Determination of binding sites of replication protein A (RPA) on the promoter of glycogen phosphorylase 2 gene in *Dictyostelium discoideum*

Bekir Çöl[1,2*] and Charles L. Rutherford[2]

[1] Biology Department, Molecular Biology Program, Faculty of Art and Sciences, Mugla University, Muğla, Turkey,
[2] Biology Department, Molecular and Cellular Biology Section, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0406, USA

ABSTRACT

Replication protein A (RPA) has multifunctional roles in the cell including DNA replication, DNA repair, recombination and regulation of transcription. RPA has been identified in numerous organisms having a highly conserved protein structure, composed of three subunits of around 70, 32, and 14 kDa. The cellular slime mold, *Dictyostelium discoideum*, is a social amoeba used as a model organism in studies including the ones aimed at understanding cell differentiation and cellular pathways involved in human diseases and this organism has a homolog of RPA (DdRPA). Glycogen phosphorylase 2 (encoded by *gp2*) is involved in cell differentiation and DdRPA has been suggested to bind a part of the *gp2* promoter as shown by gel shift assays. In this study, in vitro footprint analysis has been used to identify the binding sites of the DdRPA protein, purified from amoeba and slug stages of the organism, on the *gp2* promoter. The results indicate that DdRPA binds to the nucleotides in the C box region, the TAG box region, TA box-1 (-608bp) and upstream regions (-636bp). Thus, DdRPA behaves like a general DNA binding protein under these conditions, binding to several regions on the *gp2* promoter. However, the TAG box region has been identified as the binding region for the DdRPA from amoeba cells, but not for the DdRPA from slug cells. This suggests that DdRPA could possibly be involved in the regulation of *gp2* gene expression during cell differentiation in *Dictyostelium discoideum*.

Key Words: DNase I footprint analysis, *Dictyostelium discoideum*, *gp2* promoter, glycogen phosphorylase 2, replication protein A binding sites

INTRODUCTION

Development to multicellularity in *Dictyostelium discoideum* is achieved by the aggregation of formerly individual identical cells. Cells within the aggregates initiate alternate pathways of differentiation that results in the appearance of two distinct cell types and morphogenesis (Gross, 1994). The cell differentiation aspect of development is achieved mainly by selective gene expression, a process controlled chiefly at the level of transcription.

The cellular slime mold, *Dictyostelium discoideum*, is an excellent eukaryotic model system for investigating the events involved in cell differentiation, eukaryotic development, and aging. If a homogeneous population of amoebae is starved, it enters into a developmental process resulting in a morphologically distinct structure consisting of two different cell types. When there is enough food, the slime mold exists as single-celled vegetative amoebae, which feed on the bacteria of decaying organic materials. The amoebae cells are genetically, biochemically and morphologically identical (Loomis, 1982). Once food becomes scarce, the cells are triggered to differentiate. Eight hours after the onset of starvation, cellular aggregation is seen. Each amoebae secretes adenosine 3, 5 -monophosphate (cAMP) that chemotaxically lead the cells to aggregate. Aggregates of approximately 100,000 cells form by 10 h. This multicellular mound goes on to form the final fruiting body. The mound has two distinct cell types; prespore cells comprise about 80% of the mound, while prestalk cells account for the remaining 20% of the cells. After the slug forms, culmination occurs and cell sorting continues, with prestalk and prespore cells exchanging places. Differentiation into mature spore and stalk cells proceeds. While stalk cells are programmed to die, spore cells can remain viable despite unsuitable conditions, such as starvation and changes in temperature. As soon as the conditions become favorable, the spore cells germinate into single-celled amoebae (Loomis, 1982). Glycogen phosphorylase 2 (*gp-2*) is a developmentally regulated gene that is first expressed during the transition from growth to differentiation in *Dictyostelium discoideum* (Rutherford et al. 1997). The gene product of *gp-2* catalyzes the degradation of glycogen into glucose monomers. These monomers are then used to produce the structural end products of terminal differentiation. The promoter for the *gp2* gene has high AT rich content, which is the characteristic of non-coding regions in the organism. In the promoter region, there are seven repeated DNA regions interrupting the long stretches of A and Ts. These regions are referred to as "C boxes" (repeated twice, 5'-ACCCACT-3'), "TAG boxes" (repeated twice, 5'-TAAAAATGGA -3') and TA boxes (repeated three times, 5'-TAATTATAA-3'). The locations of the boxes

---

on the *gp2* promoter are -328 to -334; -365 to -371 for the C boxes, -437 to -461 separated by five nucleotides for the TAG boxes and on the further upstream for the TA boxes. The transcription start site is a nucleotide T at -287 upstream of the AUG methionine codon (Rutherford et al. 1997).

