Original Article

HNF-4α Regulates Expression of Human Ornithin Carbamoyltransferase through Interaction with Two Positive *Cis*-Acting Regulatory Elements Located in the Proximal Promoter

(ornithine carbamoyltransferase / OTC / HNF-4α / transcription / regulation / promoter / DNase I footprinting / EMSA)

O. LUKŠAN[1], L. DVOŘÁKOVÁ[2], M. JIRSA[1]

[1]Laboratory of Experimental Hepatology, Institute for Clinical and Experimental Medicine and Institute of Medical Biochemistry and Laboratory Diagnostics, First Faculty of Medicine, Charles University in Prague, Czech Republic
[2]Institute of Inherited Metabolic Disorders, First Faculty of Medicine, Charles University in Prague and General University Hospital in Prague, Czech Republic

Abstract. *OTC* encodes ornithine carbamoyltransferase, mitochondrial matrix enzyme involved in the synthesis of urea. The tissue-specific expression of *OTC* in the liver and intestine is dependent on the interaction of *OTC* promoter with an upstream enhancer. HNF-4 and C/EBPβ are crucial for this interaction in the rat and mouse. In the present study we focused on characterization of elements involved in the regulation of *OTC* transcription in human. Using a set of 5'-deleted promoter mutants in a reporter assay we identified two positive *cis*-acting regulatory elements located at c.-105 and c.-136 within the human *OTC* promoter. Both are essential for the transcriptional activity of the promoter itself and for the interaction with the enhancer. Protein binding at the corresponding sites was confirmed by DNase I footprinting. Electromobility shift assay with a specific competitor and anti-HNF-4α antibody identified the DNA-protein binding sites as HNF-4α recognition motifs. A third HNF-4α binding site has been found at the position c.-187. All three HNF-4α binding sites are located within 35 bp upstream of the transcription start sites at positions c.-95, c.-119 (major) and c.-169 (minor). A series of C/EBPβ recognition motifs was identified within the enhancer. Involvement of C/EBPβ and HNF-4α in the promoter-enhancer interaction is further supported by a massive DNA-protein interaction observed in the footprinting and EMSA assays. Since the *OTC* promoter lacks general core promoter elements such as TATA-box or initiators in standard positions, HNF-4α most likely plays an essential role in the initiation of *OTC* transcription in human.

Received October 22, 2013. Accepted February 14, 2014.

The study was supported by Charles University in Prague, Czech Republic, grants GAUK-251168_99910 and SVV2013/266504, and by the Ministry of Health, Czech Republic – conceptual development of research organization („Institute for Clinical and Experimental Medicine – IKEM, IN 00023001", and General University Hospital – VFN, 64165).

Corresponding author: Milan Jirsa, Institute for Clinical and Experimental Medicine (IKEM), Laboratory of Experimental Hepatology, Vídeňská 1958/9, 140 21 Prague 4, Czech Republic. E-mail: milan.jirsa@ikem.cz

Abbreviations: C/EBP – CCAAT-enhancer-binding protein, *Cis*+ – *cis*-acting regulatory element, COUP-TF – chicken ovalbumin upstream promoter-transcription factor, DPE – downstream promoter element, EMSA – electromobility shift assay, GATA1 – GATA-binding factor 1, HNF – hepatocyte nuclear factor, Inr – initiator element, OTC – ornithine carbamoyltransferase, OTCD – ornithine carbamoyltransferase deficiency, PCR – polymerase chain reaction, PIC – pre-initiation complex, Sp1 – specificity protein 1, TBP – TATA-box binding protein, TFII – general transcription factor for RNA-polymerase II, TIS – translation initiation site, TSS – transcription start site.

Introduction

Ornithine carbamoyltransferase (OTC, EC 2.1.3.3; MIM #300461) is a mitochondrial matrix enzyme catalysing the synthesis of citrulline – the second step of urea biosynthesis. The human *OTC* gene, located on chromosome Xp2l.1, comprises 10 coding exons (Lindgren et al., 1984; Hata et al., 1988). Ornithine carbamoyltransferase deficiency (OTCD; OMIM #311250) is the most common inherited urea cycle disorder (Brusilow and Horwich, 1995). Although its incidence is not high, the severity of clinical manifestations requires an early diagnosis, especially in males. Variations

in the *OTC* coding regions were thoroughly investigated (Tuchman et al., 2002; Yamaguchi et al., 2006; Storkanova et al., 2013); however, molecular diagnosis still fails in 20–25 % of cases. Molecular basis of the semi-dominant disease in some of these cases might be related to genetic defects in the regulatory elements. Recently, we have reported a first patient manifesting a mild form of OTCD caused by a point mutation in the promoter (Luksan et al., 2010).

The human *OTC* transcription start site (TSS) was first assigned to the position c.-135 (GenBank K02100; Horwich et al., 1984). Later published data suggested that transcription of human *OTC* does not initiate at a particular site but within a 70 bp region, mostly at preferred positions c.-95, c.-120, c.-150, c.-161 and c.-166 (Brusilow and Horwich, 1995). Three alternative TSS located at the positions c.-95, c.-119 and c.-169 upstream of the initiation codon were identified using rapid amplification of cDNA ends in total liver RNA (Luksan et al., 2010).

OTC is expressed predominantly in hepatocytes and epithelial cells of the intestinal mucosa (Ryall et al., 1985; Hamano et al., 1988). The mechanism of the tissue-specific pattern of *Otc* expression was first studied in the mouse. In transgenic experiments using sparse fur mice, Jones et al. (1990) showed that the rat *Otc* promoter was capable of direct liver- and intestine-specific transcription, but the expression level of the introduced gene was much lower than in the case of wild-type mouse with endogenous *Otc* expression. In our previous study (Luksan et al., 2010) we investigated the 793 bp human *OTC* promoter localized 46 bp upstream of the translation initiation site (TIS). In line with the findings by Jones et al. (1990) we observed a significant increase of the reporter luciferase activity in Hep-G2 cells, but not in HEK293 cells, achieved by subcloning a 465 bp sequence of the distal enhancer localized 9 kb upstream of TIS. This demonstrates the importance of the promoter-enhancer interaction for the regulation of human *OTC* expression.

Two positive *cis*-acting regulatory elements located around the positions -187 and -104 upstream of TIS were identified in the rat and mouse promoter. Both can bind an identical set of two members of the orphan receptor superfamily: hepatocyte nuclear factor 4 (HNF-4) and chicken ovalbumin upstream promoter-transcription factor (COUP-TF). In co-transfection experiments, HNF-4 activated transcription of a reporter gene from the *Otc* promoter, whereas competition with COUP-TF had a repressor effect (Kimura et al., 1993). Moreover, HNF-4 seems to contribute to the tissue-specific expression of *Otc* as its mRNA is present only in a limited number of tissue types such as the liver, kidney and intestine (Drewes et al., 1996). On the other hand, COUP-TF exhibits rather a ubiquitous expression pattern (Miyajima et al., 1988), and therefore competition for the binding site can result in repression of *Otc* in other tissues.

