FEBS Letters 540 (2003) 58–64 FEBS 27044

An initiator and its flanking elements function as a core promoter driving transcription of the *Hepatopoietin* gene

Yanlin Zhao, Fei Tang, Jingwei Cheng, Li Li, Guichun Xing, Yunping Zhu, Lingqiang Zhang, Handong Wei, Fuchu He*

Department of Proteomics and Genomics, Beijing Institute of Radiation Medicine, Chinese Human Genome Center at Beijing, 27 Taiping Road, Beijing 100850, PR China

Received 27 November 2002; revised 23 January 2003; accepted 26 January 2003

First published online 19 March 2003

Edited by Ned Mantei

Abstract Hepatopoietin (HPO)/ALR (augmenter of liver regeneration), as a versatile hepatotrophic growth factor and a cellular thiol oxidase, is involved in a wide variety of basic processes of various tissues, especially in liver and testis. Here, we studied the regulation of *HPO* gene expression. By sequential deletion of the *HPO* 5′-flanking region, the minimal promoter of the *HPO* gene was shown to span positions −22 to +42 relative to the transcriptional start point. Further transfection assay and mutation analysis showed that the core promoter contains a functional initiator. Interestingly, three tandem repeats of a CTGGAGGC element, surrounding the transcription start site and bound by specific nuclear factors, were found to be pivotal for the promoter activity. This initiator flanking element functions in an initiator-dependent fashion and is present in many initiator-containing genes. Taken together, our findings revealed that the initiator-like element and its flanking repeat sequence comprise a core promoter and drive the transcriptional initiation of the *HPO* gene in a combinatorial manner. The *HPO* gene promoter might represent a novel architecture for core promoters.

Key words: Core promoter; Constitutive expression; Initiator-dependent flanking element

1. Introduction

Hepatopoietin (HPO) belongs to the new HPO/ALR (augmenter of liver regeneration)/Erv1p protein family, members of which are found in eukaryotes from yeast to man and even in some double-stranded DNA viruses [1,2]. Family members share domains of approximate 100 amino acids that are ca. 30% identical and contain a CXXC motif, a feature of thioredoxin-like proteins [3].

Mammalian HPO/ALR was originally identified as a hepatotrophic growth factor that can specifically stimulate liver regeneration after partial hepatectomy in vivo and proliferation of cultured primary hepatocytes in vitro [1,4]. Later, it was found that the receptor of HPO exists on the membrane of primary hepatocytes and hepatoma cell lines [5], and that HPO might act as an autocrine growth factor to stimulate autonomous proliferation of hepatoma cells [6]. Further investigations showed that extracellular HPO stimulates proliferation of hepatoma cells via the signaling of mitogen-activated protein kinase through epidermal growth factor receptor [7], and that intracrine HPO interacts with JAB1 to modulate the transcriptional activity of AP1 [8]. In addition, it was demonstrated that HPO/ALR reduces the lytic activity of liver-resident natural killer cells and the expression level of mitochondrial transcription factor A by down-regulating interferon-γ [9,10]. Intriguingly, recent investigations have revealed that ERV1p/ALR is a cellular FAD-linked sulfhydryl oxidase that participates in a cytoplasmic pathway of disulfide bond formation [11–14]. Taken together, these characteristics indicate that HPO and its analogs take part in a wide variety of basic processes that are operational in various cells and tissues.

Since HPO participates in so many fundamental processes, the regulation mechanism of its gene expression might display significant impact on those fundamental processes. Up to now, however, there have been no reports on such a mechanism in this promising family. In the present study, we tried to elucidate the molecular mechanisms that direct the expression of the *HPO* gene. The 5′-flanking region of the human *HPO* gene was cloned and its promoter activity was characterized. The data showed that a functional initiator (Inr)-like element and its flanking repeat elements together comprise a core promoter to direct the transcriptional initiation of the *HPO* gene. In addition, the significance of *HPO*'s initiator with its flanking elements as potentially a novel type of general transcriptional apparatus is further discussed.

2. Materials and methods

2.1. Isolation of the HPO promoter region from the human genome

A 1529-bp fragment of the *HPO* promoter region was isolated from human placenta genomic DNA by polymerase chain reaction (PCR) with sense primer 5′-GAAATCCACTCCCCATATGATAGCCAGC-3′ and antisense primer 5′-CCACGTCTTGAAGTCGACGCAGGC-CCGG-3′. The amplified fragment was ligated into pGEM-T vector (Promega) to generate pGEM-T-HPO. This genomic DNA was then confirmed by sequencing and used as template of PCR in subsequent plasmid construction.

*Corresponding author. Fax: (86)-10-68171208.
E-mail address: hefc@nic.bmi.ac.cn (F. He).

Abbreviations: HPO, hepatopoietin; EMSA, electrophoretic mobility shift assay; Inr, initiator; DPE, downstream promoter element; IFE, initiator-dependent flanking element; BRE, TFIIB recognition element

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)00158-3

2.2. Plasmid constructions

A series of 5′-truncated constructs were created by PCR with combinations of the common antisense primer 5′-GCGCCTGGGCTGGCGTCGAG-3′ and different sense primers. *Xho*I and *Hin*dIII recognition sequences were incorporated into the sense and antisense primers, respectively, and PCR products were subcloned into the reporter vector pGL3-Basic (Promega) at the corresponding sites. In this way we produced the 5′-truncated plasmids pGL3-HPO(−1222/+205), pGL3-HPO(−1000/+205), pGL3-HPO(−793/+205), pGL3-HPO(−608/+205), pGL3-HPO(−416/+205), pGL3-HPO(−236/+205), pGL3-HPO(−54/+205), and pGL3-HPO(+110/+205). The numbers in parentheses indicate the boundaries of the inserts relative to a putative transcription start site, which is the purine A within the Inr-like element as described in Section 3.

