*Proc. Natl. Acad. Sci. USA*
Vol. 92, pp. 5749–5753, June 1995
Genetics

# Insertional mutagenesis and marker rescue in a protozoan parasite: Cloning of the uracil phosphoribosyltransferase locus from *Toxoplasma gondii*

(pyrimidine salvage/molecular parasitology/genetic systems/gene knock-outs)

ROBERT G. K. DONALD AND DAVID S. ROOS

Department of Biology, University of Pennsylvania, Philadelphia, PA 19104

*Communicated by Louis H. Miller, National Institutes of Health, Bethesda, MD, March 20, 1995*

**ABSTRACT** **Nonhomologous integration vectors have been used to demonstrate the feasibility of insertional mutagenesis in haploid tachyzoites of the protozoan parasite *Toxoplasma gondii*. Mutant clones resistant to 5-fluorouracil were identified at a frequency of $\approx 10^{-6}$ ($\approx 2 \times 10^{-5}$ of the stable transformants). Four independent mutants were isolated, all of which were shown to lack uracil phosphoribosyltransferase (UPRT) activity and harbor transgenes integrated at closely linked loci, suggesting inactivation of the UPRT-encoding gene. Genomic DNA flanking the insertion point (along with the integrated vector) was readily recovered by bacterial transformation with restriction-digested, self-ligated total genomic DNA. Screening of genomic libraries with the recovered fragment identified sequences exhibiting high homology to known UPRT-encoding genes from other species, and cDNA clones were isolated that contain a single open reading frame predicted to encode the 244-amino acid enzyme. Homologous recombination vectors were exploited to create genetic knock-outs at the UPRT locus, which are deficient in enzyme activity but can be complemented by transient transformation with wild-type sequences—formally confirming identification of the functional UPRT gene. Mapping of transgene insertion points indicates that multiple independent mutants arose from integration at distinct sites within the UPRT gene, suggesting that nonhomologous integration is sufficiently random to permit tagging of the entire parasite genome in a single transformation.**

The protozoan parasite *Toxoplasma gondii* is a ubiquitous human pathogen that has recently achieved considerable notoriety as an opportunistic infection associated with AIDS (1). Complications arising from long-term administration of antifolates (2) have rendered the development of new therapeutic strategies a matter of some urgency (3). Beyond its direct medical importance, *T. gondii* also offers a promising system for genetic exploration (4–6), and several strategies for molecular genetic manipulation have recently been devised (5–11). Using vectors encoding the fused dihydrofolate reductase–thymidylate synthase enzyme (DHFR–TS; ref. 12) of the parasite, engineered to contain point mutations predicted from pyrimethamine-resistant *Plasmodium falciparum* (malaria) (13–15), chromosomal integration of recombinant DNA has been observed in >5% of viable parasites (8, 10). Vectors harboring cDNA-derived "minigenes" appear to integrate exclusively via nonhomologous recombination (8), whereas constructs containing the entire 8-kb genomic locus can also integrate by homologous recombination, at frequencies proportional to the length of contiguous genomic sequence provided (10).

The extraordinarily high frequency of stable transformation observed in *T. gondii*, coupled with the parasite's relatively small genome size (16, 17) and the haploid nature of parasite tachyzoites (16, 18), suggests that vectors which integrate into the genome by nonhomologous recombination might be suitable for gene tagging by insertional mutagenesis. By analogy with other systems (19–21), insertional mutagenesis could be of tremendous value for genetic investigation of *Toxoplasma* (5).

An ideal target for testing the feasibility of insertional mutagenesis in *T. gondii* would be a nonessential gene whose inactivation can be specifically selected. Pioneering studies on nucleotide metabolism in *Toxoplasma* by E. R. Pfefferkorn and coworkers have identified several such loci (4, 22–25). In this report, we have targeted the uracil phosphoribosyltransferase (UPRT) gene, which encodes a parasite-specific enzyme offering both chemotherapeutic potential (26) and a negative selectable marker suitable for further genetic manipulation (5, 27). UPRT normally catalyzes the conversion of uracil to UMP during pyrimidine salvage in *T. gondii* (28, 29), but phosphoribosylation of fluorouracil to fluorouridine leads to the synthesis of F-dUMP and fatal inhibition of TMP synthesis. Selection for resistance to 5-fluorodeoxyuridine (FUDR; metabolized intracellularly to 5-fluorouracil) therefore permits identification of parasites in which the UPRT gene has been inactivated (22, 23). An active UPRT enzyme is not essential for parasite viability, as *T. gondii* is also capable of *de novo* pyrimidine synthesis (28, 29).

