Multiple protein – DNA interactions in the TATAA-less mouse thymidylate synthase promoter

Keith Jolliff^{3,*}, Yue Li² and Lee F. Johnson^{1,2,3,*}

Departments of ¹Molecular Genetics and ²Biochemistry and the ³Molecular, Cellular and Developmental Biology Program, The Ohio State University, Columbus, OH 43210, USA

Received March 3, 1991; Accepted April 9, 1991

ABSTRACT

The mouse thymidylate synthase (TS) promoter lacks a TATAA box and directs transcriptional initiation at multiple sites over a 60 nucleotide region. All of the sequences that are important for transcription are located within close proximity to the first initiation site. Gel mobility shift and footprinting analyses with various sequences from this region of the TS promoter identified three major protein-DNA interactions. One of these corresponds to Sp1 interacting with a non-consensus binding site downstream of the first transcriptional initiation site. Inactivation of the binding site by site-directed mutagenesis led to a 3-fold reduction in gene expression as well as a significant change in distribution of transcriptional start sites. The proteins responsible for the other two complexes (CII and CIII) do not appear to correspond to any of the common transcription factors that have been studied previously. The CII protein binds very close to the first initiation site. Inactivation of the binding site by site-directed mutagenesis had little effect on expression. The CIII protein binds immediately upstream of CII. Inactivation of this site led to a 12-fold reduction in expression, indicating that it is important for expression of the TS gene.

INTRODUCTION

Thymidylate synthase (TS) is the enzyme that catalyzes the synthesis of TMP in the de novo biosynthetic pathway. This housekeeping enzyme is essential for survival of proliferating cells and is the target enzyme for a variety of chemotherapeutic drugs (1). In mammalian cells, the level of TS gene expression is tightly regulated with respect to the proliferative state of the cell (2,3); TS enzyme and mRNA levels are low in quiescent cells but increase approximately 20-fold when the cells are stimulated to reenter the cell cycle (4,5). The increase in expression occurs at about the same time that the cells begin DNA replication, which is in line with the role of this enzyme in DNA synthesis.

To study the control of TS gene expression in more detail, we have cloned and characterized the mouse TS gene (6) and

have been analyzing the sequences that are important for expression. The gene is 12 kb in length, and the 1 kb coding region is interrupted by 6 introns. Several unusual features of the mouse TS gene were uncovered. First, the predominant mRNA encoded by the gene lacks a 3' untranslated region; the poly(A) tail immediately follows the UAA translational stop codon (7,8). Second, transcription of the gene is initiated at multiple sites over a 60 nucleotide region. The G/C-rich 5' flanking region lacks TATAA and CCAAT elements but does contain potential binding sites for a variety of other transcription factors. For example, binding sites for the cellular transcription factors Sp1 (9–11) and USF/MLTF (12,13) are located 35–50 nucleotides upstream of the first transcriptional start site (14).

To identify the promoter elements that are required for efficient expression of the TS gene, deletion studies were performed. TS minigenes which retain as few as 22 nucleotides upstream of the first transcriptional initiation site are expressed at about the same level (and retain the same pattern of transcriptional start sites) as minigenes extending to –1 kb. Therefore, the upstream Sp1 and USF elements are not essential for gene transcription. However, when the deletions extended a few nucleotides downstream of the first initiation site, the level of expression was reduced greater than 10-fold (14). These analyses indicate that sequences in close proximity to the first initiation site contains one or more important promoter elements. In the present study, we have determined that at least three proteins bind to the TS promoter in the vicinity of the first transcriptional initiation site. One of these is Sp1 whereas the other two (designated CII and CIII) do not correspond to any of the common transcriptional factors that have been identified to date. Only one of these proteins (CIII) appears to be highly important for promoter activity.

MATERIALS AND METHODS

Cell culture

The LU3-7 cell line (15) is a fluorodeoxyuridine-resistant derivative of mouse 3T6 fibroblasts (16) that overproduces TS approximately 50-fold. The cells were grown on plastic culture

* To whom correspondence should be addressed

+ Present address: Department of Medicine, Division of Dermatology, Washington University School of Medicine, 660 S. Euclid, St Louis, MO 63110, USA

dishes in Dulbecco's Modified Eagles Medium (GIBCO) supplemented with 10% calf serum (Colorado Serum), 3 μ M fluorodeoxyuridine, 1 mM uridine, and 1 mM cytidine. Large scale cultures were grown in glass roller bottles.

Preparation of nuclear extracts

Nuclear extracts were prepared from LU3-7 cells as described by Hennighausen and Lubon (17) except that the supernatant from the 100,000 $\times$ g centrifugation was dialyzed against 50 volumes of Buffer D [20 mM HEPES-KOH (pH 7.9), 20% (v/v) glycerol, 0.2 mM EDTA, 0.5 mM DTT and 0.5 mM PMSF] which contained 0.1 M KCl. Protein concentrations were determined by the method of Bradford (18) and ranged from 6–15 mg/ml. All extracts were rapidly frozen in liquid nitrogen and stored at -70°C .

