Stable integrative transformation of *Trypanosoma brucei* that occurs exclusively by homologous recombination

(gene targeting/transfection/calmodulin gene/neomycin-resistance gene/homologous integration)

JOSIANE EID AND BARBARA SOLLNER-WEBB

Department of Biological Chemistry, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205-2185

Communicated by M. Daniel Lane, December 21, 1990

ABSTRACT A calmodulin-neomycin-resistance fusion gene was introduced into *Trypanosoma brucei* by electroporation, and stably transformed cell lines were obtained. In all of the transformants, the fusion gene had integrated into the host genome at the cognate locus, evidently by homologous recombination within flanking calmodulin DNA. This unusual observation distinguishes trypanosomes as the only eukaryote other than yeast known to undergo gene targeting in essentially 100% of the stable transformants. It should now be possible to systematically manipulate the trypanosome genome, directing predetermined mutations to virtually any chromosomal locus.

Trypanosoma brucei are parasitic protozoa and the causative agent of sleeping sickness, a major disease in Africa. This organism exhibits many fascinating biological phenomena including antigenic variation (1, 2), polycistronic transcription (1, 3), trans-splicing of RNA (4, 5), RNA editing (6, 7), and stage-dependent mitochondrial activation (8). To better study these and other features of T. brucei, a stable transformation system involving homologous gene targeting was sought, for this would allow directed mutation of the genome as well as long-term expression of introduced DNA. Such homologous gene targeting has proven extremely valuable in numerous species to selectively alter genes of interest in the host chromosome. In yeast, virtually 100% of transfected DNA that integrates into the host genome does so by homologous recombination (9, 10); other species show homologous gene targeting as well, but with lower efficiency (≈0.01%–1% of the integrants in mammalian cells; refs. 11-13). Although stable transformation of another trypanosomatid, Leishmania, has been reported (14, 15), the introduced DNA remained extrachromosomal, precluding the possibility of mutating the host genome.

In this paper, we present evidence for stable integrative transformation of *T. brucei* by a fusion gene construct, CNeoC. This construct, which contains the neomycinresistance (neo^r) coding region flanked by extensive 5' and 3' segments of the *T. brucei* calmodulin locus, was designed to favor homologous targeting (16) into the chromosomal calmodulin locus. After electroporation of procyclic cultures of *T. brucei* with the excised fusion gene, G418-resistant cells were obtained. DNA analysis from seven of the resultant stable cell lines (detailed in this report) and preliminary evidence from 20 additional similarly derived lines showed that in all cases the neo^r fusion gene was integrated into the calmodulin locus of the host genome, evidently by homologous recombination.

MATERIALS AND METHODS

The CNeoC Plasmid. The CNeoC plasmid (see Fig. 1) contains the translation initiation codons of both the *T. brucei*

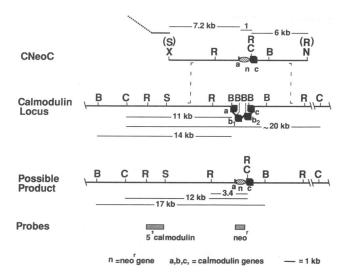
The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

calmodulin gene (17) and the neomycin phosphotransferase (neor) gene (18) in frame and 36 base pairs apart, the translation termination regions of both genes, and the calmodulin poly(A) addition region (19). It also includes extensive flanking calmodulin sequences, extending upstream to the EcoRI site 9.2 kilobases (kb) 5' to the ATG of the calmodulin a gene and downstream to the EcoRI site 6 kb beyond the ATG of the calmodulin c gene (17). In essence, the two central 0.84-kb Bgl II fragments of the calmodulin locus containing the bulk of the calmodulin a and b genes were replaced with a 0.97-kb Bcl I-Ava I fragment derived from pSV2neo (18) containing the neor coding region plus 3' EcoRI and Cla I polylinker sites. The 30 base pairs preceding the neor ATG from pSV2neo were deleted by oligonucleotide-directed mutagenesis and the EcoRI and Sal I sites indicated in Fig. 1 were replaced by Not I and Xba I sites, respectively. The cloning vector was the 2.3-kb Pvu II (converted to Not I)-EcoRI fragment of pBR322.

