Combinatorial selection of high affinity RNA ligands to live African trypanosomes

Matthias Homann and H. Ulrich Göringer*

Laboratorium für Molekulare Biologie, Genzentrum der Ludwig-Maximilians-Universität München, Am Klopferspitz 18a, 82152 Martinsried, Germany

Received January 14, 1999; Revised and Accepted March 3, 1999

ABSTRACT

African trypanosomiasis is a parasitic disease caused by a specific class of protozoan organisms. The best-studied representative of that group is *Trypanosoma brucei* which is transmitted by tsetse flies and multiplies in the blood of many mammals. Trypanosomes evade the immune system by altering their surface structure which is dominated by a layer of a variant surface glycoprotein (VSG). Although invariant surface proteins exist, they are inaccessible to the humoral immune response. Using a combinatorial selection method in conjunction with live trypanosomes as the binding target, we show that short RNA ligands (aptamers) for constant surface components can be isolated. We describe the selection of three classes of RNA aptamers that crosslink to a single 42 kDa protein located within the flagellar pocket of the parasite. The RNAs associate rapidly and with high affinity. They do not discriminate between two different trypanosome VSG variant strains and, furthermore, are able to bind to other trypanosome strains not used in the selection protocol. Thus, the aptamers have the potential to function as markers on the surface of the extracellular parasite and as such they might be modified to function as novel drugs against African trypanosomiasis.

INTRODUCTION

African trypanosomes are unicellular, unflagellated protozoan parasites. They cause African trypanosomiasis or sleeping sickness, a chronic disease in humans as well as in wild and domestic animals. An estimated 50 000 000 people world wide are at risk of such an infection, and Smith *et al.* (1) recently estimated a number of about 300 000 newly infected cases/year. Moreover, the disease in domestic animals considerably limits agricultural development in a substantial part of Africa, thus indirectly affecting an even larger number of people (2,3). None of the available therapeutic measures against the various diseases is very effective and therefore novel approaches are needed in the development of drugs against trypanosome infections (4).

African trypanosomes are transmitted by tsetse flies, and as extracellular parasites they multiply within the peripheral blood, the capillary beds and within the tissue fluids of the infected hosts. During this bloodstream life cycle stage, trypanosomes are covered with a layer of approximately 10 000 000 molecules of a glycoprotein species known as variant surface glycoprotein (VSG). VSG molecules have a molecular mass of ~60 kDa, they homodimerise and are glycosylphosphatidylinositol-anchored within the plasma membrane. The VSG surface induces a strong T-cell-independent IgM response as well as a T-cell-dependent B-cell response which elicits VSG-specific IgG (5; reviewed in 6). The parasites, however, evade the host immune system by temporarily expressing different VSG variants (7–9). This phenomenon has been termed antigenic variation and has its molecular basis in the surface presentation of structurally polymorphic N-terminal domains of the different VSG variants. Though the C-termini of the various VSG molecules are structurally conserved, they are buried within the VSG layer and thus not accessible to the humoral immune response (10–12). The trypanosome genome encodes a repertoire of about 1000 different *vsg* genes, but only one VSG is expressed at a given time. The switching frequency from one variant to the next has been estimated to range from 10^-2 to 10^-7/cell generation and seems to be a stochastic event (13,14). Thus, the VSG surface acts as an exclusion barrier for larger molecules, such as antibodies, while its variable characteristics cause the inability of the infected host to clear the infection. Besides, the surface changes complicate any vaccination approach.

Although the entire surface of bloodstream stage African trypanosomes is covered by the VSG protein, the coat nevertheless shows a highly dynamic structure. VSG molecules show rapid lateral movement within the cell membrane (15). In addition, other surface components have been identified, including invariant surface glycoproteins (ISGs), receptor complexes and transporter molecules (16–20). Some of these molecules are embedded within the VSG layer and distributed over the entire surface of the trypanosome cell body including the flagellum. Other surface components are localised to the flagellar pocket, an invagination of the plasma membrane around the base of the flagellum which functions as an exo- and endocytosis site (reviewed in 21). Moreover, the trypanosome surface is not an impenetrable casing. Molecules of lower molecular mass, such as the protease trypsin (23 kDa), have been shown to cleave VSG molecules at a site deep within the protein layer (17,22). This suggests the existence

*To whom correspondence should be addressed. Tel: +49 89 8578 2475; Fax: +49 89 8578 3810; Email: goeringe@biochem.mpg.de
of transient molecular cavities within the layer of the VSG homodimers.