Probes

To facilitate the labeling of promoter fragments for the gel mobility shift assays, subclones were generated from the 5'-flanking region of the TS gene. The plasmids pH/H-Wt, pH/H-M4, pH/H-M10, and pH/H-M14 contain the 57 bp *HinfI* to *HaeII* restriction fragment (–114 to –57) from either the wild type TS promoter or various mutated promoters (see Figure 1). The restriction fragments were blunted using T4 DNA polymerase (USB) and ligated into the *SmaI* site of pUC18.

The plasmid pPR29 was created by cloning a double-stranded synthetic oligonucleotide (PR29) (synthesized by the OSU Biochemical Instrument Center) into the *PstI* and *BamHI* restriction sites of the pBluescript SK(–) polylinker (Stratagene). The PR29 DNA insert:

5'-GAGTTTCCCAGCAGGAAGAGGCGGGCTGGTG-3'
3'-ACGTCTCAAAGGGTCGTCCTTCTCCGCCGACACCTAG-5',

consists of the TS promoter region from –99 to –71 (underlined) with *PstI* and *BamHI* cohesive ends. The plasmid pCR19 consists of a double stranded oligonucleotide (CR19) that was cloned into the *EcoRI* site of pBluescript SK(–). The CR19 DNA insert:

5'-AATTCGGCGGCCCGGAAGTTTCCCAG-3'
3'-GCCGCCGCCCTTCAAAGGGTCTTAA-5'

consists of the TS promoter region from –109 to –91 (underlined) with *EcoRI* cohesive ends.

Gel mobility shift assays

DNA probes were asymmetrically 5'-end labeled using [γ - ^{32}P]ATP and T4 polynucleotide kinase (USB). The DNA binding reactions (20 μ l) were carried out in a binding buffer consisting of 10 mM Tris-HCl (pH 7.9), 0.5 mM HEPES-KOH (pH 7.9), 50 mM KCl, 1 mM DTT, 1 mM EDTA, and 5% (v/v) glycerol. A typical reaction contained 0.5 ng (10^4 cpm) of probe, 1.5–2.0 μ g of crude nuclear protein and 1–2 μ g of poly[d(I·C)] (Pharmacia). Sp1 purified from HeLa cells (19) (generously provided by J. Kadonaga and R. Tjian) was incubated with probe in a mixture that contained 25 ng poly[d(I·C)], 60 mM KCl, 7% (v/v) glycerol, 200 μ g/ml BSA, 1.25 mM MgCl_2 , 0.1 μ M ZnSO_4 . After the incubation period was complete, 2 μ l of loading buffer (6.7 mM Tris-HCl (pH 7.9), 3.3 mM sodium acetate, 50% (v/v) glycerol, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol) was added to the reaction mixture and the DNA/protein complexes were resolved on a 1 mm thick 5% polyacrylamide gel (acrylamide:bis, 60:1) cast in running buffer (44.5 mM Tris-base, 44.5 mM boric

acid and 1.0 mM EDTA). The gels were pre-run for 1.5 hr at 15 mA. The samples were then loaded and the current was increased to 20 mA. The gels were dried and exposed to X-ray film for 10–16 hr at -70°C with an intensifying screen (DuPont Cronex).

For competition analyses, the binding reactions were assembled as described above except that the appropriate unlabeled DNA competitor was included and the labeled probe was omitted. After a 10 min incubation at ambient temperature, the labeled probe was added and incubation was continued for 20 min. The samples were then electrophoresed as described above. The competitor DNAs were either restriction fragments isolated from one of the plasmids described above or the double stranded synthetic oligonucleotides (Stratagene) shown below:

SP1 5'-GATCGATCGGGGCGGGGCGATC-3'
3'-CTAGCTAGCCCCGCCCGCTAG-5'
NF1/CTF 5'-ATTTTGGCTTGAAGCCAATATG-3'
3'-TAAAACCGAAGCTTCGGTTATAC-5'
NF-kB 5'-CTAGTGGGACTTTCACAGATC-3'
3'-GATCACCTGAAAGGTGTCTAG-5'

The E2 competitor was derived from the plasmid pE2100 (supplied by J. Nevins) and contains the Adenovirus E2 promoter sequence from –21 to –98 (20) cloned into the *EcoRI* and *HindIII* restriction sites of the vector pUC13.

OP-Cu footprinting assays

The 1,10-phenanthroline-copper ion complex (OP-Cu) footprinting was performed using the method of Kuwabara and Sigman (21) with minor modifications. Briefly, preparative gel mobility shifts were carried out as described above. The gel was then incubated with the OP-Cu chemical nuclease for 5 min under conditions previously described (21). The wet gel was exposed to X-ray film overnight at 4°C . Regions of the gel that contained the DNA/protein complexes and free probe were excised and the DNA was recovered. Equal amounts of the radioactive digestion products of the free and complexed probe were analyzed on a DNA sequencing gel. The G- and G+A-specific Maxam and Gilbert sequencing reactions (22) of the probe served as size markers. The wet gels were exposed to X-ray film at -70°C .

