MOLECULAR AND CELLULAR BIOLOGY, June 1990, p. 3243–3246
0270-7306/90/063243-04$02.00/0

Vol. 10, No. 6

A and T Homopolymeric Stretches Mediate a DNA Inversion in *Plasmodium falciparum* Which Results in Loss of Gene Expression

LAURA G. POLOGE,† DERIK DE BRUIN, AND JEFFREY V. RAVETCH*
Sloan-Kettering Institute, 1275 York Avenue, New York, New York 10021

Received 8 December 1989/Accepted 15 February 1990

Ring-infected erythrocyte surface antigen-negative isolates of *Plasmodium falciparum* demonstrate a complex DNA rearrangement with inversion of 5′ coding sequences, deletion of upstream and flanking sequences, and healing of the truncated chromosome by telomere addition. An inversion intermediate that results in the telomeric gene structure for RESA has been identified in the pathway. This inversion creates a mitotically stable substrate for the sequence-specific addition of telomere repeats at the deletion breakpoint.

The genome of *Plasmodium falciparum*, the protozoan parasite responsible for the most severe form of human malaria, is extremely flexible. It displays polymorphisms on the order of fifty to hundreds of kilobase pairs between genetically equivalent chromosomes from different parasite isolates (8, 10, 14). These chromosome length polymorphisms can involve deletions in which structural genes are lost (4, 10, 11). One well-characterized mechanism by which this class of polymorphisms is generated is the introduction of double-stranded breaks in the DNA, followed by the enzymatic addition of telomere repeats to the free 3′ ends, thereby stabilizing the foreshortened chromosome fragment to mitosis (11). A conserved sequence element has been identified at the breakpoint of multiple, independent events, suggesting a specificity to either the breakage or the healing reaction (11). This specificity has been further characterized by the detailed analysis of a subtelomeric rearrangement for chromosome 1 in the ring-infected erythrocyte surface antigen (RESA) gene.

RESA is a 155,000-M_r peptide of unknown function which is deposited onto the erythrocyte surface by the invading merozoite during asexual development (2) and which accumulates in the erythrocyte cytoplasm surrounding the gametocyte at the onset of sexual development (13). The structure of the RESA gene in a RESA$^-$ isolate has been shown to involve both an inversion of the signal-encoding exon and a large deletion. This inversion is mediated by homologous recombination between homopolymeric A and T sequences in the 5′ untranslated and intron sequences, respectively. The deletion includes DNA sequences from the middle of the signal exon 5′ to the end of the chromosome (see Fig. 2). The resulting truncated gene is followed by an abrupt transition to the *P. falciparum* telomere repeat sequence (3).

In order to determine the mechanism and generality of the RESA gene deletion and chromosome 1 rearrangement, restriction mapping was performed on DNAs from a number of *P. falciparum* isolates. Figure 1 shows *Hin*cII-digested genomic DNA probed with a RESA sequence probe which recognizes the 5′-most *Hin*cII RESA gene fragment. A 7.8-kilobase fragment was detected in the DNA from RESA$^+$ isolate FC27, as well as in DNAs from clone D10 and the nonclonal isolate FCR3. However, five FCR3-derived clones displayed a smaller *Hin*cII band and migrated as heterogeneous-sized DNA fragments. BAL 31 susceptibility of the RESA gene in total genomic DNA from one of these clones demonstrated that the gene had assumed a telomeric location in these parasites (E96; data not shown). Cloning and sequencing of the D3 mutant confirmed the previously reported RESA$^-$ gene structure and demonstrated the identical breakpoint (3; data not shown). It is