Transformation and Propagation of T. brucei. Logarithmicphase procyclic T. brucei cells, propagated in SM growth medium (20), were washed twice with phosphate-buffered saline and suspended in electroporation buffer (272 mM sucrose/7 mM potassium phosphate, pH 7.4/1 mM MgCl₂) containing Xba I- and Not I-cleaved CNeoC (20 μ g/ml). Cell suspensions (5 \times 10⁷ cells per ml, at room temperature) were electroporated three times at 250 V and 500 μ F, by using the Bio-Rad Gene Pulser and Bio-Rad cuvettes (≈50% killing). After 10 min at room temperature, the cells were transferred to SM medium to recover at 27°C, and 48 hr later, selection was initiated in cloning medium (21) supplemented with G418 (Geneticin, GIBCO; $50 \mu g/ml$) in which untransfected cells die and these early transfected cells survive but do not divide. After 5–7 days, surviving cells were counted and plated at <1surviving cell per well in 96-well plates in 100 μ l of cloning medium lacking drug. (Nonetheless, certain lines could derive from more than one original transformed cell.) Upon reaching near confluence (≈4 weeks; cell division was not observed until >1 week), the cell lines (obtained at a frequency of ≈ 2 per 10⁷ starting cells) were propagated in SM medium supplemented with G418 (30 μ g/ml). In control experiments, cells were transfected with the CNeoC plasmid in the closed circular form or with the 18-kb plasmid pR4 (22) containing T. brucei rRNA genes after linearization with Cla I.

Analyses of DNA and RNA from the Transfected Cells. Late-logarithmic phase *T. brucei* genomic DNA was extracted as described (23) and analyzed by the method of Southern (24). The DNA was resolved by electrophoresis on 0.6% agarose gels, transferred onto a GeneScreen*Plus* membrane, and hybridized to DNA probes radiolabeled by random priming (25). Pulsed-field gel electrophoresis was performed, as described in Schwartz *et al.* (26, 27), following optimal conditions for resolving ≈200-kb DNA (1% agarose, 12.5 V/cm, 25-sec pulsing time, and 24-hr run time). Total

Abbreviation: neor, neomycin resistance.



cellular RNA was prepared by the lithium chloride method (28), and poly(A)⁺ RNA was fractionated by two passes over oligo(dT)-Sepharose (29). RNA was electrophoresed on 1% formaldehyde/1.2% agarose gels and analyzed by Northern blotting on GeneScreen*Plus* membranes (29).

RESULTS AND DISCUSSION

To attempt homologous gene replacement in T. brucei, we constructed a plasmid (CNeoC) bearing an 18-kb segment of the T. brucei calmodulin locus (17) in which three of the four tandem calmodulin coding regions were deleted and the neo^r coding region was inserted, in frame, into the remaining calmodulin gene. Thus, transcription directed by the flanking 5' and 3' calmodulin sequences should allow production of neor mRNA. By following the strategy for homologous gene replacement in yeast (16) and for successful targeting in mammalian cells (11), the CNeoC plasmid was cleaved with Xba I and Not I (Fig. 1) to generate a fragment containing the neor sequences and at least 6 kb of calmodulin DNA on both ends. Procyclic T. brucei were transfected by electroporation and then selected for neor by using the drug G418. Resistant cells were indeed obtained in six out of six separate experiments. These cells have been cloned and propagated for 10 months. Plasmid cleavage is critical, since no G418-resistant cells were obtained after transfection with supercoiled CNeoC plasmid. A linearized plasmid lacking the neor gene also yielded no G418-resistant transformants. This latter result indicates that the G418 resistance is not due to mutations induced by linear DNA, as has been found to occur in Chlamydomonas (S. Dutcher, personal communication).

