ELSEVIER

Cloning and functional expression of the rabbit transferrin gene promoter

Bilal A.A. Ghareeb [a], Dominique Thépot [a,*], Corinne Delville-Giraud [b], Louis-Marie Houdebine [a]

[a] *Laboratoire de Biologie Cellulaire et Moléculaire, INRA, 78352 Jouy-en-Josas, Cedex, France*
[b] *Laboratoire de Génétique Biochimique, INRA, 78352 Jouy-en-Josas, Cedex, France*

Abstract

The transferrin gene is expressed in all mammals, mainly in the liver. A rabbit genomic library was screened using cDNA probes, and 8 kb of 5′ flanking sequence of the rabbit transferrin gene was cloned upstream of the *cat* reporter gene. The first 200 nucleotides of this promoter were sequenced. The rabbit transferrin promoter is highly homologous to the human and murine ones. Its functional activity was tested in the human hepatic cell line HepG2. Using transitory transfections in these cells, a proximal positive region, a negative region and a distal positive region located between -3.6 and -4.0 kb were identified. This distal positive region sequence is highly conserved with the the human gene enhancer sequence, and contains an HNF3α binding site, the mutation of which totally abolished its effect in HepG2 and HuH7 cell lines. The rabbit transferrin 5′ flanking sequence thus shows a promoter organization very similar to that of the human gene, and the HNF3α binding site in the distal positive region presents the same functional importance in both genes. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: HepG2; HuH7; Liver; Enhancer; Transcription; Tissue-specific expression

1. Introduction

Transferrin is the major iron-binding glycoprotein found in the serum of vertebrates and in the physiological fluids of invertebrates. It is responsible for transporting iron throughout the body from the sites of absorption to the sites of storage or utilization. The major transferrin secretory organ is the liver, but other tissues, mainly those isolated from the circulating system by a barrier, also secrete it at lower levels. In the rat, for example, the levels of transferrin mRNA are 6.5-fold higher in the liver than in the testis and the brain which are isolated by the hemato-testicular barrier and the blood–brain barrier, respectively (Idzerda et al., 1986). Much lower *Tf* gene expression is also detected in many organs such as the heart, stomach, spleen, kidney, oesophagus and salivary glands (Aldred et al., 1987). In all these tissues, the transferrin gene is regulated by numerous factors in a tissue-specific fashion.

The sequences and the transcription factors underlying the transcriptional activity of the *Tf* gene in different cell types were identified in the human gene. The liver-specific expression is governed by the $-125/+1$ promoter region and modulated by a positive $-620/-125$ upstream promoter region, a negative $-1000/-620$ region, and a strong $-3600/-3300$ enhancer region (Schaeffer et al., 1989). The $-125/+1$ liver-specific promoter region contains two binding sites called PRI and PRII. Interactions of HNF4 with PRI, and C/EBPα with PRII result in liver-specific expression (Schaeffer et al., 1993). The negative region is more precisely localized in the $-1000/-819$ region, functions as a silencer and contains a binding-site for yet unknown liver-enriched proteins (Sawaya et al., 1996). The enhancer region is organized into two domains, A and B (Boissier et al., 1991). Domain A is essential for enhancer activity and binds the HNF3α (Augé-Gouillou

* Corresponding author. Tel: 33 134652544; Fax: 33 134652241; e-mail: thepot@biotec.jouy.inra.fr

et al., 1993), while domain B alone has no enhancer activity. In contrast, this enhancer region is inactive in Sertoli cells and in brain cells, due to the absence of HNF3 activity in these cell types. The Sertoli cell-specific expression is controlled by the same $-125/+1$ promoter region (Guillou et al., 1991), but a different combination of transcription factors bind to PRI and PRII (Schaeffer et al., 1993). In brain cells (oligodendrocytes, epithelial choroid plexus cells and B103 cells), a longer $-164/+1$ promoter region with a third binding site called central region I (CRI) is necessary for cell-type specific expression. COUP-TF proteins of the C/EBP family and a protein of the CREB family bind to PRI, PRII and CRI, respectively, in these brain cells (Espinosa de los Monteros et al., 1994; Sawaya and Schaeffer, 1995). Thus, different upstream regions of the gene and different combinations of transcription factors are implicated in the tissue-specific expression of the *Tf* gene.

In addition to tissue-specific expressions, some species-specific differences of expression also exist. For instance, the distribution of cerebral expression of the *Tf* gene is species-specific, and a more recent study outlines the differences in the sites of Tf protein presence in the brain between rat and rabbit (Koeppen et al., 1995). The mammary gland represents another example of species-specific differences in *Tf* gene expression which occurs in the mouse (Chen and Bissell, 1987), rat (Grigor et al., 1990) and rabbit (Puissant et al., 1994), but not sheep (Sanchez et al., 1992) or human (Masson and Heremans, 1971) mammary glands. Moreover, *Tf* expression in the mammary gland is at least as high as in the liver in the mouse (Chen and Bissell, 1987), two-fold higher in the rat (Grigor et al., 1990) and three-fold higher in the rabbit (Puissant et al., 1994).

Thus, the rabbit appears to be an especially interesting species in which to study tissue-specific *Tf* gene expression. In order to determine *Tf* gene expression in the rabbit, we decided, in a first attempt, to investigate its regulation in transfected hepatic cell lines in which the regulation of the *hTf* gene has been well documented and to compare the two. In this paper, we describe the cloning of the rabbit *Tf* promoter and its expression in transfected hepatic cell lines.

2. Materials and methods

2.1. Cloning and sequencing of the rTf gene promoter

A New-Zealand rabbit genomic λ EMBL3 T7/SP6 library (Clontech) was first screened using rabbit transferrin (*rTf*) cDNA nucleotides 3–413 (Banfield et al., 1991) as a probe following standard procedures of hybridization on positively charged nylon membrane Hybond N+ (Amersham). Five clones were isolated. Assuming that the rabbit *Tf* gene has the same

exon/intron structure as the human *Tf* gene, this probe covered the first three exons of the gene and part of the fourth exon. Inserts specific from exon 1 (5′ primer *Tf1.1*: CCACCACCCCTGAGCTACAA; 3′ primer *Tf1.2*: CCAGGGCGGCGCAGGCCAGCA) and from exon 2 (5′ primer *Tf2.1*: GAATGATCACGAGGCCAG; 3′ primer *Tf2.2*: TAGGAGGCCTTCTTCACACAG) were amplified by PCR (1 min at 94°C, 1 min at 55°C, 1 min at 72°C; 30 cycles) and were then used as probes. Two positive clones were selected for further characterization (Fig. 1).