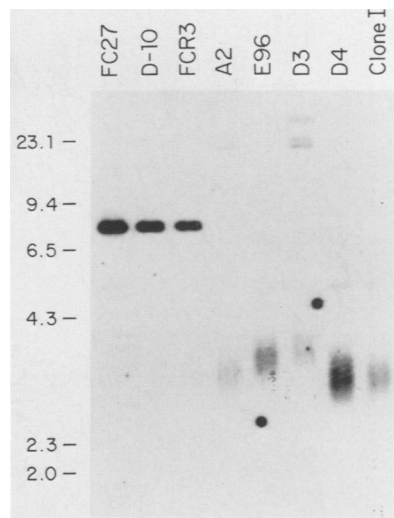

FIG. 1. Southern blot analysis of the RESA gene in DNA from multiple isolates. FC27 (RESA$^+$) is a mixed isolate derived from a patient in Papua New Guinea (9). Isolate D10 (RESA$^+$) was cloned by limiting dilution from strain FC27 (1, 8). FCR3 (RESA$^+$) was isolated from the blood of a patient from Gambia (7). Clones A2, D3, D4, and clone I (RESA$^-$) were cloned by micromanipulation from the Gambian isolate FCR3 (12). E96 (RESA$^-$) is a K$^-$ isolate that arose spontaneously from the clonal K$^+$ isolate, A2 (W. Trager, personal communication). *Hin*cII-digested DNA was probed under stringent hybridization conditions with a DNA fragment that spanned nucleotides 1919 to 1984 in the published RESA sequence (5). Sizes are shown on the left, in kilobases.

* Corresponding author.
† Present address: Department of Medical and Molecular Parasitology, New York University Medical School, New York, NY 10016.

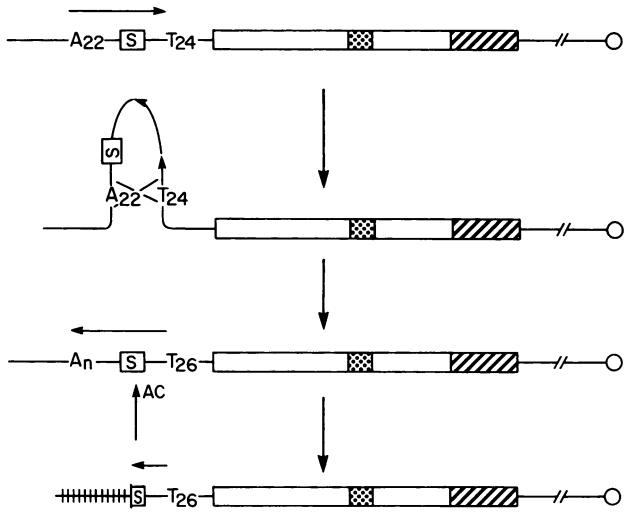

FIG. 2. Proposed mechanism for generating the RESA$^-$ structure. The RESA gene is located on chromosome 1 and is oriented 5′ to 3′, telomere to centromere. Symbols: ○, centromere; //, indeterminant number of kilobase pairs; ▭, translated sequences of the RESA gene with the nontranscribed poly(dA) and poly (dT) sequences 5′ of the signal encoding sequences (5) and second exon, respectively; ▩ and ▨, sequences encoding repeated amino acids (the direction of transcription is 5′ to 3′ from left to right); → (above the signal exon), sequence that is inverted in the pathway; +++++++, telomere repeats. The conserved CA at the breakpoint is shown in line 3 oriented 5′ to 3′, centromere to telomere.

likely, therefore, that the differences in the mobilities of the RESA$^-$ *Hin*cII fragments in Fig. 1 are due to differences in the average number of telomere repeats in the various parasite populations and not to differences in the deletion breakpoints. The breakpoint occurs following a CA dinucleotide sequence, as has been found immediately preceding the transition to telomere repeats in every gene rearrangement to a telomere in *P. falciparum* described to date (11). This suggests that a common sequence element is necessary for some step in the pathway to a viable rearrangement (Fig. 2). This observation leads us to propose a model in which an inversion, which is mediated by base pairing and a crossover event between the homopolymeric A and T sequences that flank the inverted sequences, precedes the deletion event. Subsequent DNA breakage (Fig. 2, ↑ arrow) and telomere addition generated the observed chromosome 1 structure.

