Rapid in vivo footprinting technique identifies proteins bound to the TTR gene in the mouse liver

Jovan Mirkovitch and James E. Darnell Jr.
The Rockefeller University, New York, New York 10021 USA

In vivo examination of the occupancy of DNA elements that can regulate transcription is critical to reveal which proteins actually take part in establishing and maintaining gene expression. We describe a new genomic sequencing method involving the rapid purification of relevant DNA segments from the bulk of the genomic DNA using a biotinylated riboprobe. The purified sequences are revealed by a single primer extension using *Taq* DNA polymerase. We used this technique to study the promoter and the enhancer of mouse transthyretin (TTR), a gene highly expressed in the liver. Footprints showed high liver-specific occupancy of some, but not all, of the DNA sites that had been identified as important for expression by transfection studies in hepatoma cells. In addition, several previously undetected sites were observed that bound proteins specifically in liver. These results suggest that not all demonstrable binding sites are involved in ongoing transcription and that in vivo studies may reveal additional and probably more relevant sites.

[Key Words: In vivo footprinting, transcription factors, chromatin, transthyretin]

Received September 26, 1990; revised version accepted October 29, 1990.

In recent years molecular genetic and gene transfer techniques coupled with isolation and characterization of DNA-binding proteins have begun to unravel the complexities of transcriptional control. It is now clear that most genes have multiple binding sites for proteins that presumably influence their expression, and dozens of regulatory DNA-binding proteins have been described, some found in most or all cell types and some with narrowly restricted cell distribution (Johnson and McKnight 1989; Mitchell and Tjian 1989). The actual participation of these sites and their cognate proteins is difficult to establish in the differentiated tissues of organisms that lack a detailed formal genetics. One straightforward, already described technique, in vivo footprinting, carried out with a variety of enzymes and/or chemicals that attack DNA, is of major assistance in defining which putative regulatory sequences are relevant (Ephrussi et al. 1985; Jackson and Felsenfeld 1985; Selleck and Majors 1987; Becker et al. 1989; Mueller and Wold 1989; Saluz and Jost 1989; Zhang and Gralla 1989). However, the clear identification of genomic sequences that interact in vivo with proteins of a known DNA-binding capacity requires a simple and reliable genomic sequencing technique with resolution at the nucleotide level. In vivo footprints with bacterial genes and with yeast genes at this level have been quite successful (Selleck and Majors 1987; Axelrod and Majors 1989; Huibregtsse and Engelke 1989; Sasse-Dwight and Gralla 1990). Although a number of techniques have been described for vertebrate cells, they are either not very sensitive or are difficult to carry out. The main limitation is that the sequences to be analyzed are present in a large excess of irrelevant DNA. Here, we describe a new technique that involves the purification from chemically or enzymatically treated nuclei of specific sequences away from the bulk of the genome; the purified sequences can then be copied by a single primer extension with a highly labeled probe, thus revealing the lengths of the selected fragments. The procedure is rapid, sensitive, and works with the several genomic sequences tested so far.

In this paper we report the analysis of the DNase I and copper-phenanthroline accessibility of the mouse transthyretin (TTR) gene regulatory elements in the liver, where the gene is actively transcribed, and in tissues where it is not transcribed. The enhancer and promoter regions of this gene, defined by transfection studies in hepatoma cells (Costa et al. 1986) and confirmed by transgenic animal studies (Yan et al. 1990), were highly sensitive to DNase I cleavage in liver but not in spleen, brain, or kidney nuclei. Using the new genomic sequencing technique it was shown that some, but not all, of the TTR regulatory sequences defined by gene transfer technique were bound by proteins in the adult liver. Sites binding the liver transcription factor HNF-3 (Costa et al. 1989; Lai et al. 1990) were highly protected in the promoter. New binding sites for the liver factor HNF-4 (Costa et al. 1989) and the ubiquitous factor NF-1 (for review, see Rupp et al. 1990) were also detected in the promoter. The enhancer showed protections at two sites that can bind the factor C/EBP (Costa et al. 1988b), when the promoter C/EBP sites were not
protected. These experiments identify the components that are part of the stable transcription machinery directing adult liver-specific expression of the TTR gene. We suggest that this technique applied to different cells under different conditions of gene activity should be invaluable in exploring true intranuclear protein–DNA interactions.