The genomic DNA from seven G418-resistant cell lines was characterized by hybridization to a neo^r probe (illustrated in Fig. 1). After *Eco*RI cleavage, the neo^r gene was detected in the DNA from the transfected cells (Fig. 2A, lanes designated

with lower case letters) and not from untransfected *T. brucei* (lane U). In all cases, this neo^r-containing fragment was 3.4 kb, the same size as the corresponding fragment derived from an *EcoRI*-cleaved CNeoC plasmid (Figs. 1 and 2A, lane P). When uncut DNA from the transfected cell lines was used, neo^r sequences migrated at the position of unresolved high molecular weight chromosomal DNA (Fig. 2B, lanes with lower case letters; compare with last lane). Furthermore,

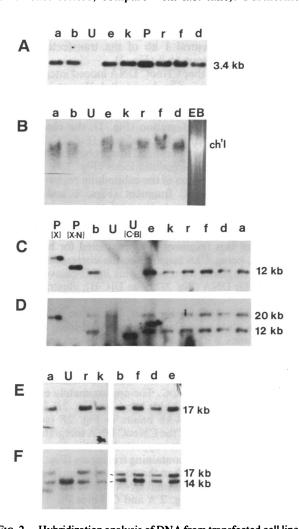


Fig. 2. Hybridization analysis of DNA from transfected cell lines. Genomic DNA prepared with minimal shearing from transfected cell lines (lanes a, b, d, e, f, k, and r) and from untransfected cells (lane U) was resolved on a 0.6% agarose gel, transferred to a GeneScreenPlus membrane, and detected by hybridization to radiolabeled probes. The neor probe was used for the analysis in A, B, C, and E. (A) Analysis of chromosomal DNA and CNeoC plasmid DNA (lane P) after cleavage with EcoRI. HindIII-cleaved \(\lambda \) DNA was the size marker. (B) Analysis of uncut DNA. On the right is an ethidium bromide stain of lane k. (C) Analysis of Cla I-digested DNA. CNeoC plasmid DNA linearized with Xba I [lane P(X); 18.5 kb] and cleaved with Xba I plus Not I [lane P(X·N); 14.2 kb] as well as untransfected cell DNA cleaved with Cla I plus Bgl II [lane U(C·B)] are also included. Sizes were determined from a semilogarithmic plot of the molecular size of the CNeoC markers and DNA standards of 19.4, 18.5, 12.2, 10.1, and 8.6 kb (BRL) vs. migration. (D) The filter in C was stripped and rehybridized with the 5' calmodulin probe, after autoradiographic confirmation of probe removal. The 4.3-kb fragment from CNeoC homologous to the 5' calmodulin probe in lane P(X·N) was below the photographed region. (In this blot the ≈20-kb region of lanes b-r did not transfer quantitatively, and lane e was 2-fold overloaded; other experiments show equal intensities of these bands.) (E) Analysis of Bgl II-cleaved DNA. Sizes were determined from the 18.5-kb and 14.2-kb CNeoC markers, as in C. (F) The filter of E was rehybridized with the 5' calmodulin probe, as described in D.

after Cla I digestion (Fig. 2C) and Bgl II digestion (Fig. 2E), all of the transformed cell lines yielded neo^r-containing fragments of the same size, 12 kb and 17 kb, respectively. Thus the recombinant neo^r gene is stably maintained as high molecular weight DNA, and its organization appears to be the same in all of these cell lines. The 12-kb and 17-kb neo^r fragment sizes (Fig. 2 C and E) are inconsistent with either head-to-head or head-to-tail concatamerization of the 14-kb CNeoC DNA, suggesting that recombination occurred with the host genome, at the same site, in all the transfected cell lines. The calmodulin locus appears the most likely target, since all but the central 1 kb of the transfected CNeoC fragment is homologous to it (Fig. 1).

To determine whether CNeoC DNA indeed integrated into the calmodulin locus, Cla I and Bgl II restriction sites flanking the CNeoC sequences were compared to those of the wild-type calmodulin locus (Fig. 2 C-F). Consistent with homologous CNeoC integration (Fig. 1), the chromosomal Cla I fragment containing the neor gene (Fig. 2C; lanes with lower case letters) also hybridized to a probe derived from the far 5' portion of the calmodulin locus (Fig. 2D). Since this probe is located upstream of the calmodulin region present in the transfected CNeoC fragment (Figs. 1 and 2D, lane P[X·N]), its hybridization to the neo^r-containing Cla I fragment indicates that these sequences were brought together in the transfected cells by recombination. Furthermore, the 12-kb size of this fragment is as predicted for homologous integration, being 1 kb larger than the corresponding 11-kb calmodulin fragment from Cla I- plus Bgl II-cleaved untransfected cellular DNA (Fig. 2D, lane U[C·B]; illustrated in Fig. 1).

