Identification and Characterization of a ran Gene Promoter in the Protozoan Pathogen Giardia lamblia*

(Received for publication, February 26, 1999, and in revised form, April 19, 1999)

Chin-Hung Sun‡§ and Jung-Hsien Tai‡‡¶

From the ‡Division of Infectious Diseases, Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan, Republic of China and §Graduate Institute of Life Sciences, National Defense Medical Center, Taipei, Taiwan, Republic of China

The promoter elements that regulate transcription initiation in Giardia lamblia are poorly understood. In this report, the promoter of the Giardia ran gene was studied using a luciferase expression plasmid pRAN-luc+ to monitor transcription efficiency. An AT-rich sequence spanning −51/−20 relative to the translation start site of the ran gene was identified and found to be required for efficient luciferase expression by deletion and mutation mapping of pRAN-luc+. The −51/−20 sequence was also sufficient for promoter activity as revealed from studies on a 32-base pair synthetic promoter derived from this region. Deletion mapping of the synthetic promoter revealed two minimal promoter elements, −51/−42 and −30/−20, sufficient for 6- and 30-fold luciferase expression above background, respectively. The transcription start sites on luc+ messenger RNA were determined by the position of the synthetic promoter in the luciferase expression plasmids as shown by primer extension experiments. Results from electrophoretic mobility shift assays revealed multiple DNA-protein complexes upon binding of nuclear proteins with either DNA strand but not the double-stranded DNA derived from the ran promoter. Our results delineate the first promoter sequence of the Giardia gene (ran), which provides an excellent model for future studies on transcription regulation in this protozoan parasite.

As a very common intestinal parasite of humans and one of the earliest diverging eukaryotic cells, Giardia lamblia is an important pathogen for biological studies (1, 2). The parasite adapts to drastic changes in host and outside environments by differentiating between actively dividing trophozoites and dormant cysts in response to extracellular stimuli (3, 4). Trophozoites emerge from cysts in the upper intestine and proliferate to cause gastrointestinal disease. Giardiasis usually manifests as self-limiting diarrhea; however, it can also lead to malabsorption and even death in children. Trophozoites differentiate into cysts in the lower intestine and are excreted to the environment.

G. lamblia exhibits both prokaryotic and eukaryotic features (1, 2). Its genome is small and compact (genome size of ~1.2 × 10⁷ bp) (1, 2). It is very interesting to study how the organism regulates its gene expression efficiently. Although intron sequences have not been reported in Giardia genes to date (5), a number of genes that code for putative components of the spliceosomal machinery have been found. The short 5′-untranslated sequence (as short as 1 nt) and the absence of 5′-cap structure of bulk Giardia messenger RNA raise the possibility that G. lamblia may exploit distinct mechanisms for translation initiation different from common strategies used by other eukaryotes (1, 7, 8).

Very little is known about the mechanism and regulation of transcription initiation in G. lamblia. Studies on gene expression during encystation and excystation of G. lamblia in vitro have shown that trophozoites, as well as cysts, produce new transcripts in response to extracellular stimuli (14–18), indicating that regulation of transcription initiation must play a pivotal role during encysting and excysting stages. DNA transfection studies have shown that a 0.8-kilobase pair DNA sequence flanking 5′ of the glutamate dehydrogenase gene (gdh) in G. lamblia is necessary and sufficient for promoter activity (11). However, DNA homology searches in this and other Giardia protein-coding genes so far examined have not revealed consensus core promoter elements used in either prokaryotes or higher eukaryotes. Instead, a number of AT-rich sequences close to or spanning the transcription initiation sites were identified by sequence alignments and suggested to be Giardia promoter elements (19–21). These postulations remain to be confirmed experimentally.

Our previous efforts in the development of an episomal stable DNA transfection system for G. lamblia utilized the promoter-containing sequence within −590 and +27 of a ras-related nuclear protein gene (ran) (13). In the present study, we further dissected the ran promoter by using DNA transfection systems to monitor transcription efficiency and start-site selection of the reporter genes in vivo. We also used electrophoretic mobility shift assay to identify potential transcription factor(s) in vitro. To the best of our knowledge, this is the first promoter sequence to be demonstrated experimentally in G. lamblia. This information will have comparative value when other promoter sequences are identified and will also provide a framework for future studies regarding the transcription regulation of G. lamblia.

EXPERIMENTAL PROCEDURES

G. lamblia Culture—G. lamblia WB axenic culture was maintained in a modified TYI-S-33 medium as described by Keister (22).

* This work was supported by National Science Council Grant NSC88-2314-B001-031 and a grant from Academia Sinica, Taipei, Taiwan, ROC. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence may be addressed: Div. of Infectious Diseases, Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan, Republic of China. Tel.: 886-2-2652-3934; Fax: 886-2-2785-8847; E-mail: bmtai@cvax.sinica.edu.tw.