In this study, we address the potential of tagging the surface of live bloodstream stage trypanosomes with molecules that bind with high affinity and specificity. If such molecules were small enough to reach and interact with invariant elements of the cell surface, binding would be independent of the expressed VSG and could be used to direct the immune response to the trypanosome surface. *In vitro* evolution methods in conjunction with combinatorial nucleic acid libraries offer effective strategies for selecting nucleic acid ligands to biological targets. The SELEX (systematic evolution of ligands by exponential enrichment) protocol (23,24) has been widely used to isolate high affinity binders, so-called aptamers, to a variety of biological targets (reviewed in 25) including complex targets such as the membranes of human red blood cells (26).

Using this methodology, we report the isolation of three classes of RNA aptamers that bind to live African trypanosomes. The binding is specific for the infective bloodstream life cycle stage of the parasites, because insect stage trypanosomes are not recognised by the selected RNAs. Individual aptamers are able to bind rapidly to the parasite surface with high affinity and specificity for a single protein located within the flagellar pocket. Although two separate selections with two cloned trypanosome strain variants were performed, the selected aptamers show no variant-specific features and furthermore are capable of binding other, unrelated trypanosome strains.

**MATERIALS AND METHODS**

**Trypanosomes**

The bloodstream life cycle stage of *Trypanosoma brucei* was grown in HMI-9 medium (27) supplemented with 10% (v/v) heat-inactivated bovine foetal calf serum (FCS). Incubation was at 37°C in 5% CO₂/95% air. The following strains were used: *T.brucet* 427-MITat serodeme, variant clones MITat 1.2 and MITat 1.4 (28); AnTat 1.1 (29); IL T A T 1.1 (30). Insect stage trypanosomes were grown at 27°C in SDM-79 medium supplemented with 10% (v/v) FCS (31).

**Oligodeoxynucleotide and RNA pool synthesis**

DNA oligonucleotides were synthesised by automated solid support chemistry using O-cyanoethyl-N,N-diisopropylphosphoramidites. The starting DNA library was a 79mer of sequence GAAT-TCAGTCGACAGCC(N)₄₀GATGACGAGATCTGCCTCCC which contained a central sequence stretch of 40 randomised nucleotides. The DNA was purified by reverse phase liquid chromatography and 1.1 mg were amplified during seven PCR cycles in a total volume of 60 ml using primers A (GC-GAACTTTAATACGACTATAGGAGACCATGTCCTCC) and B (GGTGCGAATTCGATCGAGACCCGG). Of the 7.5 mg of double-stranded PCR product, 1.5 mg were transcribed into 2.5 mg RNA (reaction volume 7.5 ml) from which 1.15 mg were used in the first round of each selection. All subsequent RNA pools were transcribed from 50–150 µg of the PCR-generated double-stranded DNA templates in a final volume of 1–1.5 ml. Reactions were performed in the presence of 5 µCi [α-³²P]-UTP (3000 Ci/mmol) in 20 mM Na x H y PO ₄, pH 7.9, 8 mM MgCl₂, 20 mM DTT, 4 mM spermidine, containing 1 mM each NTP and 500 U T7 RNA polymerase. The synthesised RNAs were extracted with phenol, precipitated with ethanol and redissolved in a buffer containing 20 mM Na x H y PO ₄, pH 7.4, 2 mM MgCl₂, 130 mM NaCl, 5 mM KCl, 20 mM glucose and 0.2 mM β-mercaptoethanol. A final purification was achieved by size exclusion chromatography on Sephadex G50 columns using the same buffer. Before each round of selection, all RNAs were pre-incubated at 37°C for 30 min.

**In vitro selection**

Trypanosomes were harvested at a cell density of 1–2 × 10⁶ cells/ml and extensively washed in binding buffer (20 mM Na x H y PO ₄, pH 7.4, 2 mM MgCl₂, 130 mM NaCl, 5 mM KCl, 20 mM glucose, 0.2 mM β-mercaptoethanol). RNA binding was performed in a final volume of 0.5–1.5 ml with ³²P-labelled pool RNA (10–20 µM) and either 5 × 10⁶ (rounds 1–7) or 2 × 10⁶ cells/ml (rounds 8–13). After an incubation for 60 min at 30°C, unbound and weakly associated RNAs were washed off by four consecutive washes with 2 ml binding buffer. Trypanosomes and bound RNAs were collected by centrifugation and the bound RNAs recovered by two phenol extractions followed by precipitation with ethanol. Isolated RNA was reverse transcribed (150 µl reaction volume, 10 µg primer B, 100 U M-MulY RT), amplified by PCR and the resulting DNA templates were transcribed into RNA which was used for the next round of selection. DNA templates from pool 12 were digested with HindIII and EcoRI and cloned into pUC118. Sequences of 53 clones of each selection were determined by dideoxy terminator sequencing.