Using electrophoretic mobility shift assays (EMSA), one of the proteins that binds to the *gp2* promoter has been described as Replication Protein A (DdRPA) (Wen, 1998). DdRPA produces a 0.40 Rf EMSA band in undifferentiated cells, and a 0.32 Rf band in differentiated cells. The 0.40 Rf band corresponds to a form of DdRPA consisting of 62 kDa, 35 kDa and 18 kDa subunits, whereas the 0.32 Rf corresponds to a form of the protein comprised of 84 kDa, 35 kDa, and 18 kDa subunits (Wen, 1998). Southwestern analysis revealed that the biggest subunits are the DNA binding domains (Wen, 1998). The gene encoding the large subunit (84 kDa) was cloned and characterized. Northern blot results indicate that DdRPA expression level was constant throughout development and the size of mRNA of all stages is the same. This size corresponds to the 84 kDa form indicating that the 62 and 84 kDa are the products of the same gene. The biggest subunit of DdRPA is transcribed as 84 kDa and then undergoes posttranslational modifications to become 62 kDa form (Wen, 1998). What nucleotides DdRPA purified from undifferentiated and differentiated cells binds has been determined in this study via DNase I footprinting analysis.

MATERIALS AND METHODS

Cell culture and harvest
Dictyostelium discoideum strain AX-3K was grown in liquid HL-5 medium (60 mM glucose, 1% Oxoid peptone, 0.5% Oxoid yeast extract, 2mM Na_2HPO_4 and 3 mM KH_2PO_4 at 21°C on a rotary shaker at 130 rpm until $5x10^6$ cells were generated. Vegetative amoebae cells grown to log phase were placed in differentiation buffer (20 mM N-morpholinoethanesulfonic acid (pH 6.8), 0.2 mM $CaCl_2$, 2 mM $MgSO_4$).

For protein purification, thirty 2 liter-flasks containing 500 ml HL5 with 10^7 cells/ml were harvested using a flow through centrifuge and the pellet was washed with 1 liter of differentiation buffer. The cells were then transferred to flasks containing 500 ml MES and placed on a rotary shaker (New Brunswick model 626) for 6hr or O/N, at 130 rpm at 21°C.

DdRPA extract preparation
Starved cells (6h or O/N) were harvested, resuspended in 1 liter of milli-Q water, divided into two and centrifuged for 5 min at 1,300 x g. The cells were then ruptured by adding 5 volumes of Buffer X (25 mM Tris (pH 7.9), 0.5 mM EDTA, 150 mM EDTA, 150 mM sucrose, 2% Nonidet P40) and incubated on ice for 15 min. The lysate was centrifuged and the resulting extract was stored at - 80°C.

DEAE Sephacel Batch Assay
The cell extract was thawed at 37°C. Phenylmethylsulfunyl fluoride (PMSF) was added to the concentration of 10 μg/ml, the mixture was stirred at 4°C for 15 min and centrifuged at 20,000 x g for 30 min. The supernatant was passed through glass wool and combined with DEAE Sephacel resin (Pharmacia, 1 g resin/ 6 ml packed cells) in a 1 liter bottle. After shaking 60 min at room temperature at 150 rpm, the suspension was poured into a column and washed with 20mM Tris buffer (pH 7.9). The resulting fractions were assayed for DNA binding activity using a gel shift assay and the 3' C probe.