Southern blot analyses, using a 164 nucleotide *Xho*I–*Bgl*II fragment as the probe (Fig. 1), were performed (Fig. 2) with standard conditions of hybridization and washing, and Hybond N+ membranes. The *Sal*I site located in the polylinker at the end of the λ23 left arm and the *Eco*47III site located in the first exon of the *rTf* gene (Fig. 1) were used to clone the 8 kb *Sal*I–*Eco*47III fragment (including *rTf* sequence from position -7916 to position $+22$) in between the *Sal*I site and the blunt-ended *Xba*I site of p*cat* BASIC Vector (Promega) to obtain the p$(-7916,+22)$*rTf-cat* plasmid.

The 10 kb *Xho*I fragment of λ23 (Fig. 1) was cloned in the *Sal*I site of pPolyIII vector and the resulting plasmid used for sequencing. The nucleotide sequence of the promoter region was determined using an Abi Prism Model 377 automated DNA sequencer (Perkin Elmer) and a Perkin Elmer Terminator kit. The two strands were sequenced. The GenBank accession No. of the *rTf* promoter is AF023263.

2.2. Functional expression of the rTf promoter

The p$(-7916,+22)$*rTf-cat* plasmid was double-digested with *Sal*I and *Spe*I, filled in using T4 DNA polymerase and self ligated to obtain plasmid p$(-5389,+22)$*rTf-cat* (Fig. 4). Plasmids p$(-5389,+22)$*rTf-cat*, p$(-5060,+22)$*rTf-cat*, p$(-3566,+22)$*rTf-cat*, p$(-2240,+22)$*rTf-cat*, p$(-1158,+22)$*rTf-cat*, p$(-410,+22)$*rTf-cat* and p$(-355,+22)$*rTf-cat*, respectively, were obtained by self ligation after a double digestion by *Sal*I and either *Asp*718 or *Sac*I or *Hind*III or *Xba*I or *Sma*I or *Xho*I, respectively (Fig. 4). Deletions p$(-4500,+22)$*rTf-cat*, p$(-4200,+22)$*rTf-cat* and p$(-4100,+22)$*rTF-cat* between p$(-5060,+22)$*rTf-cat* and p$(-3566,+22)$*rTf* were generated from p$(-5060,+22)$*rTf-cat* by the exonuclease III/mung bean nuclease deletion system (Stratagene).

Human hepatoma HepG2 cells were cultured in DMEM (GIBCO-BRL) supplemented with 10% fetal calf serum (Boehringer). Cells were plated out at 40% confluence on a 6 cm dish 1 day before transfection. Plasmids were prepared by the PEG method and introduced into cells by the calcium phosphate precipitation method. p*CMV-βgal* plasmid (Clontech) was cotransfected as an internal reference plasmid. Cells were
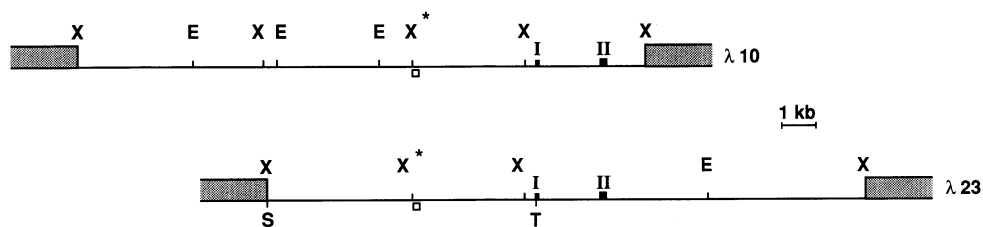
Fig. 1. Restriction map of rabbit transferrin genomic clones. Restriction maps of clones λ10 and λ23 were aligned. Grey boxes correspond to λ arms, black boxes to exons, and the white box corresponds to the 164 nucleotide probe (see Section 2.1). Roman numerals I and II correspond to the first two exons. The other exons in λ23 are not drawn. Abbreviations for the endonuclease sites are: E = *Eco*RI and X = *Xho*I. The position of the *Sal*I and the *Eco*47III restriction sites used to obtain the 8 kb fragment cloned in p*(−7916,+22)rTf-cat* are indicated by S and T.
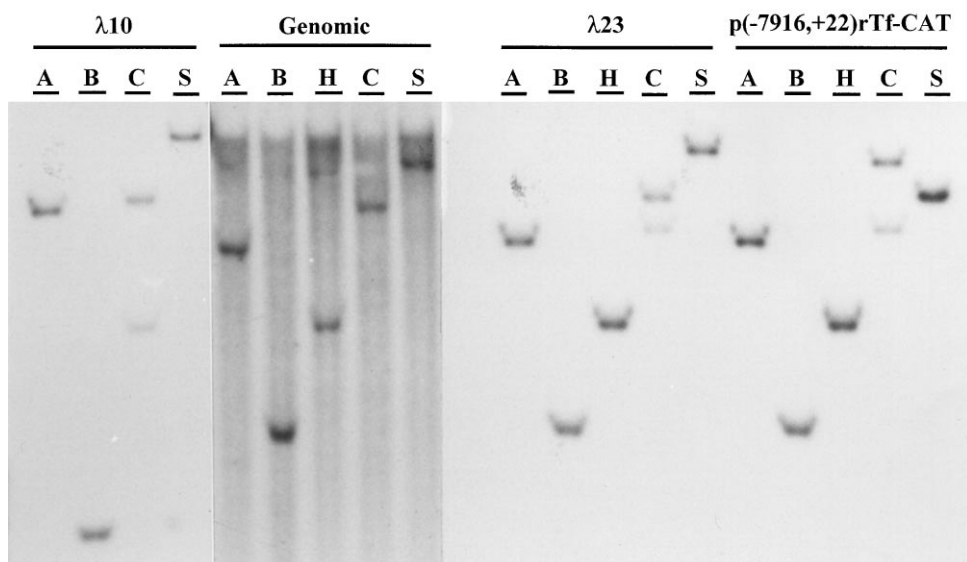