‡ The abbreviations used are: bp, base pair(s); nt, nucleotide(s); RLU, relative luciferase unit(s); op, tetracycline operator; dsDNA, double-stranded DNA; ssDNA, single-stranded DNA.
DNA Transfection and Luciferase Assay—Plasmids were delivered into *G. lamblia* by electroporation as described previously (13). For transient transfection, 30 μg of plasmid DNA was electroporated into 1.5 × 10^7 *G. lamblia* cells. For stable transfection, 100 μg of plasmid was used. Stable luciferase expression cell lines were established by selection of transfecants with 600 μg ml^-1 of geneticin as described previously (13). The luciferase activity in *G. lamblia* was assayed as described previously (13).

Plasmid Constructions—The sequences of the modified plasmids described below were verified by the automatic DNA sequencing method performed on both strands of the DNA (ABI).

The plasmids pRAN5'2 and pRANneo were previously described (13). The insert in plasmid pRAN5'-2 was amplified by polymerase chain reaction using primer pair ran5'-1 (AATAGGCTTAAGCTAGGAT-TCG) and T7 primer, and cloned into pGEM-T (Promega) to produce pRAN5'-1. The insert was excised from pRAN5'-1 by XbaI/HindIII digestion and ligated into XbaI/EcoRI-digested Bluescript KSII' together with a 1.7-kilobase pair luc+ fragment excised from pSPLuc+ (Promega) by HindIII/EcoRI digestion, resulting in pRAN3Δ. The 3’-sequence flanking the reporter gene was excised from pRANneo by EcoRI/ClaI digestion and subcloned into EcoRI/ClaI-digested pRAN3Δ, yielding pRANLuc+ (Fig. 1A).

Detailed descriptions of the plasmid construction below are available upon request. A series of pRANLuc+ mutant constructs with deletions from the 5’-sequence flanking luc+ (see Fig. 2) were made by the Erase-a-base system according to the supplier’s instructions (Promega). Targeted mutagenesis by polymerase chain reaction was used to create mutations within the −51/−1 region of pRANLuc+, and the mutated sequences in these plasmids are shown in Fig. 3. Defined dsDNA fragments derived from the −51/−20 region were inserted into pSPLuc+ or subsequently derived vectors, and the insertions are shown in Figs. 4 and 5. The luciferase expression cassettes from pAT-20F, pAT-66, pAT-20F-op, and p(−41/−20) (see Figs. 4 and 5) were excised by KpnI/XhoI and subcloned into KpnI/XhoI-digested pRANneo, which resulted in pNAT-20F, pNAT-66, pNAT-20F-op, and p(−41/−20), respectively (Fig. 6, A and C).

Nuclear Extract Preparations—Nuclear extracts were prepared as described previously but with modifications (23). Briefly, *G. lamblia* WB cells were harvested from late logarithmic phase culture and washed once with phosphate-buffered saline. The cells (3 × 10^7) were resuspended in 10 volumes of lysis buffer consisting of 10 mM HEPES (pH 7.9), 0.32 M sucrose, 0.1 mM EDTA, 0.1% Triton X-100, 10 mM β-mercaptoethanol, 20 μM leupeptin, and 2 mM MgCl₂. The lysed cells were homogenized by a Potter-Elvehjem tissue grinder (B. Braun) at the speed of 1,500 rpm for 15 min. The nuclear sediment was obtained by centrifugation at 16,000 × g and resuspended in one volume of extraction buffer consisting of 25 mM HEPES (pH 7.9), 400 mM KCl, 0.1 mM EDTA, 25% glycerol, 10 mM β-mercaptoethanol, 20 μM leupeptin, and 5 mM MgCl₂ and incubated on ice for 30 min with low-speed stirring. Soluble nuclear extract was recovered from centrifugation at 16,000 × g. The nuclear extract was dialyzed in nuclear buffer consisting of 25 mM HEPES (pH 7.9), 40 mM KCl, 0.1 mM EDTA, and 15% glycerol. The protein concentration in the nuclear extract was determined by a BCA protein quantitation kit as described by the supplier (Pierce).

Gel Retardation Assays—Oligonucleotides were labeled at the 5’-end with [γ-32P]ATP (Amersham Pharmacia Biotech) by T4 polynucleotide kinase as described by the supplier (Promega). Binding reactions were performed in a 20-μl volume containing 0.02 pmol of the probe and 1–10 μg of nuclear extract in binding buffer consisting of 50 mM Tris-Cl (pH 7.4), 50 mM KCl, 1 mM EDTA, 1 mM dithiothreitol, and 0.1 mg ml^-1 bovine serum albumin and incubated for 15 min at room temperature. Competition assays were performed by including a 200-fold molar excess of cold DNA oligonucleotides in the reaction mixtures unless otherwise specified. Glycerol was added to the reaction mixture to a 10% concentration, and the mixture was separated in a 6% acrylamide gel by electrophoresis.

