Volume 16 Number 2 1988 Nucleic Acids Research

Deletional analysis of the promoter region of the human transferrin receptor gene

John L.Casey, Bruno Di Jeso, Krishnamurthy K.Rao, Tracey A.Rouault, Richard D.Klausner and Joe B.Harford*

Cell Biology and Metabolism Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892, USA

Received August 6, 1987; Revised and Accepted December 7, 1987

ABSTRACT

Fragments of human genomic DNA corresponding to the promoter region of the gene for the transferrin receptor have been cloned upstream of the bacterial gene for chloramphenicol acetyltransferase and these constructs used to assess promoter activity following transfection into a human rhabdomyosarcoma cell line. Progressive 5' deletions as well as internal linker-substitution constructs support a critical role in gene expression of a sequence element approximately 70 bp upstream of the mRNA start site. In this region, the receptor gene was found to contain 11bp that are identical to a segment of the enhancers of polyoma virus and adenovirus. A fragment encompassing this element was shown to increase gene expression when the fragment was placed in either orientation upstream of the remainder of the transferrin receptor promoter but the same fragment did not activate an enhancer-less SV40 promoter. Removal from within the receptor promoter of three potential binding sites for the transcription factor Sp1 did not decrease the promoter's activity.

INTRODUCTION

Proliferating eukaryotic cells acquire iron through endocytosis of the carrier protein transferrin (Tf) in association with the transferrin receptor (TfR), a common component of all such cells (1-3). The Tf-TfR complex enters the cell *via* coated pits (4,5) and subsequently encounters the acidic environment characteristic of endosomes (6). Whereas the lower pH of the endosome results in dissociation of other ligands [e.g. asialoglycoproteins (7)], Tf remains receptor-associated under these conditions and recycles to the cell surface with the TfR (8,9). However, the acidic environment appears instrumental in the release of iron from Tf (1,10,11). The importance of the Tf cycle is emphasized by the requirement of diferric Tf or an alternate iron source for growth of cells in culture (10,12).

The expression of the TfR is highly regulated. The level of expression is related to the proliferative state of the cells (13-15) as well as to the induction of differentiation (16-18). Resting lymphocytes express few, if any, TfR with a rapid increase in expression occurring upon activation (19-21). Within a population of dividing cells, iron availability modulates the expression of TfR (22-24). Previously, we have demonstrated that TfR

mRNA levels are increased by iron deprivation and decreased when excess exogenous iron is supplied (25,26). Isolated nuclei from cells treated with the iron chelator desferrioxamine incorporate more UTP into TfR mRNA than do nuclei from cells treated with hemin, an iron source (26). These results were interpreted as being consistent with at least a portion of the iron regulation of TfR expression being at the transcriptional level.

We present here results of a deletional analysis of the promoter region of the human TfR gene performed by linkage of fragments of the TfR gene to the bacterial enzyme chloramphenicol acetyltransferase (CAT). Gene expression was monitored by CAT enzyme activity in a human rhabdomyosarcoma cell line transfected with these constructs and this enzyme activity was shown to correlate with CAT mRNA levels. The results indicate that an element with sequence similarity to a portion of the enhancers of certain DNA tumor viruses is a critical element for maximal activity of the TfR promoter. A fragment encompassing this element was shown to augment gene expression when the fragment was placed in either orientation upstream of the remainder of the TfR promoter. However, the same fragment was found to be incapable of activating an enhancer-less SV40 promoter. The elimination of the three best potential binding sites for Sp1 did not decrease the strength of the TfR promoter as monitored by transient expression of CAT.

MATERIALS AND METHODS

Isolation of genomic DNA containing the TfR promoter.

We subcloned a 65 bp fragment from the 5' end of the TfR cDNA clone pcD-TR1 (27) (generously provided by Drs. A. McClellend, L. Kühn, and F. Ruddle). This subcloned fragment was used by us to screen a human genomic library in phage lambda (28). One of the positive clones (termed λTR4) was found to have an approximately 15 kb insert comprised of exon 1 of the TfR gene together with approximately 12 kb of upstream sequence and approximately 3 kb of intron 1. A partial restriction map of λTR4 was generated and fragments of this clone were utilized as the sources of the TfR promoter in the plasmid constructs described below. As a reference point, we have numbered the fragments relative to the TfR mRNA start site previously determined (29) by primer extension (negative numbers representing bases 5' of the start site).

Plasmids.

The construct -900/+1200 CAT was prepared by ligation of a 2.1 kb fragment [bounded by Nde I (~-900) and Hind III (~+1200)] of λTR4 with the plasmid pSV0-CAT (30) which had been digested with Hind III and Nde I. The plasmid -900/+1200 CAT was digested with Hind III and Xba I, filled in with the Klenow fragment of DNA polymerase, and ligated to yield

-900/+66 CAT. The 5' deletions of the TfR promoter region were prepared by timed digestion with the exonuclease Bal 31 after opening the plasmid -900/+1200 CAT at an Acc I site located at approximately -400. Following fill-in and Bam HI digestion, the fragment containing the Bal 31-digested TfR promoter plus CAT gene was isolated by agarose gel electrophoresis. These fragments were ligated into the vector pUC 8 (Bethesda Research Laboratories) that had been digested with Sma I and Bam HI. The +67 to +1200 region was removed from the resultant plasmids by digestion with Xba I and partially with Hind III prior to fill-in and ligation. Following clonal selection and amplification, the end-points of the Bal 31 digestion were determined by DNA sequencing.