2.3. Mutagenesis

Oligonucleotide-directed mutagenesis with the overlap extension PCR amplification method [15] was used to mutate the major transcription start site and the three flanking CTGGAGGC repeats in the context of pGL3-HPO(−54/+205) by transversion at −21/−18 (TGGC to GTTA); −9/−6 (TGGC to GTTA); −1/−2 (TC to AG); +2/+3 (TC to GA); +4/+7 (TGGC to GTTA) and a transition at +1 (A residue to G). The mutated segments of pGL3-HPO(−54/+205) furnished with *Xho*I and *Hin*dIII sites were directly subcloned into pGL3-Basic as described above. All of the mutants were verified by sequencing.

2.4. Cell culture and transient transfection

Cells used in this experiment were maintained as monolayer cultures and grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum in a humidified 5% CO_2 incubator at 37°C. For transient transfection, cells were plated in 24-well tissue culture plates at a density of 2×10^5 cells per well in 1.2 ml of DMEM (Life Technologies). After a 24-h attachment period (50–60% confluence of cells), transfections were performed with LipofectAmine PLUS reagent (Life Technologies) following the manufacturer's instruction. Briefly, 0.4 μg of each construct of *HPO* promoter, 0.04 μg of pRL-TK and 2 μl of PLUS reagent and 2 μl of LipofectAmine in a total amount of 250 μl serum-free DMEM were mixed and added to cells. After 6 h incubation, the medium containing the LipofectAmine–DNA complex was removed and replaced by DMEM containing 10% fetal bovine serum. Cells were then cultured further for 48 h and harvested in passive lysis buffer (Promega). Luciferase assays were conducted with the Dual Luciferase Assay system (Promega) according to the manufacturer's recommendations with modification for 10 μl of cell lysate in 50 μl of firefly/*Renilla* luciferase assay reagent. To normalize the transfection efficiency, the firefly luciferase activity was divided by the *Renilla* luciferase activity from the internal control pRL-TK. All transfection and reporter assays were performed independently at least three times, each in duplicate or triplicate as shown in the figure legends. The results are presented as means ± S.D.

2.5. Preparation of nuclear extract

Nuclear extracts were prepared from HepG2 cells as described by Therrien and Drouin [16]. Protein concentration was determined according to Bradford [17]. The extracts were frozen in aliquots and stored at −70°C. The HeLa nuclear extract was purchased from Promega.

2.6. Electrophoretic mobility shift assay (EMSA)

Oligonucleotides were annealed in equimolar amounts and end-labeled with ^{32}P in the presence of [γ-^{32}P]ATP and T4 polynucleotide kinase. The oligonucleotides used in these assays were as follows (the mutated nucleotides are in lower case): wild-type Inr, 5′-AGGCTCATCTGGAGGCCGA-3′; mut −2/+2, 5′-AGGCgacgCTGGAGGCCGA-3′; mut +4/+7, 5′-AGGCTCATCgttcGGCCGA-3′; wild-type R1, 5′-TCTGGCCTGGAGGCTGACC-3′; mut R1, 5′-TCTGGCCgttcGGCTGACC-3′. The consensus oligonucleotides for AP1 (5′-CGCTTGATGAGTCAGCCGGAA-3′) and Sp1 (5′-ATTCGATCGGGGCGGGGCGAGC-3′) were purchased from Promega. For each binding reaction, a total of 35 fmol probe labeled with [γ-^{32}P]ATP was combined with 5 μg nuclear proteins in 10 mM Tris–HCl (pH 7.5), 50 mM NaCl, 0.5 mM dithiothreitol, 0.5 mM EDTA, 1 mM $MgCl_2$, 0.05 mg/ml poly(dI-dC) and 4% glycerol (v/v), in a total volume of 10 μl. Reaction mixtures were incubated for 20 min at room temperature and resolved on a non-denaturing (4%, w/v) acrylamide gel (29:1, w/w) in 0.5×TBE. For competition assay, a 100–1000-fold molar excess of various unlabeled competitor DNAs was added to the reaction mixture prior to the addition of the labeled probe. The gel was then exposed to X-ray film at −70°C with an intensifying screen.

3. Results

3.1. Promoter activity of HPO 5′-flanking region

To identify important regions regulating *HPO* gene transcription, we generated a series of sequential 5′-deletions of the *HPO* 1222-bp 5′-flanking region contiguous to the promoter and 205-bp 5′-untranslated region, fused to a luciferase reporter gene (Fig. 1A). These constructs were transiently transfected into HepG2 cells, 293 cells and HeLa cells, and the luciferase activity was determined. The data showed that the promoter region of *HPO* has transcriptional activity in all tested cell lines (Fig. 1B). Furthermore, the effect of sequential 5′-deletion on transcriptional activity was similar among the three transfected cell lines for all constructs except for pGL3-HPO(−608/+205) (Fig. 1B). This indicates that the 5′-flanking region of *HPO* gene might have a similar regulatory pattern in cell lines of various origins. For the pGL3-HPO(−608/+205) construct, the promoter activity was decreased in HepG2 cells and HeLa cells, whereas it was increased in 293 cells (Fig. 1B). Among the various constructs, deletion from −1222 to −1000 reduced the luciferase activity in all three cells. This indicates that positive regulatory elements might be present in the region from −1000 to −1222. Deletion of fragments from −1000 to −608 has less influence on transcriptional activity. However, deletion downstream of −416 resulted in a 1.5–4-fold decrease, whereas further deletion to −236 caused a 5–12-fold increase in luciferase activity (Fig. 1B). These results indicate that positive regulatory elements may exist in the region from −416 to −608 and negative regulatory elements may lie in the region between −416 and −236. In HepG2 and HeLa cells, the highest promoter activity was observed in the pGL3-HPO(−54/+205) construct. However, pGL3-HPO(+110/+205), formed by the removal of 164 bp from the 5′-end of pGL3-HPO(−54/+205), was devoid of promoter activity (Fig. 1B). These data suggest that the minimal promoter of *HPO* lies somewhere between −54 and +110.

