–385.

Haile, S., Cristodero, M., Clayton, C., and Estevez, A.M. (2007) The subcellular localisation of trypanosome RRP6 and its association with the exosome. *Mol Biochem Parasitol* **151:** 52–58.

Hendriks, I.A., D'Souza, R.C., Yang, B., Verlaan-de Vries, M., Mann, M., and Vertegaal, A.C. (2014) Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat Struct Mol Biol* **21:** 927–936.

Hu, H., Yu, Z., Liu, Y., Wang, T., Wei, Y., and Li, Z. (2014) The Aurora B kinase in *Trypanosoma brucei* undergoes post-translational modifications and is targeted to various subcellular locations through binding to TbCPC1. *Mol Microbiol* **91:** 256–274.

Klein, C.A., Droll, D., and Clayton, C. (2013) SUMOylation in *Trypanosoma brucei*. *PeerJ* **1:** e180.

Lamarche, B.J., Orazio, N.I., and Weitzman, M.D. (2010) The MRN complex in double-strand break repair and telomere maintenance. *FEBS Lett* **584:** 3682–3695.

© 2015 John Wiley & Sons Ltd,

Liao, S., Wang, T., Fan, K., and Tu, X. (2010) The small ubiquitin-like modifier (SUMO) is essential in cell cycle regulation in *Trypanosoma brucei*. *Exp Cell Res* **316:** 704–715.

Lopez-Farfan, D., Bart, J.M., Rojas-Barros, D.I., and Navarro, M. (2014) SUMOylation by the E3 ligase TbSIZ1/PIAS1 positively regulates VSG expression in *Trypanosoma brucei*. *PLoS Pathog* **10:** e1004545.

Mao, Y., Sun, M., Desai, S.D., and Liu, L.F. (2000) SUMO-1 conjugation to topoisomerase I: a possible repair response to topoisomerase-mediated DNA damage. *Proc Natl Acad Sci USA* **97:** 4046–4051.

Matic, I., Schimmel, J., Hendriks, I.A., van Santen, M.A., van de Rijke, F., van Dam, H., *et al*. (2010) Site-specific identification of SUMO-2 targets in cells reveals an inverted SUMOylation motif and a hydrophobic cluster SUMOylation motif. *Mol Cell* **39:** 641–652.

Nathan, D., Ingvarsdottir, K., Sterner, D.E., Bylebyl, G.R., Dokmanovic, M., Dorsey, J.A., *et al*. (2006) Histone sumoylation is a negative regulator in *Saccharomyces cerevisiae* and shows dynamic interplay with positive-acting histone modifications. *Genes Dev* **20:** 966–976.

Obado, S.O., Bot, C., Echeverry, M.C., Bayona, J.C., Alvarez, V.E., Taylor, M.C., and Kelly, J.M. (2011) Centromere-associated topoisomerase activity in bloodstream form *Trypanosoma brucei*. *Nucleic Acids Res* **39:** 1023–1033.

O'Shea, J.P., Chou, M.F., Quader, S.A., Ryan, J.K., Church, G.M., and Schwartz, D. (2013) pLogo: a probabilistic approach to visualizing sequence motifs. *Nat Methods* **10:** 1211–1212.

Pandya, U.M., Sandhu, R., and Li, B. (2013) Silencing subtelomeric VSGs by *Trypanosoma brucei* RAP1 at the insect stage involves chromatin structure changes. *Nucleic Acids Res* **41:** 7673–7682.

Richard, P., Feng, S., and Manley, J.L. (2013) A SUMO-dependent interaction between Senataxin and the exosome, disrupted in the neurodegenerative disease AOA2, targets the exosome to sites of transcription-induced DNA damage. *Genes Dev* **27:** 2227–2232.

Steinacher, R., Osman, F., Lorenz, A., Bryer, C., and Whitby, M.C. (2013) Slx8 removes Pli1-dependent protein-SUMO conjugates including SUMOylated topoisomerase I to promote genome stability. *PLoS ONE* **8:** e71960.

Tammsalu, T., Matic, I., Jaffray, E.G., Ibrahim, A.F., Tatham, M.H., and Hay, R.T. (2014) Proteome-wide identification of SUMO2 modification sites. *Sci Signal* **7:** rs2.

Tatham, M.H., Rodriguez, M.S., Xirodimas, D.P., and Hay, R.T. (2009) Detection of protein SUMOylation in vivo. *Nat Protoc* **4:** 1363–1371.

Ulrich, H.D. (2009) Regulating post-translational modifications of the eukaryotic replication clamp PCNA. *DNA Repair (Amst)* **8:** 461–469.

Yang, X., Figueiredo, L.M., Espinal, A., Okubo, E., and Li, B. (2009) RAP1 is essential for silencing telomeric variant surface glycoprotein genes in *Trypanosoma brucei*. *Cell* **137:** 99–109.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. Analysis of His-HA-*Tb*SUMO cell line.