HNF-4α is critical for urea homeostasis in the mouse (Inoue et al., 2002) since liver-specific HNF4α-null mice exhibited elevated serum ammonia and decreased serum urea, while only a significant decrease in protein expression within the urea cycle enzymes was that of OTC. Similarly as in the rat, two HNF-4α recognition motifs were identified within the mouse 5' flanking region at positions -111 bp and -191 bp upstream of the TIS, and HNF-4α binding was confirmed using specific antibody. Moreover, co-transfection of reporter constructs carrying the mouse *Otc* promoter with a rat HNF-4α expression plasmid increased the reporter gene activity in CV-1 cells (Inoue et al., 2002).

Nishiyori et al. (1994) demonstrated that the rat *Otc* enhancer is activated by co-operation of two liver-selective transcription factors: HNF-4 and CCAAT-enhancer-binding protein (C/EBP). Furthermore, co-transfection experiments showed that both HNF-4 and C/EBPβ are required and neither of them alone is sufficient for activation of the reconstituted enhancer in non-hepatic cells. Combinatorial operation of these two liver-enriched (but not strictly liver-specific) transcription factors leads to more restricted liver-specific transcription of *Otc* (Nishiyori et al., 1994).

In the present study we identified key elements responsible for the regulation of *OTC* transcription in human. We suggest that initiation of *OTC* transcription in the absence of the general core promoter elements such as TATA-box or initiators in standard positions is mediated directly by HNF-4α.

Material and Methods

Reporter plasmid constructs

Fourteen reporter constructs were generated from pGL3-basic vector (Promega, Mannheim, Germany) by subcloning the 465 bp *OTC* enhancer (9 constructs) and promoter variants (12 constructs) in *Sal*I and *Xho*I restriction site, respectively. Six promoter variants 685, 458, 223, 173, 133 and 81 bp in length were prepared by a series of deletions from the 5'-end of the full 793 bp promoter (Table 1). All sequences and nucleotide positions used in the study are in accordance with GenBank NG_008471.1 and NM_000531.5.

Transient transfections and reporter gene assays

Human hepatoblastoma Hep-G2 (ECACC 85011430) cells were cultivated in minimal essential Eagle's medium supplemented with L-glutamine, 10% foetal bovine serum, 100 units/ml penicillin, and 100 μg/ml streptomycin, all purchased from Sigma-Aldrich, St. Louis, MO. The cultures were maintained in humidified atmosphere at 37 °C and 5% CO_2. Transfection was performed 48 h after seeding using FuGene HD reagent (Promega) in the reagent to DNA ratio 5 : 1 according to the manufacturer's protocol. The cells cultivated in 96-well plates were co-transfected with pGL3-derived experimental plasmids (0.1 μg per well) and a *Renilla* luciferase-containing reporter plasmid pRL-TK (0.01 μg per well). Cell lysis and determination of the luciferase

Table 1. Reporter constructs used in the dual luciferase reporter assay

Construct	Promoter variant		Enhancer
	Fragment length	Position	
pGL3_793/Enh	793 bp	c.-838 .. c.-46	yes
pGL3_685/Enh	685 bp	c.-730 .. c.-46	yes
pGL3_458/Enh	458 bp	c.-503 .. c.-46	yes
pGL3_223/Enh	223 bp	c.-268 .. c.-46	yes
pGL3_173/Enh	173 bp	c.-218 .. c.-46	yes
pGL3_133/Enh	133 bp	c.-178 .. c.-46	yes
pGL3_81/Enh	81 bp	c.-126 .. c.-46	yes
pGL3_0/Enh	-	-	yes
pGL3-793/0	793 bp	c.-838 .. c.-46	-
pGL3_223/0	223 bp	c.-268 .. c.-46	-
pGL3_173/0	173 bp	c.-218 .. c.-46	-
pGL3_133/0	133 bp	c.-178 .. c.-46	-
pGL3_81/0	81 bp	c.-126 .. c.-46	-
pGL3_793(HNF4m)/Enh	793 bp	c.-838 .. c.-46	yes
pGL3_0/0	-	-	-

activity was performed according to the manufacturer's instructions using the Dual Luciferase Reporter Assay System (Promega). All measurements were performed in a Synergy 2 automatic microplate reader (BioTek, Winooski, VT).

Signal of the *Photinus* luciferase was normalized by *Renilla* luciferase luminescence in all measurements. All transfections were done in six replicates and the results were calculated from at least six separate experiments. Because of the high inter-assay variation, the transcriptional activity of each promoter variant is presented relative to the activity of pGL3_793/Enh (positive control, 100 %). Statistical evaluations were performed after logarithmic transformation using the SigmaPlot software (Systat Software, Chicago, IL). All groups passed the test for normality and analysis of variance using one-way ANOVA. Multiple comparisons were performed using the Bonferroni method.

Nuclear extract preparation

Nuclear and cytoplasmic extracts were prepared from Hep-G2 cells using the NE-PER nuclear extraction kit (Thermo Scientific, Waltham, MA) according to the manufacturer's instructions. Extracts were dialysed for 4 h against 50 volumes of buffer containing 20 mM HEPES (pH 7.9), 20% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM PMSF, and 0.5 mM DTT. Protein concentrations were determined using the Pierce BCA Protein Assay Kit (Thermo Scientific).

DNase I footprinting

The DNase I footprinting method based on the separation of fluorescently labelled products using capillary electrophoresis was derived from the literature (Wilson et al., 2001; Zianni et al., 2006). A fluorescent probe spanning the 223 bp *OTC* proximal promoter, 5'- untranslated region and a part of exon 1 was prepared by polymerase chain reactions (PCR) using 6-carboxyfluorescein (6-FAM) labelled reverse primer 5'-(6-FAM)-TTTCGAACCATGAAGTTGTGACCA-3' and forward primer 5'-AAATGAGGAGGCCAGGCAAT-3'. The 200 µl reaction mixtures containing 100 µl of the protein binding buffer (2×), 2 µg of poly(dI-dT), 133 µg (sample) or 0 µg (control) of Hep-G2 nuclear extract and BSA up to 400 µg of total protein were incubated with 600 fmoles of fluorescently labelled probe for 20 min on ice. DNA cleavage was performed with 10 µl of DNase I at the concentration of 4 U/ml (Fermentas – Thermo Scientific, Pittsburgh, PA) after addition of 20 µl of Mg^{2+}/Ca^{2+} solution (10×). The cleavage was inhibited after 5 min by addition of the DNase stop solution and supplemented with GeneScan™ 500 ROX™ (Applied Biosystems, Foster City, CA) serving as a size standard and assay control. The DNA fragments were purified by phenol-chloroform extraction followed by ethanol precipitation, dissolved in 15 µl of Hi-Di formamide (Applied Biosystems) and separated by capillary electrophoresis in the Applied Biosystems 3130 Genetic Analyzer. The composition of the buffers and solutions used was as follows: protein binding buffer (2×) – 50 mM Tris-HCl, pH 8.0, 100 mM KCl, 12.5 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, 20% (v/v) glycerol; Mg^{2+}/Ca^{2+} solution (10×) – 50 mM $MgCl_2$, 25 mM $CaCl_2$; DNase stop solution – 200 mM NaCl, 50 mM EDTA, 1% (v/v) SDS.