## MATERIALS AND METHODS

**Parasite Growth and Transformation.** *T. gondii* tachyzoites (RH strain) were maintained by serial passage in primary cultures of human foreskin fibroblasts (HFF cells) as described (5). Freshly harvested parasites were purified by filtration through a 3-$\mu$m-pore-size polycarbonate membrane (Nuclepore), concentrated by centrifugation, and resuspended in cytomix electroporation buffer (30) at a density of $\approx 2 \times 10^7$ parasites per ml. Parasites ($10^7$) were added to 20- to 100-$\mu$g plasmid DNA and subjected to a 1.5-keV (1 eV = $1.602 \times 10^{-19}$ J) electroporation shock in a 2-mm gap cuvette using a BTX model 600 electroporator (BTX, San Diego) at a resistance setting of 24 $\Omega$. After electroporation, parasites were inoculated into T flasks containing HFF cells and maintained under appropriate selection, as detailed in the text. Individual parasite clones were isolated by limiting dilution in microtiter plates containing HFF cells (5). The ability of infected cell cultures to incorporate [$^3$H]uracil into trichloroacetic acid-precipitable material was measured as described (5).

**Molecular Methods.** Construction of plasmids used for insertional mutagenesis (pDHFR-TSc3/M2M3 and pDHFR-TSc3/M2M4) has been described (8, 10). These vectors are

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.

Abbreviations: UPRT, uracil phosphoribosyltransferase; TS, thymidylate synthase; FUDR, 5-fluorodeoxyuridine; DHFR, dihydrofolate reductase; HFF, human foreskin fibroblasts.

based on Bluescript pKS− (Stratagene) and contain DHFR–TS genomic 5′ flanking sequence fused to a full-length DHFR–TS cDNA that was altered to contain the mutation $Ser^{36} \rightarrow Arg$ (designated M2) and either $Thr^{83} \rightarrow Asn$ (M3) or $Phe^{245} \rightarrow Ser$ (M4). Fifty to one-hundred micrograms of DNA was transfected either as supercoiled circular plasmid or after restriction digestion with *Hin*dIII to produce the linearized molecule diagrammed in Fig. 1. The linearized plasmid was extracted with phenol and ethanol-precipitated before transfection to remove residual restriction enzyme.

Genomic *Toxoplasma* DNA was prepared from purified tachyzoites and used in blotting experiments according to standard procedures (31). For the rescue of fragments flanking transgene insertions, DNA from transgenic parasite clones was digested to completion with appropriate restriction enzymes (as indicated in text) and diluted to ≤1 μg/ml to favor intramolecular ligation (with T4 DNA ligase; New England Biolabs). After transfection of *Escherichia coli* DH5α by electroporation, plasmids were isolated from ampicillin-resistant colonies.

Genomic and cDNA clones were isolated from bacteriophage λ libraries according to standard screening procedures (31). Genomic libraries of *T. gondii* (strain RH) DNA, partially digested with *Mbo* I or *Sau*3A and size-fractionated, were constructed in the λDASH vector (Stratagene), according to the manufacturer's instructions. cDNA libraries in λZAPII (Stratagene) were also prepared as recommended, and an additional λZAPII cDNA library was obtained from the National Institutes of Health AIDS Research and Reference Reagent Program. The complete cDNA sequence and genomic sequence spanning the UPRT locus has been deposited with GenBank data base (accession no. U10246).

For gene knockout experiments, parasites were transfected with an 11-kb genomic subclone spanning the UPRT locus (in pKS−), from which an internal 1-kb *Bgl* II fragment including two exons was removed. Parasites were electroporated with 20 μg of the resulting plasmid (R11ΔBgl) in circular or linear form, and FUDR-resistant clones were screened for the presence of wild-type and mutant alleles by PCR using the following primers: F10 (sense), 5′-TGCGTTTGCGGATGCGGAGTTCT-3′; F9 (antisense), 5′-CGGTCACCGAAATCACCAATGC-3′, resulting in the amplification of 1502-nt and/or 473-nt fragments derived from the wild-type and mutant alleles, respectively. Amplification involved 35 cycles of denaturation at 94°C for 45 sec, annealing at 58°C for 2 min, and extension at 72°C for 2 min using 2.5 units of *Taq* polymerase. Buffer conditions were optimized to include 75 mM KCl and 1.5 mM $MgCl_2$, pH 8.8 (at 25°C).