A. The linearized construct His-HA-*Tb*SUMO in the endogenous locus tagging vector pEnT6P is able to recombine through the homologous sequences with *T. brucei* genome resulting in the replacement of the endogenous *Tb*SUMO by the His-HA-tagged form of the allele. Arrows indicate the relative position of primers and PCR amplification product used to confirm the proper insertion.

B. Ethidium bromide-stained PCR products in a 1% agarose gel used to confirm the appropriate replacement of endogenous *Tb*SUMO alleles in His-HA-*Tb*SUMO S (single replacement) and D (double replacement) strains.

C. Growth of His-HA-*Tb*SUMO S and D strains compared with wild-type (WT) parasites. WT and transgenic parasites were cultured up to one month without observing significant differences in growth rate.

D. Conjugating capacity of His-HA-*Tb*SUMO in S and D strains. Whole-cell extracts were boiled in Laemmli's sample buffer immediately after harvesting, separated in a 7.5–12.5% discontinuous acrylamide gel (3×10^7 cells per lane), and analysed by Western blot. Anti-tubulin was used as a loading control. When evaluating the corresponding SUMOylation pattern using antibodies against the HA-tag, we consistently observed two main bands below the 130 and the 100 kDa pre-stained protein markers, in good agreement with previous reports (Klein *et al*., 2013; Lopez-Farfan *et al*., 2014).

E. Immunofluorescence analysis of WT, His-HA-*Tb*SUMO S and D strains. Nuclear and kinetoplast DNA were visualized by DAPI staining (blue). Representative images of anti-*Tc*SUMO, anti-HA and anti-HA/*Tc*SUMO-DAPI merged images are shown. Higher magnification of the nucleus showing anti-HA/*Tc*SUMO/DAPI and merged images is shown. Scale bars: 3 μm.

F. Tandem purification profile of His-HA-*Tb*SUMO conjugates analysed by Western blot from His-HA-*Tb*SUMO D strain compared with eluates from control purification from WT parasites. Input (3×10^7 cells per lane): Insoluble fraction after urea treatment (Insoluble) (3×10^7 cells per lane); FT Ni^{+2}, flow through fraction after nickel-affinity chromatography (equivalent to 3×10^7 cells per lane); Eluate Ni^{+2}, eluate from nickel-affinity chromatography (equivalent to 3×10^8 cells per lane); FT a-HA, flow through fraction after anti-HA affinity chromatography (equivalent to 3×10^9 cells per lane) Eluate a-HA, eluate from anti-HA affinity chromatography (equivalent to 3×10^9 cells per lane). The asterisk indicates an immunoreactive band from anti-HA antibody remnant from affinity purification.

Fig. S2. Analysis of His-HA-*Tb*SUMOK9R^{T106R} cell line.

A. Schematic representation of the purification strategy of *Tb*SUMO conjugates from His-HA-*Tb*SUMOK9R^{T106R} parasites.

B. Ethidium bromide-stained PCR products in a 1% agarose gel used to confirm the appropriate replacement of endogenous *Tb*SUMO alleles in His-HA-*Tb*SUMOK9R^{T106R} strain (see Supporting Information Fig. S1A). Genomic DNA from the double transfectant pool of His-HA-*Tb*SUMOK9R^{T106R} parasites was used as positive control.

C. Growth of His-HA-*Tb*SUMOK9R^{T106R} strain compared to WT parasites. WT and transgenic parasites were cultured up to one month without observing significant differences in growth rate.

D. Conjugating capacity of His-HA-*Tb*SUMOK9R^{T106R} strain. Whole-cell extracts were boiled in Laemmli's sample buffer immediately after harvesting, separated in a 7.5–12.5%

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

discontinuous acrylamide gel (3×10^7 cells per lane), and analysed by Western blot.

E. Immunofluorescence analysis of His-HA-*Tb*SUMOK9R^{T106R} parasites. Nuclear and kinetoplast DNA were visualized by DAPI staining (blue). Representative images of anti-HA and anti-HA-DAPI merged images are shown. Scale bars: 2 μm.

F. Ni^{+2}-resin purification profile under denaturing conditions of *Tb*SUMO conjugates from His-HA-*Tb*SUMOK9R^{T106R} strain analysed by Western blot; Input (5×10^7 cells per lane), FT Ni^{+2}, flow through after nickel-affinity chromatography (equivalent to 5×10^7 cells per lane), Eluate Ni^{+2}, eluate from nickel-affinity chromatography (equivalent to 2×10^8 cells per lane). Eluate from the second step of Ni^{+2}-resin purification after Lys-C digestion is shown.

Table S1. Proteomic data.

Supplementary Experimental procedures.

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*