DNA Affinity chromatography
Preparation and coupling the DNA target sequence to CNBr Sepharose-4B was performed as previously described (4). Briefly, 25bp double stranded oligonucleotide containing the 3' C box and surrounding sequences (5'GATCCAGTGCAAACTC<u>ACCCACT</u>CACAAT3' (the 3' C probe) was used as DNA to isolate binding proteins on an affinity column. The oligonucleotide strands were annealed and concatemerized (using a non-promoter GATC sequence added to the 5' end) and coupled to CNBr Sepharose-4B (Kampang, 1995).

Electrophoretic mobility shift assay (EMSA)
The 3' C probe was labeled using a fill-in reaction with alpha-^{32}P dATP. The probe was separated from unincorporated nucleotides by using a Biospin 6 Column (Biorad). Reaction mixture, containing the labeled probe (30,000 cpm), 2 μl of extract and 7.5 μl reaction buffer (22 mM Tris-HCl, K_2HPO_4/KH_2PO_4, Imidazole, or MES containing 8.5% glycerol, 43 mM NaCl, 4.4 mM $MgCl_2$, 4.4 mM EDTA, 2.2 mM dithiothreitol, 4.4% Nonidet P40) was incubated at room temperature for 20 min. After adding 2 μl of

loading dye (250 mM Tris-HCl (pH 7.9) 50% glycerol, 4 mg/ml bromphenol blue, 4 mg/ml xylene cyanol), the gel shift assay was performed using 5-6% polyacrylamide gels in TBE buffer. The gel was dried and autoradiographed for 10-20 hours.

DNaseI footprinting

The footprint protocol as described in Current Protocols (Ausubel, 1995) was employed with minor modifications. Template was prepared as follows. Genomic DNA was isolated from amoebae cells (Ausubel, 1995). A Bluescript clone containing 453 bp of the promoter with all boxes (-726bp to -273bp) was prepared using Qiagen Plasmid Maxi Prep Kit. Appropriate primers were used with this template **to** produce the desired fragment using the polymerase chain reaction. The products were checked with agarose gel electrophoresis and purified with the Gene Clean System (Bio 101, Vista, CA). The yield was measured with the EtBr dot quantitative analysis (Ausubel, 1995).

Primers of N (AGAACAAATAAATTATTGATTGTGTT), 9001100 (GCTATTT TTTCACAGATA CC), 460660 (AAATCCATAAAGTCCAAAAT), TAG(TAAAAATGGATAAACTAAAAATGGA), TAG COMP (TCCATTTTTAGTTTATCCATTTTTA) were used. In order to ensure only one 5' end could be radioactively endlabeled, the 5' ends of the appropriate primers were first phosphorylated with cold rATP using T4 polynucleotide kinase. One kinased and one regular primer were used for PCR reaction. Success of phosphorylation was checked using the exonuclease assay.

Optimum PCR conditions for each primer set were determined. The amplification reaction contained; 30 pmol kinased primer, 20 pmol regular primer, 5 µl 10 X PCR buffer (1), 0.08 mM dNTP, 2.5 U Taq polymerase in a final volume of 50 µl. The following program was run using Perkin Elmer thermocycler; 2.30 min at 94°C; 0.30 min at 94°C, 0.30 min at 52°C, 1 min at 72°C, for 34 cycles, and 10 min at 72°C.

For the N-9001100K primer pair, all conditions were the same except 6.5 µl 10X PCR buffer was used and Tm of 48°C was applied. The 5' end of the purified PCR product (180-300 ng/total) was radioactively endlabeled as described (Ausubel, 1995). The specific activity was measured using a scintillation counter and was generally 100,000-200,000 CPM/µl.