## RESULTS

**Insertional Mutagenesis of the UPRT Locus.** In a typical transformation experiment, starting with $10^7$ *T. gondii* tachyzoites (only $\approx 2 \times 10^6$ of which are viable by plaque assay), $\approx 10^6$ parasites survive electroporation. Using vectors containing pyrimethamine-resistant DHFR–TS genes, we have reported stable transformation frequencies of ≈6% of the $10^6$ parasites that remain viable after electroporation, yielding a total of $\approx 6 \times 10^4$ independent integrations per experiment. As noted previously (8, 10), cDNA-derived DHFR–TS vectors typically integrate by nonhomologous recombination, presumably because insufficient contiguous sequence is available for homologous recombination (the endogenous gene is fragmented by nine introns; ref. 12). Given the estimated size of the *T. gondii* genome as $\approx 8 \times 10^7$ bp of DNA (16, 17), truly random transgene integration should yield an average density of approximately one hit in any 1.3-kb segment of genomic DNA—i.e., a predicted frequency of one plaque per transformation for insertional inactivation of a gene containing 1.3 kb of essential coding and regulatory sequences. Because *Toxoplasma* is haploid throughout its life cycle, except in the unsporulated oocyst stage (16–18), a single insertion should be sufficient to inactivate a single-copy gene.

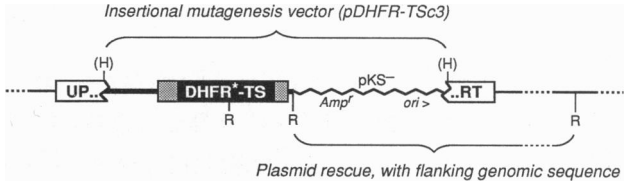


FIG. 1. Strategy for insertional mutagenesis of the *T. gondii* UPRT gene and for rescue of the tagged locus. The *UPRT* gene (open boxes) is shown disrupted by insertion of mutagenesis plasmid pDHFR-TSc3 (8), producing resistance to FUDR. pDHFR–TSc3 consists of a DHFR–TS "minigene" (genomic promoter fused to the cDNA-derived protein coding region), engineered to incorporate point mutations that confer resistance to pyrimethamine, cloned into bacterial plasmid pKS− (zig-zag line) between the *Hin*dIII and *Eco*RI polylinker sites. Insertion is depicted as if a single plasmid linearized by *Hin*dIII digestion has integrated into the UPRT gene intact, in the "sense" orientation. Total genomic DNA digested with *Eco*RI (or other polylinker enzymes that separate the DHFR–TS minigene from the bacterial vector) can be self-ligated to permit reisolation of the pKS− plasmid along with flanking UPRT genomic sequence by direct transformation of bacteria. R, *Eco*RI sites; (H), plasmid *Hin*dIII sites (assumed lost during integration).

The insertional mutagenesis and DNA rescue strategy is outlined in Fig. 1. Strain RH tachyzoites were transfected with pyrimethamine-resistance plasmids, which preferentially integrate into the parasite genome by nonhomologous recombination (8), and inoculated into HFF cells. After a single passage (≈3 days) in 1 μM pyrimethamine to select for the presence of DHFR–TS transgenes, FUDR was added to a final concentration of 5 μM. Transfection with either circular or linearized plasmid produced resistant parasites at a frequency of $\approx 10^{-6}$ (≈1 plaque per transformation)—remarkably close to the predicted frequency. No FUDR-resistant parasites were seen in mock-transfected controls or samples that received no electroporation shock. Four independent mutants (from separate experiments) were cloned by limiting dilution and assayed for UPRT activity, as described (23); in each case, <2% of wild-type parasite activity was observed (data not shown).

Mapping of pDHFR–TSc3 transgenes has shown relatively little rearrangement (8), and Southern blotting indicated that this is the case for the FUDR-resistant mutants as well (data not shown). Clones 1 and 2 were originally transfected with circular plasmid, whereas clones 3 and 4 were transfected with linearized plasmid. Clones 1–3 each harbor single-copy insertions, whereas clone 4 contains multiple tandem insertions at a single site. These observations suggested that it might be possible to recover the integrated plasmid from clone 3 along with flanking genomic DNA by a classical marker rescue strategy, as diagrammed in Fig. 1. DNA from these parasites was therefore digested with *Eco*RI (which cuts once in the transfected DNA, separating DHFR–TS sequences from the plasmid vector as diagrammed in Fig. 1) or *Nco* I (which does not cut within the transformation vector) and self-ligated in dilute solution. The integrated vector was rescued by electroporation of *E. coli* and selection for ampicillin resistance, as described under *Materials and Methods.*

A 2.1-kb flanking genomic fragment rescued along with the integrated vector from *Nco* I-digested DNA was radiolabeled and hybridized with restriction-digested DNA from two independent FUDR-resistant clones, as shown in Fig. 2. The upward shift of the *Bam*HI fragment in clones 1 and 2 clearly indicates that the same locus was tagged in both of these clones as in clone 3 (from which the probe was derived). Digestion with *Eco*RI indicates that integration occurred at distinct sites in clones 1 and 2, however, as different genomic *Eco*RI fragments were disrupted. Unrelated transgenic parasite lines

(pyrimethamine-resistant, but FUDR-sensitive; lanes marked by T) exhibit wild-type organization at this locus (compare with lanes marked RH).