**Primer Extension**—Primer extension was performed as described previously (24). Primer extensions on ran messenger RNA were performed using oligonucleotides smg8 (24) and smg9 (AATGTACTGCTTTCGAAAATTCG). Primer extensions on neo messenger RNA were performed using neo1 (CTCGGTTGCAATCCTATCTTGCAA) and neo2 (TCTCTACCCCAAGCGGCGGAGA). Primer extensions on luc+ messenger RNA were performed using luc1 (AACCAGGGCG-TATCTCCTCATAGC) and luc2 (CTTATGCAGTPGCTTCTCCAGCGGT).

**RESULTS**

**Transient DNA Transfection Assays**—A luciferase expression plasmid pRANLuc+ (Fig. 1A) was used to test ran promoter activity in subsequent experiments. The time course of luciferase expression in *G. lamblia* cells transfected with pRANLuc+ was examined in several independent experiments with reproducible results. Different batches of plasmid DNA were also tested for transcription efficiency with only small variations in luminometer readings (less than ±10%). As shown in Fig. 1B, the luciferase activity in cells transfected with pRANLuc+ was detectable (28-fold above background) at 2.5 h post-transfection and increased linearly to 880-fold above background at 20 h post-transfection. Luciferase activity began to level off at subsequent time points and reached a peak value (−1,000-fold above background) at 33 h post-transfection. Luciferase expression remained above 80% of the peak value until 56 h post-transfection. In contrast, cells transfected with pSPLuc+, in which the expression of the reporter gene is driven by a bacteriophage SP6 promoter, only exhibited background activity at all time points tested (0.1% of the peak value). Cells transfected with pGL3-control (Promega), in which luciferase gene expression is driven by an eukaryotic SV40 promoter and enhancer, exhibited 0.5% activity at peak value relative to that of pRANLuc+ at different time points tested.

**Deletion and Mutation Mapping of the ran Promoter**—In subsequent experiments, luciferase activity in the transfected cells was assayed at 30 h post-transfection. The 5’-boundary of the ran promoter was mapped by testing the transcription efficiency of a series of pRANLuc+ + deletion mutants (Fig. 2). Deletion from −595 to −267, −233, −131, −109, −93, −62, or −51 relative to the translation start site had little effect on luciferase expression. The luciferase activity was reduced to 40% of the original level with the deletion to −44. The luciferase activity decreased dramatically to 0.6% of the original level with the deletion to −19. Luciferase expression was completely abolished with the deletion to −2 (0.1%). On the other hand, cells transfected with pRAN3Δ, which lacks the 3’-flanking
sequence, exhibited ~23% of the original activity. These results reveal that the sequence spanning −51/−2 of the ran gene is required for ran promoter activity.

The −51/−1 sequence of pRANLuc+ was examined by scanning mutagenesis (Fig. 3). Luciferase activity decreased to 9, 8, and 6% of the original level with clustered mutations within −51/−42, −41/−31, and −30/−20, respectively. By contrast, the activity only decreased to 55 and 44% with clustered mutations within −19/−11 and −10/−1, respectively. In conjunction with the deletion mapping described above, these results suggest that −51/−20 is more important than −19/−1 for ran promoter activity. We then focused on −51/−20 for further mutational analysis. Luciferase activity only decreased to 27% of the original level with mutations confined to −25/−20. More extensive mutations spanning −51/−31 and −40/−20 further decreased luciferase activity to 2 and 0.3% of the original level, respectively. Mutations of the entire −51/−20 sequence led to only 0.2% of the original activity. These results show that the sequence spanning −51/−20 is necessary for optimal luciferase expression, and the sequence is composed of multiple overlapping suboptimal components.

Synthetic Promoter—A 32-bp synthetic DNA derived from −51/−20 of the ran gene was placed close to the luc+ gene in pSPLuc+ at different positions and orientations (Fig. 4). Since these plasmids lack the 3′-untranslated region of the ran gene, we used pRAN luc+3′Δ as the reference plasmid in the following experiments. The activity of pAT-20f and pAT-66 exhibited 78 and 75% promoter activity, respectively, relative to that of pRAN luc+3′Δ. The activity of pAT-AT was 2-fold that of pRAN luc+3′Δ. On the other hand, pAT-20r only exhibited 8% promoter activity relative to that of pRAN luc+3′Δ, and pAT-d only exhibited activity equivalent to pSPLuc+. Although the −51/−20 sequence could drive gene expression in either orientation, it was 10-fold weaker in the reverse orientation. Moreover, within the limit tested, the distance between the synthetic promoter and the translation start site exerts little influence on transcription efficiency.

