Multiple Guide RNAs for Identical Editing of Trypanosoma brucei Apocytochrome b mRNA Have an Unusual Minicircle Location and Are Developmentally Regulated*

We identified four different guide RNAs (gRNAs) that specify identical editing of Trypanosoma brucei apocytochrome b (CYb) mRNA, which indicates gRNA redundancy in T. brucei. All four gRNAs appear functional since they occur in chimeras, some of which contain an interesting gRNA 3’ “extension.” The gRNAs are encoded in different minicircles, rather than maxicircles as in other species. However, these gRNA genes are not between 18-base pair repeats as are the other minicircle gRNA genes in T. brucei. The three minicircles clonally contain the same gRNA genes, one of which is substantially diverged, all in the same order, indicating that they are related. CYb gRNA is less abundant in procyclic than bloodstream forms. Procyclic forms contain abundant edited CYb mRNA unlike bloodstream forms thus suggesting that CYb mRNA editing may be regulated at the level of gRNA utilization.

Most mitochondrion mRNA of kinetoplastids are edited by the insertion and deletion of uridines. Editing creates initiation and termination codons and eliminates frameshifts (reviewed in Refs. 1–3). Over 45% of the nucleotides in the mRNAs of ATPase subunit 6 (A6),1 NADH-dehydrogenase subunit 7 (ND7), and C-rich gene (CR3) are the result of editing (3). In contrast, limited editing adds 4 uridines to cytochrome oxidase II (COII) mRNA (4) and adds 34 uridines to CYb mRNA (5, 6).

Accumulation of edited mRNAs is developmentally regulated in Trypanosoma brucei, which is especially evident in the case of CYb. Edited CYb mRNA is abundant in procyclic forms, which have a functional cytochrome system. However, edited CYb mRNA is nearly undetectable in slender blood form cells (7) and is present in low abundance in stumpy blood form (8), which contain oxidoreduction activities absent in slender blood form but lack cytochromes (9, 10). The abundance of edited COII mRNA parallels that of edited CYb mRNA (8). In contrast, ND7 mRNA that is edited in its 3’ domain is primarily present in blood form cells (11).

Guide RNAs (gRNAs) are small RNAs that apparently specify the edited sequences of mitochondrion mRNAs. The gRNAs share common features: they are slightly less than 60 nucleotides in length, many have an RYAYA sequence at their 5’ end, they complement edited mRNA by a combination of Watson-Crick and G-U base pairing, and they have 10–15 post-transcriptionally added uridines at their 3’ end (12–15). Some gRNA genes are in the maxicircle in Leishmania and Crithidia (12, 16) but most have a minicircle location. Many gRNA genes have been identified in T. brucei, and all of these are encoded in minicircles (14, 17–19). Interestingly, all the T. brucei minicircle gRNA genes identified to date are on the same DNA strand and are located in “cassettes” flanked by a pair of inverted 18-bp repeat sequences (14, 17, 20). The T. brucei maxicircle sequence potentially encodes maxicircle unidentified reading frame 2 (MURF2) and COII gRNA genes similar to those found in Leishmania and Crithidia (16). However, unlike Leishmania and Crithidia, the T. brucei maxicircle sequence can not encode CYb gRNAs.

We report here the identification of at least three CYb gRNA genes in T. brucei. These gRNAs are encoded in minicircles, but they are not encoded between 18-bp repeats as are all other CYb gRNA genes found in T. brucei. These gRNAs specify identical editing of the 3’ portion of the CYb mRNA-editing domain, and their presence in gRNA/mRNA chimeras suggests that all of them are functional. Two of the gRNAs in chimeras have an unusual 20 nt non-guiding 3’ extension. Each of the three minicircles that encode the CYb gRNA genes also encode A6 and CR3 gRNAs in identical locations indicating that the minicircles are related. The gRNA genes diverge among the minicircles, so substantially in one case that its guiding capacity may be reduced. The CYb gRNAs are present in both life cycle stages but are lower in abundance in procyclic form where edited CYb mRNA accumulates. This may suggest preferential CYb gRNA utilization in this life cycle stage.

EXPERIMENTAL PROCEDURES

Trypanosome Production and Preparation of Nucleic Acids—Procyclic form, blood form, and dyskinetoplastic mutant T. brucei IsTaR 1 (clonally derived from EATRO 164) were grown as previously described (21, 22). RNA and DNA were purified as previously described (23–26). Two µg of KDNA was decatenated with 40 units of topoisomerase II at 27°C overnight prior to use for PCR (27).

DNA Oligonucleotides and DNA Sequencing—Underlined nucleotides indicate added restriction enzyme sites or promoter sequences.

