Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 6761–6765, September 1986
Biochemistry

Structure and function of a chloroplast DNA replication origin of *Chlamydomonas reinhardtii*

(nucleotide sequence/DNA secondary structure/D-loop)

MADELINE WU, J. K. LOU, D. Y. CHANG, C. H. CHANG, AND Z. Q. NIE

Department of Biological Sciences, University of Maryland Baltimore County, Baltimore, MD 21228

Communicated by J. Herbert Taylor, June 2, 1986

ABSTRACT **Chloroplast DNA replication in *Chlamydomonas reinhardtii* is initiated by the formation of a displacement loop (D-loop) at a specific site. One D-loop site with its flanking sequence was cloned in recombinant plasmids SC3-1 and R-13. The sequence of the chloroplast DNA insert in SC3-1, which includes the 0.42-kilobase (kb) D-loop region, as well as 0.2 kb to the 5′ end and 0.43 kb to the 3′ end of the D-loop region, was determined. The sequence is A+T-rich and contains four large stem–loop stuctures. An open reading frame potentially coding for a polypeptide of 136 amino acids was detected in the D-loop region. One stem–loop structure and two back-to-back prokaryotic-type promoters were mapped within the open reading frame. The 5.5-kb *Eco*RI fragment cloned in R-13 contains the 1.05-kb SC3-1 insert and its flanking regions. A yeast autonomously replicating (ARS) sequence and an ARC sequence, which promotes autonomous replication in *Chlamydomonas*, have been mapped within the flanking regions [Vallet, J.-M. & Rochaix, J.-D. (1985) *Curr. Genet.* 9, 321–324]. Both R-13 and SC3-1 were active as templates in a crude algal preparation that supports DNA synthesis. In this *in vitro* system, chloroplast DNA synthesis initiated near the D-loop site.**

Eukaryotic cells depend on multiple cytoplasmic organelles for the synthesis of large amounts of organelle proteins. The number of organelles may vary in response to changing metabolic needs of the cell. DNA molecules carrying small but vitally important sets of genes are present in the larger organelles, chloroplasts, and mitochondria. The mechanism that regulates and maintains the coexistence of nuclear genome and the indigenous genomes of organelles is intriguing, albeit poorly understood. Nuclear and organelle genomes differ in the degree of control exerted over their replication, recombination, and segregation (1–3). The replication of corn and pea chloroplast DNAs has been shown to initiate with the formation of displacement loops (D-loops) (4, 5). In *Euglena gracilis*, the origin of chloroplast DNA replication has been mapped in a polymorphic region by electron microscopic studies (6–8), and this region has been sequenced (9). The polymorphic region and its immediate vicinity are extremely A+T-rich and have the capacity to form multiple stem–loop structures. A $tRNA^{Trp}$ pseudogene and an open reading frame (ORF) potentially coding for a 41-kDa protein were located within this region. In *Chlamydomonas reinhardtii*, two putative replication origins were mapped previously in this laboratory by electron microscopic study of chloroplast DNA restriction fragments containing D-loops. One D-loop region was delimited and cloned (10, 11). Using an independent approach, Rochaix *et al.* (12) constructed plasmids that replicate autonomously in *C. reinhardtii* (ARC plasmids), by inserting random algal DNA fragments into the yeast plasmid pYe-arg4. Plasmid pYe-arg4 can replicate autonomously in yeast but not in *Chlamydomonas*. After transformation of a strain of *C. reinhardtii* lacking argininosuccinate lyase with recombinant plasmids constructed from pYe-arg4 and algal DNA, arginine prototrophy was used to select ARC plasmids. Inserts of several ARC plasmids were localized on the physical map of the chloroplast genome (12). One plasmid, ARC1, hybridized to the restriction fragment that contains the D-loop region cloned by Wang *et al.* (11). Recently, ARC1 sequence and a yeast ARS (autonomously replicating in *Saccharomyces*) sequence were mapped in the vicinity of the D-loop region (13). In the work reported here, the DNA sequence of the D-loop region and its flanking regions was determined and analyzed. An *in vitro* DNA replication system was used to study the origin functions of the cloned chloroplast DNA fragments.