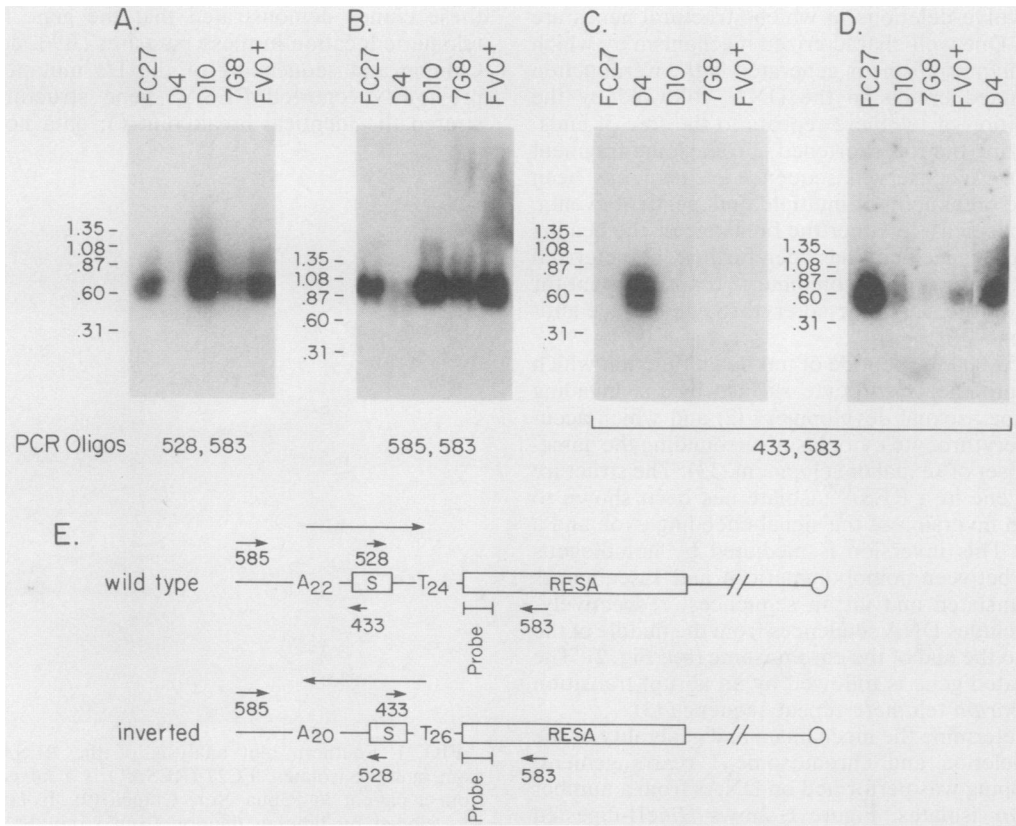

FIG. 3. PCR analysis of wild-type and rearranged RESA sequences from total genomic DNA purified from RESA$^+$ isolates FC27, D10, 7G8, and FVO$^+$ and RESA$^-$ isolate D4. (A) DNA amplification of wild-type RESA sequences; (B) DNA amplification of wild-type and inverted intermediate sequences; (C) DNA amplification of inverted sequences, including sequences organized to a telomere (exposure, 1 h, −70°C with an intensifying screen); (D) overexposure (2 days, −70°C with an intensifying screen) of panel C with D4 amplified DNA loaded onto the gel at a 1:100 dilution compared with other samples; (E) map of RESA gene (top) and inversion intermediate (bottom) on chromosome 1 with PCR oligonucleotides indicated in their 5′-to-3′ orientation. Blots were probed as described elsewhere (15). Oligonucleotide sequences are 585 (TTAAAAGCTTCCTTATTCTTG), 528 (CTCAACAATATATGGGTAC), 433 (CTATATGCATGAAAAGGTCTC), 583 (AAGT TATCAACCTTTTCG), and probe (CCATTATATCCAAGATTACCC). Other details of the map are described in the legend to Fig. 2.