Pcr

Guide RNAs (gRNAs) are small RNAs that apparently specify the edited sequences of mitochondrion mRNAs. The gRNAs share common features: they are slightly less than 60 nucleotides in length, many have an RYAYA sequence at their 5’ end, they complement edited mRNA by a combination of Watson-Crick and G-U base pairing, and they have 10–15 post-transcriptionally added uridines at their 3’ end (12–15). Some gRNA genes are in the maxicircle in Leishmania and Crithidia (12, 16) but most have a minicircle location. Many gRNA genes have been identified in T. brucei, and all of these are encoded in minicircles (14, 17–19). Interestingly, all the T. brucei minicircle gRNA genes identified to date are on the same DNA strand and are located in “cassettes” flanked by a pair of inverted 18-bp repeat sequences (14, 17, 20). The T. brucei maxicircle sequence potentially encodes maxicircle unidentified reading frame 2 (MURF2) and COII gRNA genes similar to those found in Leishmania and Crithidia (16). However, unlike Leishmania and Crithidia, the T. brucei maxicircle sequence can not encode CYb gRNAs.

We report here the identification of at least three CYb gRNA genes in T. brucei. These gRNAs are encoded in minicircles, but they are not encoded between 18-bp repeats as are all other CYb gRNA genes found in T. brucei. These gRNAs specify identical editing of the 3’ portion of the CYb mRNA-editing domain, and their presence in gRNA/mRNA chimeras suggests that all of them are functional. Two of the gRNAs in chimeras have an unusual 20 nt non-guiding 3’ extension. Each of the three minicircles that encode the CYb gRNA genes also encode A6 and CR3 gRNAs in identical locations indicating that the minicircles are related. The gRNA genes diverge among the minicircles, so substantially in one case that its guiding capacity may be reduced. The CYb gRNAs are present in both life cycle stages but are lower in abundance in procyclic form where edited CYb mRNA accumulates. This may suggest preferential CYb gRNA utilization in this life cycle stage.

EXPERIMENTAL PROCEDURES

Trypanosome Production and Preparation of Nucleic Acids—Procyclic form, blood form, and dyskinetoplastic mutant T. brucei IsTaR 1 (clonally derived from EATRO 164) were grown as previously described (21, 22). RNA and DNA were purified as previously described (23–26). Two µg of KDNA was decatenated with 40 units of topoisomerase II at 27°C overnight prior to use for PCR (27).

DNA Oligonucleotides and DNA Sequencing—Underlined nucleotides indicate added restriction enzyme sites or promoter sequences.

Pcr

Guide RNAs (gRNAs) are small RNAs that apparently specify the edited sequences of mitochondrion mRNAs. The gRNAs share common features: they are slightly less than 60 nucleotides in length, many have an RYAYA sequence at their 5’ end, they complement edited mRNA by a combination of Watson-Crick and G-U base pairing, and they have 10–15 post-transcriptionally added uridines at their 3’ end (12–15). Some gRNA genes are in the maxicircle in Leishmania and Crithidia (12, 16) but most have a minicircle location. Many gRNA genes have been identified in T. brucei, and all of these are encoded in minicircles (14, 17–19). Interestingly, all the T. brucei minicircle gRNA genes identified to date are on the same DNA strand and are located in “cassettes” flanked by a pair of inverted 18-bp repeat sequences (14, 17, 20). The T. brucei maxicircle sequence potentially encodes maxicircle unidentified reading frame 2 (MURF2) and COII gRNA genes similar to those found in Leishmania and Crithidia (16). However, unlike Leishmania and Crithidia, the T. brucei maxicircle sequence can not encode CYb gRNAs.

We report here the identification of at least three CYb gRNA genes in T. brucei. These gRNAs are encoded in minicircles, but they are not encoded between 18-bp repeats as are all other CYb gRNA genes found in T. brucei. These gRNAs specify identical editing of the 3’ portion of the CYb mRNA-editing domain, and their presence in gRNA/mRNA chimeras suggests that all of them are functional. Two of the gRNAs in chimeras have an unusual 20 nt non-guiding 3’ extension. Each of the three minicircles that encode the CYb gRNA genes also encode A6 and CR3 gRNAs in identical locations indicating that the minicircles are related. The gRNA genes diverge among the minicircles, so substantially in one case that its guiding capacity may be reduced. The CYb gRNAs are present in both life cycle stages but are lower in abundance in procyclic form where edited CYb mRNA accumulates. This may suggest preferential CYb gRNA utilization in this life cycle stage.
Fig. 1. CYb gRNA gene complementarity with edited CYb mRNA. The three gRNA genes are aligned below a portion of the edited CYb mRNA sequence with Watson-Crick base pairs indicated by vertical lines and G-U base pairs by colons. Sequence identity between gCyb(560A) and B is indicated by dots. Uraciles added to Cyb mRNA by editing are shown in lowercase. The single nucleotide difference between the gCyb(560A) and C sequences, determined from a single chimera clone shown in Fig. 4, is indicated by the G under gCyb(560B).