3.2. A minimal core promoter is present in the region from −22 to +42

The results above showed that the minimal fragment of −54 to +205 alone has high promoter activity, which indicates that a core promoter may be present in this region and may initiate transcription of the gene. To more accurately dissect the position of the core promoter, we performed reporter assays of 5′- and 3′-truncated forms starting from pGL3-HPO(−54/+205). Luciferase activity was still high when the fragment of −54 to +205 was truncated to −22 and then was lost upon further truncation to +27 from the 5′-end (Fig. 1C). The results suggest that the minimal core promoter exists in the region between −22 and +27. The deduction was confirmed by truncation from the 3′-end. The promoter activity of pGL3-HPO(−54/+42) was still 75.8% of that seen with pGL3-HPO(−54/+205) (Fig. 1C). Sequence analysis of the region revealed that it lacks a canonical TATA box, but has

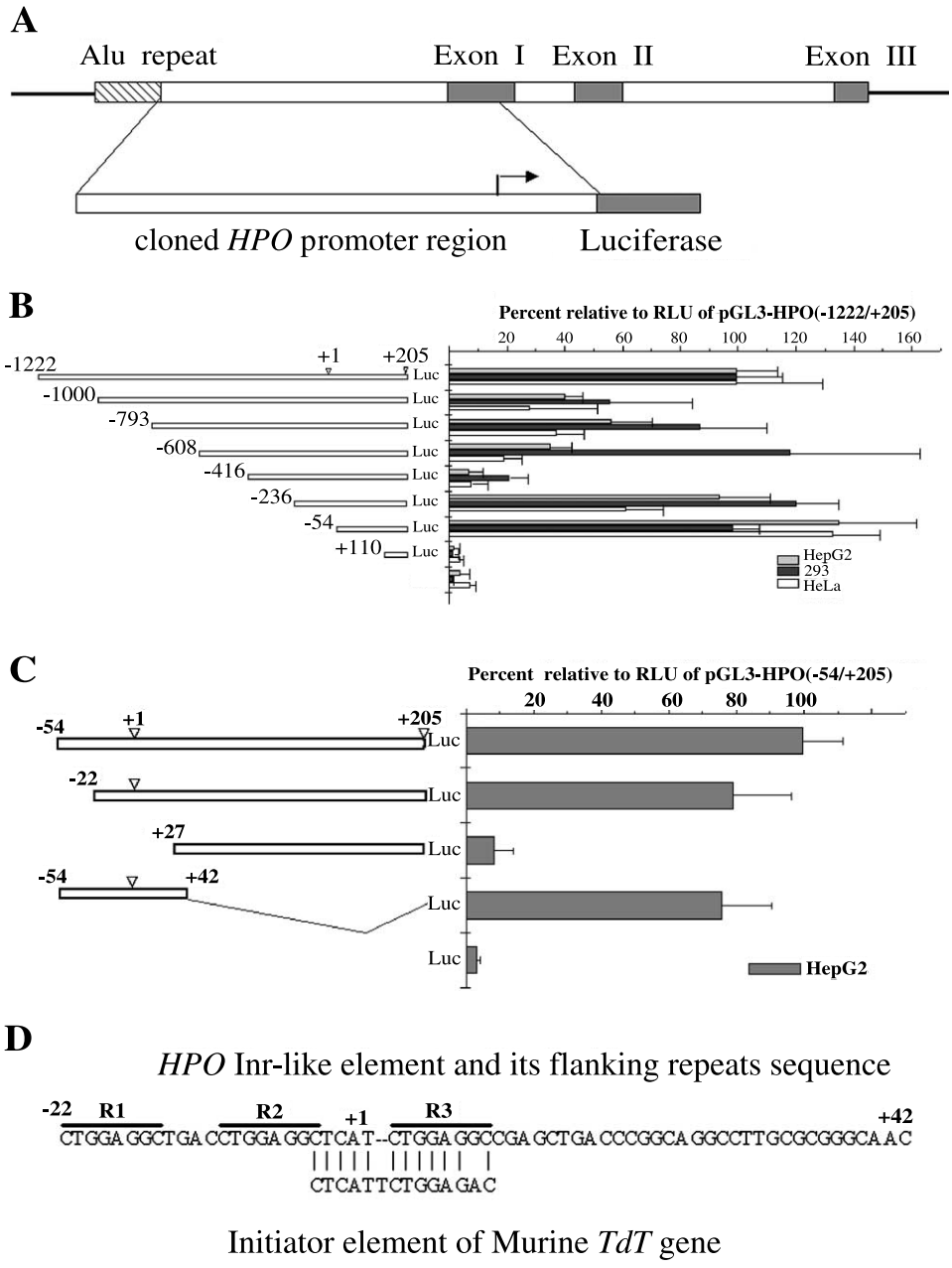

Fig. 1. Transcriptional activities of the human *HPO* promoter. A: Schematic diagram of the cloned *HPO* promoter region, showing the 5′-flanking region of *HPO* immediately downstream of an *Alu* repeat (hatched bar) to exon I (filled bar). A total of 1427 bp was cloned by PCR according to GenBank™ sequence AC005606. The purine A within Inr-like element CTCATCTGG was designated +1. B: The 5′-truncated *HPO* promoter regions shown on the left were inserted into pGL3-Basic. The reporter constructs were transfected to HepG2 (grey filled bar), 293 (dark filled bar) and HeLa (open bar) cells. Firefly luciferase activities were normalized to the *Renilla* luciferase activity of pRL-TK co-transfected as internal control, and then expressed as percent relative luciferase units (RLU) given by pGL3-HPO(−1222/+205). Results are represented as the mean ± S.D. of three independent experiments in triplicate. C: A minimal core promoter activity in the region −22 to +42. A fragment extending from −54 to +205 was truncated from either the 5′- or the 3′-proximal side (left, open bar). The constructs were transfected into HepG2 cells and luciferase activities