**Determination of binding parameters**

Uniformly ³²P-labelled pool RNAs (5 nM, ~20 000 c.p.m.) or RNA aptamers (3–5 nM, 10 000–50 000 c.p.m.) were incubated with *T.brucet* cells in 100 µl binding buffer at cell densities of 1–3 × 10⁶/ml for bloodstream stage cells and 5 × 10⁶/ml for insect stage trypanosomes. After a 30 min incubation at 37°C, the cells were washed (four times with 2 ml binding buffer), collected by centrifugation and the percentage of bound RNA was determined by scintillation counting. Average relative binding efficiencies were calculated from three independent experiments and were normalised to the binding efficiency of pool 1 RNA. Equilibrium dissociation constants were derived from experiments determining the percentage of bound RNA at varying trypanosome cell numbers. *K₅₀* values were calculated based on the Scatchard equation

\[ r [A] = n K_d \left( 1 - \frac{r [A]}{K_d} \right) \]

where [A] is the free RNA, *r* = B/[cells] and *B* is bound RNA. The number of binding sites on the trypanosome surface (*n*) was derived from the *x*-intercept of the Scatchard plot. Association rates were determined with uniformly ³²P-labelled DNA aptamer 2-16 (0.75 nM, ~50 000 c.p.m.) and varying amounts of *T.brucet* cells in 50 µl binding buffer at 37°C. Parasite cell densities varied between 2.5 × 10⁷ and 2 × 10⁸/ml corresponding to calculated target concentrations of 3–24 nM (assuming 72 000 targets/cell). Under the conditions of a large excess of one binding partner, a pseudo-first-order rate constant *kₐₜ* can be determined from

\[ k_{obs} = \frac{d [A]}{dt} \]

where [A] is the free RNA, *r* = B/[cells] and *B* is bound RNA. The number of binding sites on the trypanosome surface (*n*) was derived from the *x*-intercept of the Scatchard plot.
**Run-off transcription, 5′-end-labelling and biotinylation**

RNA aptamers were transcribed from linearised plasmid DNA templates using T7 RNA polymerase following standard procedures. DNA templates (1 µg) were transcribed in a buffer containing 20 mM Na₂HPO₄, pH 7.9, 8 mM MgCl₂, 20 mM DTT, 4 mM spermidine and 1 mM each NTP. Radioactive labelling of the RNAs was achieved by using [α-³²P]UTP (10–50 µCi) and 5 µM UTP in the transcription reaction. The DNA was digested with DNase I and RNA transcripts were phenol extracted and purified from non-incorporated NTPs by size exclusion chromatography. Aptamer RNAs were 5′-end-labelled using alkaline phosphatase and T4 polynucleotide kinase and [γ-³²P]ATP according to standard procedures.

Fluorescence labelling of aptamer 2-16 was achieved by tagging the BODIPY TMR-C₅ fluorophore (Molecular Probes) to the 5′-end of the RNA. The reaction was performed according to the manufacturer’s specification and fluorophore-conjugated 2-16 RNA was dissolved in a PBS-based binding buffer at 0.4 µg/µl.

Biotinylation of aptamer 2-16 was performed by *in vitro* transcription in a reaction mixture containing 1 mM each GTP, CTP and ATP, 0.05 mM UTP and 5–50 µM biotin-16-UTP (Boehringer Mannheim). RNA transcripts were separated from non-incorporated nucleotides by size exclusion chromatography, precipitated with ethanol and dissolved in a PBS-based binding buffer at 20 ng/µl.

**RNA structure determination**

Theoretical secondary structures were calculated using the Mfold subroutine of the GCG software package based on a free energy minimisation algorithm (Wisconsin Package v.9.1; Genetics MFold subroutine of the GCG software package based on a free energy minimisation algorithm). Theoretical secondary structures were calculated using the Wisconsin Package v.9.1; Genetics MFold subroutine of the GCG software package based on a free energy minimisation algorithm.

