A 165-Base Pair Sequence Between the Dihydrofolate Reductase Gene and the Divergently Transcribed Upstream Gene Is Sufficient for Bidirectional Transcriptional Activity*

(Received for publication, May 26, 1989)
Takashi Shimada, Hiroyuki Fujii, and Henry Lin
From the Clinical Hematology Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892

The dihydrofolate reductase gene encodes a key enzyme of one-carbon metabolism and is constitutively expressed in all cells. Recently, transcripts initiated at 89 base pairs upstream from the transcriptional initiation site of the dihydrofolate reductase gene and transcribed from the opposite strand have been identified and shown to encode for a protein with homology to a bacterial DNA mismatch repair enzyme (Fujii, H., and Shimada, T. (1989) J. Biol. Chem. 264, 10057–10064). Therefore, the two genes are organized in a head-to-head configuration separated by an 89-base pair segment. The promoter activities of this short spacer sequence were studied in a transient assay using the chloramphenicol acetyltransferase and the guanine phosphoribosyltransferase genes as reporters. A 165-base pair fragment from −111 to +54 relative to the dihydrofolate reductase initiation site was shown to be sufficient for transcriptional activity in either direction, suggesting that expression of the two divergent genes is regulated by a bidirectional promoter that may use common regulatory elements.

Dihydrofolate reductase (DHFR) is a key enzyme of one-carbon metabolism that is essential for biosynthesis of various molecules in cells. Therefore, the DHFR gene is among the housekeeping genes that are expressed at a low level in all cells and in all stages of development. However, the activity of DHFR enzyme fluctuates during the cell cycle, increasing early in S phase and declining following S phase. Regulation of DHFR gene expression appears to occur at multiple levels (1).

An interesting finding is that transcripts initiated in the 5′-flanking region of the DHFR gene but transcribed from the opposite strand have been identified in human, mouse, and Chinese hamster ovary cells (2–5). RNA analysis has shown that the opposite strand transcripts in human cells are polyadenylated, start at 89 bp upstream from the DHFR initiation site, and are composed of two species of 5.0 and 3.8 kb in length (2). In mouse cells, two types of the divergent upstream transcripts have been reported. Farnham et al. (3) have found small 180–240-nucleotides-long poly(A) divergent RNA in the nuclei of mouse 3T6 cells, whereas Kellems and colleagues (4) have detected multiple species of cytoplasmic poly(A)* RNA with 920–6000 nucleotides in several mouse cell lines. In Chinese hamster ovary cells, a 4-kb poly(A)* divergent RNA has been found (5). A potential open reading frame starts within the first exon sequence of the divergently transcribed gene in each of the three species. Therefore, the divergent upstream transcripts appear to encode a protein.

We have recently succeeded in isolation of complementary DNA clones derived from these divergent transcripts and have identified a 3.5-kb open reading frame. Although the function of this putative protein product has not been identified, computer-assisted sequence analyses have revealed significant homology of amino acid sequence between this human protein and the bacterial mutS (6) or hexA (7) proteins that are essential components of DNA mismatch correction systems (8).

The organization of the two divergent genes suggests that the promoter sequences required for expression of the DHFR gene and the upstream gene overlap with each other. In this paper, we demonstrate that a 165-base pair fragment containing the short spacer sequence between the two genes is sufficient for bidirectional transcription.

MATERIALS AND METHODS

DNA Constructions—Recombinant DNA molecules were constructed by standard methods (9). The plasmid pBRCAT was constructed by inserting a 1.7-kb HindIII-BamHI fragment, which contained the chloramphenicol acetyltransferase (CAT) coding sequence and splicing and polyadenylation signals from pSV2CAT (10), between the HindIII and BamHI sites of pBR222. Various DNA fragments from pDHFR-1.8 (11) were converted to blunt HindIII fragments and were cloned between a blunt SalI site and a HindIII site of pBRCAT. The structure of the CAT constructs used in this study is shown in Fig. 1. To make a dual marker plasmid, pGPT/CAT, a 2-kb HindIII-BamHI fragment containing a bacterial xanthine-guanine phosphoribosyltransferase (GPT) coding sequence and processing signals from pSVgpt (10) was converted to a HindIII-NdeI fragment and was inserted between a HindIII and a NdeI site of pSV2CAT. A 165-bp SalI-AvaI promoter sequence was inserted into a unique HindIII site of pGPT/CAT after blunting and addition of a HindIII linker, yielding pGPT/SACAT. The construction of the DHFR minigene DM-1 (pCN8) in which the DHFR gene is directed by a 1251-bp 5′-flanking sequence has been described previously (11). A deletional mutant of the DHFR minigene (DM(−111)) was made by removing an EcoRI (−1251)–SalI (−111) fragment from DM-1.