Linker-substitution constructs were produced by combining selected 5' and 3' deletions such that specific sequence elements of the TfR gene were replaced with 5' **GGAATTCCC** 3' arising from an Eco RI linker (New England Biolabs). To prepare 3' deletions, -900/+66 CAT was digested with Xba I (site at +62) and subjected to timed Bal 31 digestion followed by fill-in, ligation with the Eco RI linker, digestion with Eco RI, and gel isolation of the ~3 kb fragment containing the ampicillin resistance gene and origin of replication of pBR322 as well as the digested TfR promoter (-900 to variable 3' end). Self-ligation of these fragments and transformation of E. coli DH5 (Bethesda Research Laboratories) yielded a pool of plasmids containing variable 3' ends of the TfR promoter region. The ~1.2 kb Eco RI-Hind III fragment from the 5' deletions and a ~2.3 kb Hind III-partial Eco RI digest fragment containing the CAT gene and the downstream SV40 splicing and polyadenylation signals from pSV2-CAT (30) were ligated with the Eco RI-digested 3' deletions. Finally, the fragment corresponding to +67/+1200 was removed as described above. Subsequent to clonal selection and amplification the nature of each of the linker-substitution deletions was determined by DNA sequencing. These constructs were designated as w/x CAT (y/z L) where w and x are the 5' and 3' extrema respectively of the TfR fragment linked upstream of the CAT gene and the portion from y to z inclusive has been replaced by the Eco RI linker sequence described above. The construct -95/+66 CAT (-53/-42 L) was prepared from -900/+66 CAT (-53/-42 L) by digestion with Nde I (~-900) and Nru I (-95) followed by fill-in and ligation.

To prepare (-119/-54)-53/+66 CAT and (-54/-119)-53/+66 CAT, the 72 bp fragment containing TfR gene bases -119 to -54 was isolated from a polyacrylamide gel following digestion of -900/+66 CAT (-53/-42 L) with Sma I (site at -119) and Eco RI (site in the linker that replaced nucleotides -53 to -42). The isolated 72 bp fragment was cloned in both orientations between the Nde I site (at -900) and the linker Eco RI site of the plasmid -900/+66 CAT (-60/-54 L). Following clonal amplification, the orientation of the 72 bp fragment relative to the remainder of the promoter was

assessed by restriction mapping. The resultant constructs contain TfR gene nucleotides -119 to -54 as a single copy in either the native or inverted orientation linked to TfR gene nucleotides -53 to +66 and the CAT structural gene. The isolated -119/-54 fragment of the TfR promoter region was also placed in both orientations by blunt-end ligation into the Bgl II site of the plasmid pA10 CAT 2 (31) that contains an enhancer-less SV40 promoter.

DNA Sequencing.

DNA sequence was determined by cloning overlapping fragments into M13mp18 and M13mp19 (Pharmacia) and using a [^{35}S]dideoxynucleotide chain termination procedure (32). Both strands were sequenced to determine the TfR promoter region sequence shown in figure 2.

Transfections and assay of CAT expression.

The human rhabdomyosarcoma cell line RD4 (33) was maintained in Dulbecco's Modified Eagle Medium supplemented with 10% fetal bovine serum. Cells (2 x 10^6) were transfected with 20 µg of the CAT construct plasmids by the calcium phosphate precipitation method (30). After 4 hr of exposure of the cells to the precipitated DNA, the medium was replaced and at 48 hr after DNA addition the cells were lysed. Lysis was by 3 cycles of freeze-thaw and this was followed by a heating to 56°C for 10 min and centrifugation at 12,000 x g for 5 min. CAT activity was assayed by incubation of portions of the supernatant fluid with [^{14}C]chloramphenicol and acetyl CoA as previously described (30). CAT activity was quantitated by excision and liquid scintillation counting of the substrate and product regions of silica gel thin layer chromatograms that had been developed with chloroform-methanol (95:5) as solvent. Independent transfections were performed in triplicate or quadruplicate. In early attempts at transfection of the RD4 cells, we experienced variability in CAT expression from transfection to transfection with a given plasmid. This problem appeared to be eliminated by digestion of plasmid preparations with DNase-free RNase (50 µg/ml for 1 hr at 37°C) prior to preparation of the precipitates for transfection. Volumes of lysate assayed for CAT were adjusted such that the maximal [^{14}C]chloramphenicol conversion to acetylated products in a given experiment was less than 50% with equal volumes of lysate being utilized for all samples within an experiment. Aside from repetition of the precipitation/transfection, many of the comparisons have been performed with multiple plasmid preparations of the constructs. The data shown here are representative of all of the comparisons that we have done.

RNA from cells transfected with the 5' deletion series was isolated using the guanidinium isothiocyanate method and polyadenylated mRNA was selected by oligo-dT affinity chromatography (34) after extensive DNase

digestion. A Northern blot of these mRNA samples was probed with the CAT Genblock (Pharmacia) labeled with ^{32}P by random priming. The resultant autoradiographic signals were quantitated by scanning densitometry using a Hoefer GS300 densitometer.

RESULTS

Promoter activity is contained within 86 bp upstream of the TfR mRNA start site.

A clone termed λTR4 was isolated from a human genomic library utilizing as probe a 65 bp fragment from the 5' end of a previously cloned TfR cDNA (pcDTR1). The genomic clone was found by restriction mapping and Southern analysis to contain exon 1 of the TfR gene as well as approximately 12 kb of upstream sequences and 3 kb of TfR gene intron 1 (data not shown). Restriction fragments of λTR4 were cloned upstream of the structural gene for CAT and these constructs were assessed for expression of CAT enzyme after transfection into human RD4 cells. The plasmid -900/+66 CAT [numbered from the TfR mRNA start site (see Fig. 2)] was found to be expressed transiently in RD4 transfectants giving rise to CAT activity approximately 25% of that seen with the plasmid pSV2-CAT (30) that contains the promoter plus enhancer of SV40. This SV40 element is a relatively powerful mediator of gene expression and thus -900/+66 of the TfR gene appeared to possess moderately strong promoter activity. When Bal 31 exonuclease was used to delete most of this fragment (to base -86), the resultant plasmid was expressed at even higher levels (Fig. 1). We have observed reproducibly an approximately 2-fold higher expression with -86/+66 CAT than with -900/+66 CAT. Thus the promoter element within -86/+66 CAT appeared to be only 2-fold lower in strength than the SV40 promoter plus enhancer of pSV2-CAT. In other experiments, we have linked either 1700 bp or 95 bp of upstream TfR gene sequence to the human TfR cDNA and transiently expressed human TfR in murine fibroblasts. Here too the shorter upstream fragment produced approximately 2-fold higher TfR expression than the longer (data not shown). This finding together with the data from the CAT constructs suggest that an element with a negative effect on expression exists in the TfR gene between -900 and -95. To date, we have not mapped or characterized this element in more detail.