Electromobility shift assays

Double-stranded biotin-labelled probes were prepared by standard PCR or by hybridization of complementary oligonucleotides. The binding reaction was carried out in 20 µl of a mixture containing 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM $MgCl_2$, 0.5 mM EDTA,

0.5 mM DTT, 4% glycerol 1 μg of poly(dI-dT), 25 μg of nuclear extract and 300 fmoles of biotin-labelled probe. In the case of competition assays, 0, 1, 4 and 12 pmoles of the non-labelled probe [-98..-126] or a 200-fold excess of a double-stranded probe AGTGAGGGTCAA-AGTTCATATCAC, derived from the consensus HNF-4α recognition sequence, was used as a specific competitor. The samples were incubated for 20 min at room temperature, separated using non-denaturing polyacrylamide gel electrophoresis (6% polyacrylamide gel in 0.5× TBE, 100 V, 50 min) and electro-blotted at 300 mA, for 30 min on the Amersham Hybond-N nylon membrane (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). The DNA-protein complexes were fixed by baking the membrane for 2 h in a dry oven at 80 °C. The membrane was processed using the Chemiluminescent Nucleic Acid Detection Module (Thermo Scientific, Rockford, IL) according to the manufacturer's instructions and analysed with the FUJI LAS 3000 system (Fujifilm, Tokyo, Japan). In supershift assays, goat anti-HNF-4α polyclonal antibody (Santa Cruz, Dallas, TX) was pre-incubated with the nuclear extract for 20 min prior to addition of labelled probes. Detection of the anti-HNF-4α antibody was performed using HRP-conjugated polyclonal rabbit anti-goat antibody (Thermo Scientific).

Prediction of transcription factor binding sites

Prediction of transcription factor binding sites was performed with MatInspector (Genomatix Software, München, Germany) and Match (Biobase, Wolfenbuettel, Germany) bioinformatic tools using the MatBase 9.0 and TRANSFAC 7.0 database versions.

Results

In silico prediction of the main regulatory domains

The sequences of rat *Otc* and human *OTC* promoter reach about 65% similarity. Higher homology is preserved in the proximal part of the promoter up to 450 bp upstream of the TIS. DNase I footprinting and functional characterization of deletion mutants generated from the rat *Otc* 5'-flanking region uncovered four regions protected against DNase I cleavage and two HNF-4α binding motifs located 104 and 187 bp upstream of TIS (Murakami et al., 1990; Kimura et al., 1993) (Fig. 1). Homologous motifs were found in the mouse by Inoue et al. (2002). Similar binding sites located at c.-105 and c.-187 were predicted in the human sequence using MatInspector (MI) and Match (M) (Fig. 1). While the HNF-4α binding motif located at c.-105 was recognized with both tools (MI score 0.884, M score 0.883), the other located at c.-187 was predicted only with Match (score 0.910). MatInspector recognized a binding site for another member of the nuclear receptor superfamily at the same position – peroxisome proliferator-activated receptor γ (PPARγ, MI score 0.833). Both HNF-4 and PPARγ recognize similar sequence motifs. Transfection experiments revealed that preferential binding is partially determined by the sequence of the core motif, but it is much more dependent on the promoter context (Nakshatri and Bhat-Nakshatri, 1998). Therefore, the sequence motif can be considered a HNF-4α binding site as well.

Two additional potential HNF-4α binding sites located 136 bp (MI/M scores 0.923/0.746) and 588 bp (MI/M scores 0.822/0.835) upstream of TIS and one site recognized by C/EBPβ located 221 bp upstream of the TIS (M score 0.970) were predicted within the human *OTC* promoter. None of these binding sites was recognized either in the rat or in the mouse promoter because of variations in the matrix core motifs.

As mentioned above, human *OTC* lacks a particular site of transcription initiation, indicating the absence of a strong core promoter element such as TATA-box. Accordingly, in our previous study (Luksan et al., 2010) we found three alternative TSSs – two major ones starting at positions c.-95, c.-119 and a minor one starting at c.-169. Of these three only the c.-169 transcript variant implies a potential TATA-like element CATAAA located 25–30 bp upstream of the TSS; however, its regulatory role seems to be low because cutting off the corresponding region did not lead to a remarkable decrease in the promoter activity (Fig. 2, constructs 173/Enh and 133/Enh). Moreover, this fully conserved motif acts as a known HNF-4α binding site located 187 bp upstream of TIS in rodents.

Identification of cis-acting regulatory elements and delimitation of the core promoter

The transcriptional activity of six 5'-deleted promoter mutants ranging 685, 458, 223, 173, 133 and 81 bp (all with 3'-ends 46 bp upstream of TIS – see Fig. 1) was compared with the full 793 bp human promoter in pGL3 constructs containing the human enhancer sequence. The empty plasmid pGL3_0/Enh containing only the enhancer was used as a negative control to eliminate the influence of the enhancer alone. The results have shown that a 5'-deletion of 108 bp has no influence on the reporter activity (both mean and median varied around 116 % of the activity of pGL3_793/Enh), while the 5'-deletion of 335 bp leads to a decreased luciferase signal (median and mean reached 57 % and 54 %, respectively) (Fig. 2). In contrast, increased reporter activity was observed with the 223 bp promoter variant (median 146 %, mean 152 %). The response remained constant through the 173 and 133 bp variants (medians 137 % and 156 %, means 150 % and 167 % respectively), while the shortest tested variant of 81 bp showed slightly lower transcriptional activity when compared with the full promoter (median 81 %, mean 80 %) (Fig. 2). The decrease was statistically significant when compared with the 133, 173 and 223 bp promoter variants, but not significant when compared with the 458 and 685 bp variants (Fig. 2). The reporter gene activity observed with