MATERIALS AND METHODS

Construction of Recombinant Plasmids. Construction of clones R-13 and SC3-1, which contain one D-loop region of chloroplast DNA isolated from *C. reinhardtii* CC125 mt^+, was reported previously (11). Clone R-13 contains a 5.5-kilobase (kb) *Eco*RI fragment cloned in pBR325. In clone SC3-1, the 1.05-kb restriction fragment generated by *Bam*HI and *Cla* I digestions of R-13 was cloned into pBR322 that had been similarly restricted. The chloroplast DNA insert in SC3-1 does not contain any *Hae* III site. Therefore, the two flanking *Hae* III sites located at nucleotide 401 and 4344 of the pBR322 sequence were used for the construction of *Alu* I–*Hae* III (AH) subclones. Cloned DNA of SC3-1 was digested with *Hae* III, and the largest fragment, which contains the complete chloroplast DNA insert, was purified by gel electrophoresis and subsequent electroelution. The purified fragment was digested with *Alu* I, ligated with *Eco*RI linker, and used for the construction of recombinant clones with M13 mp8. After partial sequence data were obtained, the restriction sites of *Hgi*AI and *Fnu*4HI were detected by computer search. The *Hgi*AI and *Fnu*4HI subclones were generated by digestion with the respective enzyme, treatment with nuclease S1, and cloning into the *Sma* I site of M13 mp19.

Recovery of DNA fragments from agarose gel by electroelution was performed according to Maniatis *et al.* (14). Restriction enzymes, nuclease S1, and bacteriophage T4 DNA ligase were obtained from New England Biolabs, and used according to the supplier's instructions.

DNA Sequence Determination and Analyses. Both the Maxam–Gilbert chemical approach (15) and the dideoxy chain-termination technique (16) were used for sequence determination. For the Maxam–Gilbert method, electropho-

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: D-loop, displacement loop; ORF, open reading frame; ARS sequence, autonomously replicating sequence in yeast; ARC sequence, autonomously replicating sequence in *Chlamydomonas*; ddCTP, 2′,3′-dideoxycytidine 5′-triphosphate; kb, kilobase(s).

retically purified DNA fragments were used. The 3′ ends were labeled by using the large fragment of DNA polymerase I (New England Biolabs). The 5′ ends were labeled by using T4 polynucleotide kinase (Bethesda Research Laboratories). When the dideoxy chain-termination technique was used, the insert of each subclone in M13 was sequenced from both directions by using a kit from New England Biolabs. The DNA sequence was analyzed by using the IBM PC Basica DNA sequence analysis version 1.03 and the Los Alamos Sequence Analysis System. For sequence comparison, the algorithm of W. J. Wilbur and D. J. Lipman (National Institutes of Health) was used with parameters set to detect less dense similarities.

Test for Origin Function. Whether recombinant clones SC3-1 and R-13 can effectively initiate DNA replication were tested in an *in vitro* DNA replication system consisting of partially purified algal protein extracts. The algal proteins were isolated from exponentially growing cells of *C. reinhardtii* CC125 harvested during the light period of the synchronous growing cycle. A thylakoid membrane fraction was isolated and purified by a flotation procedure (17), with minor modification. Phenylmethylsulfonyl fluoride (1 mM), benzamidine (1 mM), and 6-aminohexanoic acid (5 mM) were added to the lysis buffer and flotation buffer to inhibit proteases. A high salt extract of the membrane fraction was obtained according to the extraction procedure of Orozco *et al.* (18). The equivalent high salt extract isolated from spinach chloroplasts can specifically initiate transcription of chloroplast DNA (18). In addition to the high salt extract, a soluble protein fraction is essential for DNA replication. For the preparation of a soluble protein fraction, the disrupted algal cell suspension was centrifuged at 10,000 rpm in a Sorvall SS34 rotor for 7 min at 2°C. The supernatant was collected, treated with 2.5% (vol/vol) Triton X-100 at 0°C for 30 min, and mixed with ammonium sulfate to 20% saturation. The precipitate was discarded and the supernatant was precipitated again by the addition of solid ammonium sulfate to 45% saturation. Precipitates were collected by centrifugation, resuspended in a minimal volume of extraction buffer, dialyzed, and stored at −80°C until use. The standard reaction mixture, analogous to that described by Fuller *et al.* (19), used for DNA replication assay contained the following components: Hepes (40 mM, pH 7.6); ATP (2 mM); GTP, CTP, and UTP (each 500 μM); bovine serum albumin (50 μg/ml); creatine phosphate (21.6 mM); dATP, dGTP, dCTP, and dTTP (each 100 μM, with [α-^{32}P]dTTP added at 3400 cpm per pmol of total dNTP); magnesium acetate (11 mM); creatine kinase (100 μg/ml); supercoiled DNA (2 μg/ml); and PEG 6000 [5% (wt/vol)]. Mixtures were assembled at 0°C, and reactions were initiated by addition of 4 μg of protein extracted from thylakoid membrane (isolated from 7 × 10^6 cells) and 40 μg of soluble protein to each 25-μl reaction mixture. Incubation was at 30°C for various times as indicated. The reaction was stopped by addition of EDTA (50 mM final), and total nucleotide incorporation was measured by determining radioactivity after trichloroacetic acid precipitation. After proteins were removed from the reaction mixture by phenol extraction, the replication products were analyzed by restriction endonuclease cleavage, agarose gel electrophoresis, and autoradiography. To localize the actual site of replication initiation *in vitro*, replicative intermediates were analyzed. Intermediates of decreasing extents of replication were generated by using an increasing ratio of the chain terminator 2′,3′-dideoxycytidine 5′-triphosphate (ddCTP) to dCTP in a series of reactions.