Construction of TS minigenes

The construction of pI56, the wild-type TS minigene, was described previously (14,23). It consists of 1 kb of 5' flanking sequence and 0.25 kb of 3' flanking sequence from the TS gene linked to the TS coding region at restriction sites in the first and last exons. Introns 5 and 6 were included at their normal positions in the coding region to ensure that the gene would be expressed at high levels (23). The minigene was cloned into pUC18. pI56(–0.25) is the same as pI56 except that it extends to the *SphI* site 0.25 kb upstream of the AUG codon.

The M4 mutation in the promoter region of pI56 was generated by using the method of Zoller and Smith (24) as described previously (14). The M14 and M10 mutations were created according to the method of Kunkel (25). The *SphI*-*SphI* fragment at the 5' end of the mouse TS gene, which contains 249 nucleotides of 5'-flanking region and 159 nucleotides of coding region, was subcloned into the *SphI* site of the RF form of M13mp18. A clone containing the fragment in the desired orientation was identified by restriction analysis. The phage was transferred to *E. coli* CJ236 (25) and uracil-containing single

stranded phage DNA was isolated. The oligonucleotides used for mutagenesis were CGCCTCTTCCTGTCCGGAACTTCCGG (for M14) and GCCCGCCTCTTCCTGTGGGAAAC (for M10). The promoter mutations were confirmed by DNA sequence analysis. The *SphI*-*SphI* fragments that contained the mutations were subcloned back to pI56(-0.25). The sequences of the promoter regions of the final constructions were re-confirmed by double-stranded DNA sequence analysis.

Transient transfection assays

TS minigenes (12 pmoles) were transfected into TS(-) hamster V79 cells (26) using the calcium phosphate precipitation procedure (27,28). To correct for slight differences in transfection efficiency, 0.8 pmoles (5 μ g) of pSV2cat (29) were also included in each transfection mixture. Cultures were harvested two days after transfection and assayed for TS and CAT enzyme levels as described previously (14,29). In some experiments, cytoplasmic RNA was also isolated and analyzed for TS mRNA level and the pattern of 5' termini of TS mRNA using an S1 nuclease protection assay as described previously (14).

RESULTS

Interactions of nuclear proteins with the TS promoter

The nucleotide sequence of the 5' flanking region of the mouse TS gene is shown in Figure 1. Deletion analyses have shown that the sequence between -113 and -85 (relative to the A of the AUG start codon) is very important for promoter activity (14). This region, hereafter referred to as the 'essential region,' is in the vicinity of the first transcriptional initiation site at about -90.

To examine the spectrum of proteins that interact with the essential region, gel mobility shift assays were performed. The H/H-Wt probe, which contained the TS promoter sequence from -114 to -57 (Figure 1), was used for the initial analyses. When this probe was incubated with an unfractionated nuclear extract

Figure 1. Important features of the TS promoter region. The 5'-flanking region of the TS gene is shown in boldface letters, with the approximate locations of the major (M) and minor (m) transcriptional initiation sites indicated above the sequence. The numbering of the bases is relative to the A of the AUG initiation codon (+1). The location of *AluI* (Alu), *HinfI* (Hinf), and *HaeII* (Hae) restriction sites are indicated. Potential binding sites and the consensus recognition sequences for the cellular transcription factors E2F (20,36), NF-kB (37), NF1 (38), and Sp1 (39) are also shown. The nucleotide changes introduced by site-directed mutagenesis are indicated (*). The regions of the promoter contained in the pCR19 and pPR29 plasmids, which were used to generate probes for binding assays, are indicated above the sequence.

prepared from LU3-7 mouse fibroblasts, three major complexes, designated CII, CIII and CIV, were observed (Figure 2A, lane 2). A fourth complex (CI) as well as other minor complexes were also observed in many experiments. The intensities of the minor complexes varied considerably in different nuclear preparations and have not been analyzed in detail.

To ascertain that the major complexes represented specific protein-DNA interactions, a self competition was performed in which the restriction fragment used to generate the probe was used as an unlabeled competitor. In the presence of a 100-fold molar excess of competitor, complexes CIII and CIV were undetectable, but CII was only slightly diminished (Figure 2A, lane 3). To reduce CII significantly, it was necessary to increase the competitor concentration to a 400-fold molar excess (Figure 2A, lane 4). The reason why so much competitor is required to reduce complex CII is not clear. These data suggest that all of the complexes arise from specific DNA/protein interactions.

To further localize the regions of the promoter that are interacting with the proteins in the nuclear extracts, probes corresponding to subregions of the TS promoter were used in mobility shift analyses. The probe PR29, which contained the