Results
The TTR enhancer and promoter are hypersensitive to DNase I in hepatocyte nuclei

Deletion and mutagenesis studies with the TTR gene identified a tissue-specific enhancer at -2 kb and promoter region up to 200 bases upstream of the initiation site. These elements directed high-level expression in human hepatoma cells but not in HeLa cells (Costa et al. 1986, 1988a). Transcriptional control sequences from active genes are known to be hypersensitive to nuclease cleavage within the chromatin of the nucleus (for review, see Gross and Garrard 1988). We applied this test to the TTR gene in several mouse tissues to assay the relative sensitivity to digestion of the TTR promoter and enhancer sequences. After digestion of nuclei with DNase I, the genomic DNA was purified, digested with BamHI, separated according to size, and blotted to nylon filters, and the in vivo cleavage sites were mapped by indirect end-labeling using radiolabeled probes from either end of the BamHI fragment. As shown in Figure 1, the promoter and enhancer regions were very susceptible to DNase I cleavage in liver nuclei, revealing a number of discrete hypersensitive sites numbered 1-8. The majority of the liver chromatin was sensitive in both promoter and enhancer regions. For example, using probe 1, the enhancer site 3 (most distant from the probe) disappeared at increasing DNase I concentrations while the proximal promoter sites (5-8), became more prominent. This points to the hypersensitivity of the promoter in most or all templates that are sensitive at the enhancer site. A mouse hepatoma cell line, BW1 [Clayton et al. 1985], showed a pattern similar to liver. Site 3 corresponds to the enhancer, while the promoter presented a number of hypersensitive sites (5-8) demonstrating a region of ~300 bp that is highly sensitive to DNase I. In transient assays the deletion of sites 1, 2, and 4 had little effect on the expression of TTR [Costa et al. 1986], but these sites are clearly absent in all other tissues tested. However, in transgenic animals, a gene with only the enhancer and promoter was actively transcribed in the liver but was also inappropriately expressed in the brain. In contrast, a construct containing 3 kb of upstream DNA, including hypersensitive sites 2 and 4, showed a tissue distribution similar to the TTR gene [Yan et al. 1990].

With respect to cell specificity, no hypersensitive region was detected in total spleen or brain nuclei. A weak sensitivity of the promoter in kidney nuclei was possibly due to a low level of TTR expression in mouse kidneys. TTR is highly expressed in the choroid plexus in the brain (Dickson et al. 1985; Soprano et al. 1985), but no hypersensitive site was detected, presumably because these nuclei constitute a tiny fraction of total brain nuclei. The presence of hypersensitive sites in the TTR enhancer and promoter suggests that these regions are readily accessible to transcriptional regulatory proteins, but the DNase I hypersensitivity analysis does not provide details about the occupancy of the protein-binding sites.

A simple and reliable genomic sequencing technique

We wished to develop a simple method for the analysis of DNA–protein interactions at the nucleotide level in cells with large genomes and to apply this technique first to the TTR regulatory regions in the mouse liver. A scheme was devised that involves, as a first step, the purification of relevant DNA sequences from the bulk of the genomic DNA [see Fig. 2A]. Cleaved genomic DNA was hybridized to a biotinylated riboprobe complementary to the sequence of interest. The hybrids were immobilized on streptavidin–agarose, and the purified genomic sequences were recovered by alkali denaturation. The use of a pure ribonucleotide probe that could be hydrolyzed rather than a deoxyribonucleotide probe was essential to obtain a genomic fraction free of the selecting nucleic acid. The described procedure was quantitative, and a number of different sequences can be purified from the same hybridization mixture.

A primer extension procedure was then used to reveal the cleaved products. An oligonucleotide of 25 residues containing ~10 radiolabeled nucleotides was annealed to the purified genomic sequences and further extended a single time using Taq DNA polymerase. A test to show that this simple detection procedure presents an accurate image of the population of cleaved products is shown in Figure 2B. A plasmid containing the TTR promoter was treated with dimethylsulfate (DMS) and cleaved with piperidine, and 1 ng was used as a template for primer extension [Fig. 2B, lane 2]. The G-ladder obtained on separation of DNA fragments was identical to that obtained if the same fragment was first end-labeled by a fill-in reaction and thereafter treated with G-specific chemistry [Fig. 2B, lane 1]. The two sequences presented in Figure 2B are shifted relative to each other because the label is at different positions and on different strands. However, the relative intensity of different G residues is the same in both lanes. This shows that a single round of primer extension with Taq polymerase permits the quantitative analysis of a complex population of cleaved molecules. In addition, as seen in Figure 3, lanes 1 and 2, the DNase I cleavage pattern on plasmid or genomic DNA was identical, and no artificial band appeared in the genomic lane if no DNase I was added [Fig. 3, lanes 4 and 7].