RESULTS

Nucleotide Sequence. A partial map of restriction sites of the cloned *Bam*HI–*Cla* I fragment and the strategy used to sequence this region are shown in Fig. 1. The complete nucleotide sequence of the chloroplast DNA insert in clone SC3-1 is shown in Fig. 2. The sequence is characterized by the presence of regions that are very rich in A·T base pairs. The overall A+T content is 70%. There are several regions of potential stem–loop structures. Four potential large stem–loop structures (sequences underlined in Fig. 2) are shown in Fig. 3. Computer search also indicates the presence of numerous direct repeats and mirror symmetries (data not shown). The overall picture suggests that this region could form several very complicated secondary structures. Whether this structural feature is pertinent to the initiation of replication remains to be investigated. Within this overall A+T-rich region, 11 small G+C-rich clusters, each consisting of 9–12 nucleotides and each having a G+C content >66%, were detected. The approximate locations of those G+C-rich clusters are shown in Fig 4. The sequence is also characterized by the presence of a small ORF. The amino acid sequence predicted by this ORF is shown in Fig. 2. The predicted molecular mass of the protein is 15.4 kDa, provided that the protein has the NH_2-terminus shown in Fig. 2. Seven of the 11 G+C-rich clusters are located within this ORF, which has an overall A+T content of 62.5%. The ORF also includes one entire stem–loop structure and a portion of another stem–loop structure.

The ARC consensus elements I and II (AaAAtAgCTTTttt.c.At and A.RT.AcCAAGT; R = G or A; uppercase letters indicate highly conserved residues; lowercase letters, less conserved; and periods, not conserved) and the ARS consensus sequence (5′ WTTTATRTTTW; W = A or T) (12) have not been located within the sequenced region. The entire chloroplast DNA sequence in SC3-1 has been compared with 387 plant sequences, 487 invertebrate sequences, 300 organelle sequences, and 732 bacterial sequences stored in the Genbank data bank (National Biomedical Research Foundation, Washington, DC, July 1, 1985). The computer search shows no significant sequence homology between the SC3-1 insert and bacterial-type *ori* regions or the chloroplast DNA replication origin of *Euglena gracilis*. However, statistical analyses show that some regions of the SC3-1 insert share partial homology with several interesting DNA sequences. The homologous regions and the extents of homology are shown in Fig. 4. The sources of the DNA sequence are the following: a portion of the coding region of the methylase modification enzyme of the bacterium *Providencia stuartii* (21), a portion of the upstream sequence of the H4 histone gene of *Tetrahymena thermophila* (22), a portion of the downstream sequence of the telomeric repeat of