Fig. 2. Southern blot of genomic and λ clones DNA. Digestions of 1 ng of p*(−7916,+22)rTf-cat* DNA or 7 ng of λ DNA mixed with 10 µg of salmon sperm DNA, or of 12 µg of rabbit genomic DNA were deposed on the gel, and the migration ran overnight at 40 V. After transfer to positively charged nylon membrane, DNA was hybridized with the radioactive 164 nucleotide probe (see Section 2.1). The film was exposed to the filter for 18 h. Abbreviations for the endonuclease sites are: A = *Acc*I, B = *Bgl*II, H = *Hin*dIII, C = *Sac*I and S = *Sca*I.

transfected with 6 µg of the *rTf-cat* plasmids and 2 µg of the *βgal* plasmid and harvested 72 h after transfection. Experiments were performed at least three times with two different plasmid preparations. Cells were lysed by three freeze-thawing cycles, and the supernatant was assayed for *βgal* activity following standard protocol and for CAT activity by the method of Nielsen et al. (1989).

2.3. The rTf distal positive region

The *rTf* distal positive region sequence was etablished by sequencing the p*(−5060,+22)rTf-cat* plasmid. The GenBank accession No. of the rabbit *Tf* distal positive region sequence is AF023262. Wild-type and mutated distal positive regions were amplified by PCR (1 min at 94°C, 1 min at 55°C, 1 min at 72°C, 20 cycles) using the following primers:5′ primer EWT GAGATGGTC-GACGCTCTT**G**TTTGCTCAGCTTC, or 5′ primer EMUT GAGATGGTCGACGCTCTT**A**TTTGCT-CAGCTTC; 3′ primer TTGTGGAGCTCTGAGC-CCAGCCCC). After digestion by *Sac*I and *Sal*I, the amplification products were ligated in the *Sal*I–*Sac*I sites of p*(−7916,+22)rTf-cat* to obtain p*(−3930, +22)rTf-cat* and p*Mut(−3930,+22)rTf-cat*. The amplified sequences were verified by sequencing. Human hepatoma HepG2 and HuH7 cells were cultured and transfected as described in Section 2.2.

3. Results

3.1. Cloning of the rTf gene promoter

Using a 5′ *rTf* cDNA probe, five genomic λ clones were first isolated. Since the structure (17 exons, 16 introns) of the *Tf* gene is conserved between human, chicken and medaka, the putative exon boundaries in the *rTf* cDNA were hence deduced by comparison with the human cDNA and gene. PCR primers were designed

to amplify exon 1 and exon 2, and these exons were used as probes to select clones containing promoter sequences. Two clones were selected, characterized and mapped with the two exon probes (Fig. 1). The two clones λ10 and λ23 overlapped by 12 kb. The two restriction maps were in perfect agreement for 8 kb, but diverge for the *Eco*RI restriction sites in their 5′ part. Restriction mapping using other restriction enzymes confirmed that there was no more homology between the two clones upstream of the *Xho*I site (data not shown). A short fragment near the point of divergence was used as the probe in the Southern blot experiment in order to compare the two clones with New-Zealand rabbit genomic DNA. As shown in Fig. 2, *Acc*I, *Bgl*II and *Hin*dIII restriction fragments from λ23 or p(−7916,+22)rTf-cat (a plasmid containing λ23 sequences) were similar to restriction fragments from genomic DNA, but the λ10 restriction fragment lengths were completely different. Since there is no *Sca*I restriction site upstream of the probe in the *rTf* sequence cloned in λ23 or in p(−7916,+22)rTf-cat, the length of the restriction fragments with this enzyme are different for genomic DNA than for λ23 or p(−7916,+22)rTf-cat DNA. Since a *Sac*I restriction site occurs in the probe, two bands were observed with this enzyme in λ10, λ23, or p(−7916,+22)rTf-cat digestions. The stronger band corresponding to the 3′ band is the same for λ clones and genomic DNA because the next 3′ *Sac*I restriction site lies in the first intron, but is different for p(−7916,+22)rTf-cat DNA which does not contain this intron. The weaker band corresponds to the 5′ band and thus differs from λ10 compared to λ23 and p(−7916,+22)rTf-cat. This band is too faint to be clearly visible in the genomic DNA digestion. The same results were obtained with genomic DNA from three unrelated rabbits (data not shown). Thus, clone λ23 was selected for sequencing and further experiments.

3.2. Sequence conservation in mammalian Tf promoters

Fig. 3 shows the alignment of the *hTf*, *mTf* and *rTf* promoter, 5′UTR and ATG sequences. The *rTf* sequence shows 76% identity to the *hTf* sequence and 71% to the *mTf* sequence, when the first 133 bp are considered. The *mTf* sequence shows 71% identity with the *hTf* sequence. 64% of the positions are occupied by the same nucleotide in the three sequences. By comparison, the coding sequence shows 85% identity for the first 15 codons between the three species, and the 5′UTR sequences 69% identity.

The PRI and PRII regions are highly conserved, but strong homology also exists between nucleotides located between these two regions and those just upstream of the PRII region. In contrast, only the central core of the CR region appears well conserved. Upstream of the CR region, the mouse sequence showed no further

homology with either the rabbit or the human sequences (data not shown), whilst the homology between the rabbit and human sequences continues upstream over nearly 50 nucleotides and then disappears. The only striking feature of the rabbit sequence is that the promoter is more compact, with the CR region closer to the PRII region, and the PRI region closer to the TATA box than is observed in the two other species. On the other hand, the 5′UTR sequence seems longer in the rabbit.

3.3. Functional expression of the rTf promoter

In order to localize the regions involved in the control of hepatic expression in the *rTf* gene upstream sequence, plasmids with *cat* gene under the control of various lengths of the *rTf* gene 5′ flanking sequence were constructed. These plasmids were introduced into hepatic HepG2 cells by the calcium phosphate precipitation method, and CAT activity was measured. p*SV2-cat* containing the promoter and enhancer of SV40 and p*cat BASIC* Vector without any promoter were used as positive and negative controls, respectively, in each experiment. Transfection efficiency was monitored by cotransfection of a *βgal* plasmid and measurement of the *βgal* activity.

