

# SCIENTIFIC REPORTS



OPEN

## Structural and Biochemical Characterization of Poly-ADP-ribose Polymerase from *Trypanosoma brucei*

Teemu Haikarainen<sup>1</sup>, Mariana Schlesinger<sup>2</sup>, Ezeogo Obaji<sup>1</sup>, Silvia H. Fernández Villamil<sup>2</sup> & Lari Lehtio<sup>1</sup>

*Trypanosoma brucei* is a unicellular parasite responsible for African trypanosomiasis or sleeping sickness. It contains a single PARP enzyme opposed to many higher eukaryotes, which have numerous PARPs. PARPs are responsible for a post-translational modification, ADP-ribosylation, regulating a multitude of cellular events. *T. brucei* PARP, like human PARPs-1-3, is activated by DNA binding and it potentially functions in DNA repair processes. Here we characterized activation requirements, structure and subcellular localization of *T. brucei* PARP. *T. brucei* PARP was found to be selectively activated by 5' phosphorylated and 3' phosphorylated DNA breaks. Importantly, the N-terminal region is responsible for high-affinity DNA-binding and required for DNA-dependent enzymatic activation. This module is also required for nuclear localization of the protein in response to oxidative stress. Solution structures of activating and non-activating PARP-DNA complexes were determined with small-angle X-ray scattering revealing distinct differences in their DNA-binding modes.

Poly-ADP-ribose polymerases are enzymes catalyzing a covalent addition of ADP-ribose to a target protein, to themselves, or to a growing ADP-ribose polymer. The enzyme family is widespread among eukaryotes, although multicellular eukaryotes contain more family members than lower eukaryotes<sup>1</sup>. In humans the family consists of 17 members, where PARPs are involved in a multitude of cellular processes including DNA damage repair, cell death pathways, and mitosis. Human PARPs (hPARPs)-1-3, whose catalytic activity is induced by DNA-binding<sup>2</sup> are involved in DNA-damage repair processes<sup>3</sup>. The role of hPARPs-1-3, and especially of hPARP-1, in DNA damage repair is extensively studied and a lot of interest have been devoted to discovery of PARP inhibitors as therapeutics against various forms of cancer<sup>4</sup>.

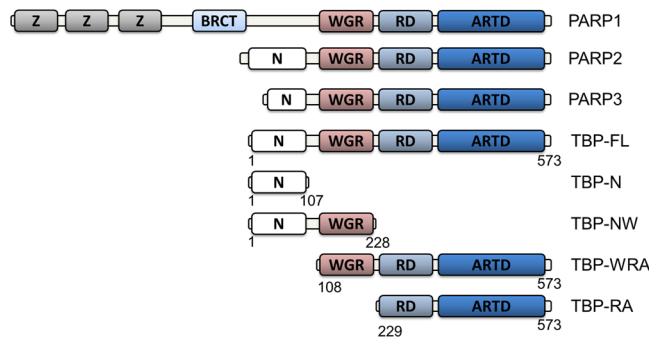
hPARP-1 is a multidomain protein consisting of DNA-binding zinc-fingers, a WGR domain necessary to DNA-dependent activation, a BRCT domain, and regulatory (RD) and catalytic (ARTD) domains responsible for poly-ADP-ribose synthesis (Fig. 1). hPARPs-2 and 3 consist of regulatory and catalytic domains and a WGR domain. Instead of the DNA-binding zinc-fingers, they have N-terminal regions which participate in DNA-binding<sup>2</sup>. hPARPs-2 and 3 were recently shown to prefer specific DNA substrates for activation suggesting their role in specific DNA-repair pathways in opposed to hPARP-1, which is activated by various DNA-damage models<sup>2,5</sup>.

*Trypanosoma brucei* is a unicellular parasitic protozoa responsible for African trypanosomiasis or sleeping sickness, a scourge in Sub-Saharan Africa for both livestock and human health. Based on sequence comparisons, the parasite contains only a single PARP enzyme (*Tb*PARP), which has a similar domain organization as hPARPs-2 and 3, including the WGR, RD, and ARTD domains (Fig. 1). It also contains a positively charged N-terminus, with higher sequence similarity to hPARP-2 than to hPARP-3.

We have previously shown that *Tb*PARP is enzymatically activated by DNA and migrates to the nucleus after a genomic stress<sup>6</sup>. Interestingly, cytotoxicity induced by hydrogen peroxide was reduced both by down-regulation and inhibition of the catalytic activity of the enzyme suggesting a role for *Tb*PARP in cell death pathways. In

<sup>1</sup>Biocenter Oulu and Faculty of Biochemistry and Molecular Medicine, University of Oulu, FI-90014, Oulu, Finland.

<sup>2</sup>National Institute for Genetic Engineering and Molecular Biology (INGEBI-CONICET), University of Buenos Aires, Buenos Aires, Argentina. Correspondence and requests for materials should be addressed to L.L. (email: [lari.lehtio@oulu.fi](mailto:lari.lehtio@oulu.fi))



**Figure 1.** *Tb*PARP constructs. Comparison of *Tb*PARP domain structure to human PARPs-2-3 and a schematic presentation of the *Tb*PARP constructs used in the study. Z, zinc finger motif; BRCT, BRCA1 C terminus domain; WGR, WGR domain; RD, regulatory domain; ARTD, catalytic domain; N, N-terminal fragment.

addition, inhibition of the catalytic activity of *Tb*PARP did not inhibit the growth of *T. brucei* procyclic and bloodstream forms<sup>6</sup>. These results are in contrast to what have been observed for the related parasite *T. cruzi*<sup>7</sup>.

Here, in effort to understand the role of PARP in the parasite, we have characterized DNA-binding and DNA-dependent activation properties of the enzyme and especially the role of the N-terminal domain in these processes. We have also studied the role of the N-terminus in the localization of the protein. Finally, we performed structural characterization of *Tb*PARP-DNA-complexes with solution scattering methods. Our results indicate a high degree of substrate selectivity in the DNA activation process and identify the N-terminus as a key domain in DNA-binding, DNA-dependent catalytic activation, and nuclear localization.

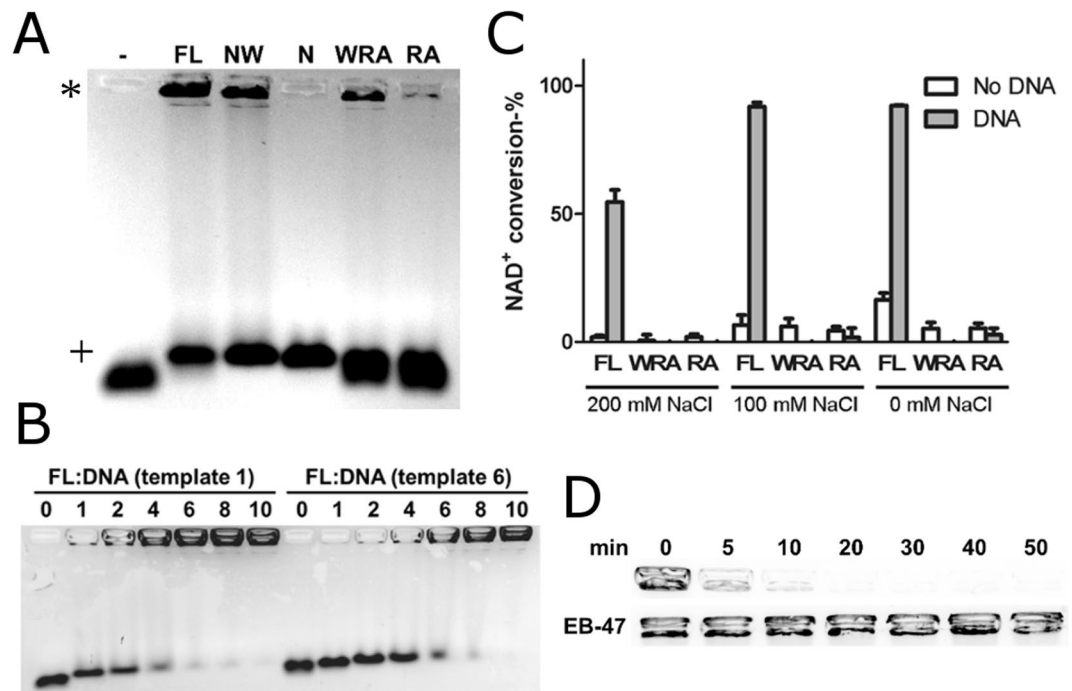
## Results

**N-terminus is required for high-affinity DNA binding and activation.** We have previously shown that *Tb*PARP is catalytically activated by damaged DNA and it most likely functions in DNA-damage response and DNA-repair like hPARPs-1-3<sup>6</sup>. While hPARP-1 utilizes three zinc-fingers for DNA binding, hPARPs-2 and 3 have a positively charged N-terminus that is utilized for DNA interaction. As the N-terminus of *Tb*PARP is rich in lysines and carries a high positive charge, we hypothesized that the DNA-binding would be mediated by this region.