The 5' calmodulin probe also detects an \approx 20-kb Cla I fragment in all the cellular DNAs derived from transfected (Fig. 2D, lanes with lower case letters) and untransfected (lane U) lines. Since trypanosomes are diploid for house-keeping loci (30), this 20-kb fragment evidently derived from the copy of the calmodulin-containing chromosome that did not recombine with CNeoC. The approximately equal intensities of the 12-kb and 20-kb bands in Fig. 2D [and of the analogous 17-kb and 14-kb bands in Fig. 2F (see below)] indicate that one copy of the CNeoC DNA integrated per cell. This conclusion is further supported by the observation that the chromosomal neo^r-containing fragments (Fig. 2 A and C) are slightly less intense than the corresponding fragment from an amount of CNeoC plasmid DNA equivalent to 1.5 copies per cellular genome (Fig. 2 A and C, lanes P).

Mapping of the transfected cellular DNA with Bgl II provided additional evidence for homologous gene targeting by the CNeoC DNA. Studies analogous to those with Cla I described above showed that the ≈ 17 -kb Bgl II fragment from the neo' cells is detected not only with the neo' probe (Fig. 2E) but also with the probe for 5' calmodulin sequences located upstream of the CNeoC region (Fig. 2F). Furthermore, this 17-kb Bgl II fragment is, as expected, 3 kb larger than the corresponding 14-kb Bgl II fragment from untransfected cells (Fig. 2F, lane U; see maps in Fig. 1). This juxtaposition of the CNeoC region in the transfected cell lines to the appropriate upstream segment of the chromosomal calmodulin locus is most consistent with the CNeoC gene having integrated into the chromosomal calmodulin locus by homologous recombination.

Since it is not practical to determine the sequence of the entire 13.2-kb region of transfected cell DNA in which the CNeoC segment could have recombined to prove that the recombination was homologous to the nucleotide, we have instead performed finer mapping to demonstrate that the integration event did not preceptibly alter the size of the junction fragment. The experiment of Fig. 2A demonstrates that no size alteration of the 3.4-kb EcoRI fragment containing the neor gene and its upstream region is detectable in the

transfected cell DNA relative to the input CNeoC DNA. Similarly, on rehybridization of this filter to a 0.24-kb (HindIII-EcoRI) probe located immediately upstream of this 3.4-kb region, a uniform 6.8-kb fragment was detected from all of the transfected cell lines as well as from the untransfected cells (data not shown). Thus, no matter where the recombination occurred within the 7.2-kb upstream region of the transfected Xba I-Not I fragment of CNeoC, it did not alter the length of the surrounding DNA by more than \approx 50 base pairs. These results, in conjunction with the observation that the pBR322 sequences present in the original CNeoC plasmid were not detected in the transfected cellular DNAs (reprobing data not shown) makes it extremely likely that the CNeoC DNA indeed integrated into the T. brucei genome by homologous recombination.

Further evidence for targeting of the CNeoC DNA into the calmodulin locus was obtained by pulsed-field gel electrophoresis (Fig. 3A). When intact transfected-cell DNA was used, the neor probe hybridized to material at the slot (Fig. 3B, lanes 2 and 4), indicating that the CNeoC fragment integrated into one of the large chromosomes; the wild-type chromosome bearing the calmodulin genes behaved identically (Fig. 3C, lane 6; compare to lanes 2 and 4). After digestion with Sfi I, the neor gene was located on an \approx 130-kb fragment (Fig. 3B, lanes 1 and 3). The 5' calmodulin probe also hybridized to an \approx 130-kb fragment in transfected (Fig. 3C, lanes 1 and 3) and untransfected (lane 5) cellular DNA. This indicates that the neor gene integrated into the same chromosome within the same Sfi I fragment that contains the endogenous calmodulin genes.