Partial minicircles containing Cyb guide RNA sequences were amplified by PCR using primer CYb-CS2 and 5'-CybN as primers. The latter primer matches the 5' end of Cyb mRNA. Denaturation was for 1 min at 94°C, annealing for 30 s, and extension for 2.5 min at 72°C. Two cycles were performed at an annealing temperature of 40°C, followed by 25 cycles at 48°C. A library representing 4 x 106 clones was screened with oligonucleotides 3'-gCyb(558) and 3'-gCyb(560A).

Chimeric and Partially Edited Cyb RNAs—Cyb gRNA/cDNA chimeras were amplified from cDNA synthesized from 2.5 μg of poly(25) form mitochondrial RNA or 5 μg of blood total RNA using Cyb-CS4 (complementary to Cyb mRNA sequence downstream of the oligonucleotide 5'-inserted U). The gRNA sequence in the anchor duplex can be predicted since it is almost invariably composed of Watson-Crick base pairs. Conserved region primers for both orientations of the minicircle were used to provide a negative control. Touchdown PCR (29), starting above the calculated Tₐ of the primer, was used since the length of the anchor duplex cannot be predicted.

RESULTS

Identification of Cyb gRNA Genes—Analysis of the T. brucei IsTaR1 maxicircle, which we have completely sequenced (34, 35), shows that it cannot encode Cyb gRNA genes, unlike the maxicircles of Leishmania tarentolae and Crithidia fasciculata (12, 16). We therefore devised a strategy to selectively amplify Cyb gRNA genes from minicircles by PCR (see "Experimental Procedures"). A 0.6-kb PCR product was obtained using a conserved minicircle sequence primer AS-MC-CS-T3 that allows amplification of a gRNA gene with the same orientation as all minicircle sequence primer SA-MC-CS-T7 for the opposite orientation (data not shown). Direct sequencing of the PCR product confirmed the presence of Cyb gRNA gene sequence but revealed sequence degeneracy at several positions (data not shown), suggesting multiple gRNA genes. Five partial minicircles cloned from the PCR product were sequenced and found to encode three different gRNAs. Five full-length minicircles isolated by screening a minicircle library with oligonucleotides 3'-gCyb(558) and 3'-gCyb(560A) specific for the gRNA genes were sequenced (see below) and contain gRNA genes 558 and 560A; the full-length 560B minicircle was not isolated although oligonucleotide 3'-gCyb(560A) can hybridize with both 560A and B minicircles. A fourth possible gRNA was detected in a chimera (see below), but its gene was not found in the cloned minicircles.

All three minicircle-encoded gRNA genes are perfectly complementary to edited Cyb mRNA over the 3' region of its editing domain (Fig. 1). The gCyb(558) gene is complementary to 42 nt of edited Cyb mRNA and predicts gRNA with a 10-bp Waton-Crick anchor duplex with unedited Cyb mRNA and the ability to specify 32 nt of edited sequence. gCyb(560A) and gCyb(560B) genes predict gRNAs that differ from each other by a single nucleotide and which are both complementary to 39 nt of edited Cyb mRNA with an 8-bp anchor duplex and 31 nt of edited sequence. The gCyb(558) gene has 75% identity to gCyb(560A) and gCyb(560B) over the region of gRNA homology to mRNA and the gCyb(560A) and B minicircles have five differences over 485 base pairs, including the mismatch in the Cyb gRNAs.

Cyb Guide RNAs—The minicircle-encoded Cyb gRNAs were detected in cellular RNA by Northern blot analysis using DNA oligonucleotide probes 3'-gCyb(558), 3'-gCyb(560A) and 3'-gCyb(560B). All three probes hybridized to an approximately...
FIG. 2. CYb gRNAs and edited and unedited mRNA in blood form and procyclic form. Panels 1–3, Northern blots of 20 μg of total blood form (B), procyclic form (P), or dyskinetoplastic (D) RNA hybridized with oligonucleotide probes 3'-gCYb(558), 3'-gCYb(560A), and 3'-gCYb(560B), specific for the gRNAs indicated. Panels 4 and 5, 10 μg of the same total RNA used in panels 1–3, hybridized with oligonucleotides CYbD and CYbR that are specific for unedited and edited CYb mRNA, respectively. Sizes of RNA molecular weight markers are indicated in nts. RNA loading and completeness of transfer were controlled by examination of ethidium bromide stained RNA on the filters (not shown).

Developmental Regulation—The Northern blot in Fig. 2 shows that edited CYb mRNA is primarily present in procyclic form RNA (panel 5). Unedited CYb mRNA is also more abundant in the procyclic form RNA (panel 4). All three oligonucleotide probes not only show that gRNA is present in the same RNA preparations from both life cycle stages but that the gRNA is more abundant in blood form than in procyclic form RNA. Densitometric analysis indicates that blood form RNA contains about two to four times more CYb gRNAs than procyclic form RNA (Table 1). In contrast, there is about five times more edited mRNA in procyclic form than blood form RNA. The procyclic form RNA also contains about twice as much unedited mRNA than blood form. Thus, there is a greater relative proportion of the gRNAs to edited and unedited mRNAs in blood form compared with procyclic form RNA.