As shown in Fig. 4, *rTf-cat* constructs were expressed in HepG2 cells at varying levels. Plasmids with 5.1 kb or 5.4 kb of the 5′ flanking sequence were expressed at the highest levels, suggesting a negatively acting region between positions −8.0 kb and −5.4 kb. Successive deletions showed a strong positive region between −4.1 kb and −3.6 kb, a negative region between −3.6 kb and −1.2 kb, and a positive region between −370 bp and the transcription start point. These results were reminiscent of those obtained with the *hTf* gene in the same cells, where an enhancer region (−3.6 kb to −3.3 kb), a negative region (−1 kb to −0.62 kb) and a positive region (−620 to −1 bp) were detected (Schaeffer et al., 1989). Note that in the rabbit, the level of *cat* expression from the first positive region was approximately 50% of that from the distal positive region, as observed in the human *Tf* gene. The most striking result was the presence of a strong distal positive region located between −4.1 kb and −3.6 kb, and consequently this region was further investigated.

3.4. Sequence conservation in the Tf distal positive region

The *rTf* distal positive region was sequenced, and the *rTf* distal positive region sequence was compared with the *hTf* enhancer sequence (Fig. 5). The two sequences were highly conserved, with 215 (73%) nucleotides identical at the same position. However, this conservation was not homogenously distributed over the whole sequence, some regions showing as much as 90% identity

```
                                              _____CR_____
HUMAN  -220  GCTTCATGTCCCTTCCCATCAACATTTCTGTGCTGGACTCCTTCCAC
MOUSE  -213  TGTGTGTGTGTGTGTGTGTGATTGTCGTGGTAGTGGATCC----AT
RABBIT -198  CCTTTGTTCCCGTTCTCAG-AGCACTTTTGTGTTGGATCCC----AC

             --T---T-----T-------A--------GT--TGGA--CC----A-


TCGCGGGTCGTCTCCAGAGCTCAGAAAATGAGGTGATCAGTGGGACGAGTAAGGAAG-GG
GCGCAGTTC-TGTCCCACAGTCAGGAAATGAGGTGATCATCAGGGCAAGAAAGGAAG-GG
GCTCCGG-----TCCAG----CAGGAAATG----TACCAGGGGGGCGAGGAAGGAAGAGT

-C-C-G------TCC------CAG-AAATG-----A-CA---GG-C-AG-AAGGAAG-G-


_____PRII_____          _____PRI_____
GGGTTGGGAGAGGGGCGATTGGGCAACCCGGCTGCACAAACACGGGAGGTCAAAGATTGC
GG--TGGATGCGGGGTGATTGGGCAATTGGACTGCGCAAACACAAGAGGTCAAAGATTGC
GGGACGGGCGAGGGCCGATTGGGCAACCGCGCTGCGCAAACACGCGAGGTCAAAGATCGC

GG---GG--G-GGG--GATTGGGCAA-----CTGC-CAAACAC--GAGGTCAAAGAT-GC


_____                                                    +1
GCCCAGCCCGCCCAGGCCGGGAATGGAATAAAGGGACGCGGGGCGCC-GGAGGCT   GCA
GCCCAGCCGGCCCCCTCCCTCCGTG-CATAAAGGCGAGCGGCGCACGGGCAGCGT   GCG
GCCTGGCCCGCCCCAACGGC------GATAAAGGGGCGCGGGGCGCGGGCAGGGT   GCG

GCC--GCC-GCCC---C----------ATAAAGG---GCGG-GC-C--G-AG--T   GC-


CAGAAGCGAGTCCGACTGTGCTCGCTGCTCAGCGCCG-CACCC-------GGAAG   ATG
CAGAAGCGGGTCGGTCTGTACTCCCCGCCTCCTCGCCA-CACAC-ACACCGAGAGG   ATG
CAGAAGCGGGTCGGTCAGCGCTCGCCGCTCCTCGCCACCACCCCTGAGCTACAAG   ATG

CAGAAGCG-GTC-G-C-G--CTC-C-GCTC--CGCC--CAC-C---------A-G   ATG
```

Fig. 3. Alignment of human, murine and rabbit transferrin sequences. By convention, the G residue experimentally determined as the transcription start site in the human gene was assigned +1 and chosen as reference for the comparison. Please note that it does not correspond to the experimentally determined transcription start site in the mouse gene (Idzerda et al., 1989). The sequences of oligonucleotides corresponding to each of the transcription factor binding sites in the human sequence (Schaeffer et al., 1993) are indicated by lines, and nucleotides the mutation of which abolished the binding of the protein are in bold. The *Eco*47III restriction site in the rabbit sequence used for the construction of plasmid p(−7916,+22)rTf-cat (see Section 2.1) is underlined.

(regions 55–119 and 265–293), and some being less well conserved. In the *hTf* gene enhancer, five regions (Ia, Ib, II, III and IV) were protected by liver nuclear extracts in DNase I footprinting experiments (Boissier et al., 1991). Human regions Ia, Ib, II, III and IV showed 85%, 95%, 82%, 53% and 71% identity, respectively, with the corresponding region in the rabbit distal positive region. In the *hTf* gene, region Ia was demonstrated to be crucial for enhancer activity (Boissier et al., 1991) and was bound by the transcription factor HNF3α (Augé-Gouillou et al., 1993). The role of the HNF3 binding site in the rabbit *Tf* distal positive region was therefore investigated.

3.5. Importance of the HNF3 binding site

Two new *rTf-cat* plasmids, p(−3930,+22)rTf-cat and pMut(−3930,+22)rTf-cat, were constructed. In p(−3930,+22)rTf-cat, the *rTf* gene 5′ flanking sequence stopped just upstream of region Ia. pMut(−3930, +22)rTf-cat is identical to p(−3930,+22)rTf-cat, except for a two-nucleotide mutation in the HNF3 binding-site of region Ia (Fig. 5) which prevents the binding of the protein (Augé-Gouillou et al., 1993). These plasmids were transfected into HepG2 and HuH7 cells as described previously. Fig. 6 shows that the mutation totally abolished the positive effect, both in HepG2 cells (Fig. 6A) and in HuH7 cells (Fig. 6B).