Because the CNeoC-transfected cell lines were selected by G418 resistance, one would anticipate that their neo^r gene is expressed. This was verified by analysis of cellular RNA (Fig. 4). An \approx 2-kb neo^r-containing transcript was indeed seen in poly(A)⁺ RNA from the transfected (Fig. 4B, lane 1) but not from the untransfected (lane 3) cells. Furthermore, the \approx 2-kb size of this transcript is as expected from the CNeoC fusion gene, in which the 1-kb neo^r region was inserted within

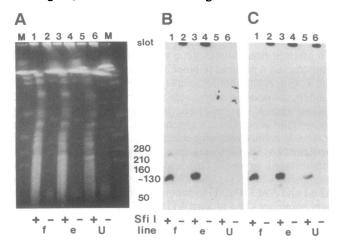


FIG. 3. Pulsed-field gel analysis of the neor gene in transfected cell DNA. Intact (-) and Sfi I-cleaved (+) chromosomal DNA was resolved on a 1% agarose gel in a pulsed electrophorectic field. Lanes: e and f, cell lines e and f, respectively; U, untransfected cells. (A) Ethidium bromide-stained gel. Sizes were determined from a semilogarithmic plot of T. brucei and yeast chromosome (BRL) sizes vs. migration. Lanes M contain molecular size markers and sizes are indicated in kb. (B) The DNA was transferred to a GeneScreenPlus membrane and hybridized with the neor probe. (C) The filter was stripped, reexposed to confirm probe removal, and rehybridized with the 5' calmodulin probe. Sfi I does not cut in the CNeoC region. The small difference in size (\approx 1 kb) in this \approx 130-kb Sfi I fragment due to CNeoC recombination would not be detectable. The faint \approx 210-kb band in B and C is due to incomplete Sfi I digestion.

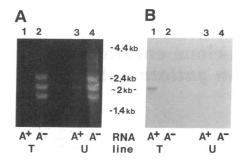


FIG. 4. Analysis of RNA. Poly(A)⁺ and poly(A)⁻ RNA (10⁸ cell-equivalents) derived from transformed (lanes T; cell line 3) and untransfected (lanes U) cells was resolved on a 1% formaldehyde/1.2% agarose gel and transferred to GeneScreenPlus. (A) Ethidium bromide staining, showing the 18S and cleaved 28S rRNAs. Standards were the 0.24- to 9.5-kb RNA ladder (BRL). (B) Hybridization with the neo^t probe.

a calmodulin gene that normally yields an ≈1-kb transcript (19). This suggests that the 5' and 3' ends of the neo^r RNA probably arose by using the normal calmodulin trans-splice and poly(A) addition signals.

We have obtained stable integrative transformation of T. brucei, exclusively by recombination into the homologous locus of the host chromosome. Seven cloned transformed cell lines are examined in this paper, and preliminary analysis (not shown) of 20 other transformed lines obtained from six experiments indicates that all 27 represent homologous integration. To date, yeast are the only other eukaryotic organisms known to undergo gene targeting with virtually 100% efficiency of homologous vs. nonhomologous recombination (9, 10, 16). Other fungi and Dictyostelium show targeting, generally at 1-30% efficiency (31, 32), and mammalian cells do also but generally at 0.01-1% efficiency (11-13). Gene targeting has proven extremely valuable in analyzing diverse biological processes, allowing the introduction of desired mutations in chosen regions of the genome. The fact that trypanosomiasis has major medical and economic impact adds to the importance of such targeting in T. brucei.

Note Added in Proof. While this manuscript was under review, three papers were published on homologous gene targeting in trypanosomatids. Cruz and Beverly (33) describe attaining homologous gene targeting in *Leishmania major* in up to nearly 100% of neof colonies. In *T. brucei*, using a tubulin/neof plasmid, Ashbroek et al. (34) demonstrate at least four independent recombination events into the tubulin locus, and by using a similar plasmid, Lee and Van der Ploeg (35) obtained a transformed cell line in which they demonstrate neof insertion by homologous recombination.

We thank C. Tschudi for providing the calmodulin gene-bearing λ phage, CMIII, and S. Hajduk for providing the procyclic *T. brucei* culture. We also thank J. Boeke, E. Mougey, S. Hajduk, P. Englund, D. Cleveland, T. Yen, D. Schwartz, S. Dutcher, R. Firtel, M.-C. Yao, and A. Fire for helpful discussions and L. Miller for typing the

manuscript. This work was supported by National Institutes of Health Grant GM34231 and the salary of J.E. was in part defrayed by the MacArthur Foundation.