Figure 2. Interactions of nuclear proteins with the TS promoter between -114 and -57. Gel mobility shift assays were performed using probes from the essential region of the TS promoter and unfractionated nuclear extracts prepared from mouse LU3-7 cells. CI, CII, CIII and CIV refer to DNA/protein complexes; FP is free probe. **Panel A:** The probe (H/H-Wt) was an *EcoRI* to *BamHI* restriction fragment isolated from the plasmid pH/H-Wt. The probe contains the TS promoter sequence from -114 to -57 plus 23 nucleotides of the pUC18 vector polylinker. Some reactions also contained unlabeled probe or the NF1 oligonucleotide as a self-competitor. Each binding reaction contained 0.5 ng (10^4 cpm) of probe and 1.5 μ g of nuclear extract (NE). Lane 1, -NE, no competitor; lane 2, +NE, no competitor; lane 3, +NE, 50 ng unlabeled H/H-Wt; lane 4, +NE, 200 ng unlabeled H/H-Wt; lane 5, +NE, 54 ng (400-fold molar excess) unlabeled NF1. **Panel B:** Three different probes were used. The H/H-Wt probe (lanes 1 and 4) was the same as the probe used in panel A. The PR29 probe (lanes 2 and 5) is a 62 nucleotide *XbaI* to *HindIII* restriction fragment isolated from pPR29 and contains TS promoter sequence from -99 to -71 plus 33 nucleotides of polylinker sequence. The CR19 probe (lanes 3 and 6) is a 42 nucleotide *PstI* to *HindIII* restriction fragment isolated from pCR19 and contained TS promoter sequence from -109 to -91 plus 23 nucleotides of polylinker sequence. Each binding reaction contained 10^4 cpm of probe. In lanes 1-3 the reaction was performed without nuclear extract while in lanes 4-6 the reaction was performed in the presence of 7 μ g of LU3-7 nuclear extract.

Figure 3. Competition assays. Mobility shift assays were performed with specific competitors. Assays included 10^4 cpm of the H/H-Wt probe and the specified amount of unfractionated LU3-7 nuclear extract (NE). The assays were performed in the presence or absence of unlabeled competitor DNAs as indicated. CI, CII, CIII and CIV refer to DNA/protein complexes; FP is free probe. *Panel A:* The probe (0.8 ng) was incubated either with nuclear extract (1.5 μ g) or with 1.0 μ l (unknown concentration) of pure Sp1. Lane: 1, +NE, -competitor; lane 2, +Sp1, -competitor; lane 3, +NE, +10.8 ng (50-fold molar excess) NF1 competitor; lane 4, +Sp1, +10.8 ng NF1 competitor; lane 5, +NE, +10.8 ng Sp1 competitor; lane 6, +Sp1, +10.8 ng Sp1 competitor. *Panel B:* The probe (1.5 ng) was incubated with 7 μ g of nuclear extract. Lane 1, -NE, -competitor; lane 2, +NE, -competitor; lane 3, +NE, +200 ng (100-fold molar excess) E2 competitor; lane 4, +NE, +50 ng (100-fold molar excess) NF-kB competitor.

promoter sequence from -99 to -71 (Figure 1) as well as polylinker sequences, was generated from the plasmid pPR29. When this probe was used in mobility shift analyses, two major bands were observed. One comigrated with CII and the other with CIV (Figure 2B). Another probe (CR19) that contained the promoter sequence from -109 to -91 (Figure 1) was generated from the plasmid pCR19. When this probe was used in mobility shift assays, two major bands were also observed. One comigrated with CII and the other with CIII (Figure 2B, lane 6). These observations indicate that complexes CII, CIII and CIV result from interactions of proteins with the TS promoter between -109 and -71. Sequences sufficient for CII binding appear to be located in the region of overlap of the two probes (-99 to -91). Sequences necessary for CIII formation appear to be upstream of -99 whereas sequences necessary for CIV formation appear to be downstream of -91.

Competition assays

The TS promoter between -109 and -71 contains several potential binding sites for known transcription factors, including NF1, NFkB, E2F and Sp1. The consensus binding sites for these factors as well as the potential sites of interaction in the TS promoter are listed in Figure 1.

When the H/H-Wt probe was incubated with pure Sp1 isolated from HeLa cells (kindly supplied by J. Kadonaga and R. Tjian), a single shifted complex was detected which comigrated with the CIV complex (Figure 3A, lanes 1 and 2). A 50-fold molar excess of an Sp1-specific oligonucleotide competed for the complex IV binding activity in crude extracts as well as the complex detected with the purified Sp1 (Figure 3A, lanes 5 and 6), whereas an NF1-specific oligonucleotide did not (Figure 3A, lanes 3 and 4).

Figure 4. OP-Cu ion footprint of Sp1 bound to the -80 GC box. **Panel A:** The binding of Sp1 to the noncoding strand was examined using pure Sp1 and 300,000 cpm of the wild-type H/H-Wt probe (labeled at the *Bam*HI site) isolated from the plasmid pH/H-Wt. Lane: 1, G-specific sequencing ladder; lane 2, G+A-specific sequencing ladder; lane 3, digestion products of the protein-DNA complex; lane 4, digestion products of free DNA. **Panel B:** The binding of complex CIV on the coding strand was analyzed using LU3-7 nuclear extracts (60 μ g) and the mutant H/H-Wt probe (230,000 cpm, labeled at *Eco*RI) isolated from the plasmid pH/H-M4 (Figure 1). This mutation prevents the formation of complex CIII, thereby facilitating the isolation of complex CIV. Lane 1, G-specific sequencing ladder; lane 2, G+A-specific sequencing ladder; lane 3, digestion products of the protein-DNA complex; lane 4, digestion products of free DNA. The sequence of the TS promoter between -114 and -57 is shown at the bottom of the Figure. The regions protected by Sp1 and by CIV are indicated.