With the mutated distal positive region, levels of CAT activity were similar to those without distal positive region.

4. Discussion

4.1. Cloning and sequence of the rTf gene promoter

During the cloning of the *rTf* gene promoter, two clones were isolated. Their restriction maps were outstandingly divergent in their 5′ region. Two explanations are possible, either a polymorphism at the *Tf* locus, or an artifact during the construction of the library. Since the rabbit genomic library (Clontech) was constructed from a pool of DNA from several rabbits, polymorphism might arise. Tf polymorphism at the protein level was known in wild rabbits, but not in New-Zealand rabbits (Zaragoza et al., 1987). This cannot rule out the possibility of a polymorphism in the 5′ distal region and not in the coding sequences, but this seems unlikely. Moreover, no nucleotide difference was found between λ23 and λ10 in the *rTf* proximal promoter sequence, or in the first intron sequence (data not shown). Thus, it is most likely that clone λ10 was rearranged, although this remains to be confirmed.

The *rTf* promoter sequence showed a greater percentage of identity to the *hTf* promoter than to the *mTf*
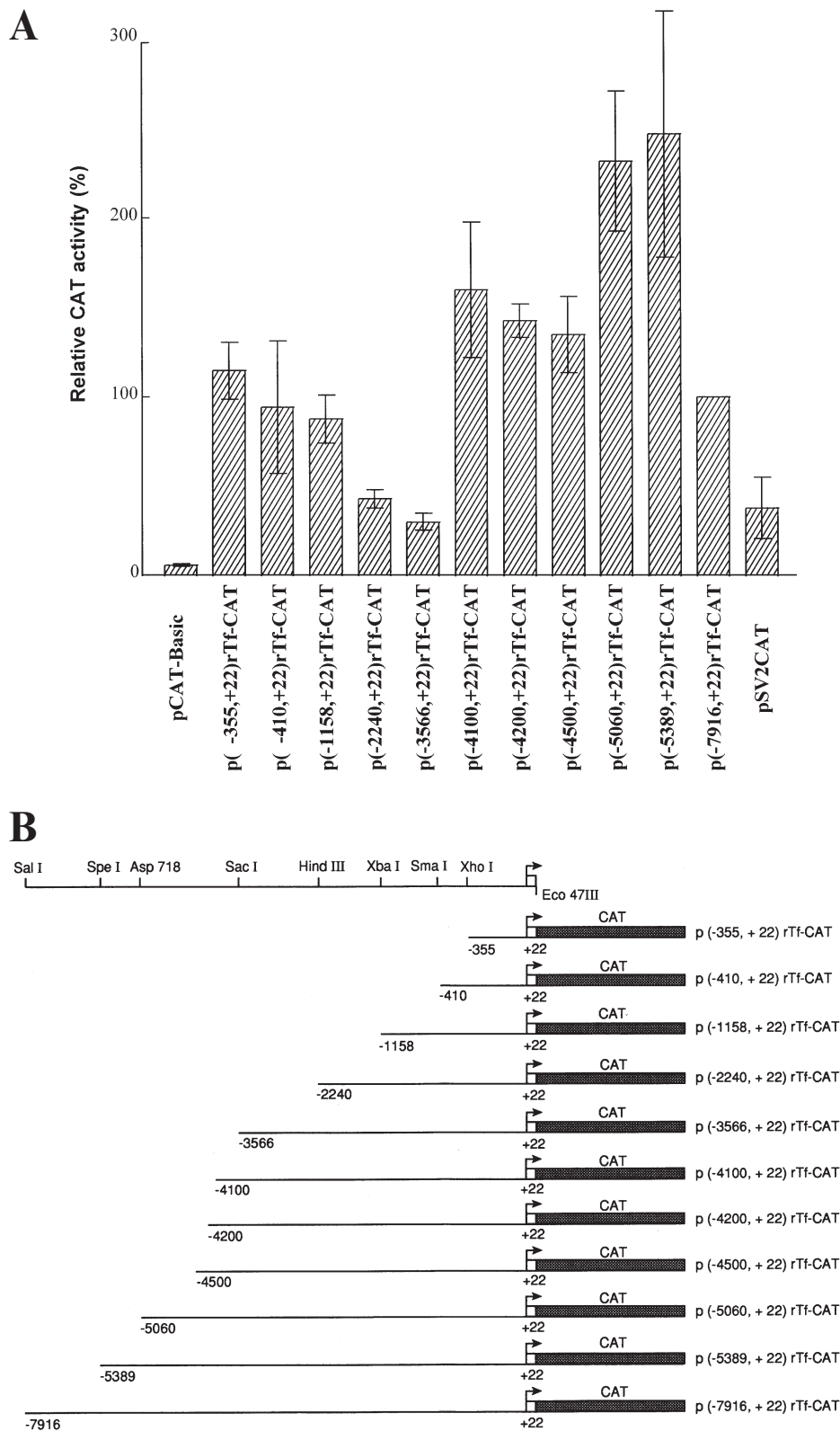
A