The TTR promoter binds a number of proteins at transcriptional regulatory sequences specifically in liver nuclei

The mouse TTR gene has been mapped thoroughly by deletion analysis and site-specific mutagenesis to define
Footprinting of TTR transcription control elements

**Figure 1.** TTR DNase I hypersensitivity upstream of the mouse TTR gene. Nuclei isolated from different tissues were digested with increasing concentrations of DNase I for 10 min on ice. Lanes 1, 2, 3, 4, 5, 6, 7, and 8 correspond to 0, 12.5, 25, 50, 100, 200, 400, and 800 μg/ml DNase I, respectively. Genomic DNA was purified, digested with BamHI, underwent electrophoresis on 0.8% agarose gels, and was blotted on nylon membranes. Hypersensitive sites were revealed by indirect end-labeling using two probes at each end of the BamHI fragment to map precisely the positions of the hypersensitive sites. Probe 1 cross-hybridized with a small fragment present in samples that were not digested with DNase I and did not appear using probe 2. Markers were the HindIII fragments of phage λ.
The genomic DNA was digested with restriction endonucleases BgIII and HaeIII, cutting, respectively, at positions -330 and -88 relative to the cap site. For each sample, ~200 μg of DNA was hybridized to a biotinylated SP6 RNA spanning the entire BgIII–HaeIII fragment, and the purified DNA segments from each strand were analyzed for protein footprints by the procedure outlined in Figure 2A. As shown in Figure 3, six regions of the TTR promoter were clearly protected from DNase I digestion in liver nuclei when compared to spleen nuclei or naked DNA. The same regions appeared protected from cleavage by copper–phenanthroline [Fig. 4A], although the protected regions were narrower in the latter case (data are summarized in Fig. 5).

With respect to the details of the liver-specific DNase I cleavage pattern, a protected region from positions -18 to -35 encompassed the TATA box and probably signifies the presence of the TATA-binding factor TFIIID [Horikoshi et al. 1988; Nakajima et al. 1988]. Protection was asymmetric around the transcription start site, the lower strand was protected from -10 to +25 [Fig. 3B], while only positions +17 to +21 were protected on the upper strand [Fig. 3A]. Surprisingly, the upper strand even presented some hypersensitive sites around the RNA initiation [cap] site (−1 and +2). The region around the cap site is sensitive to single-stranded specific modification agents and might be occupied by an RNA polymerase in isolated liver nuclei [J. Mirkovitch and J.E. Darnell, unpubl.]. Two other regions not shown previously to bind proteins in vitro were protected in liver but not in spleen nuclei. The region from -59 to -68 is protected on both strands and is flanked by hypersensitive sites. This region is probably occupied by an NF-1-like factor as it contains the sequence TGGCA (see Rupp et al. 1990). Gel retardation analysis and oligonucleotide competition for DNA–protein complex formation showed that the site at −58 to −68 can bind NF-1-like proteins in vitro [data not shown]. The sequence from −212 to −235, which can be deleted from the promoter without affecting TTR expression in transfection assays [Costa et al. 1986] was also specifically protected in liver nuclei. Nuclear extracts from a number of tissues contain multiple proteins binding specifically to this site [data not shown]. One of these proteins in hepatocytes could be HNF-4, as the sequence −212 to −235 competes efficiently for binding of purified HNF-4 protein to the site originally used to define it (the −140 to −151 region of the TTR promoter; F. Sladek, pers. comm.). A sequence similar to the −212 to −235 region is also present in the rat tyrosine aminotransferase [TAT] gene ~150 nucleotides from the cap site and appears protected in vivo only in TAT-expressing cells [Becker et al. 1987].

Earlier transient expression assays in HepG2 hepatoma cells, coupled with gel retardation analysis and in vitro DNase I protection and methylation interference studies, suggested that HNF-1, HNF-3, and HNF-4 would play important roles in liver-specific expression of TTR by binding between positions −90 and −150. In the in vivo analysis this region of the promoter displayed

at least nine protein-binding sites in the promoter and upstream enhancer regions [Costa et al. 1988a,b, 1989]. In addition to the widely distributed protein API, at least four proteins found in liver cells, but not in most other adult tissues, bind to the TTR regulatory sites. These four liver-enriched proteins include HNF-1, a POU–homeo domain protein [Courtois et al. 1987; Frain et al. 1989], CCAAT/enhancer binding protein [C/EBP], a leucine zipper protein [Costa et al. 1988b; Landschulz et al. 1988], HNF-3, a protein whose DNA-binding domain is not like any other described so far [Costa et al. 1989; Lai et al. 1990], and HNF-4, a zinc-finger protein [Costa et al. 1989; Sladek et al. 1990]. To determine the extent to which the binding sites for these proteins were occupied in hepatocytes, nuclei were isolated from mouse liver and spleen, treated with DNase I, and analyzed by the new genomic sequencing method.