were measured (filled bar). The firefly luciferase activities were normalized to the *Renilla* luciferase activity, and then were presented as percent relative to activity of pGL3-HPO(−54/+205). Results are represented as the mean ± S.D. of four independent experiments in duplicate. D: Alignment of Inr-like element between murine *TdT* core promoter and *HPO* core promoter. The Inr-like element and three CTGGAGGC repeat elements, designated Inr, R1, R2 and R3 respectively, are presented in the region from −22 to +42 of pGL3-HPO(−54/+205). The three repeat elements are indicated by lines above the sequence. The Inr-like element of *HPO* was aligned to the murine *TdT* Inr element. The putative transcription start site was denoted +1.

a sequence CTCATCTGGAGGC with homology to the Inr sequence element CTCATTCTGGAGAC in the mouse terminal deoxynucleotidyltransferase (*TdT*) gene [18]. Intriguingly, in addition to one CTGGAGGC sequence within the Inr-like element, two other tandem repeats lie upstream of the Inr-like element (Fig. 1D), suggesting that the Inr-like element and its flanking repeats might be involved in transcriptional initiation of the *HPO* gene.

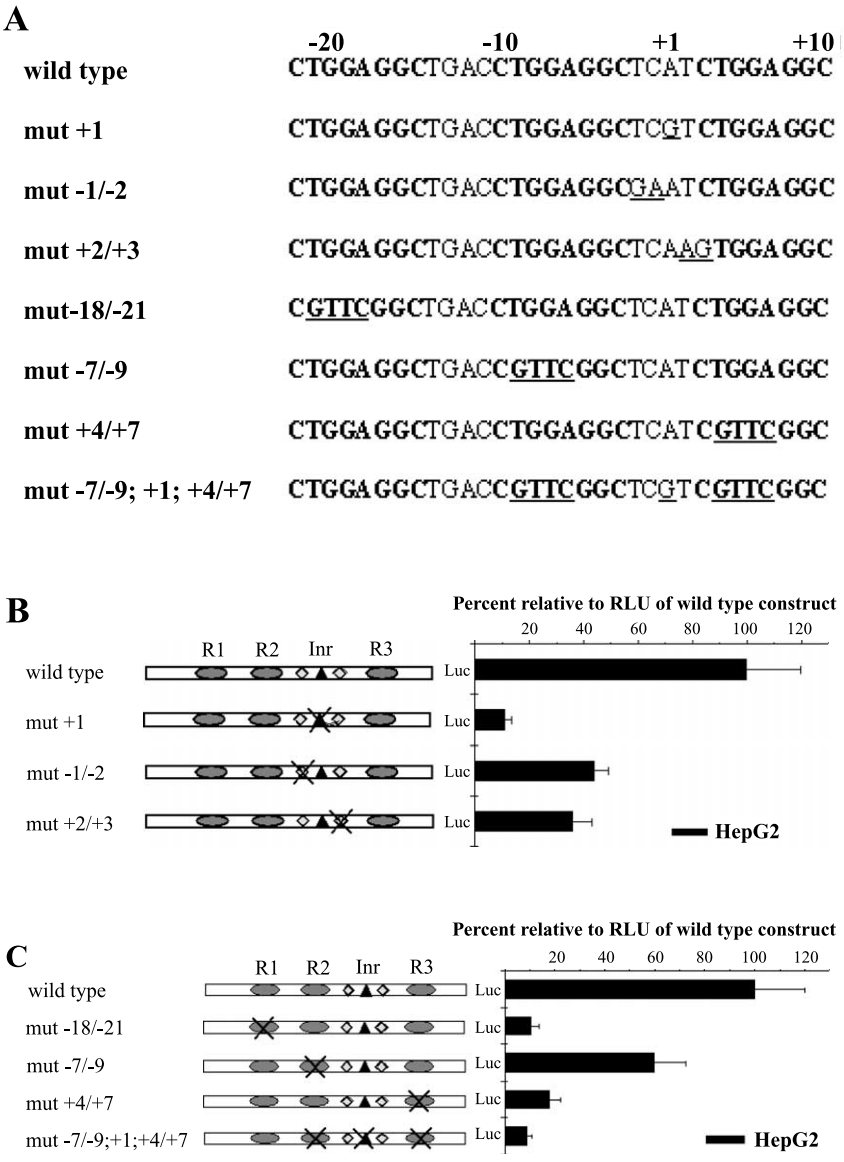

Fig. 2. Reporter assay of the functional initiator and its flanking CTGGAGGC repeat elements. A: The region from −22 to +10, which contains the Inr-like element and triple repeats, was mutated in the context of pGL3-HPO(−54/+205). The mutated nucleotides are underlined corresponding to the sites of wild-type and marked on the left. B: Effect of mutagenesis on Inr elements. The mutants of Inr were transfected into HepG2 cells and the effect of mutagenesis was represented as percent relative to wild-type pGL3-HPO(−54/+205). C: Effect of mutagenesis of three CTGGAGGC tandem repeat elements. Xs indicate mutations introduced at specific sites marked on the left. The mutants of Inr-flanking element were transfected into HepG2 cells and the effect of mutagenesis was represented as percent relative to wild-type pGL3-HPO(−54/+205). Results are represented as the mean ± S.D. of three independent experiments in triplicate.