The *in vitro* reaction contained a cocktail containing 2.5-4 ng PCR product (radiolabeled at one end), 2 µg poly dI:dC or poly dA:dT (Pharmacia) and Buffer B (10 mM bis-Tris-Cl, 2.5 mM $MgCl_2$, 1 mM $CaCl_2$, 0.1 mM EDTA, 200 mM KCl, 100 µg/ml BSA, 2 µg/ml herring sperm DNA; pH 7.0 (1). To the cocktail, purified DdRPA was added and allowed to incubate at room temperature for 20 min. The final volume was 200 µl. 0.01U DNaseI (Worthington) was added for partial digestion of the template. The reaction was stopped by adding 700 µl of stop solution (92% ethanol, 0.7 M NH_4SO_4, 5 mg yeast t-RNA), vortexing vigorously and putting into a dry ice/ethanol bath. Following the precipitation, the resuspension was centrifuged and the pellet was washed with 80% ethanol and resuspended in 6 µl formamide loading dye. 1 µl of this solution was used to measure the specific activity as determined by a scintillation counter. A 6% denaturing polyacrylamide gel was pre-run for 20-25 min at 2000 V, while the samples were denatured for 20-30 min at 90°C. An equal amount of radioactivity was loaded into each lane of the gel. Separation and autoradiography was performed as described (Ausubel, 1995). DNase I conditions were determined to be optimum at 5 min for the templates.

Marker generation

The Maxam Gilbert reaction was used to generate a G ladder from the footprinting probe (Maxam and Gilbert, 1977).

RESULTS AND DISCUSSION

DdRPA purification and DNaseI digestion conditions

The immediate aim was to ascertain the DNA binding ability of the purified replication protein A (DdRPA) and to see if there was a distinction between the binding affinities of undifferentiated and differentiated forms of the protein. DdRPA was purified as described (Wen, 1998) and the binding activity of the concentrated protein was checked with gel shift assay using the 3'C probe (Figure 1). This method was thought to be more relevant, as it provided information about the active protein present in the preparations.

DNase I digestion conditions were standardized for all templates such that there would be one cleavage event per strand (data not shown). This results in the formation of a ladder where each band is separated by one nucleotide. This has been done mostly on naked DNA, but in some cases proteins were added to observe the effect of protein on the overall efficiency of DNase I digestion (Figure 2a). The best partial digestion

conditions were chosen for footprint experiments; specifically, 5 min for the two templates (460660-9001100 PCR products and its cloned counterpart).
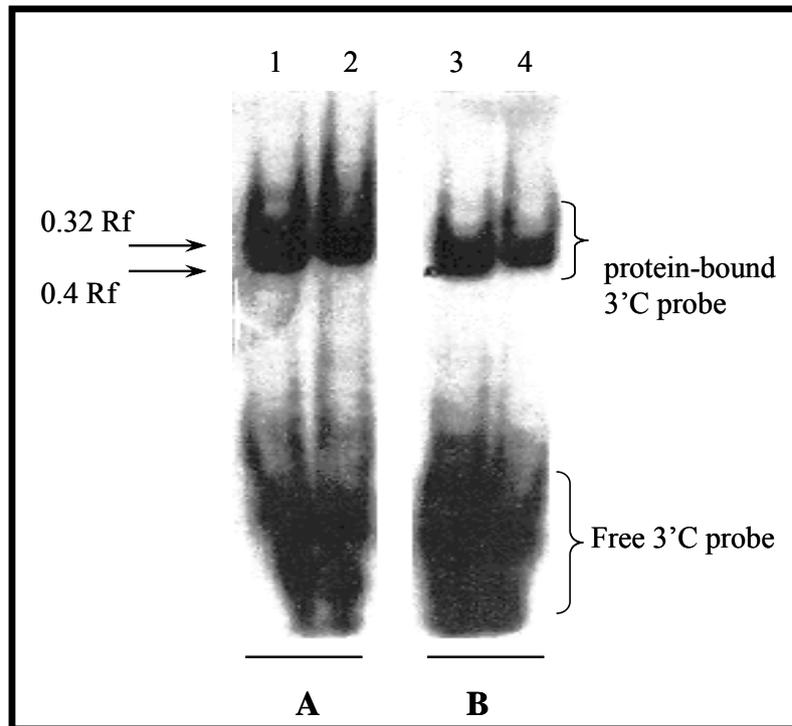
Figure 1. Gel shift assay to test the activity of purified DdRPA. DdRPA was prurified from undifferentiated (amoebae) and differentiated cells (slug) and the 3'C probe was used in the gel shift assay. Lanes 1 and 3 correspond to undifferentiated DdRPA(0.4 Rf) Lanes 2 and 4 correspond to differentiated DdRPA (0.32Rf). In A, the probe used was 1/1 and in B the probe was diluted 1/10.