**Photo-crosslinking of RNA aptamers to living cells**

Uniformly ³²P-labelled RNA aptamers (5–10 nM) were incubated with *Tbrucei* cells (2–8 × 10⁹/ml) in binding buffer at 37°C for 20 min. Following incubation, the cells were put on ice and irradiated with UV light (254 nm) for 10 min at an energy dose of 0.12 J/cm². The samples were heat denatured (95°C, 5 min), digested with a cocktail of RNase A, RNase T1 and DNase I (30 min, 37°C) and finally analysed by discontinuous SDS–PAGE. Gels were stained with Coomassie brilliant blue, dried and the ³²P-labelled proteins were visualised by autoradiography.

**Fluorescence in situ hybridisation**

Bloodstream stage trypanosomes were harvested at a cell density of 2 × 10⁹/ml and washed in binding buffer (four times). Samples of 2 × 10⁷ parasites were incubated at 37°C with either biotinylated (2 ng/µl) or 5′-BODIPY TMR-C₅-labelled aptamer RNA (10 ng/µl) in 40 µl binding buffer. After 20 min the cells were washed (twice with 2 ml binding buffer) to remove unbound RNA and put on ice. Cells were smeared onto microscope slides, briefly air dried and mounted in ProLong Antifade (Molecular Probes). In the case of biotinylated aptamer preparations, the parasites were further incubated with a 1:100 dilution of a monoclonal Cy⁵-conjugated anti-biotin antibody (Dianova). Nuclear and kinetoplast DNAs were stained with Hoechst H-33342 (Molecular Probes). Slides were examined with a digital fluorescence microscope. Images were captured and merged using IP Lab Spectrum 3.1.2c (Scanalytics Inc.). Assembly into figures and false colouring were performed using Adobe® Photoshop 3.0.3 (Adobe® Systems Inc.).

**RESULTS**

**Selection of RNAs that bind to live trypanosomes**

For the selection we used a starting pool of 10¹⁶ RNA molecules with an estimated complexity of 2 × 10¹³ unique sequences. The library contained 40 randomised nucleotides flanked by primer binding sites 21 and 24 nt in length, which together restricted the final molecular mass of the selected aptamers to a value around 25–27 kDa. In the case of the protease trypsin, this has been shown to be a molecular size still able to penetrate the VSG layer (22). Two independent selections were performed in parallel using two different *Tbrucei* variant antigen clones as target cells: MITat 1.2, a cell line which stably expresses VSG 221, and MITat 1.4, which express VSG 117 on their surface (28). These two variant strains were further chosen because of steric differences in the VSG packing on the parasite surface: they respond differently when incubated with an anti-ISG75 antibody or with the protease trypsin (17). Lastly, the crystal structure for the N-terminal domain of the MITat 1.2 VSG is known to a resolution of 2.9 Å (33).

Both parasite variants were grown as bloodstream stage trypanosomes and RNA binding was accomplished over a 60 min incubation period. Despite some RNA degradation, a significant amount of full-length RNAs remained intact after the incubation (Fig. 1A) which was processed further. In both experiments, 13 rounds of selective binding and subsequent amplification were performed. Maximal enrichment of trypansome-interacting RNAs was achieved in both selections in round 12 at which time ∼18% of the input RNAs was associated with the parasite cells (Fig. 1B). Confirmation of a reduction in the pool complexities with increasing rounds of selection was derived from analytical RNase T1 digests of the different ³²P-labelled RNA pools. Starting at rounds 8 and 9, distinct RNase cleavage products could be resolved electrophoretically, indicating the selective enrichment of defined RNA sequences (Fig. 1C).