Thus it appears that complex IV corresponds to binding of Sp1 to the H/H-Wt probe. The interaction is probably with the GC box at -80, but this experiment does not rule out the possibility that other GC-rich sequences are involved.

When a 400-fold molar excess of a double-stranded synthetic oligonucleotide corresponding to an NF1 binding site (Figure 2A, lane 5) was used as a competitor, none of the complexes that were formed with the H/H-Wt probe were reduced. Figure 3B shows that a 100-fold molar excess of an oligonucleotide containing an NF- κ B binding site (lane 4), or a 82 nt unlabeled restriction fragment that contains the E2F binding site (lane 3) from the Adenovirus E2 gene had no detectable effect on the formation of complexes CII or CIII. (Complex CIV was difficult to detect with this preparation of nuclear extract.) In a control experiment we found that a 100-fold excess of unlabeled E2F probe completely blocked the formation of the complex between E2F and the E2F probe (data not shown). Furthermore, a 300-fold molar excess of the NF- κ B competitor also had no effect on the pattern (not shown). Therefore it appears that NF1, NF-

Figure 5. The effect of point mutations on protein binding. Probes were generated by restricting the plasmids pH/H-Wt, pH/H-M4, pH/H-M10 and pH/H-M14 with *EcoRI* and end-labeling the 5'-termini. Secondary digestion with *BamHI* resulted in an 80 bp probe containing the TS promoter sequence between -114 and -57, along with 23 nt of the pUC18 polylinker. Protein-DNA interactions were measured by the gel mobility shift assay. Each binding reaction contained 10^4 cpm of probe and 3.2 μ g of LU3-7 nuclear extract (NE). *Panel A.* Mobility shift analysis of the M4 mutant probe. Lane 1, Wt probe, -NE; lane 2, Wt probe, +NE; lane 3, M4 probe, +NE; lane 4, M4 probe, -NE. *Panel B.* Mobility shift with the M14 mutant probe. Lane 1, Wt probe, -NE; lane 2, Wt probe, +NE; lane 3, M14 probe, +NE; lane 4, M14 probe, -NE. *Panel C.* Mobility shift analysis of the M10 mutant. Lane 1, Wt probe, -NE; lane 2, Wt probe, +NE; lane 3, M10 probe, +NE; lane 4, M10 probe, -NE.

kB and E2F were not responsible for the formation of any of the major complexes between proteins present in the mouse nuclear extracts and the essential region of the TS promoter.

Footprint analysis of CIV/Sp1

In a previous DNaseI footprinting study with the TS 5' flanking region, we showed that Sp1 protected the G/C-rich region located at -130 but found no evidence for protection at -80 (14). The upstream site, GGGGCGGGGT, matches the consensus binding site for Sp1 whereas the potential downstream site, GAGGC-GGGCT, does not. It is possible that protection of the downstream site was not detected because the affinity of Sp1 for the downstream site was much lower than that for the upstream site.

To avoid this problem, we performed footprint analyses using a probe that contained only the downstream Sp1 site. In addition a more sensitive 1,10-phenanthroline copper footprinting procedure (21) was used. In this procedure, the protein is incubated with the H/H-Wt probe and fractionated in a gel mobility shift assay. The chemical nuclease, 1,10-phenanthroline copper, is used to cleave the free and bound probe while they are still within the polyacrylamide gel matrix. The digestion products are then analyzed on a sequencing gel. We first analyzed the binding of purified Sp1 to the noncoding strand of the essential region of the TS promoter. Figure 4A shows that Sp1 protected a 9 bp region of the probe from -77 to -85.

We then used this approach to determine the location of the binding site of the complex IV protein. In this case, unfractionated nuclear extracts were used and the probe was isolated from the plasmid pH/H-M4. This probe is the same as the wild-type probe

Figure 6. Effects of promoter mutations on the pattern of transcriptional start sites. Cytoplasmic poly(A)⁺ mRNA from transfected ts-hamster V79 cells was isolated and 2 μ g of mRNA from each sample was analyzed by S1 nuclease protection analysis. The probe was labeled at the 5' end with 32 P at a *BamHI* site in the coding region (256 nucleotides downstream of AUG start codon) and extended to the *SphI* site 249 nucleotides upstream of the AUG start codon. The digestion products were analyzed on a 6% polyacrylamide sequencing gel and detected by autoradiography.

except for 3 nucleotide changes which abolish the formation of complex III (see below). The absence of complex III made it possible to isolate complex IV generated in the presence of the nuclear extracts and thereby verify that complex CIV was generated by Sp1. Figure 4B shows that the protein responsible for CIV protected a 10 nt region of the probe from -76 to -85. This coincides with the GC box centered at -80 and with the footprint produced by pure Sp1. There was no indication of protection in other GC-rich regions of the probe. Thus it appears that CIV is the result of Sp1 interacting with the mouse TS promoter at -80.