Footprints of Replication Protein A (DdRPA)
The results indicate that DdRPA can bind to some nucleotides in the C box region, the TAG box region, TA box-1 (-608bp) and upstream regions (-636bp) (Figures 2 and 3). In terms of the difference between footprints generated by DdRPA from undifferentiated and differentiated cells, undifferentiated DdRPA binds to the TAG box region, whereas differentiated DdRPA did not.

All footprint experiments involving DdRPA produced a footprint covering the region (CAGTGC at −347 bp relative to the translation start site) between the two C boxes (Figure 2b). This may imply that DdRPA prefers to bind to this region, rather than directly to the C boxes. The explanation for this may be that these nucleotides, not the 3'C box, were responsible for binding the protein during purification using affinity chromatography and the 3'C probe. Nevertheless, later experiments with more active protein showed that DdRPA binding could expand through the 5'C box region.

Footprint experiments were done using different templates (e.g., 460660K-9001100, N-9001100K, and 9001100K-460660) in an effort to assess the effect of template length on protein binding and to view the regions of interest on both strands of the DNA template. Also, two different nonspecific competitors (poly dI:dC and poly dA:dT) were used to interpret the nature of the binding interaction. For example, if the protein bound a specific site when poly dI:dC is used as competitor, but the footprint can no longer be detected when poly dA:dT is used, this would indicate that protein binding is facilitated by interactions with A and/or T nucleotides in the binding site.

C Box regions footprints
In Figure 2A, using increasing times of DNase I digestion (5 and 10 min), clear but small footprints are shown covering the C box region. The regions of protection are especially evident between the C boxes.

Another footprint can be seen upstream of the 5'C box, covering stretches of T's (-413 bp). In this experiment, only DdRPA from undifferentiated cells was used with poly dI:dC as nonspecific competitor.
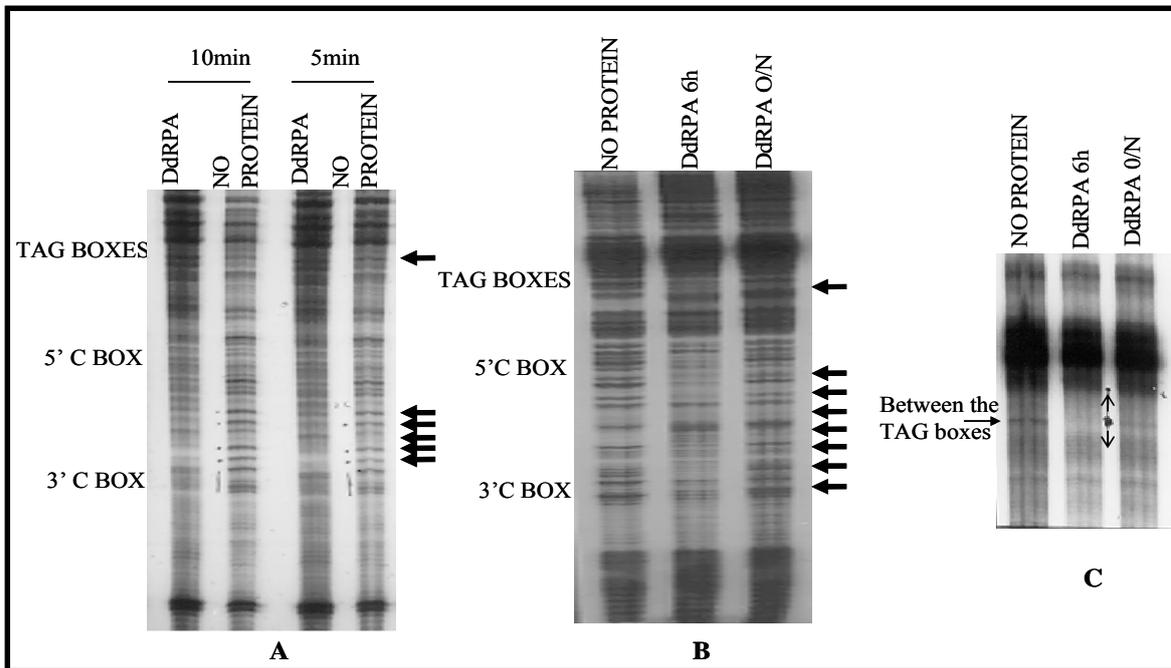
Figure 2. DNase I footprint assay to map the binding sites of RPA on C and TAG boxes. **A.** DdRPA from undifferentiated cells (6h development) was assayed. 460660K-9001100 was used as the template, poly dI:dC was used as the nonspecific competitor. The C and TAG boxes are shown, whose locations were determined by the Maxam Gilbert G reaction using the same template (data not shown). Arrows indicate the regions of protection. DNase I digestion time (5 and 10 min) is indicated on top. **B.** The difference between the footprints generated by DdRPAs purified from undifferentiated or differentiated cells (O/N, overnight development) is shown by arrows. The missing bands (footprints) are evident spanning the regions between the 3' and 5' C boxes. **C.** Footprint on TAG boxes. A small footprint is indicated by an arrow at the region between the two TAG boxes. N-9001100K was used as template, poly dI:dC was used as nonspecific competitor.