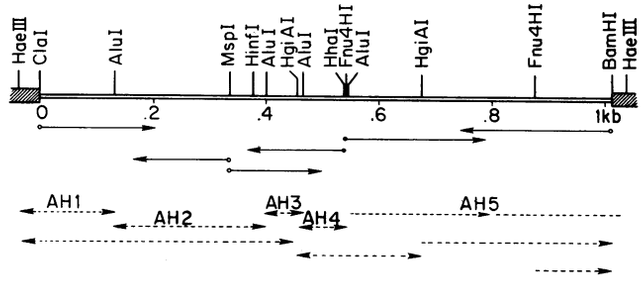

FIG. 1. Restriction map of the chloroplast DNA insert in SC3-1. Crosshatched bars represent pBR322 sequences. Strategy for determining the nucleotide sequence by the Maxam–Gilbert chemical approach (15) is represented by solid arrows. DNA was digested with *Cla* I, *Msp* I, *Hha* I, or *Bam*HI and end-labeled at its 5′ termini (●) or 3′ termini (○). Arrows indicate the direction and extent of sequence determined for each fragment. Strategy of sequencing by the dideoxy chain-termination technique (16) is represented by dashed arrows. Arrows indicate the direction of sequence determination for each fragment. The nomenclature of *Alu* I–*Hae* III subclones (AH1–AH5) is indicated on the respective dashed arrows.

```
   -190        -180        -170        -160        -150
ATC GATAGTATTG TTCATTGTAT AAAGTGTACG TACCCGTTAA GGGTACGTAC

-140        -130        -120        -110        -100        -90
ACTTTAATGC  AAGATAAACA  AAAATCAATA  CATATTACTA  GTTACTAGTA  TAAAGTACAA

-80         -70         -60         -50         -40         -30
TTGATTTCTG  TGTATTTGTA  GCTTTTAAAT  TAAATTTTTA  ATTAACTGTT  ACATAAAAAT

-20         -10                             15                  30
TTAAAATTAT  AAATAAAAAC  ATG TTA AGT CCA AAA AGA ACA AAA TTC CGT AAA CCA
                        Met Leu Ser Pro Lys Arg Thr Lys Phe Arg Lys Pro
        45                  60                  75                  90
CAC CGT GGT CAT TTA AGA GGA AAA GCA ACA CGT GGT AAT AAA ATT GTA TTT GGT
His Arg Gly His Leu Arg Gly Lys Ala Thr Arg Gly Asn Lys Ile Val Phe Gly

                105                 120                 135
GAT TTT GCA TTA CAA GCA CAA GAA CCT TGT TGG ATT ACA TCA CGT CAA ATT GAA
Asp Phe Ala Leu Gln Ala Gln Glu Pro Cys Trp Ile Thr Ser Arg Gln Ile Glu

    150                 165                 180                 195
GCC GGA CGT CGT GTT TTA ACA CGT TAT GTT CGT CGT GGT GGT AAA TTA TGG ATT
Ala Gly Arg Arg Val Leu Thr Arg Tyr Val Arg Arg Gly Gly Lys Leu Trp Ile

            210                 225                 240
CGT ATT TTC CCA GAT AAA GCT GTT ACT ATG CGT CCT GCT GGT ACT CGT ATG GGT
Arg Ile Phe Pro Asp Lys Ala Val Thr Met Arg Pro Ala Gly Thr Arg Met Gly

255                 270                 285                 300
TCT GGT AAA GGT GCA CCT GAT TAT TGG GTA GCT GTT GTA CAT CCT GGT AAA ATT
Ser Gly Lys Gly Ala Pro Asp Tyr Trp Val Ala Val Val His Pro Gly Lys Ile

    315                 330                 345                 360
TTA TAT GAA ATG CAA GGT GTA TCT GAA ACA ATT GCT AGA CAA GCA ATG CGC ATT
Leu Tyr Glu Met Gln Gly Val Ser Glu Thr Ile Ala Arg Gln Ala Met Arg Ile

            375                 390                 405
GCA GCT TAT AAA ATG CCA GTA AAA ACA AAA TTT TTA ACA AAA ACA GTG TA
Ala Ala Tyr Lys Met Pro Val Lys Thr Lys Phe Leu Thr Lys Thr Val

        420         430         440         450         460         470
ATTATTGTTA  TTAAAAATGT  TGTTTAGAAA  GAATTAATGA  TTTAACTTAC  TTAAAAAGCA

        480         490         500         510         520         530
TAATCTCAAA  TTAGAGCACA  AGTATAATTT  AAAAAATATT  TAAGAAAATT  AAGAGCATAA

        540         550         560         570         580         590
GTATTGTTTC  GCTTTGGCTC  AAAAGCCAAT  ACTAAAGTAT  ATATTACTTT  TTGTAAGTTT

        600         610         620         630         640         650
TTACTTACTC  GGTTTGTACC  AGGCAACCCT  ATAAATATAG  TAAAATGGAA  TTAAACTAGA

        660         670         680         690         700         710
TATATCTCTT  TAAGAAAGAT  TTTCTCATCA  AGGCTGCCCT  TTAACTTTAA  CCTAGAATGA

        720         730         740         750         760         770
CTAAAAGGAG  TAAGCAAATA  CCGAGAAATT  TATTTTTTCA  CTTAATGAAA  AAATAAATTT

        780         790         800         810         820         830
TATCTCTTTC  TCTTTTAAGC  ATATAAATAT  GAAGGTAAGT  AAACTCTACT  AGGGAAAAGC

        840         850         860
ATAGTGTTGA  AGGATATACT  TTCTTGGGAT  CC
```

FIG. 2. Nucleotide sequence of the 1.05-kb chloroplast DNA segment cloned in SC3-1. Sequences that can form stable stem–loop structures are underlined. Amino acid sequence of the ORF is shown under the corresponding DNA sequence. The −10 and −35 consensus region of each putative prokaryotic-type promoter is shown by an arrow above the corresponding DNA sequence, pointing in the direction of transcription.