The gel mobility shift analyses suggested that the CII protein interacts with the TS promoter between -99 and -91 whereas the CIII protein binds between -99 and -109. To confirm these possibilities, we attempted to footprint the protein bound to the CII and CIII binding sites. Unfortunately, all of our footprinting attempts using unfractionated nuclear extracts and a variety of protocols (including the 1,10-phenanthroline copper procedure described above) proved unsuccessful. We are currently attempting to isolate highly purified CII and CIII proteins to overcome this problem.

Effect of mutations in the TS essential region

To confirm the locations of the CII and CIII binding sites and to evaluate the biological significance of the DNA/protein interactions in the essential region of the TS promoter, specific mutations were introduced into the CII, CIII and Sp1(-80) binding sites by site-directed mutagenesis. To determine the effects of the mutations on protein binding, probes were generated from the altered promoters and utilized in gel mobility shift assays.

The M4 mutation (Figure 1), centered at -102 in the region recognized by the CIII binding protein, altered the sequence

Table 1. Effect of promoter mutations on TS minigene expression.

Promoter mutation ¹	TS enzyme level ²
pI56 (wt, -0.25 kb)	1.0
-CII (M14)	1.37 ± 0.16
-Sp1 (M10)	0.36 ± 0.07
pI56 (wt, -1 kb)	1.0
-CIII (M4)	0.08 ± 0.03

¹ The 5' flanking region of the minigene extends either to the *SphI* site 0.25 kb upstream of the AUG codon, or to the *XbaI* site 1 kb upstream of the AUG codon. The promoter mutations are described in Figure 1.

² The minigenes that contained the mutated promoters were transiently transfected into TS(-) hamster V79 cells along with pSV2-cat as a control gene. TS enzyme level was normalized to CAT enzyme level (to correct for differences in transfection efficiency). This value was then normalized to the value obtained with the wild type (wt) promoter, which was assigned a value of 1.0. Data represent the average (± standard deviation) of at least four independent experiments.

Figure 7. Summary of protein-DNA interactions in the mouse TS promoter. The coding strand of the 5' flanking region of the TS gene is shown. The approximate locations of the first two transcriptional initiation sites (T) are indicated. The sites of interaction of Sp1 and USF between -120 and -145 were determined by DNaseI footprinting (14) whereas the downstream Sp1 binding site was determined by chemical footprinting. The regions encompassing the CII and CIII binding sites are also shown.

CGGAA to TATAA. The creation of the TATAA sequence should not only prevent the formation of the CIII complex but should permit the interaction of the cellular transcription factor TFIID with the TS promoter (30). When a probe containing this alteration was incubated with crude LU3-7 nuclear extracts and analyzed by mobility shift assay, complex CIII was abolished (Figure 5A). This suggests that the CIII complex is due to protein-DNA interactions in the region around -102. No additional complexes that might correspond to TFIID interacting with the M4 probe were formed. This may be due to the fact that the TFIID factor does not form a particularly stable complex with the TATAA box (12).

Mutant M14 contains an alteration within the CII binding site which converts CAG (-92 to -90) to GGA (Figure 1). This alteration led to a drastic reduction in complex CII in the gel mobility shift analysis (Figure 5B), which provides further confirmation that CII binds with this region of the promoter. Finally, the Sp1(-80) mutation (M10) (Figure 1) prevented the formation of the Sp1 complex (CIV), but had no effect on the formation of the other complexes (Figure 5C).

To test the effect of these mutations on gene expression, minigenes that contained the altered promoter regions were transiently transfected into *ts*-hamster V79 cells. The cultures were harvested 48 hr later and the level of minigene expression was determined by measuring the TS enzyme level. In addition, the 5' terminal structure of the TS mRNA was examined by S1 nuclease protection assays. Inactivation of the CII binding site (mutant M14) had little effect on the level of expression of the minigene (Table 1) or on the pattern of transcriptional start sites

(Figure 6), indicating that this element does not play a major role in TS gene expression.

Inactivation of the Sp1 site at -80 led to a 3-fold decrease in gene expression (Table 1), indicating that this element is important for TS gene expression. This alteration also led to a significant change in the pattern of transcriptional start sites (Figure 6). The locations of the start sites that were used appears to be the same. However, there is a striking increase in the frequency of utilization of the upstream start sites and a decrease in use of the downstream start sites.

Inactivation of the CIII binding site (M4 mutation) had the most severe effect on gene expression. The TS enzyme level was about 12-fold lower than that observed with the wild-type minigene (Table 1). The pattern of transcriptional start sites was similar to that of the wild-type promoter except for the appearance of a new start site approximately 25 nucleotides downstream of the first T of the TATAA sequence created by this mutation. Since the TATAA sequence is usually 26–34 nucleotides upstream of the unique initiation site that is observed in most TATAA-containing promoters, and is believed to be important for specifying the location of the initiation site (31), this observation is not surprising. It appears that the introduction of a TATAA box is unable to compensate for the loss of the CIII binding site in this promoter.