DNA templates from RNA pools of round 12 were cloned and the sequences of 53 clones from each selection experiment were determined. A majority of these sequences can be grouped into two families (I and II) based on common sequence motifs (Fig. 2). Family I members are characterised by an invariant CGT and a conserved NGC base triplet which are flanked by sequences that show a high degree of co-variation, suggesting base pairing. Two of the family I sequences were found in both selections. The first one is represented by the identical aptamers 2-11 and 4-2, accounting for ∼50% of all clones in both selections. The second sequence (aptamers 2-16 and 4-10) appeared at a frequency of 24% in the selection using MITat 1.2 trypanosomes and at a frequency of 7% in the selection with MITat 1.4 parasites. Thus,
Figure 1. Selective enrichment of RNA ligands that interact with live trypanosomes. (A) Degradation kinetics of $^{32}$P-labelled pool 1 RNA (85 nt) upon incubation with either MITat 1.2 or MITat 1.4 trypanosomes. (B) Graphical representation of the percentage of trypanosome-bound RNAs in each round of selection using MITat 1.2 or MITat 1.4 cells. Binding was performed at 30°C using 5 nM RNA and a cell density of $3 \times 10^8$ cells/ml. (C) Analytical RNase T1 digestion of different RNA pools from both selections. Uniformly $^{32}$P-labelled aptamer pools were digested with RNase T1 at limiting conditions and the resulting hydrolysis products were separated in denaturing polyacrylamide gels and analysed by autoradiography. The appearance of distinct cleavage products (arrowheads) indicate a reduction in the pool complexity and thus an enrichment of a few selected sequences. The very similar digestion pattern in both selection experiments indicates the presence of similar RNA sequences.

almost 70% of the sequences were identical in both selections. The remaining family I sequences are unique to each selection and were either found in several identical isolates or as single sequences. Family II (Fig. 2) contains only five members; all are characterised by four repeats of a GGNN sequence motif. Aptamer 4-30 of that group was found in three identical clones, while all other clones were unique. Six sequences could not be assigned to either of the two families and were grouped into a separate set (family III) (Fig. 2).

**Binding efficiencies of the RNA aptamers**

Relative binding efficiencies for all aptamers were determined using both parasite variant strains as targets (Fig. 2). Each aptamer RNA shows an increased relative binding capacity compared to the starting randomised RNA pool. The values range between 2 and 40%, which corresponds to a 12- to 280-fold increase in the relative binding efficiency. Thus, weak as well as strong binders were selected in the two experiments. A more thorough inspection shows that the family I aptamers can be divided into a high affinity group which binds 150- to 280-fold stronger to either MITat 1.2 or MITat 1.4 cells and a group of RNAs with only moderate affinity (14- to 50-fold stronger binding). The family III RNAs also contain weak and strong binders, while the aptamers of family II show intermediate binding. The relative binding efficiencies range between a 50- to 90-fold increase to MITat 1.2 cells and a 50- to 180-fold stronger binding to MITat 1.4 cells. None of the RNAs is capable of discriminating between MITat 1.2 and MITat 1.4 cells, regardless of which parasite variant was used in the selection protocol. This indicates either that the RNAs interact with the structurally constant domains of the two different VSGs or that identical surface targets were selected in both SELEX experiments. In contrast, all RNAs are incapable of recognising the surface of insect stage trypanosomes (binding efficiencies $\leq 0.1\%$), indicating a bloodstream stage specificity of the aptamer/trypanosome interaction.

**RNA structure determination**

To gain insight into the folding details of the three aptamer families, we computed energy-minimised secondary structures for all RNA sequences. In addition, selected aptamers were experimentally probed with single- and double-strand-specific RNases and analysed by CD spectroscopy and UV melting. The combination of all data resulted in a secondary structure model for the family I RNAs consisting of a hairpin loop as shown in Figure 3A (left). On one side, the main helical element of the hairpin is interrupted by the invariant CGU triplet, which forms a bulge that protrudes from the helical geometry. The conserved NGC sequence is located within the apical loop sequence of the hairpin. Additionally, a pseudoknotted structure can be formed involving a sequence stretch of the 3′ primer binding sequence (PBS) (Fig. 3A, right). In the case of aptamer 2-16, the
Figure 2. DNA sequences and relative binding efficiencies of all aptamers from both selection experiments. The name of each aptamer is followed by its clone abundance (in parentheses) and the primary sequence of the randomised region. Base complementary sequences are underlined and co-varying nucleotide positions are indicated as capital letters in bold. Relative binding efficiencies were measured for MITat 1.2 and MITat 1.4 cells and were normalised to the binding efficiency of the starting RNA pool, which was set to a value of 1. Standard deviations were estimated from three independent binding experiments. Members of family I (92 out of 106 clones) contain an invariant CGT and a conserved NGC triplet (shaded boxes). Family II members (seven out of 106 clones) contain four repeats of a GGGN box (shaded in grey) located at the 3′-end of the randomised region. Family III members (seven out of 106 clones) show no similarity to either of the two families; three RNAs, however, contain stretches with similar sequences (shaded boxes). Capital letters indicate nucleotides from the 5′ and 3′ PBS involved in base pairing.