We analyzed the DNA-binding properties of the full-length protein and various fragments (Fig. 1) with EMSA using an agarose gel (Fig. 2A). We used double-stranded DNA oligonucleotide (template 1, Fig. 3) in the assay. Full-length protein (TBP-FL) and the fragment containing WGR and N-terminal domains (TBP-NW) resulted in a supershift (DNA retained in the well) of DNA on the gel. The fragment containing N-terminus only (TBP-N) did not retain DNA in the well but still resulted in a shift on the gel compared to the DNA only, indicating binding. The fragment lacking the N-terminus (TBP-WRA) had seemingly lower affinity with DNA as only a minor fraction of the DNA was supershifted on the gel. Therefore, N-terminus was not strictly required for DNA binding but appeared to increase the binding affinity. When both N-terminus and WGR domain were omitted (TBP-RA), the DNA-binding activity was almost completely abolished. With TBP-FL and TBP-NW, two shifts were observed in EMSA: a supershift and a shift corresponding to the position of TBP-N shifted DNA (Fig. 2A). We hypothesized that the smaller shift arises from interactions with the N-terminus, which might be non-specific and have higher on- and off-rate constants. We tested this by titration of the DNA with protein: when protein concentration is increased the DNA gradually shifts towards the well (Fig. 2B). Notably phosphorylated oligonucleotide (template 6, Fig. 2B) seemed to require higher protein concentrations for fully supershifting the DNA indicating higher protein:DNA stoichiometry.

The seemingly lower DNA binding affinity of TBP-WRA prompted us to test whether high affinity binding mediated by the N-terminus was required for DNA-dependent activation. We compared the DNA dependent activation of TBP-FL, TBP-WRA and TBP-RA with a fluorescence-based activity assay measuring NAD<sup>+</sup> consumption. A robust ( $\approx 40$ -fold) activation of TBP-FL was observed upon incubation with activated DNA (Fig. 2C, 200 mM NaCl). Although TBP-WRA and TBP-RA both retained basal enzymatic activity, there was no increase in enzymatic activity upon incubation with activated DNA. This suggests that the N-terminus is required for both high affinity interaction and DNA-dependent activation of the protein. Since TBP-WRA (and TBP-RA to lower extent) showed interaction with DNA on EMSA at lower salt concentration, we repeated the activity assay at conditions containing 100 mM NaCl and without salt. Higher activation of TBP-FL by activated DNA was observed at lower salt concentrations but these conditions did not result in activation of TBP-WRA or TBP-RA.

hPARPs-1-3 have been shown to be released from DNA upon automodification<sup>2,8</sup>. This is most likely a result of repulsive interaction between the negatively charged DNA and poly(ADP-ribose) (PAR). We proceeded to test whether the DNA-dependent ADP-ribosylation results in dissociation of the *Tb*PARP-DNA complex. When incubated with NAD<sup>+</sup>, the preformed TBP-FL-DNA complex rapidly dissociates from DNA (Fig. 2D, upper lane). To verify that the automodification is responsible for the dissociation, the preformed complex was incubated in the presence of EB-47, a potent *Tb*PARP inhibitor<sup>6</sup> before initiating the automodification with NAD<sup>+</sup>. Inhibition of the catalytic activity effectively prevents the dissociation of the complex (Fig. 2D, lower lane).

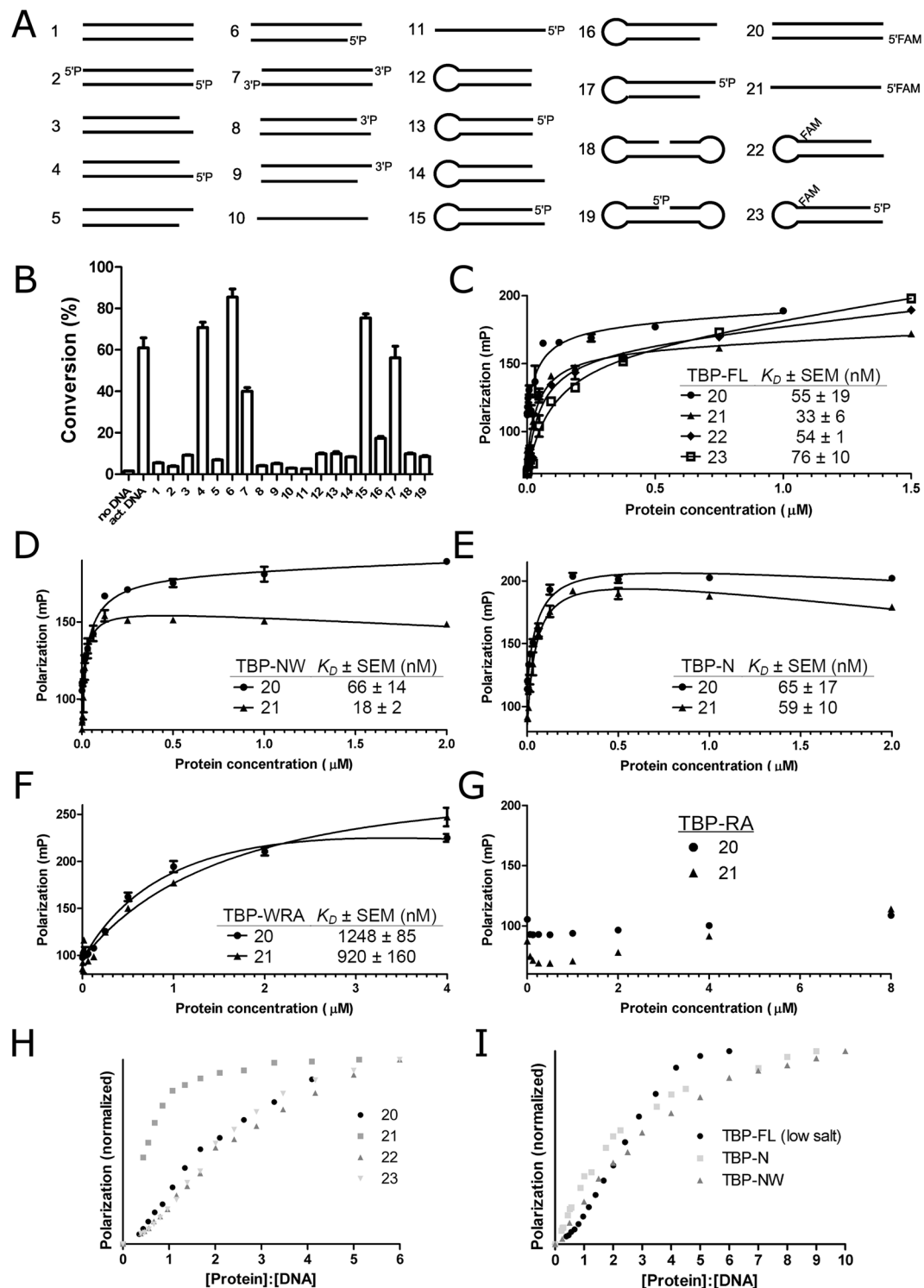


**Figure 2.** Characterization of DNA-binding properties and DNA-dependent activation of *TbPARP*. (A) The binding of *TbPARP* constructs (2  $\mu$ M) to DNA template 1 (0.5  $\mu$ M) were evaluated with agarose gel electrophoresis. (\*) Shows the position of the supershifted DNA and (+) displays the smaller DNA shift. (B) Concentration-dependent analysis of DNA binding to TBP-FL. TBP-FL was titrated against 0.5  $\mu$ M DNA (templates 1 and 6) and analyzed with agarose gel electrophoresis. (C) DNA dependent activation of TBP-FL, TBP-WRA and TBP-RA using different salt concentrations. Activated DNA was used as a template. (D) DNA dissociation from TBP-FL followed by automodification was evaluated as a function of time with agarose gel electrophoresis either without (upper lane) or with (lower lane) inhibitor EB-47 (10  $\mu$ M).