Trypanosoma brucei (23), and two short regions of an intercistronic sequence of an octopine Ti plasmid, pTi15955 (24). All homologies are significantly higher than random occurrence. Whether these homologies suggest any functional role of these regions is not known.

Structural Features of the Sequenced Region. The sequence analysis presented here, in conjunction with some functional mapping conducted previously, allowed us to compose an overall diagram (Fig. 4). The D-loop position was mapped by electron microscopy (10, 11). The ORF was located by computer search. The D-loop region almost completely overlaps the ORF. Computer search identified several regions that approximate the prokaryotic promoter in the 5′ flanking region of the ORF. However, the promoter function of this region could not be confirmed by cloning experiments. By use of the promoter-cloning, *galK* expression plasmids pKO1 and pKOTWI, two putative prokaryotic promoters arranged back to back in the chloroplast DNA insert of subclone AH4 (Fig. 1) were detected (25). Putative −10 and −35 consensus regions of these promoters are shown in Fig. 2. Both promoters were mapped within the ORF and one promoter was located entirely within the stem region of a stem–loop structure (Fig. 4).

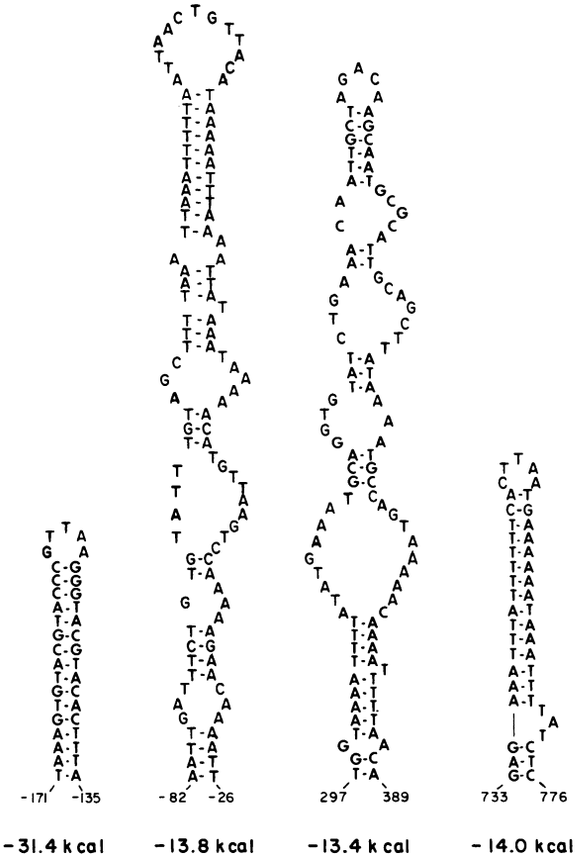

FIG. 3. Potential stem–loop structures. The Gibbs energy of each structure at 25°C was estimated by using the formula of Tinoco *et al.* (20).

Test for Origin Function. In the absence of an effective chloroplast transformation method, *in vitro* DNA replication system using algal proteins has become the method of choice to test the origin functions of clones R-13 and SC3-1. Details of this *in vitro* system will be published later. Some characteristics of the system are shown in Table 1. The incorporation of labeled dNTP into acid-insoluble material is ATP-dependent. The system is very sensitive to ethidium bromide inhibition. In *Chlamydomonas*, chloroplast replication has been reported to be sensitive to ethidium bromide (26). Under phototrophic conditions, nuclear DNA synthesis is not inhibited by ethidium bromide (27).

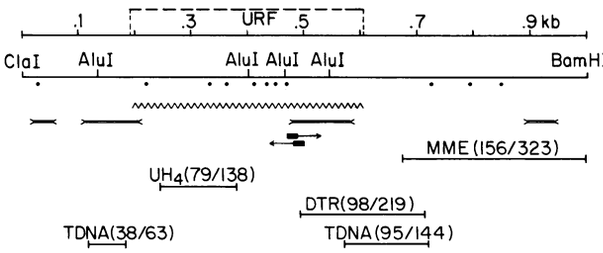

FIG. 4. Schematic representation of the sequence organization of the origin region. ∿, D-loop region; ●, G+C-rich cluster; ═, stem–loop structure; ▬►, promoter and its direction of transcription. MME (156/323) indicates that 156 nucleotides of this 323-nucleotide region are homologous to the coding region of the methylase modification enzyme of the bacteria *Providencia stuartii* (21). UH_4 represents the upstream sequence of the H4 histone gene of *Tetrahymena thermophila* (22). DTR represents the downstream sequence of the telomeric repeat of *Trypanosoma brucei* (23). TDNA represents an intercistronic sequence of the octopine Ti plasmid pTi15955 (24).