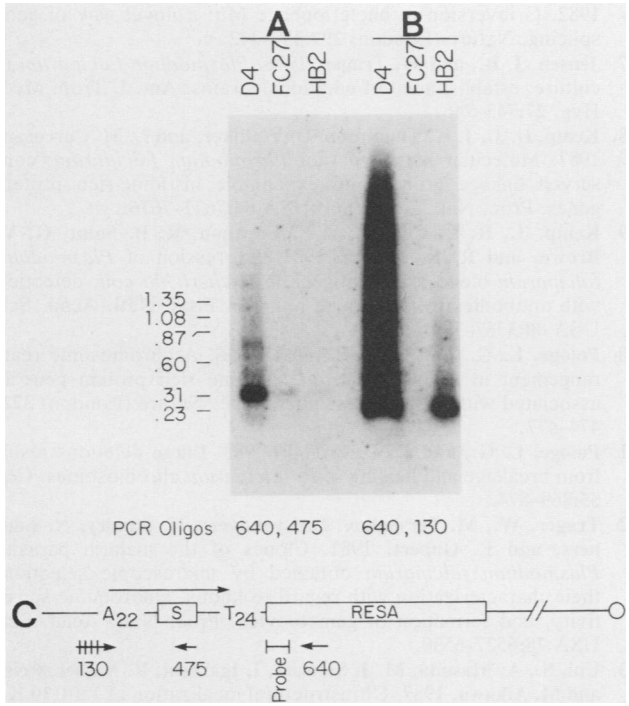

FIG. 4. (A and B) PCR analysis of rearranged RESA sequences from RESA$^-$ isolate D4 and RESA$^+$ isolates FC27 and HB2. Oligonucleotides (Oligos) used to target sequences for amplification are indicated under the autoradiogram. (A) Inverted sequences; (B) sequences rearranged to a telomere. (C) Map of the RESA$^-$ gene on chromosome 1 with PCR oligonucleotides and other details as described in the legend to Fig. 3. Oligonucleotide 475 (TATAGTAT TCCAATAATTGG) was used with oligonucleotide 640 (CTGCA GAAGCTTGAACATCTATCAGTAAA) to specifically amplify the inversion. Oligonucleotide 130, containing three copies of the C-rich telomere repeat sequence CCCTAAA, and oligonucleotide 640 were used to specifically amplify RESA sequences that had been rearranged to a telomere.

We have described similar rearrangements involving chromosome breakage and healing by telomere addition for the knob-associated histidine-rich protein gene on chromosome 2 and the histidine-rich protein II gene on chromosome 8 (11).

This model predicts the existence of an intermediate in which the signal exon and the intron sequences are inverted. To identify this intermediate, the polymerase chain reaction (PCR) was used with oligonucleotides which would specifically target the inverted structure. Total genomic DNAs from a RESA$^-$ mutant and from a number of parasite isolates in which the RESA gene appeared unrearranged were characterized.

Oligonucleotides which would amplify only the wild-type sequence (528,583), the inverted sequence (433,583), or both (585,583) were synthesized (Fig. 3E). The PCR-amplified products were detected by hybridization with an internal sequence (probe). Figure 3A shows the amplification of wild-type RESA sequences in five different geographical isolates. A 580-base-pair fragment was detected in all RESA$^+$ isolates but was absent in the RESA$^-$ isolate D4. Even upon extreme overexposure of the autoradiograms, there is no evidence of wild-type RESA sequences in this isolate. However, using oligonucleotides which flank the presumptive inverted sequences and which would therefore target both the wild-type and the presumptive inversion intermediate, we amplified a 950-base-pair DNA fragment from all of the parasite isolates tested including the D4 DNA (Fig. 3B). This observation has been verified by using a more internal 5′-flanking oligonucleotide, which generates a correspondingly smaller inversion fragment in RESA$^-$ strain D4 of the same intensity, as seen here (data not shown). To confirm the presence of an inversion in strain D4, oligonucleotides that amplify only the inversion were used (433,583; Fig. 3C). Oligonucleotide 433 corresponds to a DNA segment that is not deleted in the telomeric gene. Hybridization with an internal ^{32}P-labeled oligonucleotide detects an abundant product from the RESA$^-$ D4 DNA that corresponds to the gene in the telomeric configuration as well as to a putative inversion intermediate. Significantly, the presence of this inversion is also detected with very long exposure times in DNA from wild-type parasite populations (Fig. 3D). We chose to investigate a number of geographically distinct isolates (FC27 and D10, Papua New Guinea; 7G8, Brazil; D4, Gambia; FVO$^+$, Vietnam) to verify the generality of the RESA inversion.