CYb gRNA/mRNA Chimeras—CYb gRNA/mRNA chimeras, thought to be intermediates in the editing process, were cloned and sequenced to assess the participation of the different gRNAs in RNA editing. PCR products were produced from procyclic form and blood form RNA with primers specific for gCYb(558) or gCYb(560A, B, and C) and a primer specific for CYb mRNA. CYb gRNA/mRNA chimeras for four gRNAs were obtained (Fig. 4). Of 19 clones sequenced, eight contained gCYb(558), eight contained gCYb(560A), two contained gCYb(560B), and one contained gCYb(560C), which differs by a single base from gCYb(560A) and by two bases from gCYb(560B). The gRNAs in the chimeras are linked by their non-encoded U-tail to various sites that are normally edited in CYb mRNA. The size of the gRNA portion and the length of their U-tails varies among the different chimeras. These are characteristics that have been reported previously for chimeras (15, 36). In general, the gCYb(558) chimeras have more Us linking the gRNA and mRNA than do the gCYb(560A) chimeras. In addition, the gRNA in six of the eight gCYb(560A) chimeras are linked to the first editing site (562), while in six of the eight gCYb(558) chimeras, the gRNA is linked to a further 5' editing site. The mRNA sequence just downstream of the site of gRNA linkage is incompletely edited in clone 34 (AGtttG rather than AtttGtG), similar to incompletely edited sites in previous reports (15, 36). Surprisingly, two gCYb(558) chimeras, that are independent based on their different oligo(U) sequences, have a non-encoded G residue at the site of gRNA linkage. Furthermore, the encoded 3' portion of the gRNA in two independent chimeras, 27 and 28, extends 20 nt beyond the end of gRNA/mRNA complementarity. These two chimeras predict 70- and 79-nt gRNAs, including U-tails, which is larger than the bulk of the CYb gRNAs detected in the Northern blots, suggesting that they are present at low levels in the gRNA population in steady-state RNA.

Partially Edited CYb cDNA Clones—We sequenced cDNA clones corresponding to partially edited CYb mRNA to examine molecules in the process of editing (Fig. 5). They were prepared using primers for sequences which flank the editing domain to avoid selection of any subset of molecules. They are compared...
Fig. 4. CYb gRNA/mRNA chimeras. The chimeras are shown with the gRNA (left) and mRNA (right) sequences aligned. The 5' PCR primer sequences are indicated by the arrows and 82 nt of 3'-CYb mRNA sequence is not shown for clarity. Non-encoded Us that link the gRNA and mRNA and that correspond to gRNA U-tails are shaded. Gs not encoded by mRNA or gRNA and an incompletely edited sequence are underlined in clones 8, 12, and 34, respectively. The gRNA sequence complementary to edited CYb mRNA is double underlined. Lowercase Us indicate uridines added by editing in mRNA and lowercase Ts indicate such uridines in cDNAs. Potential editing sites 562 and 571 are indicated in the mRNA. Single base differences between gCYb(560A) and (B or C) are indicated by the arrowheads under the sequence. Chimeras 21C and 28C were cloned from blood form RNA.

with partially edited CYb cDNAs from T. brucei TREU 667 which were prepared using a 5' primer for unedited sequence within the editing domain (37). The two sets of cDNAs have similar characteristics, and clones c162 and c95 are identical to 1-22 and 2-16, respectively. However, the TREU 667 cDNAs appear to be biased toward molecules that are incompletely edited in the 3' region of the CYb-editing domain. All 43 partially edited clones are edited in a 3' to 5' direction; no cases of clones edited in the 5' but not the 3' region of the CYb-editing domain were observed. Two cDNAs (c122 and c160) go directly from edited to unedited sequence. However, most cDNAs have an incompletely edited "junction sequence" between the edited and unedited sequence (38); sites within the junction that require further editing to match the mature mRNA are interspersed among those that do not.

Twelve IsTAR1 and one TREU 667 cDNAs are fully edited through the region specified by the identified CYb gRNAs. The 3' limit of the junction of six of these clones is immediately upstream of the region specified by the CYb gRNAs. This suggests that the editing by the first gRNA is completed and editing by a subsequent gRNA has begun. Another seven cDNAs (a3-c121) may have been at a similar stage since their 3' junction limits are one or two editing sites within the region specified by the CYb gRNAs and four of these (c181 and a8) differ from completely edited RNA by only a single T. gRNAs for the 5' region may form an anchor duplex within the region specified by the identified gRNA and/or there may be multiple gRNAs for the 5' portion of the domain. Seventeen of the remaining cDNAs are incompletely edited entirely within the region specified by the identified CYb gRNAs, suggesting that they were being edited by these or related gRNAs when isolated. Six other cDNAs have junctions that span both the region specified by the identified CYb gRNAs and the region specified by subsequent guide(s).