Table 1. Requirements for DNA synthesis *in vitro*

Condition	DNA synthesis activity, %
Complete	100
Without DNA	<5
Without ATP, creatine phosphate, creatine kinase	25
Without rCTP, rGTP, and rUTP	73
Without dCTP	56
Without Mg^{2+}	31
Without soluble proteins	18
Without membrane extract	57
With ethidium bromide	
2 μg/ml	24
10 μg/ml	0

DNA synthesis activity was measured for 60 min at 30°C in the standard reaction assay as described in *Materials and Methods*; 100% equals 9.6 pmol of dNMP incorporated.

Clone R-13 contains ARC1 and ARS08 in addition to the D-loop region cloned in SC3-1. On the chloroplast genome, this is the only region in which three replication-related sequences are closely clustered. The relative positions of these functional regions and some useful restriction sites are shown in Fig. 5. The five conserved regions shared between *C. reinhardtii* and *Chlamydomonas* sp. strain WXM were mapped previously by electron microscopic heteroduplex analysis (11). The length and location of these regions are indicated by bars A–E in Fig. 5. Vallet and Rochaix (13) mapped the ARC1 element within the 0.35-kb conserved region B and indicated that the ARS08 sequence could overlap with conserved region E (Fig. 5). ARS08 is A+T-rich and can replicate autonomously in yeast. Fig. 6 shows the kinetics of the DNA synthesis activity in this *in vitro* system. Data were collected from several independent experiments. The incorporation of labeled dNTP into acid-insoluble material depended totally on addition of exogenous DNA templates. Peak incorporation was frequently detected after 50 min of incubation at 30°C, using supercoiled DNA of R-13 as template. In this system, maximal incorporation was 9 pmol of dNTP in the 25-μl reaction mixture containing 0.05 μg of template DNA, which is equivalent to 0.07 pmol of R-13. Supercoiled SC3-1 in general showed slightly lower replicating activity than did R-13. In a typical assay, 7 pmol of dNTP was incorporated into 0.15 pmol of SC3-1 template, while in a parallel assay, less than 3.7 pmol of dNTP was incorporated into 0.13 pmol of supercoiled pBR325 (Fig. 6). Autoradiographic analyses of the DNAs replicated *in vitro* when pBR325, SC3-1, and R-13 were used as template (Fig. 7, lanes 1–3) showed more incorporation of [α-^{32}P]dTTP into SC3-1 and R-13. The DNAs had not been treated with proteinase K before phenol extraction and gel electrophoresis. The relatively low intensity of the 1.65-kb band and the

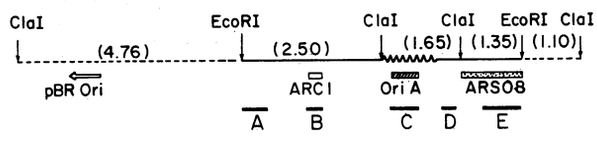

FIG. 5. Physical and functional map of clone R-13. Restriction sites of *Eco*RI and *Cla* I are shown. The lengths (in kb) of each restriction fragment generated by *Eco*RI/*Cla* I double digestion are indicated in parentheses. Broken line represents pBR325 sequence. Zigzag line represents sequence cloned in SC3-1. The five regions conserved between *C. reinhardtii* and *Chlamydomonas* sp. strain WXM mapped by Wang *et al.* (11) are indicated by bars A–E. ARC1 and ARS08 regions were mapped by Vallet and Rochaix (13). The replication origins of pBR325 and chloroplast DNA (OriA) are also indicated.

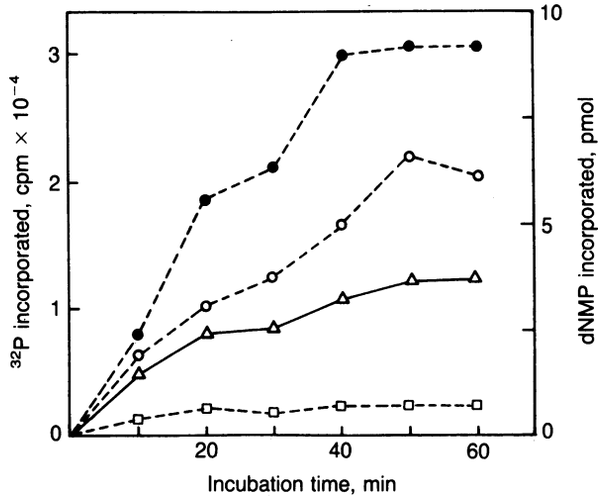

FIG. 6. Kinetics of DNA synthesis *in vitro* when R-13 (●), SC3-1 (○), or pBR325 (△) or no exogenous DNA (□) was used as template (0.05 μg per 25-μl reaction mixture).

large amount of labeled DNA retained within the loading well suggested that the 1.65-kb restriction fragment has strong protein-binding properties. In subsequent autoradiographic analyses (lanes 4–6 of Fig. 7), extensive proteinase K treatment was used prior to phenol extraction and gel electrophoresis.