Sequences that confer suboptimal promoter activities were mapped by deletion of the −51/−20 sequence in pAT-20f from either the 5′- or 3′-end (Fig. 5). Luciferase activity decreased to 17, 15, and 7% relative to that of pAT-20f in 5′-deletion from −51 down to −41, −34 and −30, respectively. No luciferase activity was observed with further deletion to −25, indicating that −30/−20 is a minimal sequence sufficient for luciferase expression at a level 30-fold above background. The activity only decreased to 83% relative to pAT-20f with insertion of a 19-bp sequence that separates the 3′- minimal sequence from its 5′ −51/−31 sequence. By contrast, the activity was reduced to ~1% relative to that of pAT-20f with 3′-deletion from −20 up to −31 or −42, indicating that −51/−42 is also sufficient for luciferase expression at a level 6-fold above background. The internal sequences such as −41/−31 and −30/−26 only exhibited background activity close to that of pSPLuc+. In conjunction with data obtained from mutagenesis studies (Fig. 3), these results imply that the sequences spanning −51/−42 and −30/−20 are two minimal promoter sequences in G. lamblia, each of which acts in concert with the adjacent AT-rich sequence to mediate higher levels of promoter activity.

Selection of Transcription Start Sites—Stable luciferase expression cell lines were established from drug selection of G. lamblia WB cells transfected with a series of pNAT plasmids. In these pNAT plasmids, the selective marker neo is driven by an episomal ran promoter (13), and luc+ is driven by various synthetic promoters (Fig. 6A). Consistent with transient assays, luciferase expression in stable cell lines harboring pNAT-66, pNAT-20f, and pNAT-20f-op exhibited similar luciferase activities (~10^8 RLU/10^7 cells), whereas cells harboring pN (−41/−20) exhibited much reduced luciferase activity (3 × 10^7 RLU/10^7 cells). The start sites of ran messenger RNA from G. lamblia WB cells were mapped by primer extension using primer smc9 as described previously (24). Consistent with earlier results, three extension products were mapped to −29, −4, and −2 of ran messenger RNA relative to the ran translation start site (Fig. 6B, lane 1). Identical results were obtained by primer smc9 (data not shown). Primer extension on neo messenger RNA from pRANneo was performed using primer neo1, which reveals a single major extension product mapped to −29 (Fig. 6B, lane 2). Identical results were obtained by primer neo2 (data not shown).

The primer extension experiments on luc+ messenger RNAs were also performed by two different primers, luc1 and luc2, with consistent results, and only those performed by luc1 are shown (Fig. 6B, lanes 3–6). Since all luc+ messenger RNAs studied here lack the ran coding sequence, the start sites described below are relative to the luc+ translation start site. Results from primer extension experiments are summarized in Fig. 6C. Primer extension on luc+ messenger RNA from pNAT-20f revealed two faint bands at −30 and −12 and two stronger

---

**Fig. 2.** Deletion mapping to define the 5′-boundary of the ran promoter. The 5′-flanking sequence (thin line), which contains an AT-rich domain (striped box) in pRANLuc+, was deleted from the 5′-end to the indicated sites. The vector sequence is shown as a thick line. Plasmid pRAN3′Δ, which lacks 3′-flanking sequence is also included. The luciferase activities in transfected cells were assayed at 30 h post-transfection. The results are the average ± S.E. of duplicate samples from three separate experiments and are shown as percentage of activity of pRANLuc+ (6615 RLU/10^7 cells taken to be 100%).
The Promoter of Giardia ran Gene

bands at −22 and +1 (Fig. 6B, lane 4). A faint band at −76 and four downstream stronger bands at −68, −52, −51, and −48 were observed on luc+ messenger RNA from pNAT-66 (Fig. 6B, lane 3), showing that the two distal start sites on the luc+ messenger RNA moved upstream by exactly 46 nt when the AT-rich promoter was moved from −20 to −66. The positions and numbers of the other proximal start sites were variable within a 13-nt window (between −12 and +1 of luc+ messenger RNA from pNAT-20f). On luc+ messenger RNA from pNAT-20f-op, a major site at −22 and a minor site at −21 were selected (Fig. 6B, lane 5). No primer extension product was seen on luc+ messenger RNA from pN(−41/−20), showing that reduced gene expression by this construct occurs at the transcrip- tion level (Fig. 6B, lane 6). It is clear that the distal start site (−29 of ran) is used to some extent in all messages except that from pNAT-20f-op, and a new start site 8 nt downstream is also used in the luc+ transcripts. Both of the consensus sites reside within the 32-bp AT-rich region. These results suggest that the position of the 32-bp AT-rich region in the promoter site is the primary determinant for start site selection.