3.3. *A functional initiator and its flanking CTGGAGGC repeats comprise the core promoter of the* HPO *gene*

To test whether the Inr-like element and the triple repeats in the *HPO* promoter region are functional, mutagenesis on the elements of interest was performed in the context of pGL3-HPO(−54/+205) and the mutants were transiently transfected into HepG2 cells. Reporter assay data (Fig. 2A) showed that the transition of A to G at site +1 reduced transcriptional activity of the promoter to about 11% relative to the activity of the wild-type promoter. Transversion of TC to GA at positions −1/−2 and TC to AG at +2/+3 also reduced the promoter activity, but less strongly than the +1 mutation (Fig. 2B). These data clearly indicate that the nucleotides between −2 and +3, and especially A at +1, are critical for efficient transcription of the *HPO* promoter in vivo. This is consistent with the findings for several other Inr-containing promoters, in which mutation at −3 to +3 drastically impaired in vivo transcription [19]. This sequence is also consistent with the preferred $PyPyA^{+1}NPyPy$ consensus sequence for a functional Inr element [20]. The Inr-proximal sequence has been found to be essential for the transcription of several Inr-containing genes [18,21]. In contrast to those genes, intriguingly, three tandem repeats of a CTGGAGGC sequence in the *HPO* core promoter were arrayed in the space of 4 bp (Fig. 1D). To assess the functional importance of the CTGGAGGC sequence for the promoter activity, we mutated the TGGA of each repeat to GTTC in the context of pGL3-HPO(−54/+205) (Fig. 2A). The transcriptional activity of those constructs was measured by transient transfection in HepG2 cells. Mutation of R1 and R3 very strongly reduced the activity of *HPO*