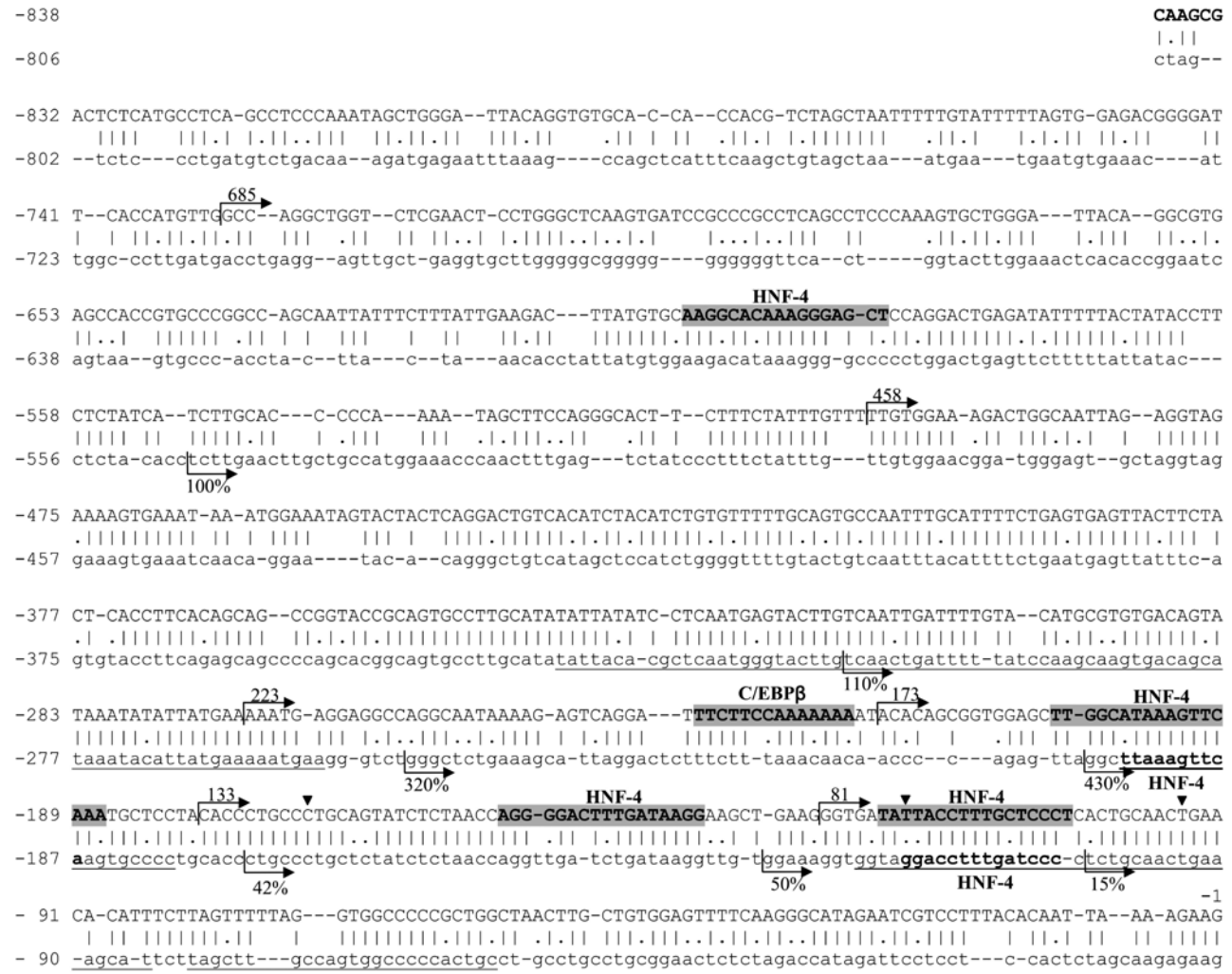
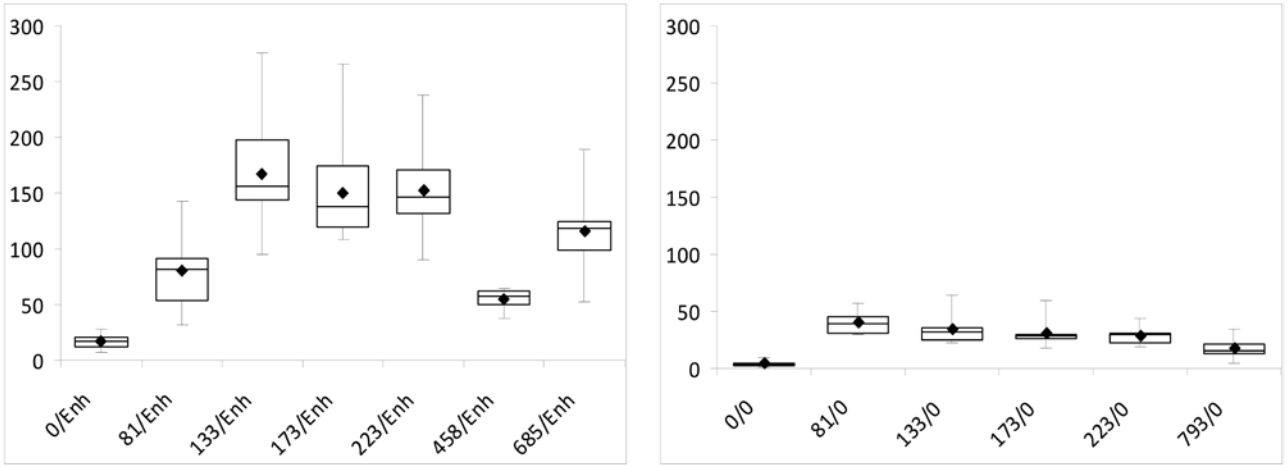


Fig. 1. Sequence alignment of the rat *Otc* (lower case) and human *OTC* (upper case) 5' flanking region. The regions protected against DNase I cleavage in the rat are underlined. The 5' ends of the rat promoter fragments studied by Murakami et al. (1990) and their relative activity in chloramphenicol acetyltransferase reporter assays are marked with arrows. Transcription factor binding sites identified as *cis*-acting regulatory elements in the rat are boldfaced. Transcription factor binding sites predicted by Match and MatInspector in human *OTC* promoter are boldfaced and highlighted in grey. Black arrows indicate human *OTC* promoter variants tested in reporter assays. Three alternative TSSs identified within the human liver mRNA are marked with black triangles.

the pGL3 vector carrying the enhancer alone was lower than 20 % of the pGL3_793/Enh and significantly lower than with all other constructs.

Based on these results we defined the human *OTC* core promoter as the 223 bp region located -46..-268 bp upstream of TIS containing three transcription initiation sites and two positive *cis*-acting elements: *Cis1*+ located within the first 81 base-pairs, reaching 83 % of the full promoter activity, and *Cis2*+ located within the upstream 52 bp sequence.

In the following experiments we evaluated the contribution of *Cis1*+ and *Cis2*+ to the transcriptional activity of the promoter alone (Fig. 2, right plot). The core promoter exhibited twice higher reporter activity than the full promoter and the level of luciferase signal remained stable within all tested constructs (81/0, 133/0, 173/0 and 223/0). Finally, the reporter activity observed with all promoter variants was significantly higher than with the negative control pGL3_0/0. The influence of *Cis2*+ was not observed in this case, indicating two possible explanations: first, *Cis2*+ regulatory element contributes only to the interaction with the enhancer whereas *Cis1*+ plays a role in both interaction with the enhancer and activation of the promoter. Second, *Cis1*+ and *Cis2*+ operate simultaneously but they contribute to formation of different mRNA variants originating from various TSSs. In such case, absence of one element responsible for formation of one transcript variant could be compensated by the action of the other element responsible for production of the other transcript originating from a different TSS. Such compensatory effect could maintain the overall transcriptional activity of the promoter. However, the compensatory capacity may reach its limit when the transcriptional activity is accelerated by the promoter-enhancer interaction, and therefore the promoter variant containing only the *Cis1*+ ele-



	81+/+	133+/+	173+/+	223+/+	458+/+	685+/+	WT0/+	81+/0	133+/0	173+/0	223+/0	WT0/0	WT+/0
81+/+		***	**	**	N	N	***	**	***	***	***	***	***
133+/+	***		N	N	***	N	***	***	***	***	***	***	***
173+/+	**	N		N	***	N	***	***	***	***	***	***	***
223+/+	**	N	N		***	N	***	***	***	***	***	***	***
458+/+	N	***	***	***		N	***	N	N	N	N	***	***
685+/+	N	N	N	N	N		***	***	***	***	***	***	***
WT0/+	***	***	***	***	***	***		***	**	N	N	***	N
81+/0	**	***	***	***	N	***	***		N	N	N	***	***
133+/0	***	***	***	***	N	***	**	N		N	N	***	**
173+/0	***	***	***	***	N	***	N	N	N		N	***	N
223+/0	***	***	***	***	N	***	N	N	N	N		***	N
WT0/0	***	***	***	***	***	***	***	***	***	***	***		***
WT+/0	***	***	***	***	***	***	N	***	**	N	N	***	

Fig. 2. Transcriptional activity of the mutant promoter variants in combination with the enhancer (left panel) and without the enhancer (right panel). The data are based on at least six independent experiments. The box represents the range between the 25th and 75th percentiles, horizontal line within the box shows the median. Whiskers indicate the minimal and maximal values. Diamonds are used for means. Multiple comparisons of variance using the Bonferroni method are presented in the table below the plots. Statistical significance is represented by asterisks: (***) indicates the differences with $P \leq 0.005$ and (**) is used for statistical significance with $P \leq 0.01$.

ment alone exhibits lower activity when compared with the variant carrying both regulatory units (Fig. 2, variants 81 and 133).