Promoter-binding Proteins—Potential nuclear factors that interact with the ran promoter were examined by electrophoretic mobility shift assay. The sequence and binding activ- ity of each oligonucleotide used in this assay are listed in Table I. The DNA-protein complexes described below were eliminated by the addition of proteinase K (50 μg ml−1) but were not affected by the addition of RNase (100 μg ml−1) in the reaction (data not shown). Ten μg of nuclear extract was used to bind 32P-labeled dsDNA probe −51/−20ds in each binding reaction, and several batches of nuclear extract were tested for DNA binding. No DNA-protein complex was detected under the conditions tested (Fig. 7A, lane 2). By contrast, a broad band was detected when a non-template strand DNA probe, −51/−20, was incubated with 1 μg of nuclear extract (Fig. 7B, lane 2), indicating the formation of DNA-protein complexes in this re- action. The DNA-protein complexes were completely eliminated by cold −51/−20 (Fig. 7B, lane 3) but were not affected by a nonspecific competitor NS (Fig. 7B, lane 4). Due to sequence redundancy in −51/−20, we also used its subunits as compet- itors and found that the complex formation was almost elimi- nated by −30/−20 (Fig. 7B, lane 7). The DNA-protein com- plexes could not be completely eliminated even with a 1,000- fold molar excess of −30/−20 (data not shown). The complex formation was not affected by the competitor sequence −51/−42 (Fig. 7B, lane 5), −41/−31 (Fig. 7B, lane 6), −45/−38, or −35/−26 (data not shown). These observations suggest that
binding of certain protein factor(s) to the ran promoter is specifically directed toward the 3'-end of the non-template strand DNA.

Two broad bands were detected when template strand DNA probe −20/−51 was used in the binding reaction (Fig. 7C, lane 2). The complex formation was completely inhibited by competitor DNA −20/−51 (Fig. 7C, lane 2) but was not affected by NS (Fig. 7C, lane 4), indicating that the template strand of the ran promoter is also the target of certain nuclear proteins. The DNA-protein complexes were almost removed when −42/−51, −31/−41, −20/−30 (Fig. 7C, lanes 5–7), or an octameric T tract oligonucleotide T8 (data not shown) was used in the competition assay. However, the DNA-protein complexes could not be completely eliminated even with a 1,000-fold molar excess of −42/−51, −31/−41, −20/−30, or T8 (data not shown). These results indicate that −20/−51 contains multiple protein binding sites that may contribute to higher affinity for certain nuclear proteins, and T-tract sequences in −20/−51 may participate in protein binding. The multiple DNA-protein complexes formed in the binding reaction were probably due to oligomerization of two or more DNA-binding proteins on this template.

We further tested whether the non-template strand DNA −30/−20 is sufficient for DNA-protein complex formation. A broad band was detected in binding of nuclear extract to the −30/−20 probe (Fig. 7D, lane 2). The complex formation was specific, since its formation was not affected by NS (Fig. 7D, lane 4) but was completely inhibited by −30/−20 (Fig. 7D, lane 3). Interestingly, the complex formation was also completely inhibited by the template strand DNA −42/−51, −31/−41, and −20/−30 (Fig. 7D, lanes 5–7), indicating that these DNA fragments bind to the same protein factor(s), and T-tract sequences in these DNA fragments may also participate in protein binding.

**DISCUSSION**

In this report, luciferase activity in transfected *G. lamblia* trophozoites was used to monitor transcription efficiency of modified ran promoter contexts. The levels of *lac* + transcripts from transient transfections were below the detection limit of the primer extension method (data not shown); however, the levels of luciferase expression correlated well with the degrees of mutations or deletions within the 32-bp AT-rich ran promoter sequence (Figs. 3 and 5). The luciferase activities in stable expressing cell lines also correlated well with the results of primer extension experiments (Fig. 6), demonstrating that alteration in luciferase activity indeed reflects the transcription efficiency of various ran promoter contexts.

A typical eukaryotic promoter site that is composed of one or more core promoter element(s) proximal to the transcription start site and additional DNA regulatory elements in the distal region usually spans a few hundred bp (25, 26). A typical promoter site in *Escherichia coli* that is composed of two core promoter elements and adjacent upstream activator binding sites usually spans ~70 bp (25, 27). By contrast, the ran promoter in *G. lamblia* apparently comprises a single AT-rich domain spanning the −51/−20 sequence of the ran gene as supported by several lines of evidence. First, progressive deletions from −590 down to −51 of the ran gene neither enhanced nor suppressed luciferase expression (Fig. 2), implying the absence of any promoter element other than −51/−20 in the 5'-untranslated region of the ran gene. Second, luciferase expression was eliminated after the entire −51/−20 sequence was deleted (Fig. 2) or mutated (Fig. 3). Smaller mutations or deletions within the −51/−20 sequence only led to reduced gene expression. The downstream sequence −19/−1, which was needed for optimal luciferase expression, was insufficient to confer significant promoter activity. Third, a 32-bp synthetic promoter spanning the −51/−20 region was necessary and sufficient to direct the expression of a downstream reporter gene (Fig. 4). The presence of an intragenic core promoter element in the ran promoter is unlikely, since the protein coding sequence of the ran gene is not required for luciferase expression in the synthetic promoter system (Figs. 4–6). It is also important to note that the 32-bp AT-rich sequence also conferred weaker promoter activity when positioned in reverse orientation (Fig. 4). Preliminary data reveal that the ran promoter has the potential to regulate transcription of an upstream gene positioned in reverse orientation relative to the ran gene. These observations indicate that the ran promoter in *G. lamblia* exhibits an unusual simple organization distinct from other common eukaryotic and prokaryotic promoters. Determination of whether such a promoter organization is common in *G. lamblia* awaits identification of more promoters in this organism.