Deletion by Bal 31 exonuclease of the 16 bp between -86 and -70 resulted in an approximately 80% decrease in promoter activity (Fig. 1). A further decrease in activity accompanied the removal of the bases between -70 and -41. We have isolated polyA$^+$ mRNA from the RD4 cells 48 hr after transfection with each of these CAT constructs and probed these RNA's for their content of CAT transcripts. In all cases, two transcripts differing by

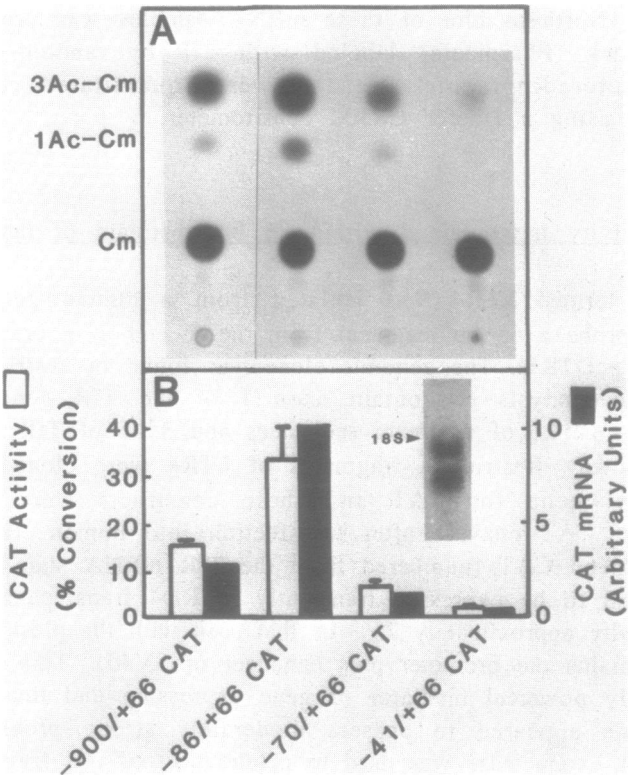


FIGURE 1. Expression of CAT constructs involving progressive 5' deletions of the transferrin receptor promoter region. CAT enzyme activity was assessed in lysates of RD4 cells 48 hr after transfection of the indicated plasmids. An autoradiograph of a representative CAT assay is shown (Panel A). Mobilities on the silica gel thin layer chromatograph of the CAT substrate [^{14}C]chloramphenicol (Cm) and the CAT products 1-acetyl-chloramphenicol (1Ac-Cm) and 3-acetyl-chloramphenicol (3Ac-Cm) are indicated. Enzyme activity was quantitated and expressed as percentage conversion of substrate (Panel B, open bars). Means and standard deviations of four independent transfections are shown. Poly A$^+$ mRNA was isolated from transfected cells and equal amounts analyzed by Northern analysis utilizing a ^{32}P-labeled CAT probe. Relative levels of CAT transcripts present were quantitated by scanning densitometry of the resultant autoradiograph (Panel B, solid bars) and found to correlate with CAT enzyme levels. The inset of panel B shows the relevant portion of the Northern blot of mRNA from RD4 cells transfected with -86/+66 CAT. The mobility of 18 S ribosomal RNA is indicated.

about 400 bases in length were seen (Fig. 1B inset). The larger of these transcripts is of the size expected for mRNA extending from the TfR mRNA start site to the SV40 polyadenylation site in the parent plasmid without

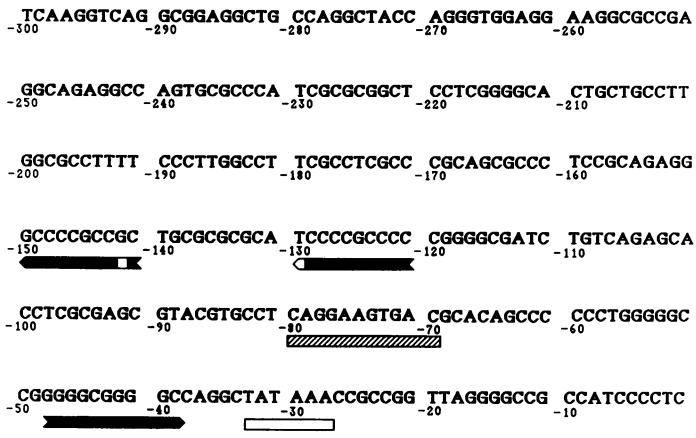


FIGURE 2. Nucleotide sequence of the promoter region of the human transferrin receptor gene. Sequences of both strands were determined by cloning overlapping fragments into M13mp18 and M13mp19 and using a dideoxynucleotide chain termination procedure (32). Numbering is assigned relative to the previously determined mRNA start site (29). Solid arrows indicate the location, orientation, and degree of match to the consensus sequence for binding the transcription factor Sp1 (34). Gaps in these arrows indicate a mismatch with the Sp1 consensus sequence. The hatched bar indicates an 11 bp identity with portions of the enhancers of polyoma virus (37,38) and adenovirus (39,40). The TATA box of the promoter is indicated by the open box.

removal of the plasmid's SV40 small T intron (located 3' of the CAT structural gene). Removal of this intron would yield a transcript of the size of the smaller CAT mRNA that we observe in the transiently expressing RD4 transfectants. The same two transcripts were observed in murine fibroblasts stably transformed with analogous CAT constructs having either the TfR promoter region or the promoter plus enhancer of SV40 (data not shown). Thus, with CAT plasmids differing in their promoters and upstream sequences, the two CAT transcripts were present. We believe that this observation together with the transcript sizes are most consistent with differential splicing rather than alternative upstream mRNA start sites. Within all of the cells transfected with the various CAT constructs involving the TfR promoter region, the ratio of the levels of these two transcripts was always similar. The sum of the two transcripts (or either one individually) correlated well with the CAT enzyme activities in the transfectants (Fig. 1B). Thus it appears that deletion of nucleotides 3' of -86 in the TfR gene resulted in lower expression of CAT mRNA and corresponding decreases in CAT enzyme because these nucleotides are critical for maximal activity of the TfR promoter.