Identification of DNA-protein interaction sites by DNase I footprinting

In a series of experiments we identified five DNase I-protected regions localized 92–102 (A), 105–114 (B), 135–145 (C), 150–172 (D) and 180–192 (E) bp upstream of TIS (Fig. 3A). The sequence motifs observed in the protected regions B and E are very close to the protected regions in the rat promoter that partially overlap with the HNF-4 and COUP-TF binding sites (Fig. 1). Another HNF-4α recognition site predicted with high matrix similarity score is located close to the protected region C.

No conventional transcription factor-binding sites were recognized within the protected region A that coincides with the c.-95 bp TSS and overlaps with a sequence motif related to initiator element Inr (MI score 0.967). The same sequence was recognized with Match as a Cap site with matrix similarity score 0.995. A general core promoter motif Inr is known to moderate transcription by direct interaction with the transcription factor II D (TFIID) complex in both TATA-containing and TATA-less promoters. Unfortunately, the role of the initiator element in the *OTC* promoter remains unclear since the strictly unidirectional Inr is located in the template strand in the antisense orientation.

Structural motifs recognized by two different transcription factors with high matrix similarity score were predicted *in silico* within the protected region D: GATA-binding factor 1 (GATA1), located 155 bp upstream of TIS (MI score 0.962, M score 0.970) and specificity protein 1 (Sp1) at position -169..-177 (M score 0.962). Whereas GATA1 is essential for erythroid and megakaryocyte development, its expression pattern is limited

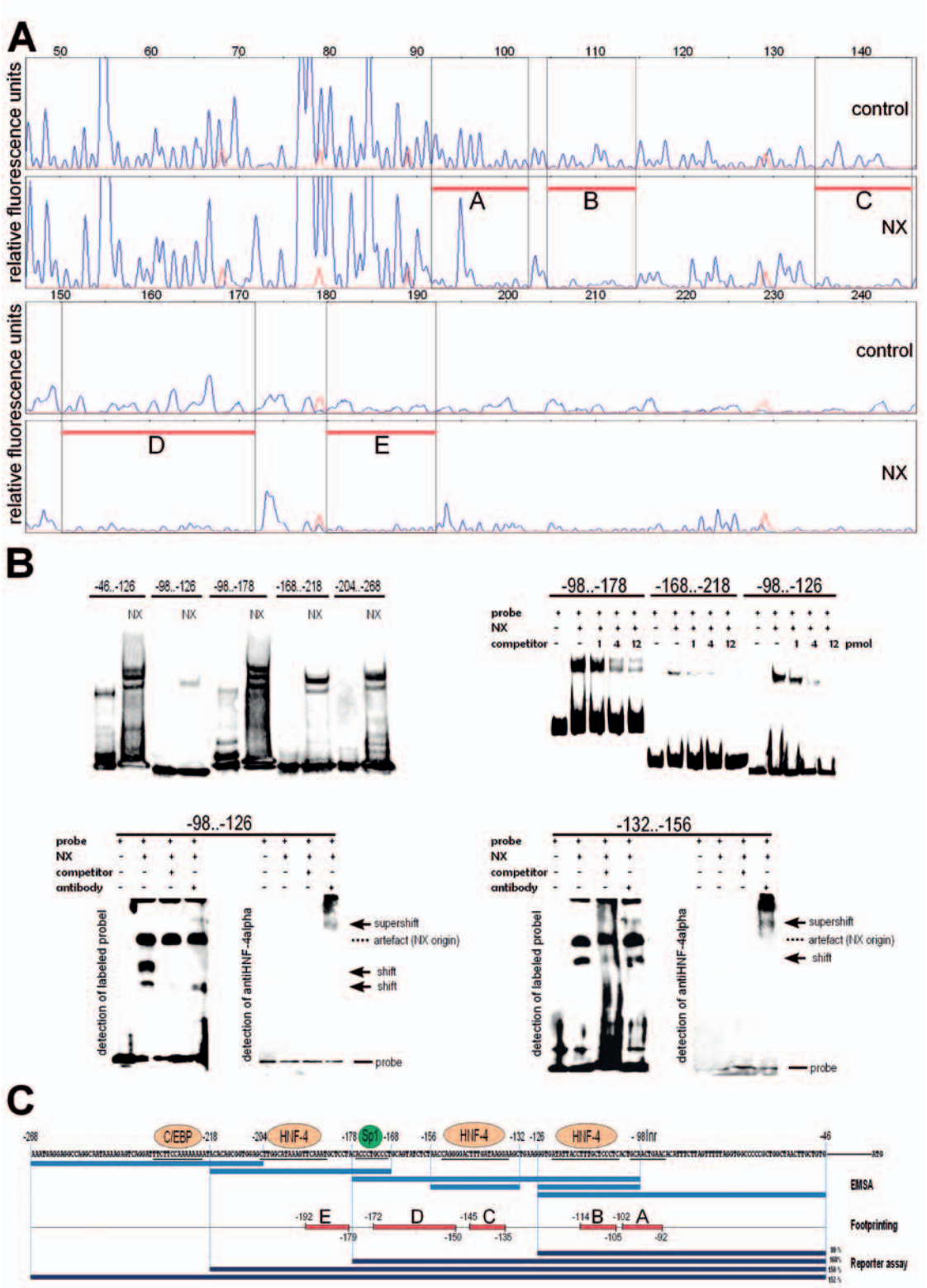


Fig. 3. Analysis of the DNA-protein interactions within the human *OTC* core promoter. **A** – DNase I footprinting. Peaks represent the fragments obtained after DNase I cleavage (in blue) and internal standard (in red). Regions protected against the endonuclease cleavage are marked by red horizontal bars. **B** – Electromobility shift assays (EMSA). Standard EMSA (upper left panel), competition assays with a specific competitor for HNF-4α (upper right panel) and specific confirmation of HNF-4α binding to the predicted sites using competitor and supershift assays (lower panels). Detection of the anti-HNF-4α antibody was performed by probing the same membrane with a secondary antibody (right electrophoreograms in lower panels). NX – incubated with nuclear extract. **C** – Schematic summary of results. Horizontal bars represent promoter fragments used in the dual luciferase reporter assay (dark blue), protected regions observed with the DNase I footprinting (red) and EMSA probes (light blue). The upper part of the figure represents transcription factors and their binding sites (underlined sequence) predicted with high score using MatInspector and Match programs.

to the haematopoietic cell lineage and its involvement in the regulation of *OTC* transcription is thus unlikely. Sp1 is rather a ubiquitous transcription factor capable of regulating basal expression in genes with TATA-less promoters (Black et al., 2001). The Sp1 recognition site overlaps the alternative TSS located 169 bp upstream of the translation origin. Since no other general core promoter has been found in a standard position related to TSS except for the TATA-like domain discussed earlier, the contribution of Sp1 to the initiation of transcription from the position c.-169 seems likely.