---

\(^2\) C.-H. Sun and J.-H. Tai, unpublished observations.
The sequence of the ran promoter is extremely AT-rich (29 out of 32 bp) and contains five short A-tracks (3–4 A residues in a run) intervened by one or two T residues and a short T tract (three T residues in a run) flanked by two A residues at the 3′-end. Results from scanning mutagenesis and deletion mapping studies show that the multiplicity of the A-tract sequence in a modified ran promoter correlates with transcription efficiency (Figs. 3 and 5). Also, mutations of the T-tract sequence in pRANluc resulted in a 4-fold reduction in luciferase expression (Fig. 3). AT-rich sequences containing two A-tracks

The Promoter of Giardia ran Gene
The promoter of Giardia ran gene

The DNA-protein interactions revealed by electrophoretic mobility shift assay. The oligonucleotides dsDNA −51/−21ds (A), non-template strand DNA −51/−20 (B), template strand DNA −20/−51 (C), and a subunit non-template strand DNA −30/−20 (D) were γ-32P-labeled (lane 1) and incubated with 10 μg (A) or 1 μg (B–D) of nuclear extract in a reaction mixture (lane 2) for 20 min at room temperature. The reaction products were separated in 6% acrylamide gels by electrophoresis. The formation of DNA-protein complexes was competed with a 200-fold molar excess of various cold DNA oligonucleotides (lanes 3–7) as listed above each lane. The residual DNA-protein complexes in the competition assays are indicated by arrowheads. NS represents a non-template strand DNA sequence derived from −71/−52 of the ran gene.

![Figure 7](image-url)

| probe | A | B | C | D |
|-------|---|---|---|---|
|       | -51/−20ds | -51/−20 | -51/−20 | -51/−20 |
| competitor | - | - | - | - |
| fold molar excess | - | - | 200 | 200 |
| nuclear extract | - | + | + | + |
|          | + | + | + | + |
|          | + | + | + | + |
|          | + | + | + | + |
|          | + | + | + | + |
|          | + | + | + | + |
|          | + | + | + | + |
|          | + | + | + | + |
|          | + | + | + | + |
|          | + | + | + | + |

(such as −51/−41; TAAAA TAAAT) or an A-tract plus a T-tract (such as −30/−20; TAAAA CTTTAA) constitute a minimal promoter in G. lamblia (Fig. 5). These observations suggest that the A-tract and T-tract sequences are two important elements of the ran promoter. Although the exact sequence requirements for these elements have not been determined by point mutations, it is interesting to note the presence of an A-tract- or T-tract-like sequences spanning the transcription start sites of several genes in G. lamblia (Fig. 8). In conjunction with a T-tract sequence-specific ssDNA-binding protein discussed below, these observations suggest that the AT-rich sequence spanning the transcription start sites described above is most likely one of the core promoter elements in G. lamblia.

The interactions among core promoter elements and associated transcription factors are primary determinants of start site selection in transcription initiation (28, 29). In G. lamblia, we found that transcription start sites selected were also determined by the 32-bp AT-rich sequence (Fig. 6B). Although some of these sites deviated from authentic ran messenger RNA start sites, they varied within a rather small range (Fig. 6C). Multiple start sites were consistently observed from primer extensions performed on ran messenger RNA. By contrast, only the distal start site was detected on neo messenger RNA (Fig. 6B, lanes 1 and 2). Interestingly, the consensus site resides within the AT-rich sequence of the ran promoter. This discrepancy may reflect differences in chromosomal (ran) and episomal (neo) environments in regulating gene expression or may be due to the involvement of certain DNA sequences downstream of the transcription start site (30). In the present study, the distal start site was still present on various lac+ messenger RNAs but with a much reduced intensity. Instead, a new preferable start site at 8 nt downstream of the distal start site appeared regardless of the positions of the synthetic promoter. These observations suggest that the AT-rich sequence is the primary determinant for start site selection. It is important to note that the two consensus start sites described above are all located within the 3’-minimal promoter element. As shown in Figs. 5 and 6, the minimal promoter element can be separated from its 5’-sequence in the ran promoter without seriously hampering the transcription efficiency or start site selection. This sequence may be a functional analogue of eukaryotic initiators (29).