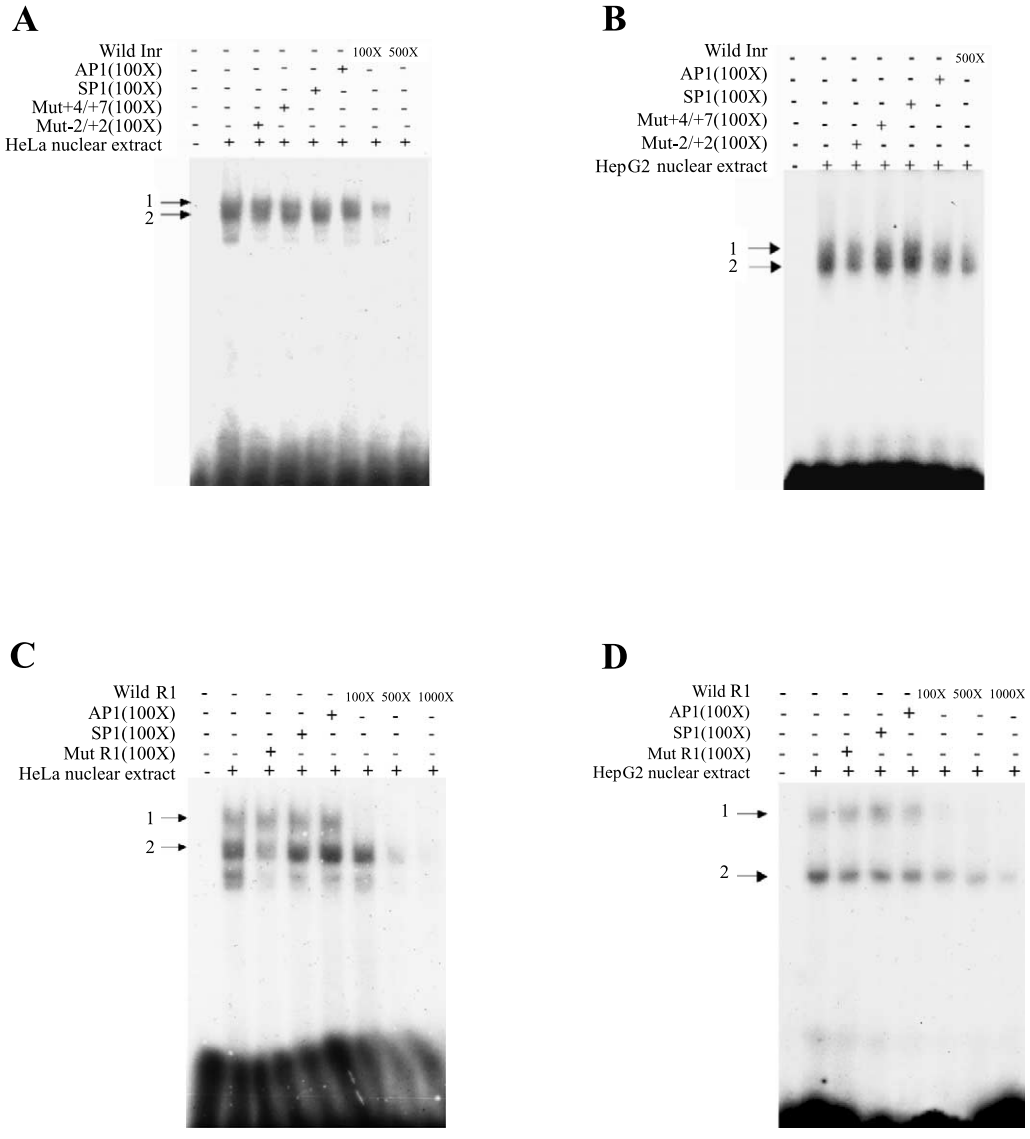

Fig. 3. Binding of specific nuclear factors to Inr and CTGGAGGC repeat elements. A: 35 fmol ^{32}P-labeled Inr probe was incubated with 5 μg HeLa nuclear extract (NE) in the absence (lane 1) or presence (lanes 2–8) of NE. For lanes 3–7, a 100-fold molar excess of the indicated cold competitor was added as follows: mutated Inr at −2/+2 (lane 3) and +4/+7 (lane 4), SP1 (lane 5), AP1 (lane 6), wild-type Inr (lane 7); for lane 8, a 500-fold molar excess of wild-type Inr cold probe was added. Arrows 1 and 2 indicate the shifted protein–DNA complexes. B: 35 fmol ^{32}P-labeled Inr probe was incubated with 5 μg HepG2 NE in the absence (lane 1) or presence (lanes 2–7) of NE. Cold competitors were added in binding reaction as shown on the left (fold excess in parentheses) before addition of labeled probe. The plus represents addition of cold competitor and minus represents no addition of competitors. Shifted complexes are indicated by arrows. C: ^{32}P-labeled R1 probe was incubated with 5 μg HeLa NE in the absence (lane 1) or presence (lanes 2–8) of NE, with addition of competitors as described above. D: ^{32}P-labeled R1 probe was incubated with 5 μg HepG2 NE in the absence (lane 1) or presence (lanes 2–8) of NE, with addition of competitors as described above.

relative to wild-type, and mutation of R2 had a lesser effect (Fig. 2C). A combination of mutated R2, R3 and the mutation of the +1 site from A to G further reduced the promoter activity to about 8% of wild-type activity (Fig. 2C). These results suggest that the CTGGAGGC repeats, especially R1 and R3, are important for optimal Inr-dependent transcription of the *HPO* gene.

3.4. *Specific nuclear factors bind to the initiator and CTGGAGGC* cis*-acting elements*

Cis-acting elements are generally bound by *trans*-acting factors. We therefore assayed HepG2 and HeLa nuclear extracts for factors specifically binding to the Inr and CTGGAGGC repeat element. We performed EMSA using ^{32}P-labeled double-stranded oligonucleotides containing the wild-type *HPO* Inr and R1 element or corresponding point mutants. For the Inr oligonucleotides, nuclear extract from HepG2 and HeLa cells both formed two gel-shifted complexes. Those complexes could be abolished by unlabeled wild-type Inr probe, but were competed neither by irrelevant probes corresponding to SP1 or AP1 binding sites, nor by oligonucleotides mutated at the −2/+2 or +4/+7 sites (Fig. 3A,B). The two shifted complexes have only a slight difference in electrophoretic mobility rate, but differ considerably in sensitivity to competition by unlabeled Inr probe. Complex 2 from HeLa nuclear extract and complex 1 from HepG2 extract could be abolished by a 100-fold molar excess of cold probe. However, complex 1 from HeLa extract and complex 2 from HepG2

extract could not, indicating that these complexes are stabler. Similar results were obtained for the CTGGAGGC repeat element. Two major complexes were formed in both binding reactions, and they displayed different sensitivities to competition by the unlabeled R1 probe (Fig. 3C,D). Complex 2 is stabler than complex 1 for nuclear extracts from HeLa or HepG2 cells. Thus, nuclear factors specifically recognize and bind to the Inr and the CTGGAGGC repeat elements. Furthermore, the similar binding patterns suggest that similar nuclear factors from HepG2 and HeLa cells may bind to both the Inr and the repeat elements.

4. Discussion

In this report we describe the identification and characterization of the TATA-less *HPO* promoter region. Through a series of sequential deletions from the 5′-end of cloned fragments of the *HPO* promoter, we demonstrated that a core promoter is present between −54 and +42 (Fig. 1B,C). Mutagenesis studies showed that this core promoter consists of a functional initiator and three CTGGAGGC tandem repeat elements (Fig. 2B,C). Both the initiator and at least one of the repeats (R1) were bound by specific nuclear factors (Fig. 3A–D). Our results indicate that Inr and CTGGAGGC elements potentially comprise a novel type of core promoter that might maintain constitutive expression of at least one transcript of the *HPO* gene.

4.1. Constitutive and inducible expression of the HPO gene

Expression profiles showed that *HPO* is widely expressed in all of the tested tissues and cell lines (unpublished data). So far, it has been demonstrated that HPO/ALR/Erv1 distributes into several subcellular compartments, including the nucleus [8], mitochondria intermembrane space [11], cell membrane [5–7], endoplasmic reticulum [22] and extracellular space [23]. In addition, it is involved in many fundamental processes including protein folding [22], mitochondrial maintenance [24,25], transmembrane signaling cascade [5,7], cellular iron homeostasis [26], and the regulation of nuclear gene expression [8,9,27]. Some of those functions may represent constitutive roles of *HPO*, such as in mitochondrion biogenesis and protein folding, which are indispensable in all cells. However, others might reflect inducible functions such as transcriptional regulation of nuclear genes and transmembrane signaling cascade, in which HPO could be conditionally induced under stresses such as partial hepatectomy, hepatitis and hepatic carcinogenesis. Those versatile functions need the *HPO* gene to be expressed both in a constitutive and in an inducible fashion. Through reporter assays of *HPO* 5′-end sequential truncations, here we found that the transcription of the *HPO* gene is initiated through an Inr-containing core promoter. Moreover, the data obtained from serial deletions of the promoter region showed that the promoter region of *HPO* mainly has more the characteristics of a housekeeping gene. Consistent with our findings, the promoters of human housekeeping genes, oncogenes, growth factors, and transcription factors are often TATA-less [28]. Our findings partly explain the mechanisms in which one transcript of the *HPO* gene is constitutively regulated. Some studies showed that expression of the *HPO* gene is higher under certain circumstances such as partial hepatectomy [29] or acute liver diseases [10] than in normal states. It is possible that some transcripts of *HPO* can be conditionally induced under stress. Further efforts are needed to elucidate the regulatory mechanism of inducible expression of *HPO*.