- The sequence includes five clones with one or two base changes.
- The sequence includes one clone with a single base change.

Figure 3. Structure models for family I (A) and family II aptamers (B) derived from sequence comparison and enzymatic probing data of 5′-32P-labelled 2-16 RNA (family I) or 5′-32P-labelled 4-29 aptamer (family II). Sequences derived from the constant 5′ and 3′ primer binding sites (PBS) are represented as shaded lines. Nucleotide accessibilities for the various enzymes used for the structure probing are given in the legend below the drawing. For the family I aptamers a pseudoknotted structure can be formed (A, right) involving a sequence stretch of the 3′-end which is supported by several of the RNase V1 hydrolysis sites. The repetitive GGGN sequence motifs of the family II aptamers are likely to fold into a G-quartet structure (34,35) as shown in Figure 3B. The four GGGN elements show different characteristics for the helical elements: a strong positive ellipticity around 270 nm and negative ellipticities at 240 and 210 nm (data not shown).

The repetitive GGGN sequence motifs of the family II aptamers are likely to fold into a G-quartet structure (34,35) as shown in Figure 3B. The four GGGN elements show different characteristics for the helical elements: a strong positive ellipticity around 270 nm and negative ellipticities at 240 and 210 nm (data not shown).
sensitivities to RNase T1 with the third GGGN motif being insensitive to the nuclease. As in the case of the family I RNAs, nucleotides of the 3′ PBS were identified as contributing to the folding of the RNAs by base pairing with a sequence stretch at the 5′-end of the randomised region. The formed helix extends the structural arrangement of the G-quartet. No consensus structure could be derived for the aptamers grouped into family III, although three of the RNAs contain short stretches with identical sequence (Fig. 2). Lastly, in contrast to the aptamers of family I, no structural differences between the strong and weak binders of families II and III could be identified.

Identification and localisation of the binding target(s)

To identify the binding target(s) on the parasite surface, we performed zero distance photo-crosslinking experiments. A representative set of results is shown in Figure 4A using the strongest binding RNA, aptamer 2-16, as the ligand. Two crosslinking products with apparent molecular masses of 42 and 29 kDa were identified. The absence of any UV irradiation or digestion with proteinase K (data not shown) after the photo-crosslinking reaction completely abolished the formation of the two products, thus identifying them as RNA–protein complexes. No radioactive signal with the apparent molecular size of VSG molecules was detected (∼60 kDa), which indicated that the VSG protein is not the target for the binding of the aptamer RNAs.

The specificity of the two crosslinks was tested in competition experiments with homologous and heterologous RNA competitors (Fig. 4A). Crosslinking to the 42 kDa protein was abolished in the presence of a 100-fold molar excess of unlabelled 2-16 RNA, whereas a 100-fold molar excess of tRNA had no effect. This indicates a specific aptamer–protein interaction for the 42 kDa protein. In contrast, the 29 kDa band shows several unusual features. First, it cannot be competed away with either competitor RNA (Fig. 4A). Second, it was only found at lower parasite cell densities ≤10⁶ cells/ml (compare Fig. 4A with B). Third, no crosslinking was achieved at incubation intervals ≥20 min (data not shown). Since the SELEX cycles involved 60 min incubations of the RNA pools with the parasite cells, the 29 kDa protein cannot be the specific aptamer binding target.

As expected from the binding data, aptamer 2-16 could be crosslinked to the same 42 kDa target in both MITat 1.2 and MITat 1.4 cells (Fig. 4B). No crosslinking was achieved with insect stage parasites, again demonstrating the specificity of the RNA binding reaction for the bloodstream life cycle stage. Weak binding aptamers could be crosslinked to the same 42 kDa polypeptide, although with lower efficiency (data not shown).