Using DdRPA from differentiated cells, footprints were detected on essentially the same regions. Experiments performed using poly dA:dT and DdRPA from differentiated cells and dI:dC appeared to give the similar result (Figure 2B). The region containing and surrounding the C box region was protected with both forms of DdRPA. A footprint on the TAG box region was detected with DdRPA from undifferentiated cells (Figure 2b).

TAG Boxes footprint
Interestingly, the only footprint difference observed between DdRPA from differentiated and undifferentiated cells was on the TAG box region. Undifferentiated DdRPA generated a footprint that appeared to bind between the boxes, whereas the differentiated form did not (Figure 2c).
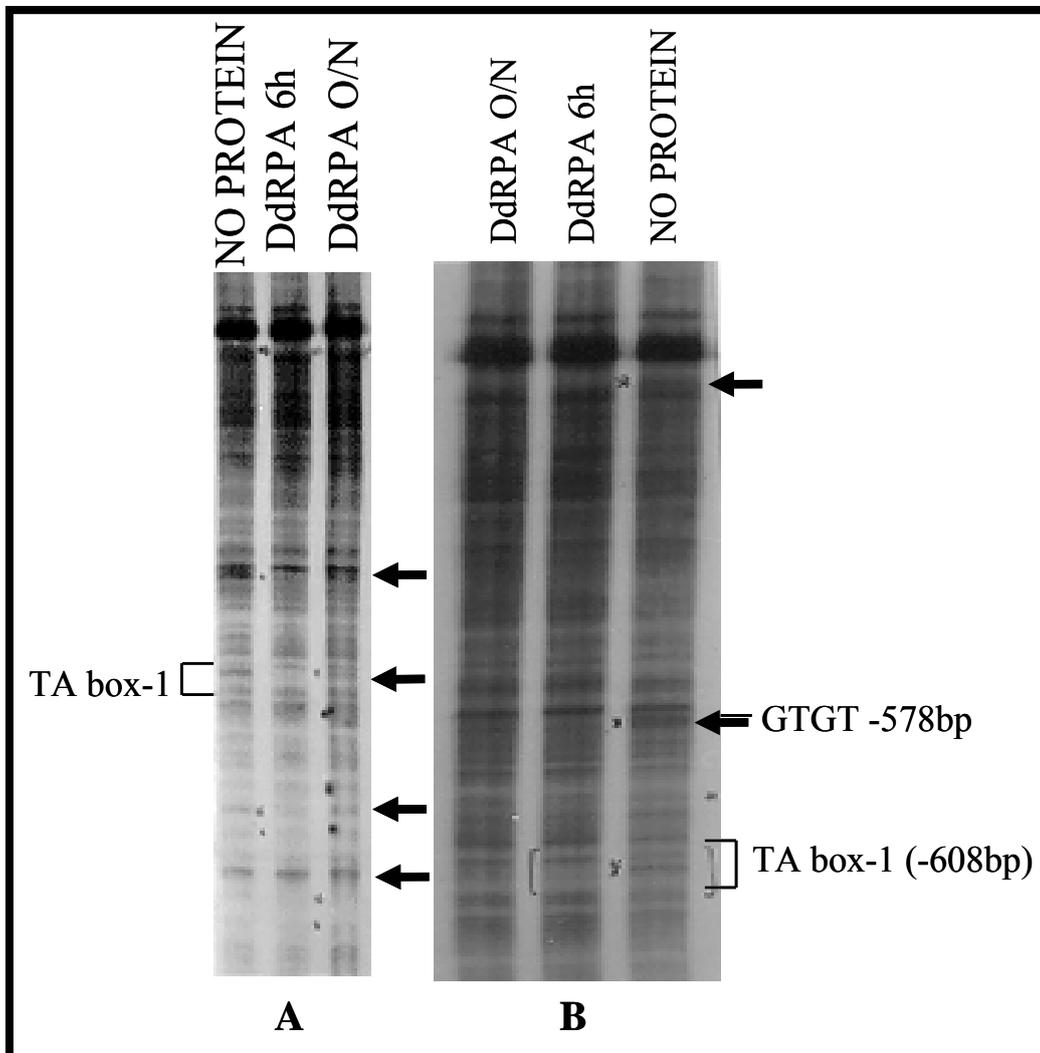
Other footprints with DdRPA
Using 9001100K-460660 as the template to target TA box-1 and upstream regions, several footprints were detected. The footprints covered TA box-1, the region between −577 and −572bp (TGTGT) and further upstream (-636bp and around). The result was the same whether poly dA:dT or poly dI:dC was used as competitor (Figure 3B and Figure 3A, respectively).