4.2. Inr together with its flanking CTGGAGGC element comprises a subset of Inr-dependent core promoters

Studies on eukaryotic promoters have identified several kinds of core promoter elements, which are characteristic DNA sequences required for promoter function [30]. The TATA box is an A/T-rich sequence located ∼25–30 nucleotides upstream of the transcriptional start site. It contains a consensus sequence, TATA(A/T)A(A/T), which is recognized by the TATA binding protein subunit of TFIID, which nucleates the formation of a pre-initiation complex [31]. The downstream promoter element (DPE), which is located 28–34 nucleotides downstream of the transcription start site in many *Drosophila* TATA-less promoters [30,32], has a consensus sequence (A/G)G(A/T)CGTG. Two of the DPE-acting factors, TFIID and NC2, may work synergistically through the DPE [33]. An upstream core promoter element, (G/C)(G/C)(G/A)CGCC, is recognized by TFIIB. This TFIIB recognition element (BRE) is located immediately upstream of the TATA box of some promoters and increases the affinity of TFIIB for the promoter [34]. The initiator is another conserved core promoter element, which contains a pyrimidine (Y)-rich core sequence $YYA^{+1}N(T/A)YY$ encompassing the transcriptional start site [19]. It is sufficient to direct accurate transcription initiation either alone or in conjunction with a TATA box or other core promoter element [31,35]. Among the binding factors of Inr, $TAF_{II}250$ is a major candidate. $TAF_{II}250$, together with $TAF_{II}150$, is widely used in both TATA-box-containing and TATA-less core promoters and mediates binding of TFIID to the Inr [36,37]. Several other factors including RNA polymerase II [38], TFII-I [39], USF [40] and YY1 [41] have been implicated in Inr function. However, they are restricted to a small subset of Inr elements [19]. These four kinds of core promoter elements, either alone or combined with others, form distinct classes of core promoters corresponding to different sets of genes. The various types of core promoter combinations, including TATA alone, TATA with DPE, TATA with Inr, TATA with BRE and Inr alone or Inr with DPE, are then recognized by distinct combinations of general transcription factors, especially distinct set of TAFs [30,42]. Now we speculate that Inr with a CTGGAGGC element – an Inr-dependent flanking element (IFE) – might comprise a novel type of core promoter to direct expression of a subset of genes including *HPO*.

HPO	CTGGAGGCTGACCTGGAGGCTCATCTGGAGGC
TdT	GAGCCCTCATTCTGGAGACACCACCT
CFI	ATCTCTCTGGATTTCAGCCAAATCTTTCAAG
PBGD	TCAGGGCTCAGTGTCCTGGTTACTGCATC
XDH Inr4	CATTGGGTAACTTTGTTTCATTTTGCTGGGAGGCGTATC
C1INH	ACCCTGGGGGACTCTCTACTCAGTCTGCAC
singled	GTTTCATTCCCACTGGAGTGCAGTT

IFE consensus sequence: CTGGNNN(N)C

Fig. 4. The IFE can lie upstream or downstream of the initiator element. Inr and IFE are present in many TATA-less genes, named on the left. The transcription start site within the initiator, usually adenine, is underlined and aligned. The IFE is in bold type. CFI, complement factor I; PBGD, human porphobilinogen deaminase; XDH, xanthine dehydrogenase; C1INH, C1 inhibitor.

The core promoter of *HPO* is TATA-less and Inr-containing. Sequence examination near Inr showed that the proximal region of Inr is G/C-rich and does not contain the consensus sequence of TATA, DPE or BRE elements. Interestingly, three tandem repeats of CTGGAGGC were found surrounding the Inr, and all of them make a great contribution to Inr-dependent transcription. More interestingly, these repeating element, lying either upstream or downstream of Inr, appear in many other Inr-containing genes, such as complementary factor I [21], TdT [18], human porphobilinogen deaminase [43], xanthine dehydrogenase [44], C1 inhibitor [45], *singled* [42], etc. (shown in Fig. 4). They have a consensus sequence CTGGNNN(N)C. Based on these results, we could postulate that the CTGGNNN(N)C sequence, as an IFE, together with Inr consists of one type of core promoter. This type of core promoter, denoted Inr^+-IFE^+, drives transcription of a subset of genes. It is clear that specific nuclear factors bind to the Inr and IFE elements. Nevertheless, further study is needed to identify those nuclear proteins, which would be helpful to understand how IFE combines with Inr to facilitate the assembly of the basic transcription machinery on the core promoter.

Acknowledgements: We thank the editor for advice on the manuscript. This work was supported by the Chinese State Key Program for Basic Research and Natural Science Foundation of China (30170450, 30170520).