Both kinetic study and autoradiographic analysis indicate that R-13 and SC3-1 are effective templates in this *in vitro* system. The higher incorporation of labeled dNTP into R-13 could either be due to the presence of multiple initiation sites or due to more effective usage of one initiation site. The initiation site of DNA replication in this *in vitro* system was investigated by use of ddCTP, which blocks chain elongation

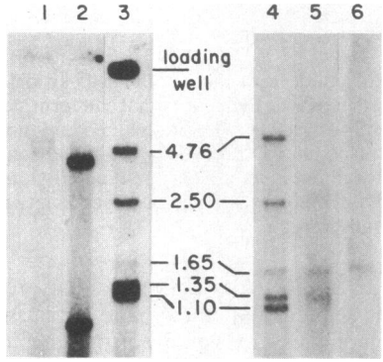

FIG. 7. Autoradiograms show the distribution of incorporated radioactivity among various restriction fragments after *in vitro* DNA synthesis. Lane 1: pBR325 was used as the exogenous template, and the labeled DNA was digested with *Eco*RI and *Cla* I. The digestion was not complete; the top band corresponds to the full-length linear molecule. Lane 2: SC3-1 was used as template, and the DNA was digested with *Bam*HI and *Cla* I. Lane 3: R-13 was used as template. The DNA was digested with *Cla* I and *Eco*RI. Each reaction mixture contained 0.05 μg of DNA template and the reaction was stopped after 40 min. Free proteins were removed by phenol extraction before restriction enzyme digestion and electrophoretic separation of restriction fragments. The low intensity of the 1.65-kb band in lane 3 is probably due to its high affinity for protein, as discussed in the test. Lanes 4–6: labeling of different restriction fragments of R-13 in the *in vitro* system, in the presence of 0, 50, or 100 μM ddCTP, respectively. Reaction mixtures contained all ingredients described in *Materials and Methods*. The reaction was stopped after 20 min. Proteins were removed by extensive proteinase K treatment (0.1 mg/ml for 1 hr) and repeated phenol/chloroform (1:1, vol/vol) and ether extractions before restriction enzyme digestion. Size of each labeled DNA fragment is indicated in kb.

at an early phase of replication (28). Replicative intermediates of decreasing extents of replication were generated by increasing the ddCTP/dCTP ratio in a series of reactions. The replication products were analyzed by restriction enzyme digestion, agarose gel electrophoresis, and autoradiography. When R-13 was used as the template, labeling of the 1.65-kb fragment generated by *Eco*RI and *Cla* I persisted up to the highest concentration of ddCTP (Fig. 7, lanes 4–6). As shown in Fig. 5, OriA and the D-loop region are located within this fragment; thus, they are the first to be replicated. ARC1 is located in the 2.50-kb fragment, and ARS08 is located in the 1.35-kb fragment. Labeling of both fragments decreases with increasing concentration of ddCTP. Since the kinetic data show that R-13 is a more effective template than SC3-1, ARC, ARS, or other adjacent sequences might play some important auxiliary role(s).

DISCUSSION

Sequence analyses, *in vitro* replication assay, and previous electron microscopic studies demonstrate that we have isolated one chloroplast DNA replication origin, which we call OriA. Structural and functional analyses of the cloned origin are essential for further study of the regulatory mechanism. The presence of extremely A+T-rich regions and the potential to form large stem–loop structures are prominent features of many replication origins (29). An interesting characteristic of this replication origin is the clustering of an ORF, a large stem–loop structure, and two back-to-back promoters within a region 408 base pairs long. The location and arrangement of these promoters suggest a possible involvement in the initiation of replication. RNA priming of DNA replication is a mechanism of general significance. The details vary in each replication system. Transcription initiated from the origin promoter might either provide primer for the inception of the DNA strand or change the configuration of the origin region transiently to favor DNA–protein interaction. In this chloroplast replication origin, both promoters are located within the stem region of a stem–loop structure that is situated within the ORF. This spatial arrangement is potentially capable of accommodating intricate interactions. In both *Escherichia coli* and bacteriophage λ, transcriptional events in the *cis* configuration with the replication origins are critical in regulating the frequency of initiation of DNA replication (30, 31).