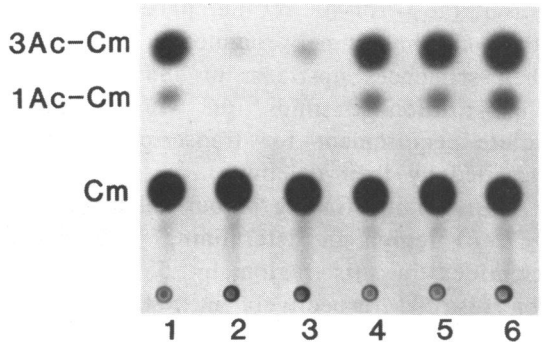
Features of the sequence of the TfR promoter region.

The sequence of 300 bases of the coding strand of the TfR promoter region is shown in figure 2. Several features are noteworthy. A TATA box is located between positions -33 and -29. The region upstream of the TATA box is GC-rich and contains several sequence motifs that have similarity to the consensus binding sequence $5'\begin{matrix}G\\T\end{matrix}\begin{matrix}G\\A\end{matrix}GGCG\begin{matrix}G\\T\end{matrix}\begin{matrix}G\\A\end{matrix}\begin{matrix}G\\A\end{matrix}\begin{matrix}C\\T\end{matrix}3'$ for the transcription factor Sp1 (35,36). The 10 bp segment of the TfR promoter region between -48 and -39 is a perfect match of this Sp1 consensus sequence. Two other Sp1 motifs, located between -130 and -121 and between -150 and -141, are 9 of 10 nucleotides matched to the consensus sequence. Weaker sequence similarities (≤8 of 10 nucleotides matched) also exist within the sequence but are not indicated in figure 2.

Between -80 and -70 of the TfR upstream sequence, there exists a sequence identical to an 11 bp segment of the enhancers of certain DNA tumor viruses including polyoma (37,38) and adenovirus (39,40). This 11 bp identity forms part of a 13 of 15 match (corresponding to TfR gene bases -80 to -66) with sequences within the α-domain of the polyoma enhancer. In adenovirus, the enhancer containing this sequence is termed the E1a enhancer. The expression of the E1a gene is under the influence of the E1a enhancer and the E1a gene product, in turn, is involved in the regulation of other early viral genes (39).

The TfR gene element with sequence similarity to viral enhancers is critical for maximal activity of the TfR promoter.

The sequence within the TfR promoter region that has similarity to portions of known enhancers is of particular interest owing to the results of the 5' deletional analysis shown in figure 1. The marked decrease in promoter activity that accompanies the digestion of -86/+66 CAT to yield -70/+66 CAT correlates with the removal of 10 of the 11 bp of the sequence recognized as being identical to the viral enhancer segments. Internal deletion constructs confirm that it is removal of these nucleotides that has the most profound effect on the TfR promoter activity (Fig. 3). Constructs in which -77/-40 or -77/-70 have been substituted by an Eco RI linker have >90% reduced promoter activity. The segments replaced in these two constructs have in common only 8 bp (i.e. -77 to -70). In the linker-substitution constructs, the fragments with substitution for nucleotides -77/-70 and for -77/-40 appeared to have similar relative activities as promoters (7% and 3% respectively) when compared to the parent plasmid -900/+66 CAT. Moreover, replacement of the TfR gene sequences outside the region of enhancer sequence similarity (i.e. -60/-54 or -53/-42) resulted in promoter activities that were at least as high as that of the complete reference plasmid. The results with the -77/-70 and the -77/-40 linker-substitution plasmids appear to be at some variance with those obtained with the plasmids generated by progressive deletion



TLC Lane	Construct	% Conversion (n=4)	Rel. Act.
1	~ -900/+66 CAT	14.9 + 0.9	100
2	~ -900/+66 CAT (-77/-40 L)	0.5 + 0.1	3
3	~ -900/+66 CAT (-77/-70 L)	1.1 + 0.2	7
4	~ -900/+66 CAT (-60/-54 L)	23.4 + 2.6	157
5	~ -900/+66 CAT (-53/-42 L)	17.2 + 1.9	115
6	-95/+66 CAT (-53/-42 L)	29.1 + 3.5	195

FIGURE 3. Expression of CAT constructs involving internal linker-substitution deletions of the transferrin receptor promoter region. RD4 cells were transfected with the indicated constructs and CAT enzyme levels in lysates of these cells were assessed (a representative autoradiograph is shown). In the linker-substitution constructs, numbers in parentheses [e.g. (-77/-40 L)] indicate the 5' and 3' boundaries of the receptor gene sequence element that has been replaced with the sequence 5' **GGAATTCCC** 3' derived from an Eco RI linker. Means and standard deviations from four independent transfections are expressed as percentage conversion of substrate and as relative activity compared to that seen with the construct -900/+66 CAT. We have found that RD4 cells when transfected with the promoter-less CAT plasmid pJYM0CAT (54) produce CAT enzyme at levels <1% of that produced following transfection with -900/+66 CAT.

in the 5' to 3' direction. A significant reduction (from approximately 20% to less than 5%) in relative promoter activity accompanied removal of the nucleotides between -70 and -41 in the constructs of figure 1. In the linker-substitution constructs, substitution of -77/-70 has a somewhat greater effect and thus this plasmid is closer in promoter activity to that containing the more extensive substitution of -77/-40. One possibility is that bases upstream of -86 modulate the contribution to promoter activity of the nucleotides in the -86/-40 region. When such upstream sequences are absent, as is the case with the 5' deletion plasmids in figure 1, both the

-86/-70 region and the -70/-40 region participate in determination of promoter strength such that some promoter activity remains after deletion of -86/-70. When sequences upstream of -86 are present, as is the case with the linker-substitution plasmids, the -77/-70 nucleotides appears to be a more absolute requirement for transcription with promoter activity more severely curtailed by their deletion.