To localize the 42 kDa binding target on the trypanosome surface we monitored the binding of aptamer 2-16 by in situ fluorescence microscopy. Direct staining with a fluorophore-conjugated 2-16 RNA preparation as well as indirect staining with a fluorophore-coupled antibody against biotinylated 2-16 RNA gave identical results. The bound RNA localised to a punctated area proximal to the kinetoplast (Fig. 5). This likely represents the flagellar pocket, the main endo- and exocytosis site of the parasite.
Quantitative characterisation of the binding reaction
As above, we used aptamer 2-16 for a quantitative analysis of the binding reaction. The dissociation constant for the interaction of 2-16 RNA with either MITat 1.2 or MITat 1.4 cells was determined in a Scatchard analysis. Varying concentrations of the radioactively labelled RNA (0.1–250 nM) were incubated with a constant number of parasite cells (5 × 10^6 cells/ml). The resulting binding curve (Fig. 6A) shows saturation characteristics demonstrating the presence of a defined number of binding sites which can be titrated with the aptamer RNA. From the Scatchard analysis of eight independent measurements with different target cell concentrations we derived a dissociation constant (K_d) of 60 ± 17 nM indicating a high affinity interaction. The target concentration at saturation corresponds to a number of 72 000 ± 18 000 binding sites/parasite cell.

The association rate of aptamer 2-16 to the surface of the parasites was determined under pseudo-first-order conditions, equivalent to a large excess of cell surface target (25 nM) over aptamer RNA (0.75 nM). Figure 6B shows a typical example of a binding kinetic which reveals an initial association phase (0–20 min) with defined first-order characteristics and a pseudo-first-order association rate (k_obs) of 0.06 min⁻¹, equivalent to a half maximal association time (t_1/2) of 12 min (t_1/2 = ln 2/k_obs). After 25–30 min, however, RNA binding slows down and a slight decrease in the association was observed at later time points (50–60 min). Hence, initial pseudo-first-order association rates (k_obs) were determined at varying target concentrations (4- to 30-fold molar excess over aptamer RNA). As expected for a bimolecular association reaction, a linear dependence of k_obs on the target concentration was identified (Fig. 4D). From the slope of the curve (k_obs = k_ass/[target]) an association rate constant (k_ass) of 4.3 × 10^5 M⁻¹ s⁻¹ was derived.

Binding to other trypanosome strains
All aptamers were unable to distinguish between the two MITat strain variants that were used in the two selections. This raised the question whether or not, non-MITat trypanosomes could also be recognised. Thus, we determined the binding efficiency of the selected RNAs to two additional Trypanosoma brucei strains, the well-characterised pleomorphic strains ILTA T 1.1 (30) and AnTat 1.1 (29). Binding efficiencies of aptamers from all three selection rounds were determined and normalised to the binding of pool 1 RNA (Table 1). Although both strains showed a 5- to 10-fold higher non-specific affinity for RNA (compared to the two MITat variants), pool 12 RNA still bound 3–4 times stronger to these two strains than pool 1 RNA. This clearly indicated that RNA species were selected that also have an affinity for the ILTA T and AnTat trypanosomes. Individual members of family I and family III bound 3–9 times stronger than pool 1 RNA and the G-quartet structures of family II showed a 10- to 25-fold stronger binding. These binding strengths result in 10–30% of the RNAs bound to the parasite surface, which are comparable values to the binding efficiencies to the MITat 1.2 and MITat 1.4 parasites.

### DISCUSSION
We demonstrate the isolation of high affinity ligands from a combinatorial RNA library to the surface of live African trypanosomes. The selection was performed with bloodstream stage trypanosomes which represent the infective life cycle stage of the parasite. Three RNA ligand families were isolated. A major group containing almost 90% of all selected RNAs was characterised by an invariant CGU sequence in addition to a highly conserved NGC base triplet. The molecules were shown to adopt an irregular hairpin structure with the invariant CGU sequence in a bulged out, single-stranded conformation and the conserved NGC sequence located within the loop of the hairpin. This likely indicates sequence-specific contacts to the binding target, especially in the case of the invariant CGU (see below). Family II aptamers share the potential of folding into G-quartet structures. Since several other selection/amplification experiments have resulted in G-quartet nucleic acid ligands (36–40), it is not clear whether this reflects a concealed selection pressure towards very stable RNA structures inherent to the SELEX protocol. However, a comparison of the RNase stability of representative RNAs of all three aptamer families in comparison to the stability of the initial RNA pool showed no significant differences (data not shown). Therefore, RNA stability was not the main selection criterion in the two experiments.