A key regulatory event in eukaryotic transcription initiation is the recognition of distinct dsDNA promoter sites by specific multimeric protein complexes (28, 29). In DNA-protein binding assays, we detected DNA-protein complexes on each of the DNA strands but not on the dsDNA fragment derived from the ran promoter (Fig. 7). Binding of a putative transcription factor...
to the dsDNA may have been unstable under our tested conditions, or such a factor was unstable in our nuclear extract preparations. Due to the lack of known dsDNA-binding protein in *G. lamblia*, we could not include a positive control in our experiments. In another pathogenic protozoan, *Trypanosoma brucei*, ssDNA-binding protein interacting with single-stranded promoter elements rather than double-stranded promoter elements was also observed (31, 32). The interactions of DNA-binding proteins to single-stranded promoter elements that influence transcription efficiency have been observed in other eukaryotic promoters (33–37). Whether ssDNA-binding protein(s) identified here also plays a role in regulating transcription initiation in *G. lamblia* is not clear from current studies. Nevertheless, it is interesting to note that a consensus T-tract sequence was present in all ssDNAs that exhibited protein-binding property (Table I), indicating that the T tract sequence may participate in protein binding on single-stranded DNA-binding proteins to single-stranded promoter elements. Our results from DNA-protein competition assays also show that residual DNA-protein complexes were still present on long ssDNA fragments (−51/−20 or −20/−51) even with a 200–1,000-fold molar excess of the short DNA fragments or T8 (Fig. 7, B and C; also data not shown), indicating that the long DNA fragments exhibit higher affinities for binding nuclear proteins than short DNA fragments or that the long fragments may provide extra protein-binding sites for additional nuclear factors. These characteristics may be responsible for enhanced transcription efficiency beyond that which can be achieved by minimal promoters. Since transcription efficiency of pm (−25/−20), which lacks the putative protein binding site on the non-template strand DNA (Fig. 3), was only reduced by 4-fold, it is tempting to speculate that binding of the putative transcription factor to the template strand DNA may be related to the recruitment of a putative transcription complex to the site of transcription initiation, and an additional protein binding site on the non-template strand DNA may stabilize this complex. These possibilities remain to be investigated in the future.

In summary, our studies on the *ran* promoter provide a model system for future investigation on the transcription machinery in *G. lamblia*, one of the earliest diverging eukaryotic single cells.

Acknowledgements—We are grateful to Chih-Yuan Yang for technical support in the preliminary study and Drs. Chun-Fen Chou and Yijuang Chen for useful suggestions and critical reading of the manuscript. We also appreciate the help of Douglas Platt in reading and correcting the manuscript.