Clearly the most striking finding of our deletional analysis is the critical nature of the -77/-70 region in determining TfR promoter strength. The removal of nucleotides in this region by 5' progressive deletion or by linker substitution resulted respectively in >80% and >90% reductions in promoter activity. Another interesting feature of the results of this analysis is the apparent dispensability of sequences with high degrees of homology to the consensus sequence for the binding of Sp1. In the plasmid -900/+66 CAT (-53/-42 L), the sequence element (-48/-39) with identity to the consensus sequence for Sp1 binding has been largely replaced. As indicated in figure 2, additional Sp1 motifs (9 of 10 nucleotides matched) exist within this construct further upstream (between -150 and -120). We deleted all three of the best matches to the consensus sequence for Sp1 binding by eliminating bases -900 to -95 using restriction enzyme digestion of the plasmid -900/+66 CAT (-53/-42 L). Rather than decreasing promoter activity, this upstream deletion resulted in an approximately 2-fold increase in promoter activity reminiscent of that seen when -900/-86 was removed in the 5' deletion series of figure 1.

A fragment of the TfR promoter region in either orientation augments gene expression in its native context.

Thus far we have demonstrated that, at a minimum, the nucleotides -77/-70 of the TfR promoter region are critical for maximal activity of the TfR promoter. These nucleotides are common to deletions in two 5' Bal 31 deletion constructs (Fig. 1) and two internal linker-substitution constructs (Fig. 3) all of which display >80% reduced promoter activities. Our sequencing indicates that the TfR promoter region contains, at -80/-70, an 11 bp identity to segments of the enhancers of certain DNA tumor viruses. These nucleotides of the TfR gene are part of a 13 of 15 match to nucleotides within the α-domain of the polyoma enhancer. Enhancers have come to be defined as cis-acting elements that activate both their own and heterologous promoters in a fashion that is relatively independent of the orientation of the enhancer element and its distance from the promoter (41). Often one enhancer orientation produces higher activity than does the other but both increase transcription relative to a promoter lacking the enhancer element.

We have isolated the 72 bp fragment that contains bases -119/-54 of the TfR promoter region along with a portion of an exogenous Eco RI linker and cloned this fragment in both orientations just upstream of base -53 of

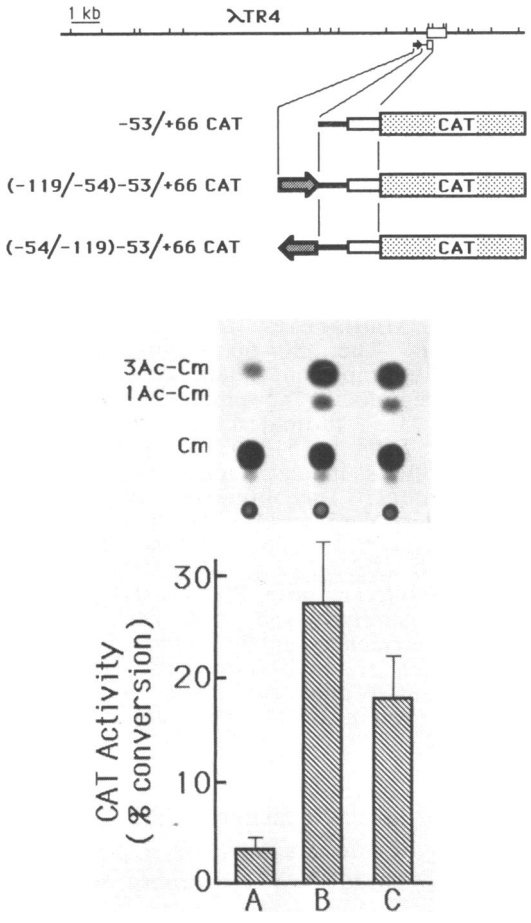


FIGURE 4. Effect of orientation of a fragment containing bases -119 to -54 of the transferrin receptor promoter region on activation of the receptor promoter. Plasmids containing TfR gene sequences -119/-54 attached in either the native (lane B) or inverted (lane C) orientation to -53/+66 CAT were compared with -53/+66 CAT (lane A) in transfections of RD4 cells. CAT enzyme levels in lysates of these cells were assessed (a representative auroradiograph is shown). Means and standard deviations from three independent transfections are expressed as percentage conversion of substrate (hatched bars). Shown above is a schematic representation of these constructs and their relationship to the insert of the parental phage clone λTR4 in which exon 1 of the receptor gene is indicated by the open box. Tic marks on the drawing of λTR4 are restriction enzyme cleavage sites, the details of which are largely unrelated to the present study.

the TfR promoter region. The effect of these insertions is to have this portion of the TfR gene in both orientations at approximately its native position relative to the remainder of the promoter. The native orientation

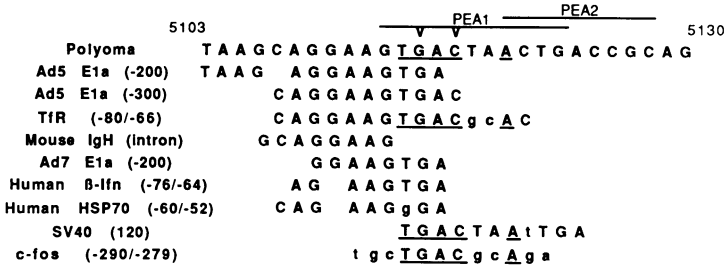


FIGURE 5. Sequence similarities to the -80/-66 element of the TfR promoter region. The -80/-66 region of the TfR promoter is compared to functionally important elements from a number of other enhancer and promoter regions. The positions of the various elements in their native contexts are indicated in parentheses. The segment encompassing nucleotides 5103 to 5130 of polyoma is used as a reference sequence. Upper case letters indicate common nucleotides and lower case letters indicate mismatches. The binding sites for PEA1 and PEA2 determined by Piette and Yaniv (44) are indicated above the polyoma sequence with critical bases 5115 and 5117 (see text) designated by arrow heads. The common bases (**TGACxxA**) in the elements of polyoma, SV40, and c-*fos* (all shown to interact with PEA1) are underlined. This motif in the the TfR sequence is similarly indicated. Data indicating the functional importance of the listed elements can be found in the following references: polyoma enhancer (37,38); adenovirus E1a enhancer (39,40); mouse immunoglobulin heavy chain enhancer (49); human ß-interferon enhancer (51); human heat shock protein HSP70 promoter (50); SV40 enhancer (48); and c-*fos* enhancer (45,46).