Relationships Among CYb Minicircles—Multiple clones of the same minicircle show low levels of base substitutions. The full-length minicircle cline MCP-5 that encodes gCYb(558) (Fig. 6A) has a sequence that is identical to that of partial minicircle clones TA-3 and TA-25 but has a total of six bp differences from full-length minicircles MCP-2 and -7 (three each, not shown). Similarly, full-length minicircle MCP-23 that encodes gCYb(660A) is identical to partial minicircle clones TA-7 and -35 but has a single bp difference from full-length minicircle MCP16 and 3 bp differences from partial minicircle clone TA-27 (not shown). This frequency of base substitution is low but higher than expected from PCR and cloning (39), suggesting sequence microheterogeneity among minicircles that encode the same gRNAs. Only a single partial minicircle clone (TA-4) was obtained for gCYb(660B) suggesting that minicircles encoding this gRNA are low in abundance in kDNA. No minicircle clones were obtained for gCYb(660C), which suggests that its minicircle may be in low abundance or that the single nucleotide difference is a PCR artifact.

Conservation of sequence and gene structure indicates that the minicircles encoding three of the different CYb gRNAs are related. Those encoding gCYb(560A) and gCYb(560B) are closely related since they differ by only five base pairs in the region cloned for both. The differences include three transitions, a base insertion, and a base deletion, indicated by symbols in the last four lines of MCP23 sequence (Fig. 6A). One transition is in the CYb gRNA gene (*), in the MCP23 sequence and a base deletion is in the CR3 gRNA gene (&).
minicircles encoding gCYb(558) and gCYb(560A) are more divergent, with 79% nucleotide sequence identity (Fig. 6A). The gRNA gene coding sequences are more conserved between these minicircles than are the other sequences except for the approximately 130-bp minicircle conserved region and part of the "adjacent region" (last line of Fig. 6A), which lies between the conserved region and cassette 3 (40). The conserved region plus the 5' adjoining 45 bp of the adjacent region have 90% identity while the remainder of the adjacent region has 50% identity. The 18-bp repeats of the gRNA coding cassettes are conserved and the ~100-bp sequence cassettes within the cassettes have 66-72% identity while sequences between the cassettes have 44-58% identity. The gRNA genes are also more conserved than the flanking region within their cassettes that is not complementary to the mRNA.

The gRNA genes in these minicircles are located in the same positions, as diagrammed in Fig 6B, but the CYb gRNAs are not encoded in cassettes. The coding sequences for gCYb(558) and gCYb(560A) are not flanked by 18-bp inverted repeat sequences, unlike all other gRNA coding sequences found thus far in T. brucei. Interestingly, the 3'-gRNA sequence that is present in gRNA/mRNA chimeras 27 and 28 (Fig. 4) is encoded by sequence that extends into the 18-bp repeat of the adjacent gRNA gene (Fig. 6, A and B). The different CYb minicircles encode A6 gRNAs in the second cassettes and CR3 gRNAs in the third cassettes (Fig. 6, A and C). The first cassettes may encode gRNAs since the sequences at the center of the cassettes are conserved between the minicircles, but the gRNAs have not been identified. Interestingly, the A6(138B) gRNA gene is substantially diverged from the A6(138A) gene. It encodes a gRNA that can form a 7 rather than 8 bp anchor duplex and it has five mismatches with A6 mRNA over the 31-nt region (exclusive of the anchor) with which gA6(138A) can form a perfect duplex (Fig. 6C). Thus, while gA6(138B) is clearly related to gA6(138A), its divergence would prevent its recognition by homology search (38) and probably affects its function as a gRNA. The gCR3(73A, B, and C) genes are very similar although gCR3(73B and C) have a nucleotide substitution 3' to the anchor duplex region that causes a mismatch with the mRNA, and gCR3(73B) has a one- and gCR3(73C) has a two-nucleotide deletion at the (gRNA) 5' end of the anchor duplex. Since A6 and CR3 gRNAs do not specify the initial editing of these mRNAs, the mRNA sequence that forms the anchor duplex with A6 and CR3 gRNAs is created by previous editing events. As a result of G-U base pairing and few cytidines in mRNA and gRNA sequence, most transitions in gRNA genes do not affect the sequence specified by gRNAs. The gCR3(73B and C) gRNAs have eight transitions and two or three transversions relative to gCR3(73A), resulting in an internal mismatch and a smaller anchor duplex. The gA6(138B) gRNA has seven transitions and six transversions relative to gA6(138A). CYb gRNAs gCYb(558A and B) have 11 transitions and two transversions relative to gCYb(558), reducing the length of guide homology by one nt at the 3' end and the length of the anchor duplex by two nt. Both transversions in CYb gRNA and the one in CR3 gRNA do not affect the sequence they specify but five of the six transversions in A6 gRNAs create mismatches with mRNA.