**Identification of the UPRT Gene.** Flanking sequences were used as probes to identify cDNA and genomic clones from wild-type *T. gondii* (strain RH) libraries. Nine cDNA clones were identified, all of which contain a single large open reading frame predicted to encode a 244-amino acid protein (GenBank accession no. U10246). The predicted protein is unambiguously related to other UPRT sequences, as shown in Fig. 3*A*. The *T. gondii* UPRT protein sequence shares 51% amino acid identity with *Saccharomyces cerevisiae* and 26% with *E. coli*. Seventy-five percent of those residues found in both *S. cerevisiae* and *E. coli* are shared by *T. gondii* as well. Conservation is particularly good near the putative phosphoribosyl pyrophosphate (PrPP) binding site (32).

As indicated in Fig. 3*B*, seven intervening sequences were identified from genomic clones. All splice junctions match the previously defined consensus for *T. gondii* (12). The 5′ termini of mature mRNAs have not been mapped, but eight different 5′ ends were observed in the nine cDNA clones sequenced, at positions between 110 and 378 nt upstream of the probable ATG initiation codon (see annotation of GenBank entry). No TATAA box or other obvious promotor sequences are apparent in the genomic DNA. All cDNA clones were polyadenylylated, but five different poly(A) sites were found between 273 and 1121 nt downstream of the TAG termination codon. Multiple 5′ ends, the lack of an identifiable TATAA box, and multiple polyadenylylation sites have also been observed in transcripts from the *T. gondii* DHFR–TS locus (A. D. Sagar, A. L. C. Moulton, and D.S.R., unpublished work). Sequencing the plasmid rescued from FUDR-resistant clone 3 revealed that integration occurred within exon IV, and integration points for insertional mutants 1, 2, and 4 were mapped by Southern blotting (data not shown) to restriction fragments shown in Fig. 3*B*.

**Inactivation of Wild-Type UPRT by Genetic Knock-Out.** Previous studies have demonstrated high-frequency homologous recombination in *Toxoplasma*, using plasmids that contain long stretches (>8 kb) of contiguous genomic sequence (10). This approach was used to create a UPRT knock-out, by

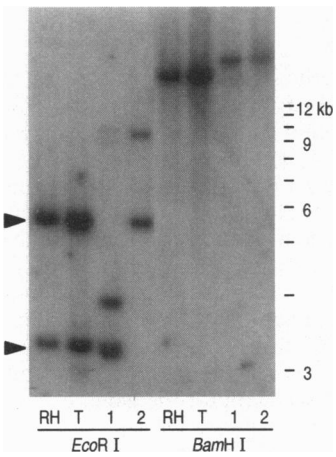


FIG. 2. Insertion at the same locus in three independent FUDR-resistant mutants. Genomic DNA flanking the vector insertion point in FUDR-resistant clone 3 was retrieved by marker rescue as described in text and hybridized with *Eco*RI- or *Bam*HI-digested genomic DNA from wild-type *T. gondii* (RH strain), an FUDR-sensitive transgenic parasite clone (T) and two independent FUDR-resistant parasite clones (1 and 2). Arrowheads indicate two *Eco*RI fragments in wild-type parasites that lie near the FUDR clone 3 insertion point; FUDR-resistant clones 1 and 2 have both inserted into this region but at different sites, as indicated by the shift of the upper and lower *Eco*RI bands, respectively.