of -119/-54 is contained in the construct (-119/-54)-53/+66 CAT whereas the inverted orientation of the fragment is represented in the construct (-54/-119)-53/+66 CAT. These two constructs were compared in RD4 transfections for promoter activity with -53/+66 CAT, that lacks the region of interest. As shown in figure 4, plasmids containing either orientation of the -119/-54 fragment yielded significantly higher (5- to 8-fold) promoter activity than -53/+66 CAT. The native orientation of the fragment gave somewhat higher activity but the difference between the two orientations was of marginal statistical significance.

Another characteristic of prototypical enhancers is the ability to activate a heterologous promoter (41). pA10 CAT 2 is a plasmid containing an SV40 promoter region that lacks the 72 bp repeats that constitute the SV40 enhancer element (31). pA10 CAT 2 has been employed as a recipient to test the ability of DNA fragments to activate the enhancer-less SV40 promoter. When the -119/-54 fragment of the TfR gene was inserted into pA10 CAT 2 and these plasmids transfected into RD4 cells, we saw no increase in CAT enzyme activity over that seen with the enhancer-less pA10 CAT 2 regardless of the orientation of the TfR fragment. Values for

pA10 CAT 2 and for plasmids having the -119/-53 fragment inserted in either orientation into the Bgl II site of pA10 CAT 2 were approximately 5% of those seen for pSV2-CAT that contains the *bona fide* SV40 enhancer element at this location (data not shown). From these data, it would appear that the -119/-54 fragment of the TfR promoter region is capable of activating gene expression in an orientation-independent fashion when the fragment is in its native context but does not display the ability to activate the promoter of SV40. Thus this fragment does not meet all of the criteria (41) that would enable it to be referred to as an enhancer.

DISCUSSION

The 5' deletion constructs (Fig. 1) and the internal linker-substitution constructs (Fig. 3) implicate nucleotides just 5' of -70 as being critical to the maximal activity of the TfR promoter. Marked decreases (>80%) in promoter strength accompany deletion or substitution of these nucleotides. Sequencing of this region (Fig. 2) reveals that the 11 bp between -80 and -70 are identical to portions of the enhancer elements found in certain DNA tumor viruses. The bases between -80 and -66 of the human TfR gene form a 13 of 15 sequence match to nucleotides within the α-domain of the polyoma enhancer. In addition, the TfR promoter region contains several regions having sequence similarity to the consensus sequence that has been identified as the binding site for the transcription factor Sp1 (35,36).

During preparation of this manuscript, another publication appeared which presented the sequence of part of the 5' region of the TfR promoter (extending to -114 of our sequence) as well as DNase footprinting of this region (42). Protection from DNase was observed in the vicinity of -70. It was suggested that this protection corresponded to a sequence (located between -72 and -63 in figure 2) having sequence similarity to a portion of the upstream region of the human dihydrofolate reductase (DHFR) gene. No reference was made to the adjacent and slightly overlapping element that we have identified at -80/-70. It is possible that these two regions form one protein binding site. We have examined the published DNase protection (42) and believe that they are consistent with this possibility. This contention is also supported, albeit indirectly, by inspection of the sequence similarities between the TfR promoter region and the α-domain of polyoma (see figure 5). This region of the polyoma enhancer is the target of at least two different cellular DNA binding proteins termed PEA1 and PEA2 (43). The binding site of PEA1 has been localized to the region 5113 to 5123 of the polyoma genome (43,44). Recently, it has been reported that PEA1 is induced upon differentiation of F9 embryonal carcinoma cells (44). Point mutations in the polyoma enhancer at positions 5115 and 5117 of the polyoma genome abolish PEA1 binding (43). In the region that we have

identified within the TfR promoter region as being critical for full promoter activity, the PEA1 binding site would correspond to TfR nucleotides 3' of -74 with the critical polyoma bases 5115 and 5117 corresponding to TfR nucleotides -72 and -70 respectively (see figure 5). It has been shown that PEA1 also interacts with the enhancer of SV40 and and that of the proto-oncogene c-*fos* (43). It was noted by these authors that the enhancers of polyoma, SV40, and c-*fos* that interact with PEA1 have in common the sequence **TGACxxA** in the PEA1 site. This nucleotide motif is also found at -73/-67 of the TfR promoter. Moreover, these seven bases of the TfR gene (**TGACGCA**) are identical to those of c-*fos*. Indeed, all of the nucleotides that c-*fos* has in common with the PEA1 site of polyoma are contained in the TfR promoter making the TfR promoter a good candidate as a target for PEA1. The results obtained with both types of our deletion constructs (5' progressive deletion and linker-substitution) are consistent with these seven bases of the TfR promoter being of critical importance in promoter activity. Both in -70/+66 CAT and in -900/+66 CAT (-77/-70 L) that have reduced activities, substantial portions of this putative PEA1 site have been eliminated. It is perhaps noteworthy that the TfR (45) and c-*fos* (46,47) genes also share the feature of being regulated in expression by serum addition to growth arrested cells.

Sequence elements with similarity to portions of the enhancers of polyoma virus and adenovirus have been implicated functionally in the enhancer of SV40 (48) and that of immunoglobulin heavy chain (49) as well as within the human HSP70 promoter (50). In addition, very similar sequences are found within the regulatory element of the gene for ß-interferon (51). We believe that our data strongly support a role for a similar sequence motif within the TfR promoter region. In addition to the deletional analysis, we have shown that a fragment of the TfR promoter region that contains the functionally implicated sequence element has one of the properties (i.e. orientation independence) characteristic of viral and cellular enhancers. The fragment -119/-54 in either the native or inverted orientation, augments promoter activity when placed upstream of the remainder of the TfR promoter. However, the same fragment does not activate an enhancer-less SV40 promoter and, therefore, does not have all of the properties of prototypical enhancers.