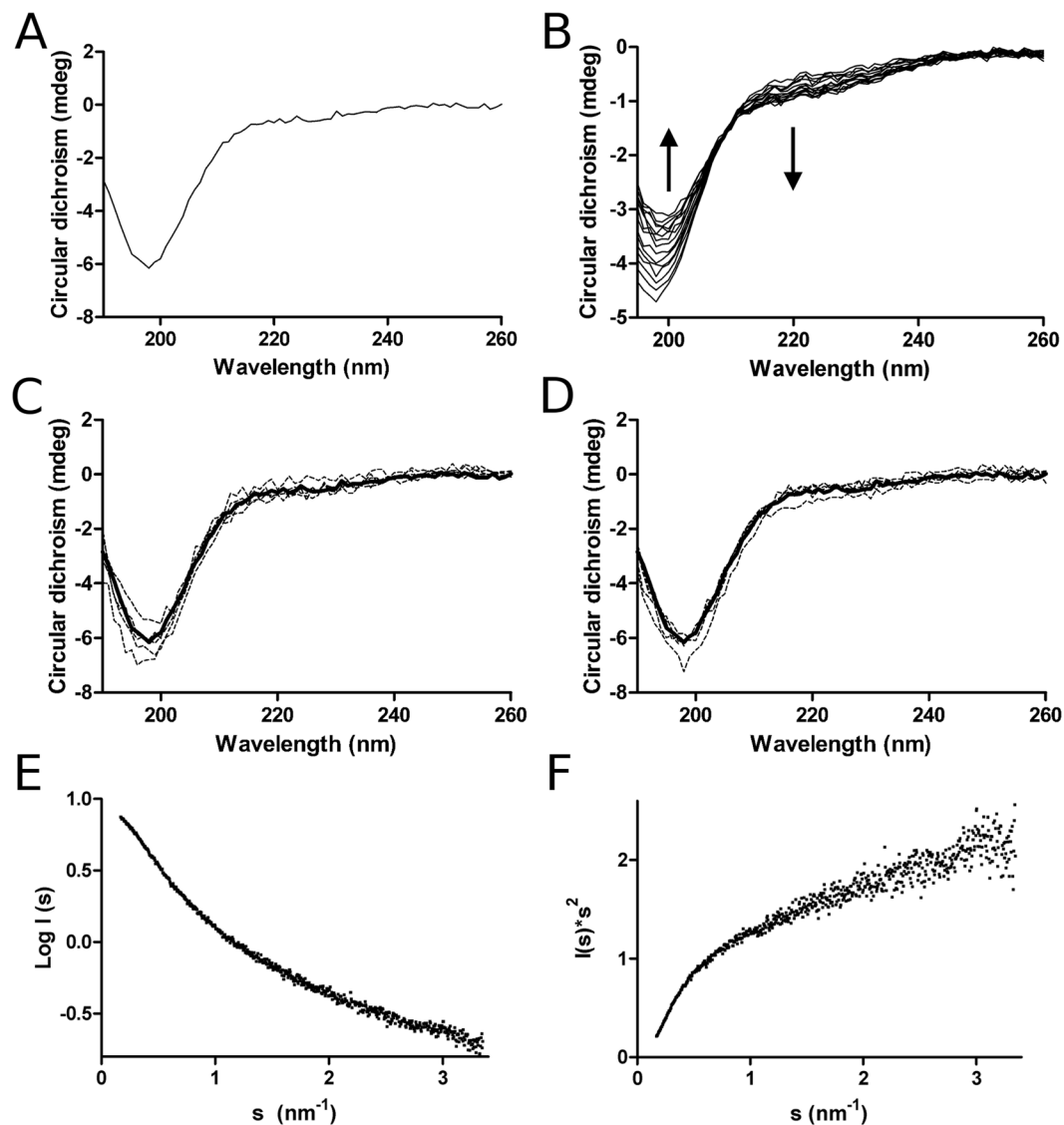
***TbPARP* is selectively activated by DNA.** hPARP-1 is activated by various DNA structures whereas hPARPs-2 and 3 appeared to be more selectively activated by DNA<sup>2,5</sup>. The selectivity of the DNA-dependent activity relates to the biological function of the proteins. Whereas hPARP-1 functions as a DNA damage sensor in the cells, hPARPs-2 and 3 appear to have more specific functions in DNA repair and maintenance of genome integrity<sup>3</sup>. In order to identify an activating DNA oligonucleotide for further studies and to get insight into the function of *TbPARP*, we tested the activation of the protein by various DNA oligonucleotides (Supplementary Table S1 and Fig. 3A,B). The oligonucleotide panel showed high selectivity in DNA substrate requirement. Non-phosphorylated DNA oligonucleotides showed little to no activation at all. In addition, the phosphorylation of the DNA changed activation properties of certain templates only. Robust activation was specifically seen with 5' phosphorylated DNA oligonucleotides with 5' or 3' overhangs. Both double-stranded and hairpin oligonucleotides with 5' or 3' overhangs combined with phosphorylation (templates 4, 6, 15, and 17) resulted in robust activation. Notably 5' phosphorylation did not change activation properties of blunt end double-stranded DNA (templates 1 and 2), single-stranded DNA (templates 10 and 11), blunt end hairpin DNA (templates 12 and 13) or dumbbell DNA (templates 18 and 19). Oligonucleotides with 3' phosphorylation behaved differently from 5' phosphorylated oligonucleotides. No activation was observed with 3' phosphorylated oligonucleotides containing 5' or 3' overhangs (templates 8 and 9). However, 3' phosphorylated blunt end oligonucleotide (template 7) resulted in moderate catalytic activation.

The DNA-binding affinities were quantified using a fluorescence polarization (FP) assay. The TBP-FL had low nanomolar affinities to dsDNA and ssDNA (templates 20 and 21) and to non-activating and activating hairpin DNA (templates 22 and 23) (Fig. 3C). Importantly, TBP-FL has similar affinities towards both activating and non-activating DNA, implying that binding affinity does not determine enzymatic activation but the protein is capable of identifying specific features on the activating DNA templates. The affinities of the fragments remain essentially unchanged with the omission of the regulatory and catalytic domains (Fig. 3D,E), as could be expected from the results from EMSA. When N-terminus is deleted (TBP-WRA), the affinity drops substantially (Fig. 3F), revealing that while WGR domain is capable of interacting with DNA, N-terminus is required for high affinity interaction. No binding was observed for TBP-RA with FP (Fig. 3G) consistent with EMSA.

Furthermore, we characterized the stoichiometry of the DNA binding. Saturating concentrations of DNA were titrated with protein and analyzed using fluorescence polarization<sup>9,10</sup>. When using dsDNA, non-activating or activating hairpin templates, an inflection point at 2:1 protein:DNA stoichiometry was observed (Fig. 3H), while ssDNA template displayed a clear 1:1 binding. With templates displaying 2:1 binding, an increasing polarization was observed after the inflection point indicating further, higher order oligomerization or the



**Figure 3.** Selectivity, affinity and stoichiometry of DNA binding. (A) DNA templates used in activation and fluorescence polarization assays. (B) Activation of TBP-FL by DNA templates measured with the fluorescence based activity assay. Activated DNA (act. DNA) was used as a control. (C–G) DNA-binding affinities for TBP-FL, TBP-NW, TBP-N, TBP-WRA, and TBP-RA were measured with fluorescence polarization assay and are presented in panels C, D, E, F, and G, respectively. A representative curve for each measurement is shown. (H) Stoichiometry of TBP-FL binding to oligonucleotides 20, 21, 22, and 23 were measured by fluorescence polarization using saturating DNA concentration (250 nM). 1:1 binding is observed with ssDNA, while 2:1 protein:DNA binding occurs with dsDNA and hairpin DNA. (I) Stoichiometry of TBP-FL in low salt (50 mM NaCl) buffer and stoichiometries of TBP-NW and TBP-N constructs using DNA oligonucleotide 20.