References

[1] He, F. and Wu, C. (1997) Chin. Sci. Bull. 42, 1505–1511.
[2] Polimeno, L., Lisowsky, T. and Francavilla, A. (1999) Ital. J. Gastroenterol. Hepatol. 31, 494–500.
[3] Kadokura, H. and Beckwith, J. (2001) Nat. Cell Biol. 3, E247–249.
[4] Francavilla, A., Hagiya, M., Porter, K.A., Polimno, L., Ihara, I. and Starzl, T.E. (1994) Hepatology 20, 747–757.
[5] Wang, G., Yang, X., Zhang, Y., Wang, Q., Chen, H., Wei, H., Xing, G., Xie, L., Hu, Z., Zhang, C., Fang, D., Wu, C. and He, F. (1999) J. Biol. Chem. 274, 11469–11472.
[6] Li, Y., Xing, G., Wang, Q., Li, M., Wei, H., Fan, G., Chen, J., Yang, X., Wu, C., Chen, H. and He, F. (2001) DNA Cell Biol. 20, 791–795.
[7] Li, Y., Li, M., Xing, G., Hu,.Z., Wang, Q., Dong, C., Wei, H., Fan, G., Chen, J., Yang, X., Zhao, S., Chen, H., Guan, K., Wu, C., Zhang, C. and He, F. (2000) J. Biol. Chem. 275, 37443–37447.
[8] Lu, C., Li, Y., Zhao, Y., Xing, G., Tang, F., Wang, Q., Sun, Y., Wei, H., Yang, X., Wu, C., Chen, J., Guan, K.L., Zhang, C., Chen, H. and He, F. (2002) FASEB J. 16, 90–92.
[9] Polimeno, L., Capuano, F., Marangi, L.C., Margiotta, M., Lisowsky, T., Ierardi, E., Francavilla, R. and Francavilla, A. (2000) Dig. Liver Dis. 32, 510–517.
[10] Tanigawa, K., Sakaida, I., Masuhara, M., Hagiya, M. and Okita, K. (2000) J. Gastroenterol. 35, 112–119.
[11] Lange, H., Lisowsky, T., Gerber, J., Muhlenhoff, U., Kispal, G. and Lill, R. (2001) EMBO Rep. 2, 715–720.
[12] Lisowsky, T., Lee, J.E., Polimeno, L., Francavilla, A. and Hofhaus, G. (2001) Dig. Liver Dis. 33, 173–180.
[13] Lee, J., Hofhaus, G. and Lisowsky, T. (2000) FEBS Lett. 477, 62–66.
[14] Senkevich, T.G., White, C.L., Koonin, E.V. and Moss, B. (2000) Proc. Natl. Acad. Sci. USA 97, 12068–12073.
[15] Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989) Gene 77, 51–59.
[16] Therrien, M. and Drouin, J. (1993) Mol. Cell. Biol. 13, 2342–2453.
[17] Bradford, M.M. (1976) Anal. Biochem. 72, 248–254.
[18] Smale, S.T. and Baltimore, D. (1989) Cell 57, 103–113.
[19] Smale, S.T. (1997) Biochim. Biophys. Acta 1351, 73–88.
[20] Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B. and Smale, S.T. (1994) Mol. Cell. Biol. 14, 116–127.
[21] Paramaswara, B. and Minta, J.O. (1999) Gene 237, 71–80.
[22] Sevier, C.S., Cuozzo, J.W., Vala, A., Aslund, F. and Kaiser, C.A. (2001) Nat. Cell Biol. 3, 874–882.
[23] Hoober, K., Sheasley, S.L., Gilbert, H.F. and Thorpe, C. (1999) J. Biol. Chem. 274, 22147–22150.
[24] Becher, D., Kricke, J., Stein, G. and Lisowsky, T. (1999) Yeast 15, 1171–1181.
[25] Lisowsky, T. (1992) Mol. Gen. Genet. 232, 58–64.
[26] Shao, Q. and Mei, H.H. (1993) Sheng Li Xue Bao 45, 387–394.
[27] Polimeno, L., Margiotta, M., Lisowsky, T., Azzarone, A., Ierardi, E., Frassanito, M.A., Francavilla, T. and Francavilla, A. (2000) Dig. Liver Dis. 32, 217–225.
[28] Zhang, M.Q. (1998) Genome Res. 8, 319–326.
[29] Yang, X.M., Hu, Z.Y., Xie, L., Wu, Z.Z. and He, F.C. (1997) Sheng Li Xue Bao 49, 557–561.
[30] Zhou, T. and Chiang, C. (2001) J. Biol. Chem. 276, 25503–25511.
[31] Burley, S.K. and Roeder, R.G. (1996) Annu. Rev. Biochem. 65, 769–799.
[32] Burke, T.W. and Kadonaga, J.T. (1996) Genes Dev. 10, 711–724.
[33] Willy, P.J., Kobayashi, R. and Kadonaga, J.T. (2000) Science 290, 982–985.
[34] Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D. and Ebright, R.H. (1998) Genes Dev. 12, 33–44.
[35] Kutach, A.K. and Kadonaga, J.T. (2000) Mol. Cell. Biol. 20, 4754–4764.
[36] Verrijzer, C.P., Chen, J.L., Yokomori, K. and Tjian, R. (1995) Cell 81, 1115–1125.
[37] Wu, C.H., Madabusi, L., Nishioka, H., Emanuel, P., Sypes, M., Arkhipova, I. and Gilmour, D.S. (2001) Mol. Cell. Biol. 21, 1593–1602.
[38] Carcamo, J., Buckbinder, L. and Reinberg, D. (1991) Proc. Natl. Acad. Sci. USA 88, 8052–8056.
[39] Roy, A.L., Carruthers, C., Gutjahr, T. and Roeder, R.G. (1993) Nature 365, 359–361.
[40] Du, H., Roy, A.L. and Roeder, R.G. (1993) EMBO J. 12, 501–511.
[41] Usheva, A. and Shenk, T. (1994) Cell 76, 1115–1121.
[42] Kutach, A.K. and Kadonaga, J.T. (2000) Mol. Cell. Biol. 20, 4754–4764.
[43] Beaupain, D., Eleouet, J.F. and Romeo, P.H. (1990) Nucleic Acids Res. 18, 6509–6515.
[44] Clark, M.P., Chow, C.W., Rinaldo, J.E. and Chalkley, R. (1998) Nucleic Acids Res. 26, 2813–2820.
[45] Zahedi, K., Prada, A.E. and Davis III, A.E. (1994) J. Biol. Chem. 269, 9669–9674.