B

Fig. 4. (A) Histogram of the observed CAT activity in extracts of HepG2 cells transfected by *rTf* and control plasmids. Control plasmids are described in Section 3.3. Activities were determined as described in Section 2.2, and were expressed relative to that expressed with the p*(−7916, +22) rTf-cat* plasmid (100%). (B) Representation of the *rTf-cat* plasmids containing 5′ deletions of the *rTf* upstream region from −7916 to +22 bp. The restriction sites used and the position of the 5′ end of the deletions are indicated.

```
                                                                      ====
GGTCAGGCAGAGGACACTGGGTGGGGGACAGGAGGAACTGCTCAAAACAGACAGAGGCTC  +60
*  * ***** **   *** ***** *  *  ****** *   * ****** ******
GACACGGCAGGGG-TGGGGGGGGGGGGGGAGGCACCAACTGC----AGCAGACACAGGCTC -3927

         _____Ia_____            _____Ib_____
TTTGTTTGCTTTGCTTCTGTGTCAACTGGGCAACATTTGGAAACAACAAATATTGGTTCA  +120
*********  ********* **  ** ****************************
TTTGTTTGCTCAGCTTCTGTGTAAATTGGGCAACATTTGGAAACAACAAATATTGGTTCT -3867

                                  _____II_____
GAGGCCCACTGCTTTCTTACCCACCTCCTGCTGGTCAGCTTTTCCAGCTTTCCTGCACGT  +180
**  * *** *** *** *  ****** *** ***** *****   *** *** ********
GAAGTCCAGTGCCTTCCTGCCCACCCCCTTCTGGGCAGCTTCCCCAACTGCCCTGCACGT -3807

          _____III_____
ACACACAAGCGCACGTATTTCTTTAAGGGGCCAACTCGTGG--TTGAAATTTGAGAGTTTT  +238
 *******   *   *  **** **** *  *   *  ***   **  ** ***** ***
CCACACAATGCCCTCTGCTTCTCTAAGAGGCCAGCCCTGGGGCTGGAAGAGGAGAGCTTT -3747

_____IV_____
GCCCTGTTTGCGGTGAAAGATTTGGC-TCATGCTTGGGTTGGTCTAGAACTGGAAGCTT  +296
 *****    ****  * * ***  *  ********** ******** ****** ***
ATTCTGTTCCTGGTGCAGGTTTTTGGGTCATGCTTGTGTTGGTCTGGAACTGAAAGTCC -3688
```

Fig. 5. Alignment of human enhancer sequence (above) and rabbit transferrin distal positive region sequence (below). An asterisk indicates the identity of a nucleotide in both sequences. Sequences protected from Dnase I digestion in human sequence (Boissier et al., 1991) are indicated by lines and roman numerals. The two nucleotides in the HNF3α binding site, the mutation of which abolished the binding, are in bold.
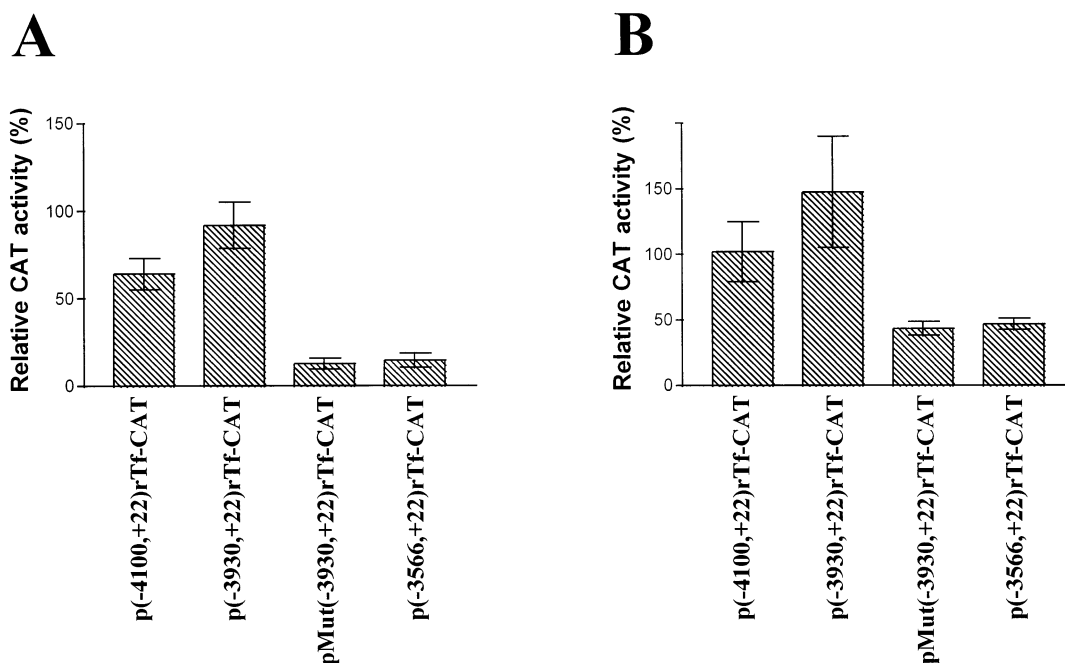

Fig. 6. Histogram of the observed CAT activity in extracts of HuH7 (A) or HepG2 (B) cells transfected by *rTf* plasmids. Activities were determined as described in Section 2.2.

promoter. This is somewhat surprising, as rabbits are lagomorphs which, like mice, are classified near to rodents. However, some recent molecular phylogeny data support the hypothesis that the rabbit is more closely related to the human than to the mouse (Graur et al., 1996). The mouse gene does show some features in common with the rabbit gene in the structure of the *Tf* promoter (a longer leader sequence, four nucleotides lacking in the 3′ side of the CR region), whilst others are in common with the human gene (more nucleotides between the TATA box and the PRI region, and between

PRII and CR regions), and TG repeats upstream of the CR region are specific. However, it is not possible to establish a clear phylogenetic relationship between the three species from the study of only a few nucleotides of a single gene.

In the *hTf* promoter, the two regions controlling liver-specific expression are PRI and PRII which bind transcription factors of the steroid hormone receptor superfamily (HNF4 and COUP-TF) and of the C/EBP family, respectively. These two regions are highly conserved in the three mammalian genomes we have

considered in this report. PRII is also conserved between the chicken and salmon genomes, whereas PRI is clearly recognizable in sequences from chicken but is less well conserved in sequences from salmon (Kvingedal, 1994). However, a CTGCRCAAACAC sequence between the PRI and PRII regions, and a CAGRAAATG and a AAGGAAG sequences between the PRII and CR regions, are as strongly conserved in the three mammalian genomes examined as the PRI and PRII regions, although no function has been attributed to them so far.