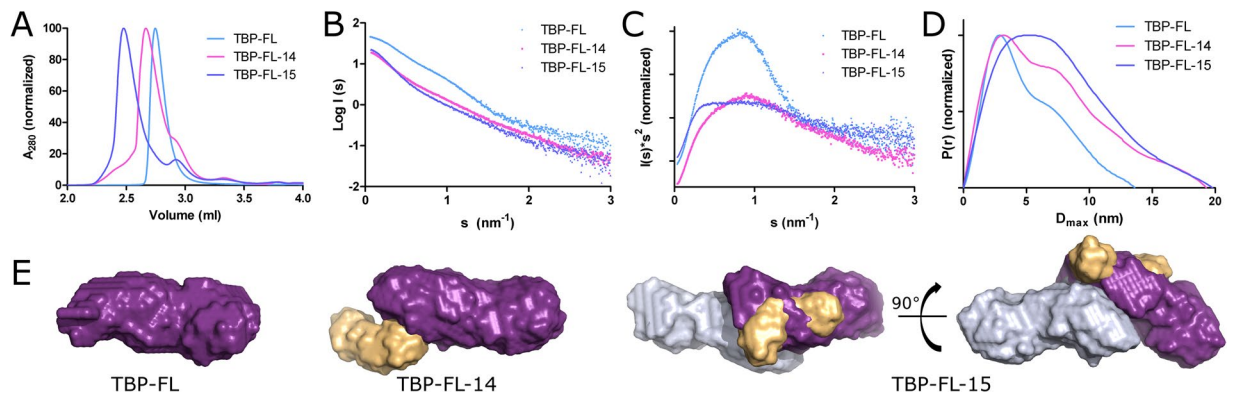
**Figure 4.** Structural properties of TBP-N. (A) CD spectra of TBP-N measured from 190 to 260 nm at 20 °C. (B) Thermal denaturation spectra (190–260 nm) of TBP-N from 20 °C to 90 °C. The arrows indicate the spectral shifts upon increasing temperature. (C,D) Effect of DNA binding to the secondary structure of TBP-N. 21 bp dsDNA (C) and 21 bp ssDNA (D) was titrated to TBP-N with different molar ratios ranging from 2:1 to 1:16 protein:DNA ratio. Bold line shows the spectra of the native protein and thin lines the DNA titrations. (E) SAXS scattering curve of TBP-N. (F) Kratky plot of TBP-N showing the disordered state of the domain.

presence of additional, lower affinity binding sites for DNA in the protein. These interactions could be due to apparent non-specificity of DNA-binding mediated by the N-terminal domain. This was tested by measuring the DNA-binding stoichiometry of TBP-N and TBP-NW fragments. Indeed, these fragments displayed no clear inflection points in the titration indicating adventitious DNA binding (Fig. 3I). Similar non-specific interaction could be seen with TBP-FL when lower salt concentration was used in the assay (Fig. 3I).

**N-terminus is structurally disordered.** While hPARP-1 interacts with DNA *via* its three zinc-fingers, no domain structure for the DNA binding N-termini of hPARP-2-3 or TbpPARP have been identified. Many DNA binding proteins harbor a disordered, positively charged terminus responsible for DNA binding<sup>11</sup>. We sought to analyze whether the N-terminus of TbpPARP would contain regular secondary structure. CD spectra of TBP-N (Fig. 4A) showed features of a disordered polypeptide without significant secondary structure<sup>12</sup>. To verify this observation a thermal denaturation experiment was conducted, which revealed no major changes in the spectra upon heating supporting the disordered nature of the fragment (Fig. 4B). This was further demonstrated with small-angle X-ray scattering (SAXS) analysis: the hyperbolic-like Kratky plot displays the features of a protein having essentially a random coil structure in solution (Fig. 4E,F)<sup>13</sup>.

Some DNA-binding domains undergo transition to a more ordered structure upon binding to DNA<sup>14</sup>. This coupled folding and binding may involve only few residues, or in some cases an entire protein domain. To test





**Figure 5.** Solution structures of *TbpPARP* and *TbpPARP*-DNA complex. HPLC chromatograms (A), SAXS scattering curves (B), Kratky plots (C),  $P(r)$  functions (D) and *ab initio* reconstructions of TBP-FL and TBP-FL-DNA complexes with DNA templates 14 and 15 (E). The DNA templates in the complexes are colored yellow, and the two *TbpPARP* monomers are colored in purple and grey in the *ab initio* model of TBP-FL-15 complex.

Protein	$D_{\max}$ (nm)	Porod volume (nm <sup>3</sup> )	$MW_{\text{exp}}$ (kDa)	$MW_{\text{calc. } ab \text{ initio}}$ (kDa)	$MW_{\text{calc. Porod}}$ (kDa)	NSD $\pm$ std
TBP-N	12.3	25	11	18	15	$0.703 \pm 0.028$
TBP-FL	13.6	106	65	71	63	$0.631 \pm 0.049$
TBP-FL-14	19.3	120	85	86	71	$0.665 \pm 0.030$
TBP-FL-15	19.8	266	150	131	156	$0.802 \pm 0.026$

**Table 1.** Summary of SAXS analysis.  $MW_{\text{exp}}$  = expected molecular weight,  $MW_{\text{calc. Porod}}$  = calculated molecular weight obtained by dividing the Porod volume by 1.7,  $MW_{\text{calc. } ab \text{ initio}}$  = calculated molecular weight obtained from DAMAVER, NSD = normalized spatial discrepancy for *ab initio* modelling.

whether significant folding of the N-terminus takes place upon DNA binding, the interaction of TBP-N with both dsDNA and ssDNA oligonucleotides was characterized with CD. The spectra of the TBP-N-DNA complexes did not show major changes when compared to the spectra of the protein alone indicating no major folding upon DNA binding (Fig. 4C,D).

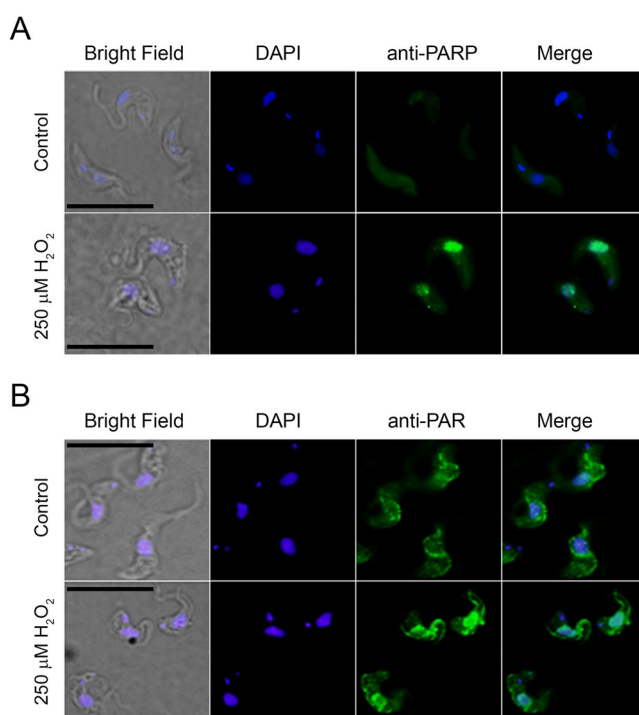
**Structural changes upon activation.** The observation that DNA binding induces protein oligomerization prompted us to study the process in more detail with SAXS (Fig. 5). We noticed that the complexes of TBP-FL with non-activating and activating DNA (templates 14 and 15, the difference of a 5' phosphate) elute with different retention volumes, indicating significantly higher particle size for the latter (Fig. 5A). Unlike the stoichiometries from fluorescence polarization, the protein-DNA complexes eluted as single peaks from the HPLC, indicating that at these conditions stable oligomers are formed. We proceeded to characterize the complexes with SAXS. We first characterized TBP-FL and found that the monomeric enzyme is characterized by an Rg of 4.04 nm with maximum dimension of 13.6 nm (Tables 1 and 2). The molecular mass derived from the SAXS data (63 and 71 kDa), is very close to the calculated monomeric weight of the protein (65 kDa). The elevated baseline at high  $s$  in the Kratky plot and the extended conformation derived from the pair-distance distribution ( $P(r)$ ) function could arise from the disordered regions in the protein and multiple domains, respectively (Fig. 5C,D).

TBP-FL in complex with non-activating DNA has a molecular mass consistent with a 1:1 complex (Table 1). The maximum particle dimension ( $D_{\max}$ ) is highly increased compared to TBP-FL indicating a flexible, non-compact particle. A two-phase *ab initio* reconstruction of the complex indicates that the binding of DNA takes place at one end of the protein molecule, likely at the N-terminus, with little to no interaction with the rest of the protein. The bound DNA extends outward from the protein leading to an elongated particle resulting in increased flexibility as seen in the Kratky plot (Fig. 5C) and extended  $D_{\max}$  (Fig. 5D).

When binding to activating DNA, the molecular mass of the complex suggests 2:1 protein:DNA ratio (Table 1) agreeing with the fluorescence polarization results (Fig. 3H) and EMSA titrations (Fig. 2B) which suggested a higher protein:DNA stoichiometry with activating DNA. The maximum particle dimension ( $D_{\max}$ ) is highly increased compared to TBP-FL in agreement with a larger particle size. The shape of the  $P(r)$  function indicates compaction of the particle in comparison to the TBP-FL (Fig. 5D). The two-phase *ab initio* reconstruction of the complex resulted in a TBP-FL dimer with DNA bound between the two protein monomers (Fig. 5E). The majority of the DNA appears to be complexed with one protein monomer, while the other monomer seems to mainly interact with another protein monomer. This implies the formation of a non-symmetrical complex, where the second monomer might not get activated by DNA-binding.