DISCUSSION

We have identified three different gRNAs that specify identical editing of the 3' portion of the editing domain of CYb mRNA in T. brucei and a possible fourth based upon gRNA
Fig. 6. Cyb gRNA minicircles. A. Alignment of MCP5, MCP23, and TA4 minicircle sequences. The conserved region that is retained among T. brucei minicircles is underlined, with the universal 13-nt sequence conserved among all minicircles double underlined. PCR primer sequences are identified by the labeled arrows above the sequence and 18-bp inverted repeat sequences, and their orientations are identified by underlining with unlabeled arrows. The gRNA/mRNA complementarities and the Cyb gene are shaded and labeled. The differences between MCP23 and the TA4 partial minicircle sequences are indicated in MCP23 by the following symbols: * is G in MCP23/A in TA4; & is deletion; $ is Cm; # is NG; @ is deletion. Base identities are dots and base deletions are dashes. B, organization of Cyb gRNA minicircles. The conserved region is shaded with the universal 13-nt sequence in black. Inverted repeats are indicated by solid arrows, gRNA genes by stippled arrows, and the cassettes are boxed. The thin line in TA4 represents uncloned sequence. The minicircles encode gA6, gCyb, and gCR3 gRNAs as indicated. C, A6 and CR3 gRNA
sequence contained in a chimera. These gRNAs all specify the same edited mRNA sequence despite substantial sequence divergence, and all four participate in editing based on their occurrence in chimeric gRNA/mRNA molecules. Interestingly, in some chimeras, the gRNAs contain 20 nt that extend beyond the region of gRNA/mRNA complementarity. The CYb gRNAs are encoded in minicircles on the same strand as all other gRNAs identified in *T. brucei*. However, the CYb gRNAs are not encoded in the cassettes defined by flanking inverted 18-bp repeats. The characteristics of the gRNAs and partially edited CYb mRNAs suggest that these gRNAs direct editing of the first seven sites of the CYb editing domain. A second set of gRNAs probably edits the remaining six further 5' sites. This pattern differs substantially from *C. fasciculata* and *L. tarentolae* where two unique gRNAs are encoded in the maxicircle which can edit four or five 3' sites and 10 5' sites of the CYb-editing domain, respectively (12, 16). Our identification of multiple, minicircle-encoded CYb gRNAs indicates that *T. brucei* has a gRNA redundancy and diversity not found in other species.

The minicircles encoding the three CYb gRNAs appear to have diverged from a common molecule based on substantial conservation of sequence homology and gRNA gene order. Additional minicircle sequence diversity is also suggested by the sequence microheterogeneity among minicircle clones. Except for the minicircle conserved region, the gRNA coding sequences are more conserved than intergenic sequences (72–81% versus 44–58%). The 14 nt differences over 41 positions between the CYb gRNA genes do not alter gRNA complementarity to edited CYb mRNA sequence, as a consequence of G-U base pairing. This is also true for CR3 gRNAs, except for a single mismatch. However, the divergence of 12 nt over 41 positions in an A6 gRNA results in five mismatches with the edited mRNA, probably preventing it from directing editing to the mature mRNA sequence, although it could direct editing of a few sites. This A6 gRNA could not be identified by computer analysis (38) that we routinely use for this purpose. It was identified by its minicircle location and homology to an identified gRNA. Divergence of this type may account for some of the 33% of gRNA genes we have been unable to identify in minicircle sequences.

Thus, selective pressure appears to conserve gRNA genes that specify mature mRNA sequence, as reflected in the bias toward transitions, which are less likely to alter the editing, rather than transversions (85% of mutations in gCYb and gCR3 are transitions, versus 55% for gA6 and about 25% for intergenic regions). The redundancy resulting from several different gRNAs for the same or overlapping sequences and the potential tolerance of gRNA/mRNA mismatch may account for the lack of gRNA processing. Occurrence of a non-encoded G at an identical position in two independent chimeras seems unlikely to be an artifact of PCR and cloning or even addition in *vivo* by terminal uridylyltransferase. Non-encoded G residues have been previously found in A6 chimeras (15). It is possible that they might represent a step in editing considering that GTP is involved in group I intron splicing (41).

Edited CYb mRNA is abundant in procyclic forms but not (slender) blood forms (7). However, free CYb gRNAs are present in lower abundance in procyclic form than in blood form, and the proportion of CYb gRNA to unedited CYb mRNA is lower still. This indicates that the presence of edited CYb mRNA is not regulated by the presence or absence of CYb gRNAs during development. These levels suggest the possibility that accumulation of edited mRNA may be regulated at the level of gRNA utilization, perhaps at the level of the editing complex. Two ND8 gRNAs are also less abundant when the mRNA is preferentially edited, similar to the CYb gRNAs (19). The lack of apparent gRNA abundance differences of 3'-ND7 gRNAs previously reported (17) may be because these gRNAs do not edit the most 3', and hence initial, editing sites as do the CYb gRNAs. In addition, the edited ND7 mRNAs are preferentially present in blood forms which can contain partially differentiated stumpy and intermediate stage cells that may exhibit some characteristics of procyclic forms (8), and there is a smaller difference in abundance of these edited mRNAs between life cycle stages compared with CYb mRNA. The lower abundance of CYb gRNA in procyclic form could also indicate that gRNA is consumed during the editing process. The length of gRNAs seen in chimeras is variable (independent of the U-tail length), with truncations of up to half of the guiding region of the gRNA (15, 36). This variation of gRNA size in CYb and other gRNA/mRNA chimeras may also be a manifestation of gRNA consumption. The finding of 20 nt of 3'-gRNA sequence that does not complement edited mRNA is unique to CYb mRNA to date.