Analysis of the DNA-protein complexes by EMSA

The DNA-protein interactions observed by DNase I footprinting were verified in a set of electromobility shift assays (EMSA). Six double-stranded biotin-labelled probes were designed to cover the regions containing *cis*-acting regulatory elements, DNase I-protected regions or *in silico* predicted transcription factor-binding sites (see Fig. 3C). Interaction with the HepG2 nuclear extract led to an obvious shift: three shifted bands were observed with the probes covering the regions located at positions -46..-126 and -98..-178, two bands with probes spanning the regions -168..-218 and -204..-268 and one zone was detected using a probe corresponding to the sequence localized -98..-126 upstream of the TIS.

The probe [-98..-126] spanning 29 bp was designed to cover the proximal HNF-4α binding site. The motif is conserved in human and rodents and overlaps the region B protected against the DNase I cleavage. The site is also contained in the [-46..-126] probe carrying the *Cis1*+ regulatory element and in the [-98..-178] probe. Both probes exhibited a different band shift pattern than the [-98..-126] probe, indicating formation of additional DNA-protein complexes (Fig. 3B). One of these complexes formed with the probe [-46..-126] may involve the intermediates formed by the factors binding to the Inr element. The pattern obtained with the probe [-98..-178] likely involves complexes with other HNF-4α and Sp1 predicted to bind the sequences of protected regions C and D.

The probe [-168..-218] was designed to overlap the conserved HNF-4α binding site, found also in the rat around the position -187 bp upstream of TIS, and the DNase I-protected region E. The probe [-204..-268] covered the distal part of the core promoter region, where a sequence motif recognized by C/EBPβ was predicted.

Specific interaction of HNF-4α with probes [-46..-126] and [-98..-178] was examined in competition assays using a non-labelled probe [-98..-126]. The reaction mixture containing both biotin-labelled and non-labelled probe [-98..-126] was used as a control indicating the specific manner of the competition. Repression of nuclear factor binding was observed with all three probes. A gradual decrease of two differentially shifted bands was observed with the probe [-98..-178] carrying two motifs recognized by HNF-4α (Fig. 3B).

Confirmation of specific interaction between HNF-4α and the predicted binding sites located 105 bp and 136 bp upstream of TIS was achieved using probes [-98..-126] and [-132..-156] (Fig. 3C). Addition of a specific competitor (a non-labelled probe corresponding to the consensus sequence of HNF-4α recognition motif) in 200-fold excess led to a significant decrease in the shifted band intensity. Supershifted bands of the DNA-protein-antibody complexes were observed after incubation with the anti-HNF-4α antibody. The presence of the anti-HNF-4α antibody in these complexes was confirmed using HRP-conjugated secondary antibody: the immunoreactive zones co-localized with the supershifted bands (Fig. 3B).

Mutation in Cis1+ element abolished HNF-4α binding

To confirm the role of the *Cis1*+ element as a HNF-4α binding site we generated a mutant 793(HNF4mut) in which the matrix sequence CAAAGG was replaced by CAccaG and we then compared the transcriptional activity of both variants combined with the enhancer in a set of reporter assays (Fig. 4A). The activity of the pGL3_793(HNF4mut)/Enh construct was more than two times lower than in the case of normal promoter and the decrease was highly significant in Student's *t*-test. When compared with the reporter gene activity obtained with other variant constructs, the pGL3_793(HNF4mut)/Enh reached a similar level as the core promoter variants without enhancer and it was about two times lower than pGL3_81/Enh (mean and median reached 41 % and 34 % of pGL3_793/Enh, respectively). Interestingly, in spite of the disruption of the HNF-4α binding site, the GL3_793(HNF4mut)/Enh activity was partially preserved, indicating the ability of the upstream element(s) to compensate for the function of the disrupted *Cis1*+.

Discussion

We have shown that the minimal promoter responsible for both the promoter activity and activation of the enhancer is located within the region -46..-126 bp upstream of TIS. As the shortest reported transcription variant starts at the position c.-95, the minimal promoter contains one strong positive *cis*-acting regulatory element *Cis1*+ located within 30 bp upstream of the TIS. The motif is highly conserved in vertebrates and the affinity of HNF-4 to this region has been clearly demonstrated in the rat (Kimura et al., 1993) and mouse (Inoue et al., 2002). Accordingly, we found that the *Cis1*+ spanning region is protected against DNase I cleavage, and experimentally confirmed binding of HNF-4α at this site (Fig. 3B).

The *Cis2*+ element was identified within the region between -127 and -178 bp upstream of the TIS overlapping with the protected regions C and D. *Cis2*+ is involved in the interaction with the distal enhancer, but it does not directly influence the promoter activity alone

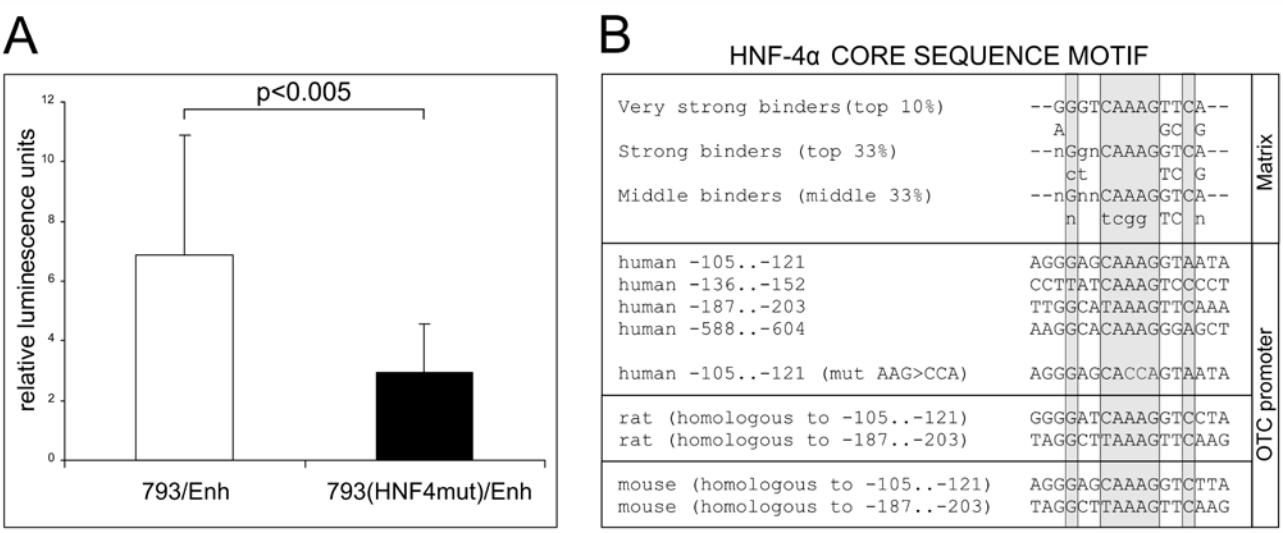


Fig. 4. A – Functional analysis of the wild-type (793) and mutated (793(HNF4mut)) *OTC* promoter, carrying a mutation in the proximal binding site for HNF-4α, both subcloned to pGL3 vector containing the enhancer. The data are presented as mean + SD. B – Sequence comparison of the *in silico* predicted human HNF-4α binding sites and their rat and mouse homologues with the matrix sequence motifs published by Fang et al. (2012). Grey boxes show the core sequence motif. "n" indicates equal incidence of all four nucleotides. Letters below the sequences in the "Matrix" section represent common (uppercase) or probable but not common (lowercase) alternatives.