It has been noted that promoter elements and enhancers overlap both physically and functionally (41). Each type of element appears to be composed of a modular arrangement of shorter sequence motifs. For example, deletion analysis has identified several critical motifs within the SV40 enhancer that individually have little stimulatory activity even though a very powerful enhancer is produced by their association (48). The molecular mechanism that underlies this cooperativity remains obscure. The multiplicity of motifs within the SV40 enhancer may account

for the ubiquity of its activity. In contrast, the polyoma enhancer, which shares only some of the motifs found in the SV40 enhancer, is restricted in activity to certain cell types (37,52). A modular organization of sequence motifs within promoters and enhancers is capable of significantly broadening the combinitorial possibilities in transcriptional control. Our data demonstrate that the -119/-54 fragment of the TfR promoter region contains an element that is critical for maximal activity of this promoter. Moreover, this fragment augments gene expression irrespective of orientation when added to the additional TfR promoter elements contained within -53/+66 CAT. Yet the same fragment is not sufficient to activate the SV40 promoter of pA10 CAT 2. This suggests that the TfR motif contained in the -119/-54 fragment may be used within the TfR promoter in conjunction with another motif that is lacking in the plasmid pA10 CAT 2. If this is the case, our data would indicate that the second motif lies 3' of nucleotide -53 of the TfR promoter region. We are currently investigating this hypothesis.

The involvement of the Sp1 binding sites in TfR expression as judged by our deletion analysis is unclear. DNase protection data has been presented by others (42) indicating that the putative Sp1 site located at -48/-39 (Fig. 2) is a region that is relatively resistant to DNase as would be expected if the region were associated with a DNA binding protein like Sp1. We had anticipated that this putative Sp1 site was most likely involved in TfR expression owing to its complete match to the consensus sequence and its proximity to the TATA box within the promoter region. In the early promoter of SV40 and in the thymidine kinase promoter of Herpes simplex virus, more than one Sp1 binding motif exist. In both cases, it appears that location relative to the TATA box is a crucial determinant of function (36). Nonetheless, the Sp1 site at -48/-39 of the TfR gene could be largely replaced in the construct -900/+66 CAT (-53/-42 L) without a decrease in promoter activity. Our sequencing revealed that two other sequence elements highly matched (9 of 10 nucleotides) to the Sp1 consensus sequence exist within the TfR promoter region 5' of nucleotide -120 (Fig. 2). It was conceivable that these alternative Sp1 sites were being utilized in expression of -900/+66 CAT (-53/-42 L). However, strong promoter activity was still observed when these sites were also deleted. Thus, the construct -95/+66 CAT (-53/-42 L) contains none of the three best matches to the Sp1 motif and yet manifests promoter activity that is about half as strong as that contained in the reference plasmid pSV2-CAT. It should be noted that within the construct -95/+66 CAT (-53/-42 L), there are other sequence elements similar ($\leq$8 of 10 nucleotides matched) to the Sp1 consensus sequence. Indeed, a portion of the linker used in this construct together with adjacent nucleotides forms a new 8 of 10 match to the consensus sequence. This new motif is oppositely oriented relative to the

native -48/-39 element and contains a substitution in nucleotide 4, a hypercritical base of the consensus sequence (36). The multiplicity of potential Sp1 binding sites within the TfR promoter region and the apparent ability of Sp1 to interact with sequences that deviate from the currently accepted consensus sequence make it difficult to rule out Sp1 involvement in TfR expression despite the apparent dispensability in our deletions/substitutions of the three regions most similar to the Sp1 consensus sequence. More direct experiments will be required to assess whether, for example, the construct -95/+66 CAT (-53/-42 L) is capable of interacting with Sp1.

We have identified nucleotides of the TfR promoter necessary for maximal TfR expression. In addition to their involvement in "basal" expression, these sequences are candidates for participation in the regulatory processes that modulate TfR expression in response to iron availability and in response to states of cellular proliferation and/or differentiation. We do not believe that the TfR promoter region is the major site of iron-mediated modulation of TfR expression. This conclusion is based on a construct that we have made between the SV40 promoter and the full-length human TfR cDNA. When murine fibroblasts are stably transformed with this construct, the human TfR expressed in these cells remains sensitive to up-modulation by iron chelation and down-modulation by an iron source. In addition, we have found that removal of DNA corresponding to the 3' untranslated portion of the TfR mRNA ablates iron regulation even when the TfR promoter region is present (53). Similar observations have been documented recently by Owen and Kühn (45). Interestingly, they reported that deletion of the 3' untranslated sequences did not affect the decrease in TfR expression that accompanies arrest of cell growth and suggested that at least two mechanisms are involved in TfR regulation. The regulation of TfR expression in growth arrest was observed in a construct having only 114 bp of TfR upstream sequence. Therefore, it remains a possibility that the nucleotides implicated by our deletion analysis are involved in the regulation of TfR expression during growth arrest.

*To whom correspondence should be addressed

REFERENCES

1. Aisen, P. and Listowsky, I. (1980) Annu. Rev. Biochem. 49, 357-393.
2. Hanover, J.A. and Dickson, R.B. (1985) In Receptor-Mediated Endocytosis. (I. Pastan & M.C. Willigham, Eds.) pp.131-161. Plenum Press. New York, NY.
3. Trowbridge, I.S., Newman, R.A., Domingo, D.L. and Sauvage, C. (1984) Biochem.Pharmacol. 33, 925-932.
4. Bleil, J.D. and Bretcher, M.S. (1982) EMBO J. 1, 351-355.