While most partially edited CYb cDNAs were probably being edited in the 3' region of the editing domain under the direction of the gRNAs described here or in the 5' region with a second set of gRNAs when isolated, a few cDNAs are incompletely edited over a sequence spanning these two regions. This may reflect editing by a gRNA outside its homology to mature mRNA, as previously proposed (38). Alternatively, there may be a set of gRNAs that specify the editing of this overlapping region, since several gRNAs for substantially overlapping regions have been found recently (17, 42). Another possibility is that these cDNAs reflect inappropriate editing by gRNAs that would specify complete editing of another site (43).

**Multiple CYb gRNAs**

---

2 R. A. Corell, G. R. Riley, P. J. Myler, and K. Stuart, unpublished results.
The *T. brucei* maxicircle has potential gRNA genes for COII and MURF2 based on sequence homology (16), but their occurrence in cellular RNA or chimeras has not been confirmed. Some gRNA genes including those for CYb are in the maxicircle in other kinetoplastids. The occurrence of CYb gRNAs in minicircles but not maxicircles in *T. brucei* raises the possibility that no gRNAs are encoded on the maxicircle in this species. In addition, the CYb gRNA genes have an unusual minicircle location. They are not in a cassette defined by 18-bp inverted repeats, unlike all other minicircle gRNAs of *T. brucei*. Leishmania or Crithidia minicircles do not contain the 18-bp repeats, and thus they are not required for the production of gRNA in these related protozoa. The significance of 18-bp repeats, and the ability to be capped may result from RNA ligase (45) action on the 5′-RYAYA sequence and a specific location of 18-bp repeats. The occurrence in cellular RNA and MURFs based on sequence homology (16), but their conservation suggests a selective pressure for an important function.

**Trypanosoma** and **Leishmania** gRNAs can be capped using guanylyltransferase, implying they are primary transcripts (13, 44). Most *T. brucei* gRNA genes have a 5′-RYAYA sequence at about 32 bp from the 18-bp inverted repeat (14, 42, 44), which has been suggested as the region of transcription initiation (17, 44). However, no gRNA promoter has been identified, and the ability to be capped may result from RNA ligase (45) and 5′-phosphate-polyribonucleotide kinase (46) action on the gRNAs, which could add monophosphates and di-and triphosphates, respectively, to polyribonucleotide 5′ ends. Nevertheless, the location of the capped gRNA genes indicates that the presence of a 5′-RYAYA sequence and a specific location of 18-bp repeats are not essential for gRNA expression.

**Acknowledgments**—We thank Tony Morales, Jeanette Bertles, and Ken Wilson for excellent technical assistance, Drs. Peter J. Myler, Hsueh Shu, Laurie Read, Uli Göringer, Donna Koelsowsky, and Randy Howard for thoughtful discussions and reading the manuscript, and Susan Schultz for help in manuscript preparation.