DISCUSSION

We have used a combination of mobility shift assays with specific regions of the TS promoter and a variety of competitor sequences to analyze the proteins that interact with the mouse TS promoter in the vicinity of the first transcriptional initiation site. We have confirmed the locations of the binding sites by footprint analyses and/or by site-directed mutagenesis of the binding sites. Our results demonstrate that at least 3 different proteins interact with the mouse TS promoter region near the first initiation site. One of these proteins, which interacts with a G/C rich region that is slightly downstream of the first initiation site, corresponds to the transcription factor Sp1. The other two factors, CII and CIII, interact with the promoter immediately upstream of the first transcriptional initiation site. In an earlier study, we showed that there were binding sites for the transcription factors USF/MLTF and Sp1 between -120 and -145. Thus at least 5 proteins are capable of binding to the TS promoter within the 60 nucleotide region from -85 to -145. These interaction sites are summarized in Figure 7.

The biological significance of the various DNA-protein interactions was investigated by inactivating the binding sites and determining the level of expression of the modified gene in a transient transfection assay. These studies revealed that some of the interactions were important for expression whereas others appeared to be irrelevant. Inactivation of the CIII binding site or the downstream Sp1 site decreased the level of expression by 12-fold or 3-fold, respectively, whereas inactivation of the CII binding site had no apparent effect on gene expression. We showed previously that inactivation of the upstream Sp1 binding site reduced the level of expression 3-fold but inactivation of the USF/MLTF site had no apparent effect on expression (14).

Although it is possible that the CII and USF/MLTF binding sites are fortuitous, two alternative explanations should also be considered. The first is related to the fact that TS gene expression is tightly regulated in response to growth-stimulation. The

transient transfection assay used in the present studies determined the level of expression of the gene construction in exponentially growing cells. It is possible that the CII and/or the USF/MLTF sites play an important role in regulating TS gene expression in response to growth-stimulation. The second is related to the fact that the TS gene is an essential gene that must be expressed in a wide variety of proliferating cells. The binding sites that are irrelevant for expression in the hamster V79 cells may be more important for expression in other cells, especially those that may contain low levels of Sp1 or CIII proteins.

The CII and CIII binding sites are located in a region that contains potential recognition sites for a variety of DNA binding proteins (Figure 1). However, competition assays indicated that the CII and CIII proteins do not correspond to any of these previously studied proteins. We did note that a stretch of 8 nucleotides (CCGGAAGT) within the region encompassing the CIII binding site are identical to a 8 nucleotide sequence within an important transcriptional regulatory element (the β element) in the genes for the mouse ribosomal proteins L32 and L30 (32,33). All three of these genes lack a TATAA box and are ubiquitously expressed. Unlike the TS gene, however, the ribosomal protein genes have a single transcriptional initiation site. Furthermore, inactivation of the β binding site only leads to a 2–3-fold reduction in promoter activity whereas inactivation of the CIII element leads to a 12-fold reduction in TS gene expression. In addition, it was recently shown that the Su2 element, which is an important positive transcriptional element of the bidirectional promoter of the mouse Surf-1 and Surf-2 genes, also contains the same 8 nucleotide sequence (34). The ubiquitously expressed Surf-1 and Surf-2 genes are similar to the TS gene in that they lack TATAA boxes and initiate transcription at multiple sites. It will be interesting to determine if the same factor is responsible for controlling the transcription of all of these housekeeping genes as well as other ubiquitously expressed genes that lack a TATAA element.

The TS promoter is unusual in that transcription is initiated at multiple locations across a 60 nucleotide region. The biochemical basis for this phenomenon is not known. Although the absence of a TATAA box may in part be responsible, some TATAA-deficient promoters are capable of initiating transcription at a unique site (32). We showed previously that elimination of the upstream Sp1 and USF/MLTF binding sites had no effect on the pattern of initiation sites. In fact, the entire promoter upstream of –113 (which is only 23 nucleotides upstream of the first initiation site) could be deleted without altering the pattern (14). Therefore the diversity of initiation sites is not due to a series of upstream protein-DNA interactions, each of which is responsible for a single initiation site. In the present study, we have inactivated three protein binding sites that are in close proximity to the first transcriptional initiation site. Inactivation of the CII and CIII binding sites had no apparent effect on the distribution of initiation sites (except for the introduction of an additional site due to the formation of a TATAA box with the mutation that inactivated the CIII binding site). However, inactivation of the downstream Sp1 site led to a significant increase in the frequency of utilization of the more upstream sites and a decrease in use of the downstream sites. One explanation for these observations is that one of the DNA-protein interactions (possibly the CIII complex) or the DNA sequence itself is responsible for establishing the locations of the initiation sites. Alternatively, there may be additional downstream transcriptional regulatory elements that have not yet been identified that are

responsible for the multiple initiation sites. The results of the S1 nuclease protection assays are consistent with the possibility that the binding of Sp1 at –80 (which is near the first initiation site) reduces the access of RNA polymerase to the upstream initiation sites and enhances the frequency of initiation at the downstream sites. Similar observations have been made with the hamster DHFR promoter (35). To directly address this issue, it will be necessary to purify the regulatory proteins to homogeneity and analyze the interactions between this factor and the transcriptional machinery in an in vitro transcription system.

ACKNOWLEDGMENTS

We thank Drs. Mark Muller and Joe DiDonato and members of our laboratory for many helpful discussions, Melissa Graham for skilled technical assistance, Drs. Joe Nevins and Robert Tjian for providing reagents, and Martin Korb for noting the similarity between the CIII and ribosomal protein β elements. These studies were supported by grants from the National Institute for General Medical Sciences (GM29356), and the National Cancer Institute (CA16058).