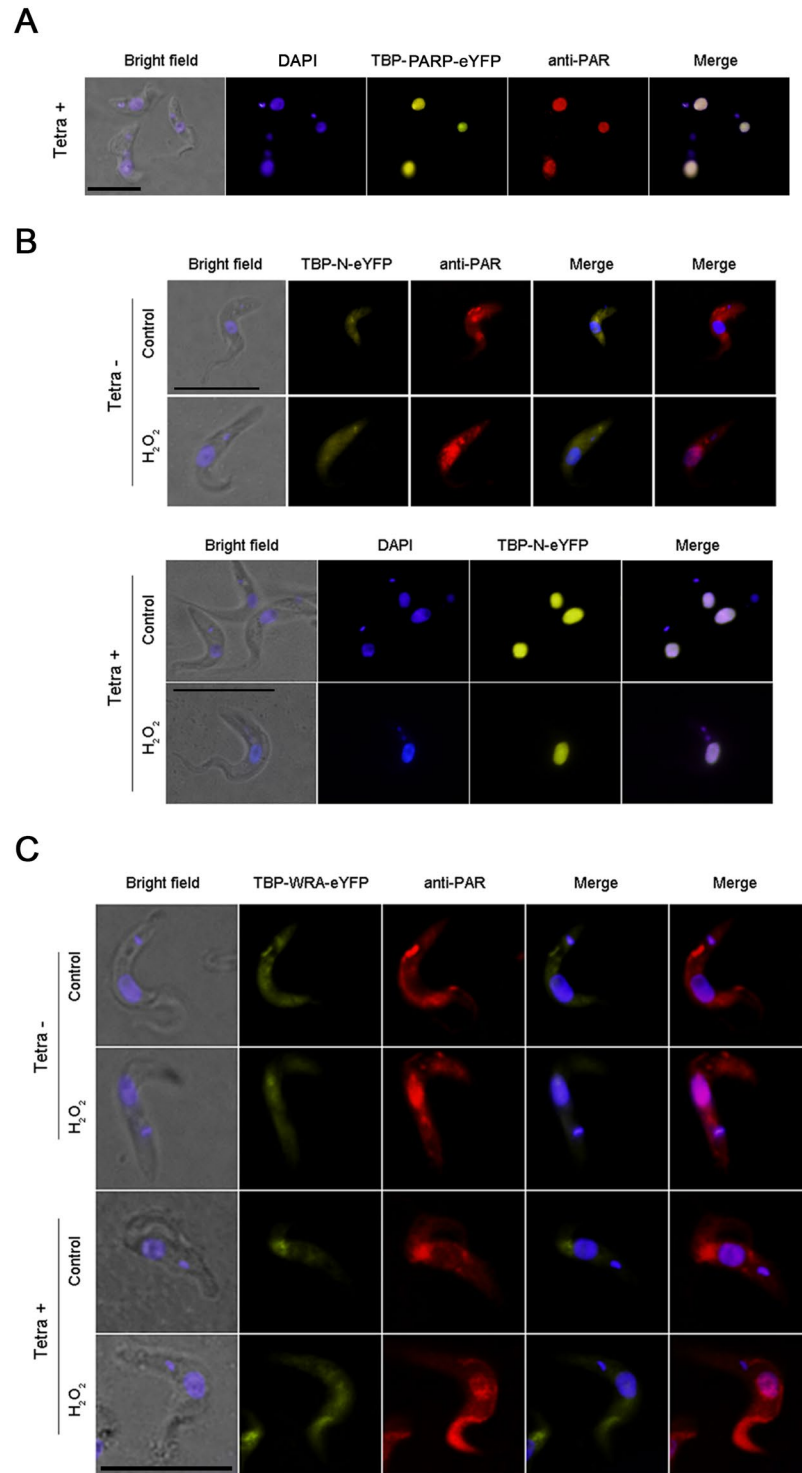
Protein	I(0)	$R_g$ from Guinier analysis (nm)	$R_g$ from GNOM analysis (nm)
TBP-N	8.49	3.46	3.67
<i>4.8 mg/ml</i>	<i>8.57</i>	<i>3.49</i>	<i>3.73</i>
<i>7.1 mg/ml</i>	<i>8.28</i>	<i>3.57</i>	<i>3.71</i>
TBP-FL	4.60	4.04	4.74
<i>2.4 mg/ml</i>	<i>4.59</i>	<i>4.11</i>	<i>4.12</i>
<i>3.5 mg/ml</i>	<i>4.56</i>	<i>4.01</i>	<i>4.13</i>
TBP-FL-14	19.77	5.45	5.65
<i>9.2 mg/ml</i>	<i>18.59</i>	<i>5.52</i>	<i>5.72</i>
<i>13.8 mg/ml</i>	<i>20.39</i>	<i>5.76</i>	<i>5.99</i>
TBP-FL-15	23.82	5.78	5.89
<i>6.0 mg/ml</i>	<i>22.00</i>	<i>5.89</i>	<i>5.95</i>
<i>9.0 mg/ml</i>	<i>24.30</i>	<i>6.06</i>	<i>6.12</i>

**Table 2.** SAXS data processing. Forward scattering ( $I(0)$ ) and radius of gyration ( $R_g$ ) is shown for both merged data and data for low and high concentrations (in italics).  $I(0)$  values have been normalized to concentration.



**Figure 6.** *TbPARP* (A) and *PAR* (B) localization was detected with anti-*TbPARP* and anti-*PAR* polyclonal antibodies, respectively, in wild type bloodstream parasites. DAPI was used to identify nuclear and kinetoplastid DNA. Merge panels show that both *TbPARP* and *PAR* are localized to the nucleus upon hydrogen peroxide treatment. Bar: 5  $\mu$ m.

**N-terminus is required for nuclear localization.** We have previously shown that *TbPARP* is localized to the nucleus in response to oxidative stress in procyclic parasites<sup>6</sup>. Here the nuclear localization of *TbPARP* was investigated in the bloodstream form of the parasite with indirect immunofluorescence microscopy. Similar to procyclic form of the parasite, *TbPARP* was found mainly in the cytoplasm, but after treatment with  $H_2O_2$  it was localized in the nucleus (Fig. 6A). Also *PAR* was found to be localized in the cytoplasm in untreated cells, and strong induction in its production was seen in the nucleus after the  $H_2O_2$  treatment (Fig. 6B). On the other hand, procyclic transgenic lines which over-expressed the fusion protein *TbPARP*-eYFP presented the heterologous enzyme in the nucleus, even in absence of a genotoxic insult (Fig. 7A). To further characterize the nuclear localization we proceeded to test whether the N-terminus is required for the nuclear import. Proteins that are transported to nucleus *via* the importin- $\alpha/\beta$  complex are recognized by their nuclear localization signals (NLS) by importin  $\alpha$ <sup>15</sup>. NLS typically consists of short stretches of lysines and arginines making the *TbPARP* N-terminus a good candidate for the location of the NLS.



**Figure 7.** Localization of fusion proteins in transgenic procyclic parasites overexpressing full-length protein (TBP-FL-eYFP) (A); or only the N-terminus (TBP-N-eYFP) (B) or the protein without the N-terminus (TBP-WRA-eYFP) (C) subjected to hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) treatment. PAR localization was also revealed with anti-PAR polyclonal antibody. DAPI was used to identify nuclear and kinetoplast DNA. The protein overexpression was induced by tetracycline. The merge panels show localization of TBP-FL-eYFP and TBP-N-eYFP to the nucleus irrespective of hydrogen peroxide treatment, while TBP-WRA-eYFP is not localized to the nucleus even after treatment with hydrogen peroxide. Bar: 5 μm.

To test the role of the N-terminus in protein localization we used procyclic parasites expressing the TBP-N-eYFP and TBP-WRA-eYFP constructs under the control of tetracyclin inducible promoter. TBP-N-eYFP



fusion protein appeared in the nucleus in 2-day tetracycline induced procyclic parasites, even in absence of H<sub>2</sub>O<sub>2</sub> treatment, unlike the uninduced control parasites (Fig. 7B). Notably, the TBP-WRA is localized in the cytoplasm without the H<sub>2</sub>O<sub>2</sub> treatment and does not appear in the nucleus even after H<sub>2</sub>O<sub>2</sub>-induced damage (Fig. 7C). Together the localization of the eYFP constructs suggest that the N-terminus is responsible for the nuclear localization of TbpPARP.