5. Pearse, B.M.F. 1982. Proc. Natl. Acad. Sci. USA 79, 451-455.
6. van Renswoude, J., Bridges, K.R., Harford, J.B. and Klausner, R.D. (1983) Proc. Natl. Acad. Sci. USA 79, 6186-6190.
7. Harford, J.B., Wolkoff, A.W., Ashwell, G. and Klausner, R.D. (1983) J. Cell Biol. 96, 1824-1828.
8. Dautry-Varsat, A., Ciechanover, A. and Lodish, H.F. (1983) Proc. Natl. Acad. Sci. USA 80, 2258-2262.
9. Klausner, R.D., Ashwell, G., van Renswoude, J., Harford, J.B. and Bridges, K.R. (1983) Proc. Natl. Acad. Sci. USA 80, 2263-2266.
10. Klausner, R.D., van Renswoude, J., Kempf, C., Rao, K., Bateman, J.L. and Robbins, A.R. (1984) J. Cell Biol. 98, 1098-1101.
11. Rao, K., van Renswoude, J., Kempf, C. and Klausner,R.D. (1984) 160, 213-216.
12. Barnes, D. and Sato, G. (1980) Cell 22, 649-655.
13. Chitambar, C.R., Massey, E.J. and Seligman, P.A. (1983) J. Clin. Invest. 72, 1314-1325.
14. Larrick, J.W. and Creswell, P. (1979) J. Supramolec. Struct. 11, 579-586.
15. Trowbridge, I.S. and Omary, M.B. (1981) Proc. Natl. Acad. Sci. USA 78, 3039-3043.
16. Delia, D., Greaves, M.F., Newman, R.A., Sutherland, D.R., Minowada, J., Kung, P. and Goldstein, G. (1982) Int. J. Cancer 29, 23-31.
17. Hu, H.Y., Gardner, J. and Aisen, P. (1977) Science 197, 559-561.
18. Omary, M.B., Trowbridge, I.S. and Minowada, J. (1980) Nature (London) 286, 888-891.
19. Galbraith, R.M. and Galbraith, G.M.P (1981) Immunol. 44, 703-710.
20. Neckers, L.M., Yenokida, G. and James, S.P. (1984) J. Immunol. 133, 2437-2441.
21. Sutherland, R., Delia, D.,Schneider, C., Newman, R., Kemshead, J. and Greaves, M. (1981) Proc. Natl. Acad. Sci. USA 78, 4515-4519.
22. Ward, J. H., Kushner, J.P. and Kaplan, J. (1982) J. Biol. Chem. 257, 10317-10323.
23. Pelicci, P.G., Tabillo, A., Thomopoulos, P., Titieux, M., Vainchenker, W., Rochant, H. and Testa, U. (1982) FEBS Lett. 145, 350-354.
24. Mattia, E., Rao, K., Shipiro, D.S., Sussman, H. and Klausner, R.D. (1984) J. Biol. Chem. 259, 2689-2692.
25. Rao, K.K., Shapiro, D.S., Mattia, E., Bridges, K. and Klausner, R.D. (1985) Molec. Cell. Biol. 5, 595-600.
26. Rao, K., Harford, J.B., Rouault, T., McClelland, A., Ruddle, F.H. and Klausner, R.D. (1986) Molec. Cell. Biol. 6, 236-240.
27. Kühn, L.C., McClelland, A. and Ruddle, F.H. (1984) Cell 37, 95-103.
28. Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G.K and Efstratiadis, A. (1978) Cell 15, 687-701.
29. Schneider, C., Kurkinen, M. and Greaves, M. (1983) EMBO J. 2, 2259-2263.
30. Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) Mol. Cell Biol. 2, 1044-1051.
31. Scholer, H.R. and Gruss, P. (1984) Cell 36, 403-411.
32. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
33. McAllister, R.M., Melnyk, J., Finkelstein, J.Z., Adams Jr., E.C. and Gardner., M.B. (1969) Cancer 24, 520-526.
34. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, New York.
35. Briggs, M.R., Kadonaga, J.T., Bell, S.P. and Tjian, R. (1986) Science 234, 47-52.

36. Kadonaga, J.T., Jones, K.A. and Tjian, R. (1986) Trends Biochem. Sci. 11, 20-23.
37. Herbomel, P., Bourachot, M. and Yaniv, M. (1984) Cell 39, 653-662.
38. Melin, F., Pinon, H., Reiss, C., Kress, C., Montreau, N. and Blangy, D. (1985) EMBO J. 4, 1799-1803.
39. Hearing, P. and Shenk, T. (1983) Cell 33, 695-703.
40. Hearing, P. and Shenk, T. (1986) Cell 45, 229-236.
41. Serfling, E., Jasin, M. and Schaffner, W. (1985) Trends Genet. 1, 224-230.
42. Miskimins, W.K., McClelland, A., Roberts, M.P. and Ruddle, F.H. (1986) J. Cell Biol. 103, 1781-1788.
43. Piette, J. and Yaniv, M. (1987) EMBO J. 6, 1331-1337.
44. Kryszke, M.-H., Piette, J. and Yaniv, M. (1987) Nature (London) 328, 254-256.
45. Owen, D. and Kühn, L.C. (1987) EMBO J. 6, 1287-1293.
46. Treisman, R. (1985) Cell 42, 889-902.
47. Treisman, R. (1986) Cell 46, 567-574.
48. Zenke, M., Grundstrom, T., Matthes, H., Wintzerinth, M., Schatz, C., Wildeman, A. and Chambon, P. (1986) EMBO J. 5, 387-397.
49. Gillies, S., Morrison, S.L., Oi, V.T. and Tonegawa, S. (1983) Cell 33, 717-728.
50. Wu, B., Kingston, R. and Morimoto, R.I. (1986) Proc. Natl. Acad. Sci. USA 82, 629-633.
51. Goodbourn, S., Burstein, H. and Maniatis, T. (1986) Cell 45, 601-610.
52. Ruley, H.E. and Fried, M. (1983) J. Virol. 47, 233-237.
53. Casey, J.L., Di Jeso, B., Rao, K., Klausner, R.D. and Harford, J.B. (1988) Proc. Natl. Acad. Sci. U.S.A., in press.
54. Siebenlist, U., Durand, D.B., Bressler, P., Holbrook, N.J., Norris, C.A., Kamoun, M., Kant, J.A. and Crabtree, G.R. (1986) Mol. Cell. Biol. 6, 3042-3049.