Tf gene expression occurs in the mammary gland in mouse (Chen and Bissell, 1987), rat (Grigor et al., 1990) and rabbit (Puissant et al., 1994), but not in human (Masson and Heremans, 1971). Two hypotheses can explain this difference. Either mouse and rabbit, but not human mammary glands express a transcription factor allowing *Tf* gene expression in this organ, or the human gene lacks a sequence susceptible to bind a putative mammary gland-specific transcription factor. If the latter explanation is true, such a sequence is not apparent in the promoter region since no sequence common to the rabbit and murine gene and absent in the human gene was found. However, this does not rule out this hypothesis, as such a sequence might be located in a regulatory region other than the proximal promoter. Since the expression of the transgene in the mammary gland has never been investigated in transgenic animals with a *hTf* 5′ flanking region, the first hypothesis can neither be confirmed nor dismissed, and it remains an open question.

4.2. Functional expression of the rTf promoter

In the hepatic HepG2 cells, the 5′ deletion analysis revealed the existence of at least three regulatory regions in the *rTf* 5′ flanking sequence: a positive region, a negative region and a strong distal positive region region. These three regions were observed clearly in all independent experiments. In contrast, in some experiments, the CAT activity was higher after transfection with p(−4500, +22)rTf-cat than after transfection with p(−5060, +22)rTf-cat, even if the mean values of all the experiments showed the opposite result. We thus cannot conclude that the difference between p(−4500, +22)rTf-cat and p(−5060, +22)rTf-cat is significant. A negative region between −5.4 kb and −8.0 kb is more likely, because it was observed in all but one experiment.

With the human gene, the CAT activity directed by the p*SVcat* plasmid was 4–5-fold higher than the maximal CAT activity directed by the *hTf-cat* plasmids. In our experiments, the maximal CAT activity directed by the *rTf-cat* plasmids was 5-fold higher than those directed by the p*SVcat* plasmid. The low level of CAT activity with p*SVcat* plasmid results from the competition for transcription factors between p*SVcat* and

p*CMV-βgal* promoters. This is especially true for the ubiquitous transcription factor Sp1, which is known to bind to both promoters. The use of *pCMV-βGal* plasmid as an internal reference plasmid presents the problem of possible interactions between *Tf* and viral regulatory elements. An Sp1 binding site was described in the *hTf* upstream region from −273 to −266 (Adrian et al., 1996) and recombinant human Sp1 protein is known to bind this site. The precise identity of the liver protein binding to this Sp1 site, however, has not been clearly demonstrated since the two DNA–protein complexes did not exactly comigrate and the anti-human Sp1 antibody did not recognize the DNA–liver protein complex. It is thus possible that, in liver, this site was bound by another member of the Sp1 family, such as Sp2, Sp3, Sp4 or BTEB. Hence, the protein which binds to this Sp1 site in the *hTf* promoter remains unknown and the functional importance of this site in the *hTf* gene expression is not demonstrated. The presence of such a site in the *rTf* promoter also remains to be shown. The general structure of the *rTf* and *hTf* promoters seems roughly similar and the activity of the *hTf* promoter was etablished without cotransfection of a reference plasmid. This suggests that major interactions between *rTf* and viral regulatory elements did not occur. In order to completely rule out this possibility, a control experiment was performed in the absence of the *pCMV-βGal* plasmid. The presence of the proximal positive region between +22 bp and −370 bp, the negative region between −1.2 kb and −3.6 kb, the strong distal positive region between −3.6 kb and −4.1 kb, and the negative region between −5.4 kb and −8.0 kb was confirmed, as well as the importance of the HNF3 binding site in the distal positive region (results not shown).

The *rTf* gene shows the same structural organization of the regulatory regions as the human *Tf* gene. The order of the three regions is the same, and only the precise location of each region differs slightly. The rabbit distal positive region is located 400 bp further upstream than in the human. This distal positive region sequence is as highly conserved between both species as the proximal promoter sequence. Moreover, the HNF3α binding site of the *rTf* and *hTf* distal positive regions plays an essential role. The distal positive region is thus a new example of a flanking evolutionarily conserved sequence or FECS (Phinney et al., 1995). These FECS were demonstrated to encode transcriptional regulatory functions and perhaps other functions (Phinney et al., 1995). The characterization of FECS appears to be a good approach with which to identify transcriptional control regions that fail to function in transient expression assays, particularly those important for expression in transgenic mice where chromatin structure is essential. Discrepancies in the measured activities between experiments using a transient expression assay and those in

transgenic mice are very common. In the case of the albumin gene, for instance, an enhancer located beween 8.5 and 10.4 kb upstream of the transcription initiation site was essential for high liver expression in transgenic mice, when 68 bp with only an HNF1 binding site and a TATA box were sufficient in hepatoma cell lines. For the *hTf* gene, there is some evidence that expression in the liver of transgenic mice requires a longer 5′ flanking region than transient expression in transfected cells where 125 bp are sufficient (Guillou et al., 1991). A $-139/+50$ fragment of *mTf* is sufficient for expression in the liver of transgenic mice (Idzerda et al., 1989), but a 125 bp *hTf* 5′ flanking region fails to support liver expression in transgenic mice (Cassia et al., 1997) and a $-152/+46$ fragment of this gene leads to a hepatic expression of the transgene at least 1000-fold lower than those obtained with a $-622/+46$ fragment (Adrian et al., 1990). The reasons why the $-152/+46$ fragment and the $-622/+46$ fragment give such different results in transgenic mice, whereas the $-125/+39$ fragment and the $-620/+39$ fragment give similar results in transient experiments in hepatic cells, are poorly understood.

In order to identify other conserved sequences between the *rTf* and the *hTf* genes, we undertook to sequence the entire 8 kb fragment of *rTf* promoter, and we are currently looking for the regions important in the mammary-specific expression, both in transgenic mice and in the transient expression assay system in transfected mammary cells.

4.3. Conclusion

(1) The rabbit *Tf* 5′ flanking sequence shows the same promoter organization as the human gene, with a proximal positive region, a negative region and a distal positive region region.

(2) The first 200 bp of the gene are well conserved between human, mouse and rabbit.

(3) The rabbit distal positive region is highly homologous to its human counterpart and the HNF3α binding site presents the same functional importance.

Acknowledgement

References

Adrian, G.S., Bowman, B.H., Herbert, D.C., Weaker, F.J., Adrian, E.K., Robinson, L.K., Walter, C.A., Eddy, C.A., Riehl, R., Pauerstein, C.J., Yang, F., 1990. Human transferrin. Expression and iron modulation of chimeric genes in transgenic mice. J. Biol. Chem. 265, 13344–13350.