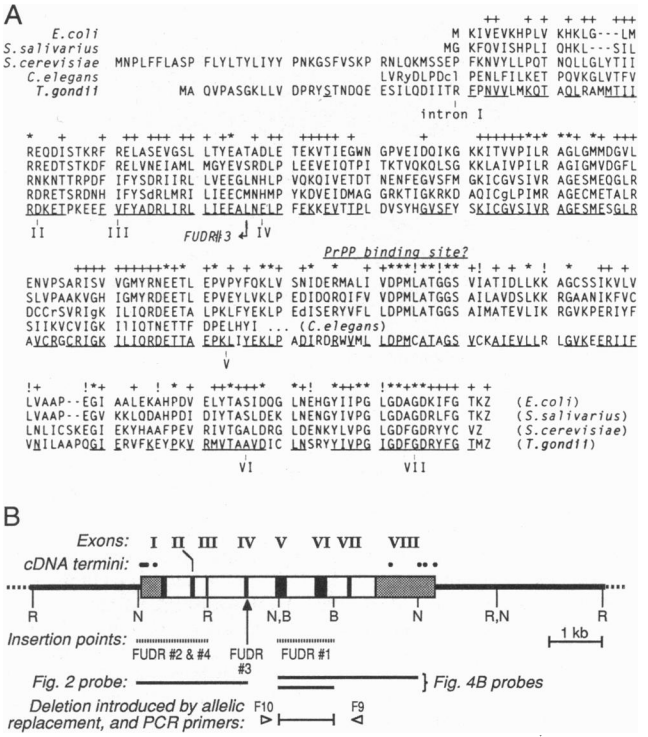


FIG. 3. Protein coding sequence and genomic organization of *T. gondii* UPRT. (*A*) Predicted protein sequence alignments for the UPRT genes of *E. coli* [Protein Identification Resource (PIR) accession no. P25532]; *Streptococcus salivarius* (GenBank accession no. L07793); *S. cerevisiae* (PIR no. P18562; lowercase letters indicate discrepancies between the PIR sequence and GenBank no. S57516); *Caenorhabditis elegans* (GenBank accession no. Z14695—an expressed sequence tag of unknown function; lowercase letters indicate uncertain sequence); and *T. gondii* (this report; GenBank no. U10246). The putative phosphoribosyl pyrophosphate (PrPP) binding site (32) is indicated. *, Residues conserved in all known sequences; +, residues conserved in most known sequences; !, residues conserved in all sequences except *T. gondii*. Underlined residues in the *T. gondii* sequence are conserved in at least one other species. The insertion point for FUDR-resistant clone 3 is shown; positions of intervening sequences in the *T. gondii* gene are indicated by Roman numerals. (*B*) Genomic map of the *T. gondii* UPRT locus; scale bar = 1 kb. Coding regions are shown as filled boxes, noncoding sequences are shown as shaded boxes, and intervening sequences are shown as open boxes. Dots indicate the 5′ and 3′ ends of cDNA clones. Restriction sites are as follows: B, *Bgl* II; N, *Nco* I; R, *Eco*RI. The vector insertion point for FUDR-resistant mutant 3 was precisely mapped near the 3′ end of exon IV (see *A*); insertion points for three additional FUDR-resistant mutants were mapped to the indicated restriction fragments by Southern analysis. A deletion lacking the 1.0-kb *Bgl* II fragment was used to knock out the endogenous UPRT locus by homologous recombination; positions of primers used to screen for UPRT deletions (Fig. 4*A*) are indicated by open arrowheads. Clones spanning the entire region shown were used to rescue activity by transient transformation (Fig. 5).

transfecting wild-type parasites with a plasmid containing the 11-kb genomic region shown in Fig. 3*B*, modified to eliminate the internal 1-kb *Bgl* II fragment indicated. This deletion removes part of exon V and all of exon VI (including the predicted phosphoribosyl pyrophosphate binding site; see Fig. 3*A*).

Wild-type parasites were transfected with circular or linearized plasmid DNA containing the knock-out construct, selected for resistance to 5 μM FUDR, and cloned by limiting dilution. Drug-resistant mutants arose at a frequency of $\approx 3 \times 10^{-3}$ for circular transfectants and $7 \times 10^{-4}$ using linearized plasmid (determined by direct plaque assay). As shown in Fig. 4*A*, 18 clones were analyzed for the presence of wild-type or internally deleted UPRT genes by PCR, using primers flanking

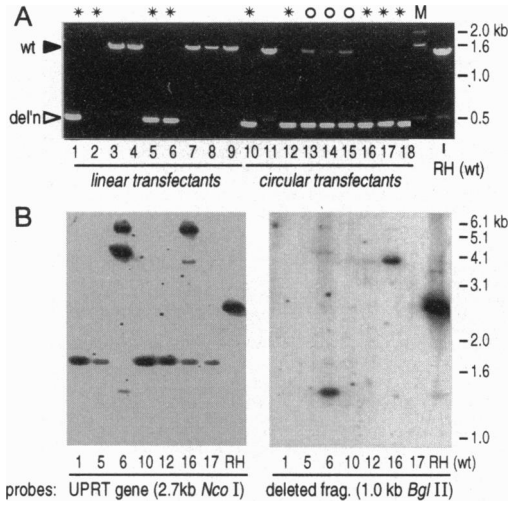


FIG. 4. Deletion of the UPRT gene by homologous gene replacement. Parasites were transfected with a genomic subclone spanning the UPRT locus, from which a 1.0-kb *Bgl* II fragment had been eliminated. DNA was isolated from individual FUDR-resistant parasite clones derived from transfections with linear (clones 1–9) or circular plasmid (clones 10–18). See map shown in Fig. 3*B* for the location of primers and probes used in screening by PCR and hybridization analysis. (*A*) PCR amplification (see *Materials and Methods* for details of primers and conditions) was used to identify a 1502-nt fragment from the wild-type (wt) locus (filled arrowhead), and/or a 473-nt fragment from the deletion (del'n) mutant (open arrowhead). These primers also occasionally amplify an unrelated fragment of >0.5 kb; see wild-type control. *, Replacement of wild-type gene with deletion mutant; o, pseudodiploids containing both wild-type and mutant alleles (10). RH, Wild-type RH-strain parasite DNA; M, DNA markers (BRL). (*B*) Hybridization of *Nco* I-digested DNA with probes spanning the UPRT deletion (*Left*) or containing only the targeted 1-kb *Bgl* II fragment (*Right*). Five of seven clones (nos. 1, 5, 10, 12, and 17) have completely lost an essential portion of the UPRT gene. Lane numbering corresponds to clones in *A*.