Transfection—HeLa cells (approximately 0.5 × 106 cells/20-cm dish) were transfected with 10 μg of CAT plasmids and 10 μg of pRSVgpt (10). Two days after transfection, the cells were lysed in 200 μl of 50 mM tris(hydroxymethyl)aminomethane buffer, pH 8.0, by three cycles of freezing and thawing. The activities of CAT (10) and GPT (12) were determined as described previously. The CAT activities of various constructs were normalized for equal transfection efficiency by use of the GPT activity as a control.

RNA Analyses—Whole cell RNA was prepared by the guanidinium isothiocyanate method (13). The RNase protection assay (2) and S1
nuclease protection analyses (11) were done as described previously. The ES-CP probe used for RNA protection was a 790-bp EcoRI-PvuII fragment derived from the ES-CAT plasmid and contained a 630-bp EcoRI-SstII promoter sequence, a 10-bp ClaI-HindIII plasmid linker sequence, and a 150-bp HindIII-PvuII CAT sequence. This fragment was subcloned into a pGEM4 vector, and an RNA probe was made by use of T7 polymerase.

RESULTS AND DISCUSSION

Promoter activities of the sequence upstream from the DHFR gene were studied using the CAT gene as a reporter (Fig. 1). The CAT activities of various constructs were normalized for equal transfection efficiency using the activity of RSV-gpt (10) that was co-transfected as an internal control. First, we checked the activity of the promoter of the DHFR gene. A 1.3-kb DHFR promoter sequence containing the 1251-bp 5' flanking sequence and 54 bp of the DHFR untranslated fragment was subcloned into a pGEM4 vector, and an RNA probe for nuclease protection analyses (11) were done as described previously. The ES-CP probe used for RNA protection was a 790-bp EcoRI-PvuII fragment derived from the ES-CAT plasmid and contained a 630-bp EcoRI-SstII promoter sequence, a 10-bp ClaI-HindIII plasmid linker sequence, and a 150-bp HindIII-PvuII CAT sequence. This fragment was subcloned into a pGEM4 vector, and an RNA probe was made by use of T7 polymerase.

First, we checked the activity of this construct. The activity of this construct was about 5% of that of pSV2CAT. The activity of the DHFR-CAT construct was designated to 100%. A 5'-deletion mutant was made by deleting an EcoRI region was linked to the CAT gene (DHFR-CAT). In a previous study using the DHFR minigene, the 1251-bp 5'-flanking sequence was shown to be sufficient for expression of the DHFR gene in Chinese hamster ovary cells (14). The CAT activity of this construct was about 5% of that of pSV2CAT. The activity of the DHFR-CAT construct was determined to be 100%. A 5'-deletion mutant was made by deleting an EcoRI fragment from SstII (-111) to AuaII (+54). SA-CAT retained 42% of the activity of DHFR-CAT.

Next, we studied the activity of the promoter directing the divergently transcribed upstream gene. A 636-bp fragment from SstII (-111) to EcoRI (+525) was linked to the CAT gene at the SstII site (ES-CAT). This construct showed a CAT activity comparable to DHFR-CAT. The SA fragment from SstII (-111) to AuaII (+54) was linked to the CAT gene at the SstII site, yielding AS-CAT. This construct gave 88% of the activity of the DHFR-CAT control. Therefore, the 165-bp SA fragment is sufficient for bidirectional promoter activity. This was confirmed by the experiment using a dual marker plasmid pGPT/SA/CAT in which SstII and AuaII sites of the SA fragment were linked to the CAT and GPT genes, respectively (Fig. 1B). Both CAT and GPT activities were easily detected in transfected HeLa cells.