REFERENCES

1. Adam, R. D. (1991) *Microbiol. Rev.* 55, 706–732
2. UPO, J., and UPOF P. (1998) *BioEssays* 20, 256–263
3. Gillin, P. D., Reiner, D. S., and McCaffery, J. M. (1996) *Annu. Rev. Microbiol.* 50, 679–705
4. Lujan, H. D., Mowatt, M. R., and Nash, T. E. (1997) *Microb. Mol. Biol. Rev.* 61, 294–304
5. Smith, M. W., Aley, S. B., Sogin, M., Gillin, F. D., and Evans, G. A. (1998) *Mol. Biochem. Parasitol.* 95, 267–280
6. Mowatt, M. R., Weinbach, E. C., Howard, T. C., and Nash, T. E. (1994) *Exp. Parasitol.* 78, 85–92
7. Yu, D. C., Wang, A. L., Botka, C. W., and Wang, C. C. (1998) *Mol. Biochem. Parasitol.* 96, 151–165
8. Sacher, A. B., and Buratowski, S. (1997) *Trends Biochem. Sci.* 22, 189–192
9. Yu, D. C., Wang, A. L., Wu, C. H., and Wang, C. C. (1995) *Mol. Cell. Biol.* 15, 4867–4872
10. Yu, D. C., Wang, A. L., and Wang, C. C. (1996) *Mol. Biochem. Parasitol.* 83, 35–49
11. Yee, J., and Nash, T. E. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 5615–5619
12. Singer, S. M., Yee, J., and Nash, T. E. (1998) *Mol. Biochem. Parasitol.* 92, 59–69
13. Sun, C. H., Chou, C. F., and Tai, J. H. (1998) *Mol. Biochem. Parasitol.* 92, 123–132
14. Lujan, H. D., Mowatt, M. R., Conrad, J. T., Bowers, B., and Nash T. E. (1995) *J. Biol. Chem.* 270, 29007–29013
15. Mowatt, M. R., Lujan, H. D., Cotten, D. B., Bowers, B., Yee, J., Nash, T. E., and Stubble, H. J. (1995) *Mol. Microbiol.* 19, 953–963
16. Que, X., Svard, S. G., Meng, T. C., Hetarko, M. L., Aley, S. B., and Gillin, F. D. (1996) *Mol. Biochem. Parasitol.* 81, 101–110
17. Ward, W., Alvarado, L., Rawlings, N. D., Engel, J. C., Franklin, C., McKerrow, J. H. (1997) *Cell* 83, 437–444
18. Hetarko, M. J., McCaffery, J. M., Svard, S. G., Meng, T. C., Que, X., and Gillin, F. D. (1998) *Exp. Parasitol.* 88, 172–183
19. Hilario, E., and Gogarten, J. P. (1996) *Biochim. Biophys. Acta* 125, 94–98
20. Horigent, D. V., and Marshall, J. (1995) *Nucleic Acids Res.* 23, 2945–2953
21. Bruderer, T., Whehrli, C., and Kohler, P. (1996) *Mol. Biochem. Parasitol.* 77, 225–233
22. Kneller, D. B. (1983) *Trans. R. Soc. Trop. Med. Hyg.* 77, 487–488
23. Dignam, J. D. (1990) *Methods Enzymology* 182, 194–202
24. Chen, L. M., Chern, Y., Ong, S. J., and Tai, J. H. (1994) *J. Biol. Chem.* 269, 17297–17304
25. Waayvyl, B. (1988) *CRC Crit. Rev. Biochem.* 23, 77–120
26. White, R. J., and Jackson, S. P. (1992) *Trends Genet.* 8, 284–288
27. Estrem, S. T., Gaal, T., Ross, W. and Gourse, R. L. (1992) *Proc. Natl. Acad. Sci. U. S. A.* 95, 9761–9766
28. Verrijzer, C. P., and Tjian, R. (1996) *Mol. Biochem. Parasitol.* 21, 338–341
29. Hamsun, M. (1998) *Microbiol. Mol. Biol. Rev.* 62, 465–503
30. Ince, T. A., and Scotto, K. W. (1995) *J. Biol. Chem.* 270, 30249–30252
31. Brown, S. D., and Van der Ploeg, L. H. (1994) *Mol. Biochem. Parasitol.* 65, 109–122
32. Vanhemalle, L., Pays, A., Tehabi, P., Alexandre, S., and Pays, E. (1995) *Mol. Cell. Biol.* 15, 5599–5606
33. Bemark, M., Olsson, H., Heinegard, D., and Leanderson, T. (1998) *J. Biol. Chem.* 273, 18881–18890
34. Tomenaga, T., and Levens, D. (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 5830–5835
35. Michelotti, E. F., Tomenaga, T., Krutzsch, H., and Levens, D. (1995) *J. Biol. Chem.* 270, 19494–19499
36. Tomenaga, T., and Levens, D. (1995) *J. Biol. Chem.* 270, 4875–4881
37. Macdonald, G. H., Itoh-Lindstrom, Y., and Ting, J. P. (1995) *J. Biol. Chem.* 270, 5327–5333
38. Yee, J., and Dennis, P. P. (1992) *J. Biol. Chem.* 267, 7539–7544
39. Ky, P. L., Khanna, K. K., Manning, P. A., and Mayr, H. (1993) *Mol. Biochem. Parasitol.* 58, 247–258
40. Alonso, R. A., and Peattie, D. A. (1992) *Mol. Biochem. Parasitol.* 50, 95–104
41. Alonso, R. A., and Peattie, D. A. (1992) *Mol. Biochem. Parasitol.* 56, 27–28
42. Kirk-Mason, K. E., Turner, M. J., and Chakraborty, P. R. (1993) *Mol. Biochem. Parasitol.* 66, 37–50
43. Murtagh, J. J., Mowatt, M. R., Lee, C. M., Lee, F. S. J., Mishima, K., Nash, T. E., Moss, J., and Vaughan, M. (1992) *J. Biol. Chem.* 267, 9654–9662
44. Roger, A. J., Svard, S. G., Towar, J., Clark, C. G., Smith, M. W., Gillin, F. D., and Sogin, M. L. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 229–234
Identification and Characterization of a ran Gene Promoter in the Protozoan Pathogen *Giardia lamblia*
Chin-Hung Sun and Jung-Hsiang Tai

*J. Biol. Chem.* 1999, 274:19699-19706.
doi: 10.1074/jbc.274.28.19699

Access the most updated version of this article at http://www.jbc.org/content/274/28/19699

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 44 references, 18 of which can be accessed free at http://www.jbc.org/content/274/28/19699.full.html#ref-list-1