In the crude *in vitro* replication system we used, the extent of repair synthesis could not be estimated. Our argument for the *de novo* initiation of DNA replication in this *in vitro* system is based upon the kinetic studies, the autoradiographic analyses of newly replicated DNAs, and the mode of incorporation of radioactive dTTP in the presence of ddCTP. The kinetic studies indicated that the extent of incorporation was template-specific. In the presence of increasing concentration of ddCTP, the preferential incorporation of radioactive dTTP into a short restriction fragment containing OriA again suggested that a significant amount of incorporation was due to replication initiated at a specific site.

We are grateful to Dr. Samuel H. Wilson and Mr. Jamie Cuticchia for computer analysis, to Mr. Gregg W. Silk for critical reading of the manuscript, and to Mr. X. M. Wang and Mr. Pat McAllister for contributing some preliminary sequence data. This work was supported by National Science Foundation Grant PCM 81-10118, United States Department of Agriculture Grant 85-CRCR-1-1761, and a grant from the Center of Marine Biotechnology (COMB) of the University of Maryland. C.H.C. is a recipient of a COMB fellowship.

1. Sager, R. (1972) *Cytoplasmic Genes and Organelles* (Academic, New York).
2. Gillham, N. W. (1978) *Organelle Heredity* (Raven, New York).
3. Birky, C. W., Jr. (1984) *Science* **222**, 468–475.
4. Kolodner, R. D. & Tewari, K. K. (1975) *J. Biol. Chem.* **250**, 4888–4895.
5. Kolodner, R. D. & Tewari, K. K. (1975) *Nature (London)* **256**, 708–711.
6. Koller, B. & Delius, H. (1982) *EMBO J.* **1**, 995–998.
7. Ravel-Chapuis, P., Heizman, P. & Nigon, V. (1982) *Nature (London)* **300**, 78–81.
8. Schlunegger, B., Fasnacht, M., Stutz, E., Koller, B. & Delius, H. (1983) *Biochim. Biophys. Acta* **739**, 114–121.
9. Schlunegger, B. & Stutz, E. (1984) *Curr. Genet.* **8**, 629–634.
10. Waddell, J., Wang, W. & Wu, M. (1984) *Nucleic Acids Res.* **12**, 3843–3856.
11. Wang, W., Chang, C. H., Waddell, J. & Wu, M. (1984) *Nucleic Acids Res.* **12**, 3857–3872.
12. Rochaix, J.-D., van Dillewijn, J. & Rahire, M. (1984) *Cell* **36**, 925–931.
13. Vallet, J.-M. & Rochaix, J.-D. (1985) *Curr. Genet.* **9**, 321–324.
14. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
15. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
16. Sanger, F., Nicklen, S. & Coulson, A. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
17. Chua, N. H., Blobel, G., Siekevitz, P. & Palade, G. E. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 1554–1558.
18. Orozco, E. M., Jr., Mullet, J. E. & Chua, N. H. (1985) *Nucleic Acids Res.* **13**, 1283–1302.
19. Fuller, R. S., Kaguni, J. M. & Kornberg, A. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 7370–7374.
20. Tinico, I., Jr., Borer, P. N., Dangler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nature (London) New Biol.* **246**, 40–41.
21. Walder, R. Y., Walder, J. A. & Donelson, J. E. (1984) *J. Biol. Chem.* **259**, 8015–8026.
22. Bannon, G. A., Bowen, J. K., Yao, M. C. & Gorovsky, M. A. (1984) *Nucleic Acids Res.* **12**, 1961–1975.
23. Blackburn, E. H. & Challoner, P. B. (1984) *Cell* **36**, 447–457.
24. Barker, R. F., Idler, K. B., Thompson, D. V. & Kemp, J. D. (1983) *Plant Mol. Biol.* **2**, 335–350.
25. Wu, M., Kong, X. F. & Kung, S. D. (1986) *Curr. Genet.*, in press.
26. Fletchtner, V. R. & Sager, K. (1973) *Nature (London) New Biol.* **241**, 277–279.
27. Keller, S. J. & Ho, C. (1981) *Int. Rev. Cytol.* **69**, 157–190.
28. Conrad, S. E., Wold, M. & Campbell, J. L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 736–740.
29. Kornberg, A. (1982) *Supplement to DNA Replication* (Freeman, San Francisco), pp. 145–152.
30. Lother, H., Kolling, R., Kucherer, C. & Schauzu, M. (1985) *EMBO J.* **4**, 555–560.
31. Tanaka, M. & Hiraga, S. (1985) *Mol. Gen. Genet.* **200**, 21–26.