**REFERENCES**

1. Feagin, J. E. (1991) *J. Biol. Chem.* 266, 19373–19376
2. Solner-Webb, B. (1992) *Nature* 356, 743–744
3. Stuart, K. (1991) *Annu. Rev. Microbiol.* 45, 327–344
4. Benne, R., van den Burg, J., Brakenhoff, J. P., Sloof, P., Van Boon, J. H., and Trapman, J. C. M. (1986) *Cell* 48, 819–826
5. Feagin, J. E., Shaw, J. M., Simpson, L., and Stuart, K. (1987) *Molecular Biology of Mitochondria and Chloroplasts. Meeting at Cold Springs Harbor Laboratory* 113 (Abstract).
6. Feagin, J. E., Shaw, J. M., Simpson, L., and Stuart, K. (1988) *Proc. Natl. Acad. Sci. U. S. A.* 85, 539–543
7. Feagin, J. E., Jasmer, D. P., and Stuart, K. (1987) *Cell* 49, 337–345
8. Feagin, J. E. and Stuart, K. (1988) *Mol. Cell. Biol.* 8, 1255–1265
9. Bienen, E. J., Saric, M., Pollariz, G., Grady, R. W., and Clarkson, A. B., Jr. (1991) *Mol. Biochem. Parasitol.* 45, 165–192
10. Clarkson, A. B., Jr., Bienen, E. J., Pollariz, G. K., and Grady, R. W. (1989) *J. Biol. Chem.* 264, 17770–17776
11. Koelsowsky, D. J., Bhat, G. J., Perrollaz, A. L., Feagin, J. E., and Stuart, K. (1986) *Cell* 62, 301–307
12. Blum, B., Bakahara, N., and Simpson, L. (1990) *Cell* 60, 189–198
13. Blum, B. and Simpson, L. (1990) *Cell* 60, 391–397
14. Pollard, V. W., Rohrer, S. P., Michelotti, E. F., Hancock, K., and Hajduk, S. L. (1986) *Cell* 63, 725–730
15. Read, L. K., Corell, R. A., and Stuart, K. (1992) *Nucl. Acids Res.* 20, 2341–2347
16. van der Spek, H., Arts, G. J., Zwaal, R. R., van den Burg, J., Sloof, P., and Benne, R. (1991) *EMBO J.* 10, 1217–1224
17. Koelsowsky, D. J., Riley, G. R., Feagin, J. E., and Stuart, K. (1992) *Mol. Biol. Cell.* 12, 2043–2049
18. Read, L. K., Myler, P. J., and Stuart, K. (1992) *J. Biol. Chem.* 267, 1123–1128
19. Souza, A. E., Myler, P. J., and Stuart, K. (1992) *Mol. Biol. Cell.* 12, 2100–2107
20. Jasmer, D. P. and Stuart, K. (1986) *Mol. Biochem. Parasitol.* 18, 321–331
21. Stuart, K. (1983) *J. Cell. Biol.* 93, 13–26
22. Stuart, K., Gobright, E., Jenkins, L., Milhausen, M., Thomaslough, L., and Agabian, N. (1984) *J. Parasitol.* 70, 747–754
23. Chomcynski, P. and Sacchi, N. (1987) *Anal. Biochem.* 162, 156–159
24. Rohrer, S. P., Michelotti, E. F., Torri, A. F., and Hajduk, S. L. (1997) *Cell* 60, 625–632
25. Feagin, J. E., Abraham, J. M., and Stuart, K. (1988) *Cell* 53, 413–422
26. Stewart, R. J., Lodes, M. J., Myler, P. J., and Stuart, K. D. (1991) *Nucl. Acids Res.* 19, 398–4008
27. Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., and Matick, J. S. (1991) *Nucl. Acids Res.* 19, 298
28. Corpas, F. (1988) *Nucl. Acids Res.* 16, 1081–1080
29. Bus, F. J., Lodes, M. J., Myler, P. J., and Stuart, K. D. (1991) *Nucl. Acids Res.* 19, 298
30. Triglia, T., Peterson, M. G., and Kemp, D. J. (1988) *Nucl. Acids Res.* 16, 8186–8196
31. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory. Cold Spring Harbor, New York*
32. Fourney, R. M., Miyakoshi, J., Day, R. S., and Paterson, M. C. (1988) *Focus* 10, 5–7
33. Myler, P. J., Glick, D. L., Feagin, J. E., Morales, T. H., and Stuart, K. (1993) *Nucl. Acids Res.* 21, 687–694
34. Simpson, L., Neckelmann, N., de la Cruz, V. F., Simpson, A. M., Feagin, J. E., Jasmer, D. P., and Stuart, K. (1979) *J. Biol. Chem.* 262, 1692–1696
35. Blum, B., Sturm, N. R., Simpson, A. M., and Simpson, L. (1991) *Cell* 65, 543–550
36. Decker, C. J. and Solner-Webb, B. (1990) *Cell* 61, 1001–1011
37. Koelsowsky, D. J., Bhat, G. J., Read, L. K., and Stuart, K. (1991) *Cell* 67, 537–546
38. Lundberg, K. S., Shoemaker, D. D., Adams, M. W. W., Short, J. M., Sorge, J. A., and Mathur, E. J. (1991) *Gene* (Amst.) 108, 1–6
39. Jasmer, D. P. and Stuart, K. (1986) *Mol. Biochem. Parasitol.* 18, 257–269
40. Ceci, T. R. (1990) *Annu. Rev. Biochem.* 59, 543–565
41. Corell, R. A., Feagin, J. E., Riley, G. R., Strickland, T., Guderson, J. A., Myler, P. J., and Stuart, K. (1992) *Nucl. Acids Res.* 20, 819–826
42. Sturm, N. R., Maslov, D. A., Blum, B., and Simpson, L. (1992) *Cell* 70, 476–497
43. Pollard, V. W. and Hajduk, S. L. (1991) *Mol. Cell. Biol.* 11, 1668–1675
44. Greer, C. L., Peebles, C. L., Gegenheimer, P., and Abelson, J. S. (1986) *Cell* 43, 537–546
45. Spencer, E., Loring, D., Hurwitz, J., and Monroy, G. (1978) *Proc. Natl. Acad. Sci. U. S. A.* 75, 4793–4797