In order to discriminate between the putative inversion intermediate and RESA sequences that were inverted and healed by telomere addition, oligonucleotide 475 was employed. This oligonucleotide corresponds to sequences that are deleted in the RESA$^-$ telomeric gene and can amplify sequences of only an inverted, undeleted intermediate. Specific amplification of the inversion intermediate, which resulted in a 290-base-pair fragment in RESA$^-$ isolate D4 and at a diminshed level in the RESA$^+$ isolates FC27 and HB2, was observed (Fig. 4A). Amplification of the inverted, telomere-associated DNA fragment was accomplished through the use of an oligonucleotide primer which corresponds to the telomere sequence and an internal RESA sequence (130,640). Amplification of the telomere-associated sequences in these three isolates resulted in an abundant species that was derived from the major RESA species in the D4 DNA, which migrates as a smear because of the amplification of many hundreds of telomeres. Significantly, the HB2 isolate displayed a minor population of sequences in which the RESA gene is associated with a telomere; no telomerically located RESA gene was found in DNA from the FC27 isolate.

These studies indicate that an inversion intermediate can be detected and is likely to precede the deleted, healed form, as predicted in the model proposed above (Fig. 2). A previous model (3) for generating the RESA$^-$ gene structure which proposed an unresolved recombination event with nuclease digestion and healing instead of an intermediate is therefore incorrect. The inversion appears to be a required first step in this DNA rearrangement. If specificity resided in the cleavage reaction, then healing could generate a RESA$^-$ telomeric gene in the absence of inversion. We have never detected a RESA$^-$ telomeric gene with the signal exon in an uninverted orientation. The inversion rearranges the conserved CA dinucleotide to a centromere-proximal position. Breakage and healing of the inverted sequence then result in a mitotically stable, truncated chromosome. These results indicate a sequence specificity for the healing reaction, which always includes a CA dinucleotide. Whether similar sequence specificity is required for the breakage event is still undetermined.

These results demonstrate a novel pathway by which a gene can be inactivated in *P. falciparum*. The generality of this mechanism is currently under investigation. DNA inversion has been shown to regulate gene expression in *Salmonella typhimurium* and bacteriophage Mu. Inversion of pro-

moter-containing segments in *S. typhimurium* switches between the production of two types of flagellar antigens (16), whereas inversion of 3′ coding sequences outside the promoter in bacteriophage Mu modulates the expression of tail fiber protein genes and thereby broadens its host range (6). It is intriguing to speculate that signal exon inversions mediated by flanking homopolymers of A and T, perhaps including transcriptional control signals, may occur as a biological switch to modify gene expression in *P. falciparum*.

This work was supported by U.S. Army grant DAMD 17-89-Z-9003 and the World Health Organization. J.V.R. is a scholar in molecular parasitology supported by the Burroughs-Wellcome Fund.

Technical assistance was provided by Mary Anne Carroll, and secretarial support was provided by Karen Yates.