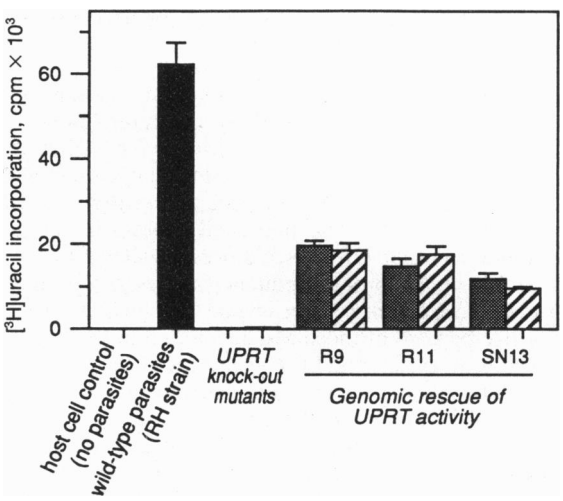


FIG. 5. UPRT inactivation in knock-out mutants and restoration of activity by transient transformation with the cloned genomic locus. Two independent UPRT–knock-out mutants (clones 1 and 17 from Fig. 4) were transfected with plasmids containing wild-type genomic DNA spanning the UPRT locus, and $10^5$ parasites were inoculated into confluent cultures of HFF cells in 24-well plates at 37°C. Five microcuries (1 Ci = 37 GBq) of [5,6-$^3$H]uracil was added 20 hr after infection, and incorporation was assayed 4 hr later by trichloroacetic acid precipitation (5). Solid bar, wild-type parasites; shaded bars, UPRT knock-out clone 1; hatched bars, UPRT knock-out clone 17. R9, 9-kb UPRT genomic clone corresponding to the two left-most *Eco*RI fragments in Fig. 3*B*; R11, entire 11-kb genomic clone from Fig. 4*B*; SN13, genomic clone encompassing the two left-most *Eco*RI fragments in Fig. 3*B* and extending 4 kb further upstream. Error bars indicate SEMs in triplicate samples. No uracil incorporation was detected in UPRT knock-out clones, but transfection with the wild-type gene restores activity. In the absence of stably transformed clones, restoration of UPRT activity relies on transient expression assays, accounting for the lower activity than seen in wild-type parasites.

the *Bgl* II deletion. The expected deletion (with no trace of residual wild-type genes) was found in four of nine parasite clones derived from electroporations with linearized plasmid and in five of nine clones transfected with circular plasmid (asterisks). Perfect allelic replacement of the endogenous locus with the deletion mutant—without insertion of vector DNA—was confirmed for five of the seven clones tested by Southern hybridization (Fig. 4*B*).

Because the selection scheme used requires loss of UPRT activity, PCR bands of apparently normal size are presumed to harbor mutations or rearrangements that cannot be detected at this level of resolution. The need to disrupt the UPRT gene to produce FUDR resistance also explains the relatively low frequency of UPRT targeting relative to previous reports of homologous recombination at the DHFR–TS locus (10): the most common mechanism—precise reciprocal crossing-over between a circular plasmid and the homologous genomic locus at a single site (duplicating the region)—would not normally inactivate UPRT activity. It is interesting to note that three clones transfected with circular plasmid (but none of the linear transfectants) contain PCR bands corresponding in size to both wild-type and deleted UPRT (lanes 13–15 in Fig. 4*A*), consistent with reciprocal cross-over associated with a frameshift or other mutation.

UPRT deletion mutants exhibit no growth defect in normal culture (or in mice; data not shown), but the ability of these parasites to incorporate radiolabeled uracil was completely abolished, as shown in Fig. 5. Transfection of UPRT knock-out mutants with constructs containing the wild-type locus restored uracil incorporation, however, confirming that this locus is both necessary and sufficient for UPRT activity. Transient expression provided 16–32% of wild-type incorporation—essentially complete restoration of activity, given that approximately half of the parasites are killed during electroporation and only ≈60% of the viable parasites exhibit transient expression of transgenes (8).

## DISCUSSION

This report describes the successful use of insertional mutagenesis to identify a genetic locus in *T. gondii*. These parasites are haploid throughout their life cycle, except for the short-lived unsporulated oocyst (18). Thus insertional inactivation of a single gene in the parasite tachyzoite form—which can readily be cultured *in vitro* (5)—is sufficient to produce an identifiable phenotype. Note, however, that the haploid nature of these parasites precludes the identification of genes essential for tachyzoite survival by insertional mutagenesis. In conjunction with classical and somatic cell genetic approaches (4) and more recently developed molecular genetic and mapping techniques (5, 6), insertional mutagenesis renders *T. gondii* a uniquely accessible system for the genetic analysis of a protozoan parasite.