To demonstrate that expression of the CAT gene is regulated by the DHFR upstream SA sequence, RNA from transfected cells was analyzed by an RNA protection assay (Fig. 2). RNA from HeLa cells transfected with AS-CAT was hybridized with the antisense ES-CP probe and digested with RNase. Although several discrete bands were seen in lanes 3 and 4, only two fragments with sizes of 320 and 180 nucleotides appeared to be specific for AS-CAT-transfected cells. The 180-nucleotide protected fragment could be produced if the transcript is initiated 20 bp upstream from the promoter-CAT gene junction. This initiation site corresponds to position -89 relative to the DHFR major initiation site, which is the normal initiation site of the divergent upstream transcripts (2). The 320-bp nucleotide fragment seemed to reflect the limit of homology between the probe and read-through transcripts or transcripts initiated from upstream sites within the pBR322 vector sequence.

Detection of transcripts from SA-CAT by the RNA protection assay was unsuccessful. The CAT activity of the SA-CAT construct was approximately half that of the AS-CAT construct (Fig. 1, panel A), and our interpretation is that the concentration of the AS-CAT transcripts was probably under the limit of detection under the conditions used. It is generally difficult to detect CAT transcripts directed from a weak

![Diagram](https://via.placeholder.com/150)

**Fig. 1. Characterization of the bidirectional promoter.** A, organization of the transcriptional units of the DHFR gene and the divergently transcribed upstream (DU) gene and CAT activities of the bidirectional promoter. Open boxes and a closed box represent the exons of the DHFR gene and the upstream gene, respectively. Nucleotide numbers are relative to the initiation site of the DHFR gene. E, EcoRI; S, SstII; A, AoaII. The CAT activities were normalized for equal transfection efficiency using the GPT activity as an internal control. Values are expressed relative to the activity of the DHFR-CAT plasmid. B, structure of the dual marker plasmid, GPT/SA/CAT, and activities of the bidirectional promoter. GPT and CAT activities were determined after a 60-min incubation at 37 °C using 10 μl of lysate in a 50-μl reaction mixture. S, SstII; A, AoaII; H, HindIII; X, xanthine; XMP, xanthine monophosphate; Chl, chloramphenicol; Ac-Chl, acetylated chloramphenicol.

![Diagram](https://via.placeholder.com/150)

**Fig. 2. Nuclease protection analyses.** A, RNase protection analysis. RNA from HeLa cells transfected with the CAT constructs was hybridized with the antisense ES-CP RNA probe and treated with RNase. Lanes: 1, probe alone; 2, pBR-CAT (20 μg of total cellular RNA); 3, AS-CAT (20 μg); 4, AS-CAT (100 μg); 5, end-labeled MspI-digested pBR322 marker. B, S1 nuclease protection assay. RNA from Cos cells transfected with the DHFR minigenes was hybridized with the end-labeled SD probe and treated with S1 nuclease. Lanes: 6, DM-1 (10 μg); 7, DM(-111) (10 μg); 8, RNA from methotrexate-resistant HeLa cells (HeLaR, 10 μg); 9, empty lane; 10, end-labeled MspI-digested pBR322 marker. The numbers indicate fragment length in nucleotides.
promoter, even though the CAT activity may be readily detectable. To overcome this problem, we constructed a new replicating DHFR minigene (DM(-111)), in which the DHFR gene is directed by a 111-bp 5'-flanking sequence. Because this plasmid contains the SV40 origin, the DHFR minigene can be amplified in monkey kidney Cos cells. RNA from Cos cells transfected with the DHFR minigene was analyzed by an S1 nuclease protection assay using a 270-bp SauI (-111)-Ddel (+152) fragment end-labeled at the Ddel site (SD probe). DHFR transcripts from the endogenous DHFR genes in methotrexate-resistant HeLa cells were detected as a cluster of protected bands around 150 nucleotides in length under our conditions (Fig. 2, lane 8). Transcripts initiated at the normal cap site were also detected when Cos cells were transfected with the DHFR minigenes directed by the 1.3-kb long promoter (DM-1) as well as the 111-bp short promoter (DM (-111)) (Fig. 2, lanes 6 and 7). Therefore, 111 bp of 5'-flanking sequence are sufficient for correctly initiated DHFR transcription. These results indicate that all the elements required for normal transcription in both directions are located in the 165-bp fragment.