LITERATURE CITED

1. **Anders, R. F., G. V. Brown, and A. Edwards.** 1983. Characterization of an S antigen synthesized by several isolates of *Plasmodium falciparum*. Proc. Natl. Acad. Sci. USA **80**:6652–6656.
2. **Brown, G. V., J. G. Culvenor, P. E. Crewther, A. E. Bianco, R. L. Coppel, R. B. Saint, H.-D. Stahl, D. J. Kemp, and R. F. Anders.** 1985. Localization of the ring-infected erythrocyte surface antigen (RESA) of *Plasmodium falciparum* in merozoites and ring-infected erythrocytes. J. Exp. Med. **162**:774–779.
3. **Cappai, R., M.-R. van Schravendijk, R. F. Anders, M. G. Peterson, L. M. Thomas, A. F. Cowman, and D. J. Kemp.** 1989. Expression of the RESA gene in *Plasmodium falciparum* isolate FCR3 is prevented by a subtelomeric deletion. Mol. Cell Biol. **9**:3584–3587.
4. **Corcoran, L. M., K. P. Forsyth, A. E. Bianco, G. V. Brown, and D. J. Kemp.** 1986. Chromosome size polymorphisms in *Plasmodium falciparum* can involve deletions and are frequent in natural populations. Cell **44**:87–95.
5. **Favaloro, J. M., R. L. Cappel, L. M. Corcoran, S. J. Foote, G. V. Brown, R. F. Anders, and D. J. Kemp.** 1986. Structure of the RESA gene of *Plasmodium falciparum*. Nucleic Acids Res. **14**:8265–8277.
6. **Giphart-Gassler, M., R. H. A. Plasterk, and P. van de Putte.** 1982. G inversion in bacteriophage Mu: a novel way of gene splicing. Nature (London) **297**:339–342.
7. **Jensen, J. B., and W. Trager.** 1978. *Plasmodium falciparum* in culture: establishment of additional strains. Am. J. Trop. Med. Hyg. **27**:743–746.
8. **Kemp, D. J., J. K. Thompson, D. Walliker, and L. M. Corcoran.** 1987. Molecular karyotypes of *Plasmodium falciparum*: conserved linkage groups and expendible histidine-rich protein genes. Proc. Natl. Acad. Sci. USA **84**:7672–7676.
9. **Kemp, J., R. L. Coppel, A. F. Cowman, R. B. Saint, G. V. Brown, and R. F. Anders.** 1983. Expression of *Plasmodium falciparum* blood stage antigens in *Escherichia coli*: detection with antibodies from immune humans. Proc. Natl. Acad. Sci. USA **80**:3787–3791.
10. **Pologe, L. G., and J. V. Ravetch.** 1986. A chromosome rearrangement in a *P. falciparum* histidine rich protein gene is associated with the knobless phenotype. Nature (London) **322**:474–477.
11. **Pologe, L. G., and J. V. Ravetch.** 1988. Large deletions result from breakage and healing of *P. falciparum* chromosomes. Cell **55**:869–874.
12. **Trager, W., M. Tershakov, L. Lyandvert, H. Stanley, N. Lanners, and E. Gubert.** 1981. Clones of the malaria parasite *Plasmodium falciparum* obtained by microscopic selection: their characterization with regard to knobs, chloroquine sensitivity, and formation of gametocytes. Proc. Natl. Acad. Sci. USA **78**:6527–6530.
13. **Uni, S., A. Masuda, M. J. Stewart, I. Igarashi, R. Nussenzweig, and M. Aikawa.** 1987. Ultrastructural localization of 150/130 Kd antigen in sexual and asexual blood stages of *Plasmodium falciparum*-infected human erythrocytes. Am. J. Trop. Med. Hyg. **36**:481–488.
14. **van der Ploeg, L. H. T., M. Smits, T. Ponnudurai, A. Vermeulen, J. H. E. Meuwissen, and G. Langsley.** 1985. Chromosome-sized DNA molecules of *Plasmodium falciparum*. Science **229**:658–661.
15. **Zeff, R. A., J. Gopas, E. Steinhauer, T. V. Rajan, and S. G. Nathanson.** 1986. Analysis of somatic cell H-2 variants to define the structural requirements for class I antigen expression. J. Immunol. **137**:897–903.
16. **Zieg, J., and M. Simon.** 1980. Analysis of the nucleotide sequence of an invertible controlling element. Proc. Natl. Acad. Sci. USA **77**:4196–4200.