The high frequency of nonhomologous recombination observed using cDNA-derived DHFR–TS vectors (8, 10) provides the basis for insertional mutagenesis experiments. Although the molecular basis for the remarkable efficiency of stable transformation using these vectors remains unclear (particularly in light of the lower frequencies observed with other selection schemes; refs. 9 and 11), these results are validated by the ability to identify genes by insertional mutagenesis.

A practical insertional mutagenesis system requires not only high-frequency nonhomologous recombination but also that

chromosomal integration occurs at random or, at least, that integration sites be sufficiently broadly dispersed to permit insertion into any gene. Previous studies reported that multiple independent transgenes integrated at different loci on different chromosomes (8). In this report, four clones that independently marked the same locus integrated into at least three distinct sites (Fig. 3*B*). Ultimately, however, the best indication that transgene integration occurs essentially at random throughout the parasite genome is the success of the insertional mutagenesis strategy itself. In addition to identifying the UPRT locus at frequencies close to that expected for random integration (see below), we have recently succeeded in identifying the hypoxanthine-xanthine-guanine phosphoribosyltransferase and adenosine kinase loci, by insertional mutagenesis followed by selection in 6-thioxanthine (25) or adenine arabinoside (24), respectively (R.G.K.D. and D.S.R. unpublished work). These results indicate that the feasibility of insertional mutagenesis is not restricted to the UPRT locus and, therefore, suggest that a single transformation results in saturation mutagenesis of the entire parasite genome.

It is difficult to precisely determine the expected frequency for insertional mutagenesis at the UPRT locus—even assuming that transgene integration occurs completely at random throughout the $8 \times 10^7$ bp genome—because the effective target size is uncertain. The UPRT gene spans ≈7 kb, but the effect of insertion within an intron remains unknown. The UPRT coding sequence is only 732 nt. If we assume that integration anywhere within the entire gene would disrupt UPRT function, a single transformation (producing $\approx 6 \times 10^4$ stable transformants; refs. 8 and 10) should yield ≈5 FUDR-resistant mutants; if the effective target consists of coding sequence only, ≈0.5 plaque would be expected per experiment (see *Results* for a more detailed description of integration frequencies). The observed frequency falls between these extremes.

In contrast to transposon mutagenesis (19), where the presence of multiple transposons throughout the genome can complicate analysis, the ability of insertional mutagenesis to integrate at a single specific site greatly facilitates identification of the affected locus (this is fortunate, as back-crosses in *Toxoplasma* are cumbersome; ref. 18). Moreover, although stable transgene integration in many insertional mutagenesis systems results in gross rearrangement of the transfected plasmid (20, 21), the lack of rearrangement seen in *T. gondii* transgenics (8) permits rapid identification of the tagged locus by self-ligation of total genomic DNA and marker rescue in bacteria, obviating the need for laborious library construction and screening. Transformation with linearized (as opposed to circular) vectors facilitates this approach in three ways: (*i*) integration is less likely to disrupt plasmid sequences necessary for bacterial replication, (*ii*) internal vector restriction sites that separate the bacterial plasmid from the selectable marker can be used for plasmid rescue, and (*iii*) mapping of restriction sites in the flanking DNA suitable for plasmid rescue is more easily done. We have occasionally observed cases of multicopy transgene insertion at a single locus (ref. 8; and additional unpublished results), but it should be possible to design more complex restriction strategies capable of retrieving the insertion junctions via plasmid rescue in these transformants as well.

The UPRT gene targeted in these studies encodes a parasite-specific enzyme with potential as a target for chemotherapeutic design (26). Although the enzyme is not essential for parasite survival (and hence provides a suitable target for insertional mutagenesis), subversive substrates such as fluorouracil serve as potential lead compounds. Analogous strategies have previously proved effective in the design of parasiticidal agents (33), as with the use of allopurinol for Chagas' disease and leishmaniasis (34, 35). The UPRT gene also provides a negative selectable marker for further genetic studies. Combined positive/negative selection vectors should permit the generation of genetic knock-outs or allelic replacements by "hit-and-run" mutagenesis (27), suitable even for loci where no direct selection is available. Such strategies offer the added attraction of removing the transfection vector from the final product—a particular advantage if pyrimethamine resistance is used for positive selection. The knock-out mutants isolated in the course of this study provide an appropriate host for these experiments.