(Fig. 2). Sequence motifs recognized by nuclear factors HNF-4α and Sp1 were predicted *in silico* in this region (Fig. 3C). Binding of HNF-4α was confirmed in the band shift competition assay where two bands obtained with the probe [-98..-178] and one band obtained with the probe [-132..-156] decreased proportionally to the increased amount of specific competitor (Fig. 3B). The presence of two differentially shifted bands with the probe [-98..-178] can be explained by formation of two different complexes, probe-HNF4 and HNF4-probe--HNF4, as the probe carries two HNF-4α binding sites and the ratio of both complexes remained constant even when repressing their formation by a specific competitor. A direct proof of HNF-4α binding at this site was obtained by detection of supershifted complex DNA--protein-antibody (Fig. 3B).

Binding of Sp1 could also explain transcription from the alternative *OTC* TSS at the position c.-169, but it cannot explain the role of *Cis2+* in the interaction with the enhancer.

Unlike the two HNF-4α binding sites discussed above, the motif located 187 bp upstream of the translation origin didn't exhibit significant influence on transcriptional activation of *OTC* (Fig. 2). Nevertheless, the matrix sequence is highly conserved and it differs from the rodent sequence only in one nucleotide located outside of the matrix core. The region partially correlates with the protected region E and interaction of HNF-4 with that site has been confirmed in rodents (Kimura et al., 1993; Inoue et al., 2002). The motif overlaps with the probe [-168..-218], which competed for HNF-4α with the HNF-4α-specific competitor in EMSA. Therefore, it is obvious that the HNF-4α binding site is active, despite that its contribution to the overall promoter function is negligible when compared with the *Cis1+* and *Cis2+* elements. The low contribution is further supported by observations in the mouse (Inoue et al., 2002).

Fang et al. (2012) recently reported a genome-wide study evaluating matrix sequence motifs in a large group of HNF4-specific genes using high-throughput ChIP-seq. Motifs recognized by 1371 specific sequences were statistically evaluated based on their binding affinity. Matrices preferred by top 10 % (strongest), top 33 % (strong) and middle 33 % (middle) binders originating from the Fang's study were compared with all four *in silico* predicted HNF-4α binding sites and their rat and mouse homologues (Fig. 4B). The matrix sequence is highly conserved and practically all compared regions could be considered as at least middle-affinity binders to HNF-4α.

Activation of the enhancer in the rat is associated with the interaction of promoter-bound HNF-4 with the enhancer-bound C/EBPβ (Nishiyori et al., 1994). A similar mechanism involving synergistic co-operation of HNF-4 and C/EBPα, both bound to the promoter, has been suggested for the human apolipoprotein B gene (Metzger et al., 1993). Indeed, we predicted a binding site for C/EBPβ in the *OTC* core promoter between positions c.-221 and c.-234, and the results of EMSA revealed two shifted bands indicating formation of DNA--protein complexes within this region. In contrast, cutting off the region located -219 bp to -268 bp upstream of the TIS did not change the level of the reporter activity in luciferase assays.

In silico analysis of the proximal 200 bp part of the enhancer region revealed a set of six repeated sequence motifs corresponding to the C/EBPβ binding site.

Accordingly, the DNase I cleavage of the proximal part of the enhancer was decreased and a shifted band was observed in the gel retardation assay after incubation of the proximal enhancer fragment (data not shown). Interaction of the liver enriched HNF-4 factors bound to the *OTC* core promoter with the set of periodically occurring liver-specific C/EBP-like proteins thus may contribute to the tissue specificity of *OTC* expression in the liver, as was seen in the rat model by Nishiyori et al. (1994).

The distribution of HNF-4α recognition sites within the *OTC* core promoter region exhibits an important feature: all three HNF-4α binding motifs are located within 35 bases upstream of the TSSs (Fig. 1). Moreover, the sequence of 223 bp proximal promoter lacks any general core promoter elements such as TATA-box, Inr, the downstream promoter element (DPE), the TFIIB recognition element (BRE), the motif ten element (MTE), downstream core element (DCE), and the X core promoter element (XCPE) located at standard positions upstream of the TSSs, thereby suggesting the importance of HNF-4α in the initiation of *OTC* transcription. Initiation of transcription mediated by DNA-dependent RNA-polymerase II is dependent on the assembly of functional pre-initiation complex (PIC) consisting of general transcription initiation factors TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH (Myer and Young, 1998; Hahn, 2004).

The transcriptional machinery and the PIC formation has been well described in promoters equipped with general core promoter elements, but it has been poorly studied in promoters lacking these elements. However, few indices suggesting the role of HNF-4α in the transcription from TATA-less and Inr-less promoters can be found in the literature. HNF-4α was shown to activate transcription *in vitro* by facilitating assembly of the pre-initiation complex through direct physical interaction with TFIIB. The interaction of TFIIB with DNA-bound HNF-4α led to TFIIB-mediated recruitment of TBP. Furthermore, HNF-4α participated in the assembly of downstream basal factors into the PIC (Malik and Karathanasis, 1996). A similar mechanism comprising both the stable complex activator-TFIIB-TBP assembly and involvement in the association of downstream basal transcription factors TFIIE and TFIIF was observed later in oestrogen receptor as another member of the nuclear hormone receptor superfamily (Sabbah et al., 1998). In a recent study, systematic characterization of the HNF4α-TFIID link revealed that the HNF4α DNA-binding domain binds directly to the TBP and, through this interaction, can target TBP or TFIID to promoters containing HNF4α-binding sites (Takahashi et al., 2009).

In conclusion, our data strongly suggest the essential role of HNF-4α in the regulation of *OTC* transcription in human. The regulatory mechanism may involve physical interaction with TFIIB, recruitment of TBP and stabilization of TFIID to a particular site in the promoter followed by the assembly of the pre-initiation complex.

References

Black, A. R., Black, J. D., Azizkhan-Clifford, J. (2001) Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer. *J. Cell. Physiol.* **188**, 143-160.

Brusilow, S. W., Horwich, A. L. (1995) Urea cycle enzymes. In *The Molecular and Metabolic Bases of Inherited Disease,* eds. R. Scriver, A. L. Beaudet, W. S. Sly, D. Valle, pp. 1187-1232, McGraw-Hill, New York.