REFERENCES

1. Danenberg, P.V. (1977) *Biochim. Biophys. Acta* 473, 73–92.
2. Conrad, A.H. (1971) *J. Biol. Chem.* 246, 1318–1323.
3. Navalund, L.G., Rossana, C., Muench, A.J. and Johnson, L.F. (1980) *J. Biol. Chem.* 255, 7386–7390.
4. Jenh, C.-H., Geyer, P.K. and Johnson, L.F. (1985) *Mol. Cell. Biol.* 5, 2527–2532.
5. Ayusawa, D., Shimizu, K., Koyama, H., Kaneda, S., Takeishi, K. and Seno, T. (1986) *J. Mol. Biol.* 190, 559–567.
6. Deng, T., Li, D., Jenh, C.-H. and Johnson, L.F. (1986) *J. Biol. Chem.* 261, 16000–16005.
7. Jenh, C.-H., Deng, T., Li, D., DeWille, J.W. and Johnson, L.F. (1986) *Proc. Natl. Acad. Sci. USA* 83, 8482–8486.
8. Harendza, C.J. and Johnson, L.F. (1990) *Proc. Natl. Acad. Sci. USA* 87, 2531–2535.
9. Dynan, W.S. and Tjian, R. (1983) *Cell* 32, 669–680.
10. Dynan, W.S. and Tjian, R. (1985) *Cell* 35, 79–87.
11. Dynan, W.S. and Tjian, R. (1985) *Nature (London)* 316, 774–778.
12. Sawadogo, M. and Roeder, R.G. (1985) *Cell* 43, 165–175.
13. Carthew, R.W., Chodosh, L.A. and Sharp, P.A. (1985) *Cell* 43, 439–448.
14. Deng, T., Li, Y., Jolliff, K. and Johnson, L.F. (1989) *Mol. Cell. Biol.* 9, 4079–4082.
15. Rossana, C., Rao, L.G. and Johnson, L.F. (1982) *Mol. Cell. Biol.* 2, 1118–1125.
16. Todaro, G.J. and Green, H. (1963) *J. Cell Biol.* 17, 299–313.
17. Hennighausen, L. and Lubon, H. (1987) *Meth. Enzymol.* 152, 721–735.
18. Bradford, M.M. (1976) *Anal. Biochem.* 72, 248–254.
19. Kadonaga, J.T. and Tjian, R. (1986) *Proc. Natl. Acad. Sci. USA* 83, 5889–5893.
20. Kovessdi, I., Reichel, R. and Nevins, J.R. (1987) *Proc. Natl. Acad. Sci. USA* 84, 2180–2184.
21. Kuwabara, M.D. and Sigman, D.S. (1987) *Biochemistry* 26, 7234–7238.
22. Maxam, A.M. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* 74, 560–564.
23. Deng, T., Li, Y. and Johnson, L.F. (1989) *Nucleic Acids Res.* 17, 645–658.
24. Zoller, M.J. and Smith, M. (1984) *DNA* 3, 479–488.
25. Kunkel, T.A. (1985) *Proc. Natl. Acad. Sci. USA* 82, 488–492.
26. Nussbaum, R.L., Walmsley, R.M., Lesko, J.G., Airhart, S.D. and Ledbetter, D.H. (1985) *Am. J. Hum. Genet.* 37, 1192–1205.
27. Wigler, M., Sweet, R., Sim, G.K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S. and Axel, R. (1979) *Cell* 16, 777–785.
28. DeWille, J.W., Jenh, C.-H., Deng, T., Harendza, C.J. and Johnson, L.F. (1988) *J. Biol. Chem.* 263, 84–91.
29. Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) *Mol. Cell. Biol.* 2, 1044–1051.
30. Buratowski, S., Hahn, S., Guarente, L. and Sharp, P.A. (1989) *Cell* 56, 549–561.

31. Breathnach, R. and Chambon, P. (1981) *Ann. Rev. Biochem.* 50, 349–383.
32. Atchison, M.L., Meyuhas, O. and Perry, R.P. (1989) *Mol. Cell. Biol.* 9, 2067–2074.
33. Hariharan, N., Kelley, D.E. and Perry, R.P. (1989) *Genes Dev.* 3, 1789–1800.
34. Lennard, A.C. and Fried, M. (1991) *Mol. Cell. Biol.* 11, 1281–1294.
35. Blake, M.C., Jambou, R.C., Swick, A.G., Kahn, J.W. and Azizkhan, J.C. (1990) *Mol. Cell. Biol.* 10, 6632–6641.
36. Kovesdi, I., Reichel, R. and Nevins, J.R. (1986) *Cell* 45, 219–228.
37. Sen, R. and Baltimore, D. (1986) *Cell* 46, 705–716.
38. Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelly, T.J. and Tjian, R. (1987) *Cell* 48, 79–89.
39. Kadonaga, J.T., Jones, K.A. and Tjian, R. (1986) *Trends Biochem. Sci.* 11, 20–23.