We thank Dr. E. R. Pfefferkorn and members of the Roos laboratory for many helpful discussions. This work was supported by Grants AI-28724 and AI-31808 from the National Institutes of Health. D.S.R. is a National Science Foundation Presidential Young Investigator, with support from the MacArthur Foundation and Merck Research Laboratories, and a Burroughs Wellcome New Investigator in Molecular Parasitology.

1. Luft, B. J. & Remington, J. S. (1992) *Clin. Infect. Dis.* **15,** 211–222.
2. Tenant-Flowers, M., Boyle, M. J., Carey, D., Marriott, D. J., Harkness, J. L., Penny, R. & Cooper, D. A. (1991) *AIDS* **5,** 311–315.
3. Laughon, B. E., Allavdeen, H. S., Becker, J. M., Current, W. L. & Feinberg, J. (1991) *J. Infect. Dis.* **164,** 244–251.
4. Pfefferkorn, E. R. (1988) *MBL Lect. Biol.* **9,** 479–501.
5. Roos, D. S., Donald, R. G. K., Morrissette, N. S. & Moulton, A. L. C. (1994) *Methods Cell Biol.* **45,** 23–65.
6. Boothroyd, J. C., Kim, K., Pfefferkorn, E. R., Sibley, L. D. & Soldati, D. (1995) *Methods Mol. Genet.* **6,** 1–29.
7. Soldati, D. & Boothroyd, J. C. (1993) *Science* **260,** 349–352.
8. Donald, R. G. K. & Roos, D. S. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 11703–11707.
9. Kim, K., Soldati, D. & Boothroyd, J. C. (1993) *Science* **262,** 911–914.
10. Donald, R. G. K. & Roos, D. S. (1994) *Mol. Biochem. Parasitol.* **63,** 243–253.
11. Sibley, L. D., Messina, M. & Niesman, I. R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5508–5512.
12. Roos, D. S. (1993) *J. Biol. Chem.* **268,** 6269–6280.
13. Peterson, D. S., Walliker, D. & Wellems, T. E. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 9114–9118.
14. Cowman, A. F., Morry, M. J., Biggs, B. A., Cross, G. A. M. & Foote, S. J. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 9109–9113.
15. Tanaka, M., Gu, H.-M., Bzik, D. J., Li, W.-B. & Inselburg, J. (1990) *Mol. Biochem. Parasitol.* **39,** 127–134.
16. Cornelissen, A. W. C. A., Overdulve, J. P. & van der Ploeg, M. (1984) *Parasitology* **88,** 531–533.
17. Sibley, L. D., LeBlanc, A. J., Pfefferkorn, E. R. & Boothroyd, J. C. (1992) *Genetics* **132,** 1003–1015.
18. Pfefferkorn, E. R. & Pfefferkorn, L. C. (1980) *Exp. Parasitol.* **50,** 305–316.
19. Berg, D. A. & Spradling, A. C. (1991) *Genetics* **127,** 515–524.
20. Feldman, K. A. (1991) *Plant J.* **1,** 71–82.
21. Tam, L. W. & Lefebvre, P. A. (1993) *Genetics* **135,** 375–384.
22. Pfefferkorn, E. R. & Pfefferkorn, L. C. (1977) *Exp. Parasitol.* **42,** 44–55.
23. Pfefferkorn, E. R. (1978) *Exp. Parasitol.* **44,** 26–35.
24. Pfefferkorn, E. R. & Pfefferkorn, L. C. (1977) *J. Parasitol.* **6,** 364–370.
25. Pfefferkorn, E. R. & Borotz, S. E. (1994) *Exp. Parasitol.* **79,** 374–382.
26. Iltzsch, M. H. & Tankersley, K. O. (1993) *Biochem. Pharmacol.* **48,** 781–791.
27. Mortensen, R. (1993) in *Current Protocols in Molecular Biology*, eds. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (Greene and Wiley, New York), Suppl. 23, 9.15.1.
28. Schwartzman, J. D. & Pfefferkorn, E. R. (1981) *J. Parasitol.* **67,** 150–158.
29. Iltzsch, M. H. (1993) *J. Eukaryotic Microbiol.* **40,** 24–28.
30. v. d. Hoff, M. J. B., Moorman, A. F. R. & Lamers, W. H. (1992) *Nucleic Acids Res.* **20,** 2902.
31. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
32. Kern, L., DeMontigny, J., Jund, R. & LaCroute, F. (1990) *Gene* **88,** 149–157.
33. Ullman, B. & Allen, T. E. (1995) in *Molecular Approaches to Parasitology*, eds. Boothroyd, J. C. & Komuniecki, R. (Wiley–Liss, New York), pp. 123–141.
34. Gallerano, R. H., Sosa, R. R. & Marr, J. J. (1990) *Am. J. Trop. Med. Hyg.* **43,** 159–166.
35. Martinez, S. & Marr, J. J. (1992) *N. Engl. J. Med.* **326,** 741–744.