- 1. Borst, P. (1986) Annu. Rev. Biochem. 55, 701-732.
- 2. Pays, E. & Steinert, M. (1988) Annu. Rev. Genet. 22, 107-126.
- Imboden, M. A., Laird, P. W., Affolter, M. & Seebeck, T. (1987) Nucleic Acids Res. 15, 7357-7368.
- Murphy, W. J., Watkins, K. P. & Agabian, N. (1986) Cell 47, 517-525.
- 5. Sutton, R. & Boothroyd, J. (1986) Cell 47, 527-535.
- 6. Simpson, L. & Shaw, J. (1989) Cell 57, 355-366.
- 7. Weiner, A. M. & Maizels, N. (1990) Cell 61, 917-920.
- 3. Vickerman, K. (1985) Br. Med. Bull. 41, 105-114.
- Hinnen, A., Hicks, J. B. & Fink, G. R. (1978) Proc. Natl. Acad. Sci. USA 75, 1929–1933.
- Orr-Weaver, T. L., Szostak, J. W. & Rothstein, R. J. (1981) *Proc. Natl. Acad. Sci. USA* 78, 6354-6358.
- 11. Thomas, K. R. & Capecchi, M. R. (1987) Cell 51, 503-512.
- Sedivy, J. & Sharp, P. (1989) Proc. Natl. Acad. Sci. USA 86, 227-231.
- Jasin, M., Elledge, G., Davis, R. & Berg, P. (1990) Genes Dev. 4, 157-166.
- Laban, A., Tobin, J. F., Curotto de la Faille, M. A. & Wirth,
 D. F. (1990) Nature (London) 343, 572-574.
- Kapler, G. M., Coburn, C. M. & Beverly, S. M. (1990) Mol. Cell. Biol. 10 1084-1094.
- 16. Rothstein, R. (1983) Methods Enzymol. 101, 202-209.
- Tschudi, C., Young, A. S., Ruben, L., Patton, C. L. & Richards, F. (1985) Proc. Natl. Acad. Sci. USA 82, 3998-4002.
- 18. Beck, E., Ludwig, G., Auerswald, E., Reiss, B. & Shaller, H. (1982) Gene 19, 327-336.
- 19. Tschudi, C. & Ullu, E. (1988) EMBO J. 7, 455-463.
- 20. Cunningham, I. (1977) J. Protozool. 24, 325-329.
- 21. Brun, R. & Shoenenberger, M. (1979) Acta Trop. 36, 289-292.
- White, T., Rudenko, G. & Borst, P. (1986) Nucleic Acids Res. 14, 9471–9489.
- Firtel, R. A., Cockburn, A., Frankel, G. & Hershfield, V. (1976) J. Mol. Biol. 102, 831-852.
- 24. Southern, E. (1975) J. Mol. Biol. 98, 503-517.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab., Cold Spring Harbor, NY) pp. 131-132.
- Schwartz, D. C., Smith, L. C., Baker, M. & Hsu, M. (1989)
 Nature (London) 342, 575-576.
- 27. Schwartz, D. C. & Cantor, C. R. (1984) Cell 37, 67-75.
- Auffray, C. & Rougeon, F. (1980) Eur. J. Biochem. 107, 303-314.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1988) Current Protocols in Molecular Biology (Wiley, New York) pp. 4.5.1-3, 4.9.1-7.
- Gibson, W. C., Osinga, K. A., Michels, P. A. M. & Borst, P. (1985) Mol. Biochem. Parasitol. 16, 231-242.
- 31. Case, M. E., Schweizer, M., Kushner, S. R. & Giles, N. H. (1979) Proc. Natl. Acad. Sci. USA 76, 5259-5263.
- 32. De Lozanne, A. & Spudich, J. (1987) Science 236, 1086-1091.
- Cruz, A. & Beverly, S. M. (1990) Nature (London) 348, 117– 173.
- 34. Ashbroek, A. L., Ouellett, M. & Borst, P. (1990) Nature (London) 348, 174-175.
- 35. Lee, M. & Van der Ploeg, L. (1990) Science 250, 1583-1587.