Drewes, T., Senkel, S., Holewa, B., Ryffel, G. U. (1996) Human hepatocyte nucler factor 4 isoforms are encoded by distinct and differentially expressed genes. *Mol. Cell. Biol.* **16**, 925-31.

Fang, B., Mane-Padros, D., Bolotin, E., Jiang, T., Sladek, F. M. (2012) Identification of a binding motif specific to HNF4 by comparative analysis of multiple nuclear receptors. *Nucleic Acids Res.* **40**, 5343-5356.

Hahn, S. (2004) Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* **11**, 394-403.

Hamano, Y., Kodama, H., Yanagisawa, M., Haraguchi, Y., Mori, M., Yokota, S. (1988) Immunocytochemical localization of ornithine transcarbamylase in rat intestinal mucosa. Light and electron microscopic study. *J. Histochem. Cytochem.* **36**, 29-35.

Hata, A., Tsuzuki, T., Shimada, K., Takiguchi, M., Mori, M., Matsuda, I. (1988) Structure of the human ornithine transcarbamylase gene. *J. Biochem.* **103**, 302-308.

Horwich, A. L., Fenton, W. A., Williams, K. R., Kalousek, F., Kraus, J. P., Doolittle, R. F., Konigsberg, W., Rosenberg, L. E. (1984) Structure and expression of a complementary DNA for the nuclear coded precursor of human mitochondrial ornithine transcarbamylase. *Science* **224**, 1068-1074.

Inoue, Y., Hayhurst, G. P., Inoue, J., Mori, M., Gonzalez, F. J. (2002) Defective ureagenesis in mice carrying a liver-specific disruption of hepatocyte nuclear factor 4α (HNF4α). *J. Biol. Chem*. **277***,* 25257-25265.

Jones, S. N., Grompe, M., Munir, M. I., Veres, G., Craigen, W. J., Caskey, C. T. (1990) Ectopic correction of ornithine transcarbamylase deficiency in sparse fur mice. *J. Biol. Chem.* **265**, 14684-14690.

Kimura, A., Nishiyori, A., Murakami, T., Tsukamoto, T., Hata, S., Osumi, T., Okamura, R., Mori, M., Takiguchi, M. (1993) Chicken ovalbumin upstream promoter-transcription factor (Coup-Tf) represses transcription from the promoter of the gene for ornithine transcarbamylase in a manner antagonistic to hepatocyte nuclear factor-IV (Hnf-4). *J. Biol. Chem.* **268**, 11125-11133.

Lindgren, V., de Martinville, B., Horwich, A. L., Rosenberg, L. E., Francke, U. (1984) Human ornithine transcarbamylase locus mapped to band Xp21.1 near the Duchenne muscular dystrophy locus. *Science* **226**, 698-700.

Luksan, O., Jirsa, M., Eberova, J., Minks, J., Treslova, H., Bouckova, M., Storkanova, G., Vlaskova, H., Hrebicek, M., Dvorakova, L. (2010) Disruption of OTC promoter-enhancer interaction in a patient with symptoms of ornithine carbamoyltransferase deficiency. *Hum. Mutat.* **31**, E1294-1303.

Malik, S., Karathanasis, S. K. (1996) TFIIB-directed transcriptional activation by the orphan nuclear receptor hepatocyte nuclear factor 4. *Mol. Cell. Biol.* **16**, 1824-1831.

Metzger, S., Halaas, J. L., Breslow, J. L., Sladek, F. M. (1993) Orphan receptor Hnf-4 and Bzip protein C/Ebp-α bind to overlapping regions of the apolipoprotein-B gene promoter and synergistically activate transcription. *J. Biol. Chem.* **268**, 16831-16838.

Miyajima, N., Kadowaki, Y., Fukushige, S., Shimizu, S., Semba, K., Yamanashi, Y., Matsubara, K., Toyoshima, K., Yamamoto, T. (1988) Identification of two novel members of erbA superfamily by molecular cloning: the gene products of the two are highly related to each other. *Nucleic Acids Res.* **16**, 11057-11074.

Murakami, T., Nishiyori, A., Takiguchi, M., Mori, M. (1990) Promoter and 11-kilobase upstream enhancer elements responsible for hepatoma cell-specific expression of the rat ornithine transcarbamylase gene. *Mol. Cell. Biol.* **10**, 1180-1191.

Myer, V. E., Young, R. A. (1998) RNA polymerase II holoenzymes and subcomplexes. *J. Biol. Chem.* **273**, 27757-27760.

Nakshatri, H., Bhat-Nakshatri, P. (1998) Multiple parameters determine the specificity of transcriptional response by nuclear receptors HNF-4, ARP-1, PPAR, RAR and RXR through common response elements. *Nucleic Acids Res.* **26**, 2491-2499.

Nishiyori, A., Tashiro, H., Kimura, A., Akagi, K., Yamamura, K., Mori, M., Takiguchi, M. (1994) Determination of tissue specificity of the enhancer by combinatorial operation of tissue-enriched transcription factors. Both HNF-4 and C/EBP ß are required for liver-specific activity of the ornithine transcarbamylase enhancer. *J. Biol. Chem.* **269,** 1323-1331

Ryall, J., Nguyen, M., Bendayan, M., Shore, G. C. (1985) Expression of nuclear genes encoding the urea cycle enzymes, carbamoyl-phosphate synthetase I and ornithine carbamoyl transferase, in rat liver and intestinal mucosa. *Eur. J. Biochem.* **152,** 287-292

Sabbah, M., Kang, K. I., Tora, L., Redeuilh, G. (1998) Oestrogen receptor facilitates the formation of preinitiation complex assembly: involvement of the general transcription factor TFIIB. *Biochem. J.* **336**, 639-646.

Storkanova, G., Vlaskova, H., Chuzhanova, N., Zeman, J., Stranecky, V., Majer, F., Peskova, K., Luksan, O., Jirsa, M., Hrebicek, M., Dvorakova, L. (2013) Ornithine carbamoyltransferase deficiency: molecular characterization of 29 families. *Clin. Genet.* **84**, 552-559.

Takahashi, H., Martin-Brown, S., Washburn, M. P., Florens, L., Conaway, J. W., Conaway, R. C. (2009) Proteomics reveals a physical and functional link between hepatocyte nuclear factor 4α and transcription factor IID. *J. Biol. Chem.* **284**, 32405-32412.

Tuchman, M., Jaleel, N., Morizono, H., Sheehy, L., Lynch, M. G. (2002) Mutations and polymorphisms in the human ornithine transcarbamylase gene. *Hum. Mutat.* **19**, 93-107.

Wilson, D. O., Johnson, P., McCord, B. R. (2001) Nonradiochemical DNase I footprinting by capillary electrophoresis. *Electrophoresis* **22**, 1979-1986.

Yamaguchi, S., Brailey, L. L., Morizono, H., Bale, A. E., Tuchman, M. (2006) Mutations and polymorphisms in the human ornithine transcarbamylase (OTC) gene. *Hum. Mutat.* **27**, 626-632.

Zianni, M., Tessanne, K., Merighi, M., Laguna, R., Tabita, F. R. (2006) Identification of the DNA bases of a DNase I footprint by the use of dye primer sequencing on an automated capillary DNA analysis instrument. *J. Biomol. Tech.* **17**, 103-113.