## Discussion

Unicellular protozoan parasite, *T. brucei* has only one PARP<sup>16</sup>, which is activated by DNA<sup>6</sup>. This is in agreement with PARPs in other trypanosomatids<sup>16,17</sup>. The characterized DNA-dependent PARPs function in DNA repair pathways, and have been found to be essential in mice, where PARP-1/2 knockout is embryonically lethal<sup>18</sup>. Previously, we found that *T. cruzi* amastigotes are sensitive to PARP inhibition<sup>7</sup>. In contrast, TbpPARP is not necessary for the viability of *T. brucei*<sup>6,19</sup> indicating differences in the roles of PARPs between the parasites.

Although TbpPARP is able to bind various DNA substrates it requires specifically DNA templates with single-strand overhangs combined with 5' phosphorylation for robust activation. Activation was also observed with 3' phosphorylation, but only when present at a blunt end of the DNA. Notably, the binding affinities do not change when binding to activating or non-activating DNA, implying the DNA structure does not affect the binding process. Same behavior has been observed also for human DNA-dependent PARPs<sup>2</sup>. N-terminus acts as a central module for both DNA-binding and DNA-dependent catalytic activation. This is in contrast to hPARPs-2-3, in which N-terminus is not necessary for DNA-dependent activation<sup>2,5</sup>. It seems that the N-terminus functions as a non-selective affinity enhancer in the binding event and the activation is determined by specific features of the DNA template interacting, possibly, with the WGR domain.

Structurally disordered, positively charged tails are common in DNA-binding proteins and are often utilized as affinity tuners in protein-DNA-interactions. CD and SAXS analysis revealed that N-terminus of TbpPARP is structurally disordered. Disordered N-terminus has also recently been observed for hPARP-2<sup>5,20</sup>. Importantly, no folding of the N-terminus was observed in response to DNA binding. This further supports the non-selective DNA-binding mode of the N-terminus. Accordingly, the selectivity for the DNA-dependent activation should arise from the WGR domain, the other domain capable of DNA binding. Indeed, WGR domain has been found essential for the DNA-dependent activity of hPARPs-2-3<sup>2</sup>.

Despite extensive studies on hPARPs-1-3, it is not clear whether the enzymes interact with DNA as monomers or dimers. hPARP-1 does not require specific DNA templates for activation and it has been found to form 1:1 complexes with single-strand break and blunt-end DNA<sup>21,22</sup>, and 2:1 complexes with DNA overhangs<sup>23,24</sup>. We studied the stoichiometry TbpPARP-DNA interaction with FP, and found that the protein interacts with ssDNA as a monomer but either oligomerizes upon binding to dsDNA or has additional lower affinity binding sites for dsDNA. Protein oligomerization or the presence of multiple DNA binding sites was observed when binding to both activating and non-activating DNA templates. However in EMSA (Fig. 2B) higher protein concentrations were required for complexation of activating DNA than non-activating DNA, indicating oligomerization of the protein when binding to activating DNA. When the protein-DNA complexes were separated by size-exclusion chromatography, stable complexes formed unlike in FP measurements. This is likely due to different conditions, e.g. higher salt concentration used in SEC. The purified complexes were analyzed by solution scattering methods, which indicated dimerization of the protein upon binding to activating DNA, while non-activating DNA bound to the protein with 1:1 stoichiometry. In the two-phase *ab initio* model two protein molecules were complexed with one activating DNA molecule. The DNA is primarily complexed to one PARP monomer, and it extends also outside the N-terminus, interacting possibly with the WGR domain. The second protein monomer seems to be mainly complexed with the other protein and not significantly interacting with DNA. In this case, the binding of additional protein monomers could be mediated by protein-protein interactions between the PARPs rather than with the DNA, as suggested by the binding of the second monomer in the SAXS model (Fig. 5).

TbpPARP is localized mostly in the cytosol but upon genotoxic stimulus it moves to the nucleus. N-terminus is required for this process as its deletion abolishes nuclear localization. In addition, N-terminus alone is directed to nucleus even in the absence of genotoxic stimulus. No canonical nuclear localization signal has been identified in *T. brucei* but the N-terminus contains several lysine clusters that may be used for nuclear localization. According to nuclear localization upon DNA damage (oxidative stress) TbpPARP appears to have a role in DNA repair processes. Based on the specific DNA template requirement for activation, it likely is involved in certain repair pathways like hPARP-2-3 and does not function as a global DNA repair protein like hPARP-1. Therefore the parasite seems to lack a PARP-1-like DNA damage sensor. The indication that TbpPARP is not essential for the parasite, even in the presence of genotoxic stimulus<sup>6</sup> indicates it does not participate in DNA double-strand break (DSB) repair through homologous recombination (HR) pathway. As nonhomologous end joining (NHEJ) is absent in trypanosomatids, DSBs are solely repaired through HR. Microhomology-mediated end joining (MMEJ) is usually viewed as a backup mechanism for NHEJ but in *T. brucei* it dominates end joining. However, little is known about this repair pathway in *T. brucei*<sup>25</sup>. This raises the possibility of the role of TbpPARP in MMEJ pathway as supported by the strict DNA template requirement for enzymatic activation but this remains to be addressed in future studies.

## Methods

**Molecular cloning and DNA oligonucleotides.** TbpPARP constructs (GenBank:DQ679800) TBP-RA and TBP-N (Fig. 1) were cloned by PCR extension cloning to pNIC28-Bsa4 vector (N-terminal 6xHis-tag with a TEV-protease cleavage site) and TBP-FL, TBP-WRA and TBP-NW constructs were cloned by PCR extension cloning to pNH-trxt vector with an N-terminal 6xHis-tag and thioredoxin tag with a TEV-protease cleavage site. Activated DNA was prepared by incubating calf thymus DNA (Sigma) with DnaseI<sup>26</sup>. All DNA oligonucleotides were purchased from IDT (Integrated DNA Technologies, Inc.).

**Protein expression.** Proteins were expressed in *E. coli* Rosetta 2 (DE3) cells grown in Terrific broth auto-induction media (Formedium) supplemented with 8 g/l glycerol, 50 µg/ml of kanamycin and 34 µg/ml of chloramphenicol. The cells were grown at 37 °C until OD<sub>600</sub> reached 1.0 and the temperature was lowered to 20 °C for protein expression. After 20 hours the cells were collected, suspended in lysis buffer (50 mM HEPES pH 7.4, 0.5 M NaCl, 10% glycerol, 10 mM imidazole 0.5 mM TCEP) and stored at –20 °C.

**Protein purification.** Cell suspension was supplemented with 0.5 mg/ml lysozyme (Sigma-Aldrich), 20 µg/ml DNaseI (Roche) and 0.1 mM Pefabloc SC (Sigma-Aldrich). Cells were lysed by sonication and the cell debris was cleared by centrifugation (31000 × g, 45 min at 4 °C). Supernatant was filtered through a 0.45 µm filter and loaded on a 1 ml HisTrap Chelating HP column (GE Healthcare) charged with Ni<sup>2+</sup>. The column was subsequently washed with lysis buffer, wash buffer 1 (50 mM HEPES pH 7.4, 0.5 M NaCl, 10% glycerol, 25 mM imidazole, 0.5 mM TCEP) and wash buffer 2 (50 mM HEPES pH 7.4, 1 M NaCl, 10% glycerol, 10 mM imidazole, 0.5 mM TCEP). Protein was eluted by elution buffer (50 mM HEPES pH 7.4, 500 mM NaCl, 10% glycerol, 500 mM imidazole, 0.5 mM TCEP).

For TBP-FL, TBP-NW and TBP-N the NaCl concentration of the buffer was diluted to 300 mM with 20 mM HEPES pH 7.0, 10% glycerol, 0.5 mM TCEP. The protein solution was loaded on to a 5 ml HiTrap Heparin HP column (GE Healthcare) and the column was washed with 20 mM HEPES pH 7.0, 300 mM NaCl, 10% glycerol, 0.5 mM TCEP. The protein was gradient eluted against 20 mM HEPES pH 7.0, 1 M NaCl, 10% glycerol, 0.5 mM TCEP.

For TBP-WRA the NaCl concentration was diluted to 100 mM with 0.1 M Tris pH 8.5, 10% glycerol, 0.5 mM TCEP and pH of the solution was adjusted to 8.5 with NaOH. The protein was loaded on to a 1 ml Q HP column (GE Healthcare) and the column was washed with 20 mM Tris pH 8.5, 100 mM NaCl, 10% glycerol, 0.5 mM TCEP. The protein was gradient eluted against 20 mM Tris pH 8.5, 1 M NaCl, 10% glycerol, 0.5 mM TCEP.

The fusion-tags were cleaved with TEV-protease at 4 °C overnight<sup>27</sup>. The solutions were loaded again on a Ni<sup>2+</sup> charged 1 ml HisTrap Chelating HP column and the proteins were collected from the flowthrough.