All of the data explained thus far is summarized in Table 1. In summary, under these conditions DdRPA binds several sites on the gp-2 promoter. However, it is possible that DdRPA may be involved in protein-protein interactions that could augment its specificity. This statement is based on the results involving crude nuclear extracts (Çöl B. et al., manuscript in preparation).

Table 1. Summary of DdRPA footprints

REGIONS OF FOOTPRINTS		UNDIFFRENTIATED REPLICATION PROTEIN A (uDdRPA)	DIFFERENTIATED REPLICATION PROTEIN A (dDdRPA)
C BOXES	**dA:dT**	+	+
	dI:dC	+	+
TAG BOXES	**dA:dT**	+	-
	dI:dC	+	-
TA BOX-1 -295bp and around-361	**dA:dT**	+	+
	dI:dC	+	+

Figure 3. DNase I footprint results using the 9001100K-460660 as template. Poly dA:dT was used as nonspecific competitor. DNase I digestion time was 5 minutes for each lane. Footprints are indicated by arrows. Poly dA:dT or poly dI:dC in A was used as the nonspecific competitor. DNase I digestion time was 5 minutes for each lane. TA box-1 and other sequences are shown. **A** and **B** used the same template but different competitors (poly dI:dC in A). A footprint on TA box1 is evident as shown.

DISCUSSION

Replication protein A from different organisms has been found to have functional roles in DNA replication, repair and recombination, as well as in the regulation of transcription (5, 11). These RPA's have 3 subunits, a feature that is evolutionarily conserved (Wold, 1997). Some of the subunits are thought to be involved in interactions with other proteins (Fanning et al. 2006). We propose that RPA can act as a replication protein in undifferentiated cells and as a transcription factor during the differentiation process in *Dictyostelium discoideum*. Determining the DNA binding properties of these two forms of DdRPA on the *gp-2* promoter are therefore important in understanding how this transition may be accomplished. We have mapped in this study the potential regulatory sites bound by the DdRPA protein on the *gp2* promoter of *Dictyostelium discoideum*, using in vitro footprint analyses.