Adrian, G.S., Seto, E., Fischbach, K.S., Rivera, E.V., Adrian, E.K., Herbert, D.C., Walter, C.A., Weaker, F.J., Bowman, B.H., 1996. YY1 and Sp1 transcription factors bind the human transferrin gene in an age-related manner. J. Gerontol. A. Biol. Sci. Med. Sci. 51, B66–B75.

Aldred, A.R., Dickson, P.W., Marley, P.D., Schreiber, G., 1987. Distribution of transferrin synthesis in brain and other tissues in the rat. J. Biol. Chem. 262, 5293–5297.

Augé-Gouillou, C., Petropoulos, I., Zakin, M.M., 1993. Liver-enriched HNF-3 alpha and ubiquitous factors interact with the human transferrin gene enhancer. FEBS Lett. 323, 4–10.

Banfield, D.K., Chow, B.K., Funk, W.D., Robertson, K.A., Umelas, T.M., Woodworth, R.C., MacGillivray, R.T., 1991. The nucleotide sequence of rabbit liver transferrin cDNA. Biochim. Biophys. Acta 1089, 262–265.

Boissier, F., Augé-Gouillou, C., Schaeffer, E., Zakin, M.M., 1991. The enhancer of the human transferrin gene is organized in two structural and functional domains. J. Biol. Chem. 266, 9822–9828.

Cassia, R., Besnard, L., Fiette, L., Espinosa de los Monteros, A., Avé, P., Py, M.C., Huerre, M., de Vellis, J., Zakin, M.M., Guillou, F., 1997. Transferrin is an early marker of hepatic differentiation and its expression correlates with the postnatal development of oligodendrocytes in mice. J. Neurosci. Res. 50, 421–432.

Chen, L.H., Bissell, M.J., 1987. Transferrin mRNA level in the mouse mammary gland is regulated by pregnancy and extracellular matrix. J. Biol. Chem. 262, 17247–17250.

Espinosa de los Monteros, A., Sawaya, B.E., Guillou, F., Zakin, M.M., de Vellis, J., Schaeffer, E., 1994. Brain-specific expression of the human transferrin gene. Similar elements govern transcription in oligodendrocytes and in a neuronal cell line. J. Biol. Chem. 269, 24504–24510.

Graur, D., Duret, L., Gouy, M., 1996. Phylogenetic position of the order Lagomorpha. Nature 379, 333–335.

Grigor, M.R., McDonald, F.J., Latta, N., Richarson, C.L., Tate, W.P., 1990. Transferrin-gene expression in the rat mammary gland. Independence of maternal iron status. Biochem. J. 267, 815–819.

Guillou, F., Zakin, M.M., Part, D., Boissier, F., Schaeffer, E., 1991. Sertoli cell-specific expression of the human transferrin gene. Comparison with the liver-specific expression. J. Biol. Chem. 266, 9876–9884.

Idzerda, R.L., Behringer, R.R., Theisen, M., Huggenvik, J.I., McKnight, G.S., Brinster, R.L., 1989. Expression from the transferrin gene promoter in transgenic mice. Mol. Cell. Biol. 9, 5154–5162.

Idzerda, R.L., Huebers, H., Finch, C.A., McKnight, G.S., 1986. Rat transferrin gene expression: tissue-specific regulation by iron deficiency. Proc. Natl. Acad. Sci. USA 83, 3723–3727.

Koeppen, A.H., Dickson, A.C., McEvoy, J.A., 1995. The heterogeneous distribution of brain transferrin. J. Neuropathol. Exp. Neurol. 54, 395–403.

Kvingedal, A.M., 1994. Characterization of the 5′ region of the Atlantic salmon (*Salmo salar*) transferrin-encoding gene. Gene 150, 335–339.

Masson, P.L., Heremans, J.F., 1971. Lactoferrin in milk from different species. Comp. Biochem. Physiol. 39, 119–129.

Nielsen, D.A., Chang, T.C., Shapiro, D.J., 1989. A highly sensitive, mixed-phase assay for chloramphenicol acetyltransferase activity in transfected cells. Anal. Biochem. 179, 19–23.

Phinney, D.G., Tseng, S.W., Ryder, K., 1995. Complex genetic organization of JunB: multiple blocks of flanking evolutionarily conserved sequence at the murine and human loci. Genomics 28, 228–234.

Puissant, C., Bayat-Sarmadi, M., Devinoy, E., Houdebine, L.M., 1994. Variation of transferrin mRNA concentration in the rabbit mammary gland during the pregnancy–lactation–weaning cycle and in cultured mammary cells. A comparison with the other major milk protein mRNAs. Eur. J. Endocrinol. 130, 522–529.

Sanchez, L., Lujan, L., Oria, R., Castillo, H., Perez, D., Ena, J.M., Calvo, M., 1992. Synthesis of lactoferrin and transport of transferrin in the lactating mammary gland of sheep. J. Dairy Sci. 75, 1257–1262.

Sawaya, B.E., Aunis, D., Schaeffer, E., 1996. Distinct positive and negative regulatory element control neuronal and hepatic transcription of the human transferrin gene. J. Neurosci. Res. 43, 261–272.

Sawaya, B.E., Schaeffer, E., 1995. Transcription of the human transferrin gene in neuronal cells. Nucleic Acids Res. 23, 2206–2211.

Schaeffer, E., Boissier, F., Py, M.C., Cohen, G.N., Zakin, M.M., 1989. Cell type-specific expression of the human transferrin gene. Role of promoter, negative, and enhancer elements. J. Biol. Chem. 264, 7153–7160.

Schaeffer, E., Guillou, F., Part, D., Zakin, M.M., 1993. A different combination of transcription factors modulates the expression of the human transferrin promoter in liver and Sertoli cells. J. Biol. Chem. 268, 23399–23408.

Zaragoza, P., Arana, A., Amorena, B., 1987. Relationship between rabbit transferrin electrophoretic patterns and plasma iron concentrations. Anim. Genet. 18, 51–62.