For TBP-WRA the NaCl concentration was diluted to 100 mM with 50 mM HEPES pH 7.0, 10% glycerol, 0.5 mM TCEP. The protein was loaded on to a 1 ml SP HP column (GE Healthcare) and washed with 50 mM HEPES pH 7.0, 100 mM NaCl, 10% glycerol, 0.5 mM TCEP. The protein was gradient eluted against 50 mM HEPES pH 7.0, 1 M NaCl, 10% glycerol, 0.5 mM TCEP.

The proteins were further purified with HiLoad 16/600 Superdex 75 (TBP-N, TBP-NW and TBP-RA) or HiLoad 16/600 Superdex 200 (TBP-FL and TBP-WRA) pre-equilibrated with gel filtration buffer (20 mM HEPES pH 7.4, 0.5 M NaCl, 10% glycerol, 0.5 mM TCEP). The purified proteins were analyzed with SDS-PAGE (Supplementary Fig. S1), flash-frozen in liquid N<sub>2</sub> and stored at –70 °C.

**Activity assay.** Assay was conducted as reported earlier<sup>6,7</sup>. Briefly, the enzymatic reactions were carried on a 96-well plate in quadruplicates. After the enzymatic reaction 20 µl of 20% acetophenone in ethanol and 20 µl of 2 M KOH were added to the plate and after 10 minutes of incubation 90 µl of formic acid was added. The fluorescence intensity was measured after 20 minutes using Tecan Infinity M1000 with excitation/emission wavelengths of 372 nm and 444 nm, respectively. Buffer in all the assays was 0.1 M Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub> pH 7.2, 0.5 mM TCEP, 0.5 mg/ml BSA, 200 mM NaCl. The concentration of activated DNA (calf thymus DNA treated with DNase)<sup>26</sup> in the assay was 10 µg/ml and concentration of the oligonucleotides was 2.5 µM.

**Electrophoretic mobility shift and automodification assays.** For electrophoretic mobility shift assay (EMSA) 0.5 µM DNA was incubated with 2 µM protein in 50 mM Tris pH 7, 0.1 mM TCEP, 0.1 mM EDTA, 5% glycerol, 100 mM NaCl at room temperature for 1 hour. The automodification assay was performed by incubating 2 µM TBP-FL with 500 nM DNA in 0.1 M Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub> pH 7.2, 0.5 mM TCEP, 0.5 mg/ml BSA, 200 mM NaCl with and without 10 µM EB-47 at room temperature for 1 hour. Automodification reaction was initiated by the addition of 100 µM NAD<sup>+</sup> at different time points. EMSA samples were resolved on 0.5% agarose gel with 1×TBE buffer at 4 °C at 80 V for 80 minutes and automodification assay samples were analyzed on 0.8% agarose gel with 1×TBE buffer at 4 °C at 80 V for 1 h. The gels were stained with 1× GelRed (Biotium) or Midori Green Direct (Nippon Genetics) for automodification assay and EMSA, respectively.

**Fluorescence polarization assay.** Fluorescence polarization was measured in a 96-well U-bottom black polypropylene plates (Greiner BioOne) in a final volume of 180 µl. The reaction was carried out in 10 mM HEPES pH 8, 0.1 mM TCEP, 0.1 mM EDTA, 175 mM NaCl and contained 5 nM of labelled DNA and various concentrations of protein. In the case of stoichiometric titrations, saturating (250 nM) concentration of labelled DNA was used. The plate was incubated at 25 °C with shaking at 300 rpm using PST-60 HL plus Thermo Shaker (Biosan, Riga, Latvia) for 80 min. Fluorescence polarization was measured on a Tecan Infinity M1000 with excitation/emission wavelengths of 475 and 520 nm, respectively. Three independent measurements (two with TBP-WRA) were fitted separately with a one-site binding model using GraphPad Prism version 5.04 for Windows (GraphPad Software).

**Circular dichroism spectroscopy.** Circular dichroism (CD) spectra of TBP-N were recorded with Chirascan™ CD Spectrometer (Applied Photophysics Ltd) using a wavelength range from 190–260 nm. Measurements were performed at 20 °C in a 1 mm-path length quartz cuvette in 10 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub> pH 7.2, 150 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. The data were analyzed using Pro-Data™ Software Suite (Applied Photophysics Ltd). Thermal denaturation spectra were measured from 20 °C to 90 °C for every 5 °C. DNA titrations were performed from 2:1 to 1:16 protein:DNA molar ratios.

**Small-angle X-ray scattering.** Prior to measurements, the protein–DNA-complexes were purified with size-exclusion chromatography using KW403-4F HPLC column (Shodex) in 20 mM HEPES pH 7.0, 300 mM

NaCl, 0.5 mM TCEP. The data were collected on BM29 beamline at ESRF (Grenoble, France) using protein concentrations between 1–9 mg/ml. The data were analyzed using the ATSAS suite<sup>28</sup> (Table 2). Data processing and  $R_g$  determination were done with PRIMUS<sup>29</sup>. The Guinier plots were displayed with ScÅtter (Supplementary Fig. S2) ([www.bioisis.net](http://www.bioisis.net)). Distance distribution functions (Pr) and maximum distances ( $D_{max}$ ) were determined with GNOM<sup>30</sup> and Porod volumes were determined with DATPOROD<sup>28</sup>. Molecular weight estimate was obtained by dividing the Porod volume by 1.7. *Ab initio* model of the TBP-FL was generated with DAMMIF<sup>31</sup> and 20 individual models were averaged with DAMAVER<sup>32</sup>. *Ab initio* models of the DNA complexes were created with MONSA<sup>33</sup>. 10 independent MONSA modellings were performed and averaged using DAMAVER. The theoretical scattering curves for DNA used in MONSA modelling was generated with CRY SOL<sup>34</sup>.

**Parasite cultures.** Cells of the procyclic form of *T. brucei* strain 29–13<sup>35</sup> were cultured at 28 °C in SDM-79 (Bioscience) supplemented with 10% (v/v) FCS and 0.002% hemin. Cells of the bloodstream form of *T. brucei* strain 427 90–13<sup>35</sup> were cultured at 37 °C in HMI-11 (Iscore's Modified Dulbecco's Medium (Invitrogen)), 100 mg/l sodium pyruvate, 136.1 mg/l hypoxanthine, 38.7 mg/l thymidine, 28.22 mg/l bathocuproinedisulfonic acid, 181.8 mg/l L-cysteine, 3.024 mg/l sodium carbonate, 196 µM β-mercaptoethanol supplemented with 10% (v/v) FCS.

**Procyclic transgenic lines.** TBP-N and TBP-WRA were sub-cloned into the expression vector p2216<sup>36</sup> to overexpress TBP-N-eYFP and TBP-WRA-eYFP fusion genes, respectively. The constructs were linearized with NotI and transfected into procyclic *T. brucei* strain 29–13 by electroporation. Briefly, 10<sup>9</sup> cells were harvested, washed once with cytomix buffer (120 mM KCl, 0.15 mM CaCl<sub>2</sub>, 10 mM K<sub>2</sub>HPO<sub>4</sub>, 25 mM HEPES (pH 7.6), 2 mM EDTA, 5 mM MgCl<sub>2</sub>), and suspended in 0.5 ml of the same buffer containing 10 µg of the construct. Electroporation was carried out in a 2-mm cuvette using the Gene Pulser (Bio-Rad) with parameters set as follows: 1.5 kV voltage, 25 microfarads capacitance and 2 pulses. The cells were transferred to 10 ml of antibiotics-supplemented SDM-79 medium immediately after electroporation and incubated at 28 °C for 24 h. The transfectants were then selected under 20 µg/ml zeocin until stable cell lines were grown up after about 3 weeks of continuous incubation. For induction of overexpression of fusion proteins, cells were cultured in the above medium containing 1.0 µg/ml tetracycline. The presence of the over-expressing transgenes was confirmed by Southern-blot and protein expression was determined by Western blot analysis.

**Immunolocalization assays.** Wild type bloodstream cultures were grown in HMI-11 medium for 24 h up to a density of 5 × 10<sup>5</sup> parasites/ml, and treated with 250 µM of H<sub>2</sub>O<sub>2</sub> for 10 minutes in the same culture media. Transgenic procyclic cultures were grown in SDM-79 medium up to a density of 5 × 10<sup>6</sup> parasites/ml and treated with 500 µM of H<sub>2</sub>O<sub>2</sub> for 10 minutes.