### Table 1. Relative binding efficiencies of RNA aptamers to ILTAT 1.1 and AnTat 1.1 trypanosomes

<table>
<thead>
<tr>
<th>Aptamer family</th>
<th>RNA</th>
<th>AnTat 1.1</th>
<th>ILTAT 1.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-11</td>
<td>5.3 ± 2.6</td>
<td>8.3 ± 3.0</td>
<td></td>
</tr>
<tr>
<td>2-16</td>
<td>2.4 ± 0.5</td>
<td>3.5 ± 1.5</td>
<td></td>
</tr>
<tr>
<td>4-25</td>
<td>6.5 ± 2.2</td>
<td>9.7 ± 1.1</td>
<td></td>
</tr>
<tr>
<td>4-13</td>
<td>6.0 ± 0.3</td>
<td>7.0 ± 0.4</td>
<td></td>
</tr>
<tr>
<td>Family II</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-5</td>
<td>11.4 ± 3.7</td>
<td>14.4 ± 3.7</td>
<td></td>
</tr>
<tr>
<td>2-10</td>
<td>17.1 ± 4.0</td>
<td>21.2 ± 12.0</td>
<td></td>
</tr>
<tr>
<td>4-29</td>
<td>17.7 ± 3.0</td>
<td>24.7 ± 4.8</td>
<td></td>
</tr>
<tr>
<td>Family III</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-18</td>
<td>3.9 ± 1.3</td>
<td>6.6 ± 1.6</td>
<td></td>
</tr>
<tr>
<td>2-29</td>
<td>4.0 ± 0.9</td>
<td>5.1 ± 1.4</td>
<td></td>
</tr>
<tr>
<td>4-19</td>
<td>6.0 ± 2.1</td>
<td>8.2 ± 0.4</td>
<td></td>
</tr>
</tbody>
</table>

All aptamers can be subgrouped into high and low affinity binders. Differences between the two groups could only be manifested in the case of the family I RNAs. Strong binders are likely to adopt a stable pseudoknot structure which involves in part the conserved NGC base triplet of the hairpin loop sequence. However, further experiments are required to substantiate this hypothesis. In the case of the other two families, the binding strength differences must be due to other structural criteria. Among several possibilities, the length of the individual stem structures, a possible coaxial stacking of the helices and/or the spacing between conserved elements might account for this effect. Nevertheless, all tested RNAs specifically crosslink to the same 42 kDa protein which is present on the surface of both the MITat 1.2 and MITat 1.4 parasites. This is surprising for two reasons. First, given the complexity of the cell surface, one might have expected the selection of aptamers to different surface targets similar to what has been described by Morris et al. (26), using human erythrocyte ghosts. Second, none of the aptamers interacts with the most abundant polypeptide on the trypanosome surface, the VSG protein. Although we cannot exclude the
possibility that the 42 kDa protein displays an epitope dominance as described by others (41–43), we attribute both phenomena to the unique situation of the trypanosome cell surface. The VSG layer is not only acting as a physical barrier, it also shows dynamic properties which enhance its protective function. VSG-bound antibodies, for instance, have been shown to be eliminated from the trypanosome surface by a rapid endocytotic turnover of the entire VSG layer through the flagellar pocket (44,45). Thus, any VSG-bound RNA was probably removed during the selection via this recycling reaction and, consequently, stable surface binding might have been restricted to the less abundant invariable surface elements.

Since the binding of the aptamers is restricted to the flagellar pocket of the parasites it is tempting to speculate that the 42 kDa protein might be ESAG 7, a subunit of the transferrin receptor. ESAG 7 has a molecular mass of 42 kDa, is located within the flagellar pocket and is exclusively expressed during the bloodstream stage of the parasite (46,47). In contrast, we estimated a copy number of about 70,000 molecules/trypanosome. In summary, the selection of high affinity nucleic acid ligands presented in this study extends the range of applications of the SELEX methodology to living cells as binding targets. The identified RNA aptamers show high affinity binding and specificity for a single protein and as such they have the potential to be used as diagnostic as well as therapeutic tools. Their ability to bind to an invariant element on the trypanosome surface opens up the possibility of side-stepping the antigenic variation of the VSG coat. Binding to the flagellar pocket, the main site for endocytotic and exocytotic events, might further allow for a specific targeting of toxic components to intracellular compartments of the parasite.

ACKNOWLEDGEMENTS

We thank U. F. Müller for discussion and input in the quantification of the binding data. E. Weyer-Stingl is thanked for her help recording the CD spectra, M. Engstler and M. Boshart for oligodeoxynucleotide synthesis. The work was supported by a grant from the German Research Foundation (DFG) to H.U.G. Paul are thanked for critically reviewing the manuscript. We thank U. F. Müller for discussion and input in the quantification of toxic components to intracellular compartments of the parasite.

REFERENCES