The protein containing a 62 kDa subunit may act in replication, while the 84kDa isoform may play a role in transcription during differentiation. The DNA binding function of DdRPA resides in this large subunit of the protein (Wen, 1998). While purifying this protein, a pH dependent *in vitro* conversion was discovered. At pH 7.0 and 7.5, the 84 kDa was observed. When the pH was decreased to 6.0, this subunit was shifted to 62 kDa form and the change was irreversible (Wen, 1998). It is likely that the 84 kDa subunit was cleaved by an unidentified protease in a pH dependent manner. Other studies have suggested that the pH of the cytoplasm increases upon cell differentiation *of Dictyostelium discoideum.* Perhaps a general increase in cytoplasmic pH inhibits pH-dependent protease. DdRPA containing the 84 kDa subunit may then be translocated into to the nucleus where it can act as a transcription factor. Western data indicated that DdRPA becomes enriched in the nucleus of differentiated cells (Favis, unpublished data). In addition, using southwestern analyses, no DdRPA was found in the nucleus of undifferentiated cells (amoebae), but the nucleus of differentiated cells (slug) contained DdRPA. Therefore, this data implies that the 84 kDa subunit contains a signal responsible for translocating the protein to the nucleus, whereas the 62 kDa subunit does not. It is likely that the additional 22kDa (84kDa-62kDa) provides the necessary signal to pass through the nuclear pore complex. Perhaps DdRPA containing the 62 kDa subunit can only enter the nucleus following breakdown of the nuclear envelope during prophase of mitosis and then acts as a replication protein.

The footprint results of this study reveal that DdRPA can bind to the C and TAG box regions. There is no dramatic difference between the bindings of the two forms to the C boxes; however, a difference in binding to the TAG box region was discovered. The undifferentiated form (62 kDa) produced a footprint, while the differentiated form (84 kDa) did not. It may be possible that the C box and TAG boxes are the sites that undifferentiated DdRPA binds to achieve replication in proliferating cells. However, differentiated DdRPA may bind to the C box regions and not the TAG box region in the process of transcription during cell differentiation. There is possibly another factor that competes for the TAG box region during transcription suggested by findings of southwestern analyses (Wen, 1998).

With regard to the small footprints obtained from DdRPA, it is likely that purified DdRPA can not produce large footprints. Protein-protein interactions may be required to generate sufficient specificity in binding. Because regions surrounding the C box regions also footprint, this is a likely possibility. In fact, there is evidence from other organisms that the second subunit (35kDa) of RPA is involved in protein-protein interactions (Fanning et al. 2006).

Future experiments that need to be done to clarify the role of DdRPA on the *gp2* gene expression involve ectopic expression of wild type and mutant forms of the protein, knocking out the gene, and overexpressing the protein in the organism and analyzing *gp2* expression in such conditions.

REFERENCES

Ausubel F.M. (1995). Current Protocols in Molecular Biology 2[nd ed] . John Wiley& Sons, Inc.

Fanning, E., V. Klimovich, and A. R. Nager. (2006). A dynamic model for replication protein A (RPA) function in DNA processing pathways. Nucleic Acids Res 34: 4126-4137.

Gross, J. D. (1994). Developmental Decisions in *Dictyostelium discoideum*. Microbiological Reviews 58: 330-351.

Kampang, P. (1995). Virginia Polytechnic Institue and State University. Purification of the proteins that bind to the *gp-2* promoter in *Dictyostelium discoideum.*

Kelman, Z. (2000). DNA replication in the third domain (of life). Curr Protein Pept Sci 1: 139-154.

Loomis, W. F. 1982. The Development *of Dictyostelium discoideum*. Academic Press, New York.

Maxam, A and Gilbert W. (1977) A new method of sequencing DNA. Proceedings of the National Academy of Sciences 74 (2): 560-564.

Rutherford, C. L., O. Selmin, and S. PetersWeigel. (1997). Temporal regulation of the Dictyostelium glycogen phosphorylase 2 gene. Biochimica et Biophysica Acta-Gene Structure and Expression 1351: 111-125.

Wen, X., P. Khampang, and C. L. Rutherford. (1998). The glycogen phosphorylase-2 promoter binding protein in Dictyostelium is replication protein A. Journal of Molecular Biology 284: 903-913.

Wold, M. S. (1997). Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annu Rev Biochem 66: 61-92.