Parasites were fixed with 3.8% (W/V) formaldehyde in PBS at 4 °C, permeabilized with fresh PBS-0.1% Triton X-100 and blocked at room temperature for 1 h. PAR was detected with 1:500 rabbit polyclonal anti-PAR antibody (BD) followed by 1:500 Alexa Fluor 488 goat anti-rabbit IgG antibody (Invitrogen). *Tb*PARP was detected with 1:100 rabbit polyclonal anti-*Tb*PARP antibody (GenScript), followed by 1:500 Alexa Fluor 594 goat anti-rabbit IgG antibody (Invitrogen). Fusion proteins TBP-PARP-eYFP, TBP-N-eYFP and TBP-WRA-eYFP were localized with fluorescence of eYFP. Excess of antibody was removed by 3 × 5 min washes in PBS, and nuclear and kinetoplast DNA stained with 2 µg/ml DAPI (Sigma). Coverslips were washed with distilled water and mounted in Mowiol and then visualized using an Olympus BX41 microscope.

## References

- Hottiger, M. O., Hassa, P. O., Lüscher, B., Schüler, H. & Koch-Nolte, F. Toward a unified nomenclature for mammalian ADP-ribosyltransferases. *Trends Biochem. Sci.* **35**, 208–219 (2010).
- Langelier, M.-E., Riccio, A. A. & Pascal, J. M. PARP-2 and PARP-3 are selectively activated by 5' phosphorylated DNA breaks through an allosteric regulatory mechanism shared with PARP-1. *Nucleic Acids Res.* **42**, 7762–7775 (2014).
- Beck, C., Robert, I., Reina-San-Martin, B., Schreiber, V. & Dantzer, F. Poly(ADP-ribose) polymerases in double-strand break repair: Focus on PARP1, PARP2 and PARP3. *Exp. Cell Res.* **329**, 18–25 (2014).
- Sonnenblick, A., de Azambuja, E., Azim, H. A. Jr. & Piccart, M. An update on PARP inhibitors—moving to the adjuvant setting. *Nat. Rev. Clin. Oncol.* **12**, 27–41 (2015).
- Obaji, E., Haikarainen, T. & Lehtiö, L. Characterization of the DNA dependent activation of human ARTD2/PARP2. *Sci. Rep.* **6**, 34487 (2016).
- Schlesinger, M. *et al.* Disrupted ADP-ribose metabolism with nuclear Poly (ADP-ribose) accumulation leads to different cell death pathways in presence of hydrogen peroxide in procyclic *Trypanosoma brucei*. *Parasit. Vectors* **9**, 173 (2016).
- Vilchez Larrea, S. C. *et al.* Inhibition of poly(ADP-ribose) polymerase interferes with *Trypanosoma cruzi* infection and proliferation of the parasite. *PLoS One* **7**, e46063 (2012).
- Murai, J. *et al.* Differential trapping of PARP1 and PARP2 by clinical PARP inhibitors. *Cancer Res.* **72**, 5588–5599 (2012).
- Updegrove, T. B., Correia, J. J., Chen, Y., Terry, C. & Wartell, R. M. The stoichiometry of the *Escherichia coli* Hfq protein bound to RNA. *RNA N. Y. N.* **17**, 489–500 (2011).
- Wojtuszewski, K., Hawkins, M. E., Cole, J. L. & Mukerji, I. HU binding to DNA: evidence for multiple complex formation and DNA bending. *Biochemistry* **40**, 2588–2598 (2001).
- Vuzman, D. & Levy, Y. Intrinsically disordered regions as affinity tuners in protein-DNA interactions. *Mol. Biosyst.* **8**, 47–57 (2012).
- Ranjbar, B. & Gill, P. Circular dichroism techniques: biomolecular and nanostructural analyses- a review. *Chem. Biol. Drug Des.* **74**, 101–120 (2009).
- Rambo, R. P. & Tainer, J. A. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **95**, 559–571 (2011).
- Dyson, H. J. & Wright, P. E. Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **12**, 54–60 (2002).
- Goldfarb, D. S., Corbett, A. H., Mason, D. A., Harreman, M. T. & Adam, S. A. Importin alpha: a multipurpose nuclear-transport receptor. *Trends Cell Biol.* **14**, 505–514 (2004).
- Fernández Villamil, S. H. *et al.* TcPARP: A DNA damage-dependent poly(ADP-ribose) polymerase from *Trypanosoma cruzi*. *Int. J. Parasitol.* **38**, 277–287 (2008).

17. Podestá, D., García-Herreros, M. I., Cannata, J. J. B., Stoppani, A. O. M. & Fernández Villamil, S. H. Purification and properties of poly(ADP-ribose)polymerase from *Crithidia fasciculata*. Automodification and poly(ADP-ribosyl) ation of DNA topoisomerase I. *Mol. Biochem. Parasitol.* **135**, 211–219 (2004).
18. Ménissier de Murcia, J. *et al.* Functional interaction between PARP-1 and PARP-2 in chromosome stability and embryonic development in mouse. *EMBO J.* **22**, 2255–2263 (2003).
19. Alsford, S. *et al.* High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Res.* **21**, 915–924 (2011).
20. Riccio, A. A., Cingolani, G. & Pascal, J. M. PARP-2 domain requirements for DNA damage-dependent activation and localization to sites of DNA damage. *Nucleic Acids Res.* **44**, 1691–1702 (2016).
21. Eustermann, S. *et al.* The DNA-binding domain of human PARP-1 interacts with DNA single-strand breaks as a monomer through its second zinc finger. *J. Mol. Biol.* **407**, 149–170 (2011).
22. Langelier, M.-F., Planck, J. L., Roy, S. & Pascal, J. M. Structural basis for DNA damage-dependent poly(ADP-ribosyl) ation by human PARP-1. *Science* **336**, 728–732 (2012).
23. Ali, A. A. E. *et al.* The zinc-finger domains of PARP1 cooperate to recognize DNA strand breaks. *Nat. Struct. Mol. Biol.* **19**, 685–692 (2012).
24. Pion, E. *et al.* Poly(ADP-ribose) polymerase-1 dimerizes at a 5' recessed DNA end *in vitro*: a fluorescence study. *Biochemistry (Mosc.)* **42**, 12409–12417 (2003).
25. Glover, L., McCulloch, R. & Horn, D. Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes. *Nucleic Acids Res.* **36**, 2608–2618 (2008).
26. Aposhian, H. V. & Kornberg, A. Enzymatic synthesis of deoxyribonucleic acid. IX. The polymerase formed after T2 bacteriophage infection of *Escherichia coli*: a new enzyme. *J. Biol. Chem.* **237**, 519–525 (1962).
27. van den Berg, S., Löfdahl, P.-A., Hård, T. & Berglund, H. Improved solubility of TEV protease by directed evolution. *J. Biotechnol.* **121**, 291–298 (2006).
28. Petoukhov, M. V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 (2012).
29. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. *PRIMUS*: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277–1282 (2003).
30. Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503 (1992).
31. Franke, D. & Svergun, D. I. *DAMMIF*, a program for rapid *ab-initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42**, 342–346 (2009).
32. Volkov, V. V. & Svergun, D. I. Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* **36**, 860–864 (2003).
33. Svergun, D. I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **76**, 2879–2886 (1999).
34. Svergun, D., Barberato, C. & Koch, M. H. J. *CRY SOL*—a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.* **28**, 768–773 (1995).
35. Wirtz, E., Leal, S., Ochatt, C. & Cross, G. A. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **99**, 89–101 (1999).
36. Kelly, S. *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.* **154**, 103–109 (2007).

## Acknowledgements

This work was funded by Biocenter Oulu, Academy of Finland (grant No. 287063 and 294085 for LL and grant No. 266922 for TH) and Sigrid Jusélius Foundation. This work was supported by Agencia Nacional de Promoción Científica y Tecnológica (Argentina), Consejo Nacional de Investigaciones Científicas y Técnicas (Argentina), Universidad de Buenos Aires (Argentina) and EMBO Short Term Fellowship. MS is a fellow of the CONICET. Small-angle X-ray scattering experiments were performed at ESRF (Grenoble, France) on beamline BM29. We are grateful to Local Contacts at ESRF for providing assistance in using the beamline. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under BioStruct-X (Grant agreement No. 283570). The use of the facilities of the Biocenter Oulu Proteomics and protein analysis core facility is gratefully acknowledged.

## Author Contributions

T.H. measured and analyzed *in vitro* protein data and wrote the manuscript; M.S. carried out the cell-based assays and wrote the manuscript; E.O. collected SAXS data, contributed to SAXS data analysis and performed EMSA assays; S.H.F.V. and L.L. analyzed the data and wrote the manuscript. All authors contributed to the final version of the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-03751-4

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017