In this study, we have demonstrated that 165 bp of the DHFR gene promoter are sufficient for bidirectional transcription. This is the minimal bidirectional promoter sequence so far characterized. This bidirectional promoter is enhancer-independent and is active in both directions in vivo (2). The sequence of the bidirectional promoter defined in this study is shown in Fig. 3. There are 4 GC boxes in the same orientation in this sequence. The GC box is thought to be a consensus sequence for nuclear factor Sp-1 binding sites and to function in an orientation-independent manner (15). Another point is that this promoter sequence lacks the classical TATA box in each strand. There is an AT-rich sequence, AAATA, 27 bp upstream from the DHFR initiation site, but deletion of this AT-rich sequence did not change the promoter activity nor the transcriptional start site in either direction. The major function of the TATA box is thought to involve the accurate positioning of the start of transcription (16). In the case of the human DHFR gene, however, transcripts in either direction from the TATA-less promoter have a single major start site (2, 11). The mechanism by which RNA polymerase II can recognize and read the correct strand is not known. An additional speculation, in view of our results, is that the TATA box may have a role in determining the direction of transcription.

Another feature of this sequence is its richness in guanosine and cytosine, amounting to 74% of the nucleotides in the 165-bp fragment. The CpG residues in an HpaII or HhaI recognition site within this fragment were shown to be totally unmethylated in the endogenous gene (17). In addition, these CpG residues became selectively demethylated when an in vitro methylated DHFR minigene was introduced into Chinese hamster ovary cells (14). Therefore, the DHFR bidirectional promoter sequence appears to be a methylation-free island (18). Recently, Lavia et al. (19) have reported that a randomly selected methylation-free island in mouse chromosomal DNA contains the origin of bidirectional transcription although the function of both transcripts has not been identified. The presence of multiple GC boxes, the lack of the TATA box, and richness in guanosine and cytosine residues appear to be common features of the promoter of the housekeeping genes. It should be of interest to see whether the promoters of other housekeeping genes have bidirectional activity as well.

Recently, it has been shown that the promoters of several tissue-specific genes have bidirectional activity. The basement membrane-specific a1(IV) and a2(IV) collagen genes are found in a head-to-head arrangement separated by approximately 130 base pairs in both human and mouse chromosomes (20-22). The spacer region shows a symmetry of putative functional elements (20). A transient assay using the CAT gene as a reporter showed that a 780-bp fragment containing the short spacer sequence had no detectable transcriptional activity in either direction. However, when the 8-kb segment of the first intron of the a1(IV) chain gene was included in the CAT constructs, the 780-bp fragment was able to drive bidirectional transcription. Therefore, transcription of these collagen genes seems to be regulated by a bidirectional promoter in a totally enhancer-dependent manner (21). The upstream sequences of the chicken skeletal a-actin gene (23) and the rat insulin II gene (24) have also been shown to have bidirectional promoter activities when tested with a heterologous marker gene. Divergent transcripts from the upstream region, however, have not been identified in vivo in either case.

Linton et al. (25) have recently cloned and characterized cDNAs of mouse divergent transcripts. The length of the mouse divergent transcripts is highly heterogenous, and two separate bidirectional promoters seemed to regulate both DHFR and divergent gene expression.

Acknowledgment—We would like to thank Dr. A. W. Nienhuis for continued encouragement and valuable suggestions.

REFERENCES


FIG. 3. Nucleotide sequence of the bidirectional promoter that directs transcription in both directions. The arrows indicate start sites of the DHFR transcripts and the divergent upstream transcript (DUT). Four GC boxes are boxed. Nucleotide numbers are relative to the initiation site of the DHFR gene.
A Bidirectional Promoter of the DHFR Gene