Genomic DNA was purified from normal or DNase I-treated nuclei from liver or spleen and digested further with restriction endonucleases BgIII and HaeIII, cutting, respectively, at positions -330 and -88 relative to the cap site. For each sample, ~200 μg of DNA was hybridized to a biotinylated SP6 RNA spanning the entire BgIII–HaeIII fragment, and the purified DNA segments from each strand were analyzed for protein footprints by the procedure outlined in Figure 2A. As shown in Figure 3, six regions of the TTR promoter were clearly protected from DNase I digestion in liver nuclei when compared to spleen nuclei or naked DNA. The same regions appeared protected from cleavage by copper–phenanthroline [Fig. 4A], although the protected regions were narrower in the latter case (data are summarized in Fig. 5).
Footprinting of TTR transcription control elements

-330 BgMI

-123 456 789 10

W&9 -ZA

a complex pattern of protection and hypersensitive sites. A strong protection from -91 to -107 flanked by a strong hypersensitive site at -108 corresponds almost exactly to the proximal HNF-3 site defined by the footprint of partially purified HNF-3 (Costa et al. 1989). A second site at -128 to -140 was also occupied in liver nuclei, and again this site has been shown to bind HNF-3 in vitro. These results argue strongly for the stable binding of the protein HNF-3 in hepatocytes. On the other hand, the “core” sequence for HNF-1 at -115 to -130, which binds HNF-1 in vitro (Courtois et al. 1988), was not protected in vivo. In this region there is also a sequence [-112 to -135] that is protected from DNase I cleavage by C/EBP in vitro (K.G. Xanthopoulos, pers. comm.). Finally, deletion analysis from upstream sequences showed that the region -140 to -151 was necessary for efficient transcription in transient assays, and the sequence could bind the factor HNF-4 (Costa et al. 1989). However, there was no sign of protection in the -140 to -150 region in vivo. Thus, although analysis of the -90 to -150 regions by in vitro binding assays suggested the possible involvement of HNF-1, C/EBP, HNF-3, and HNF-4, the footprint obtained from in vivo data suggests that only HNF-3 sites are occupied in this region in adult liver nuclei. Similarly, the region -177 to -195 that enhanced expression in transfected hepatoma cells by a factor of 2 (Costa et al. 1986) and that can interact with C/EBP in vitro (K.G. Xanthopoulos, pers. comm.) did not show any clear protection in liver nuclei. However, this is not due simply to artifactual loss of C/EBP from putative nuclear sites because other C/EBP-binding sites appear protected in the TTR enhancer (see below).

For spleen nuclei, the concentration of DNase I had to be raised -10 times to digest the TTR promoter to an extent similar to liver nuclei. This shows that the TTR promoter is organized in a much less accessible chromatin conformation in spleen nuclei. No particular pattern of protection or hypersensitivity was seen for the TTR promoter in spleen nuclei even at the higher levels of DNase I. This could result from the presence of different cell types in spleen, whereas ~90% of the liver DNA is derived from hepatocytes. Alternatively, this region of the genome could be organized into a less accessible state with nonspecific DNA-binding proteins (histones or other proteins) not specifically positioned.

Figure 3. Sensitivity to DNase I of the TTR promoter in liver nuclei, spleen nuclei, and naked DNA. [A] DNase I cleavage pattern on the upper strand (mRNA-like) in cloned plasmid DNA [lanes 1 and 10], liver genomic DNA [lanes 2 and 3], liver nuclei [lanes 4–6], and spleen nuclei [lanes 7–9] using primer 1 that corresponds to positions +88 to +64. All samples were digested in the same buffer for 1 min at 25°C with no DNase I [lanes 4 and 7], 1 µg/ml [lane 2], 2 µg/ml [lanes 5 and 10], 5 µg/ml [lane 3], 10 µg/ml [lane 6], 40 µg/ml [lane 8], and 80 µg/ml [lane 9]. Protections in liver nuclei are shown by solid bars; enhancements are shown by stars. Numbers are relative to the start site. [B] The DNase I cleavage pattern of the lower strand of the same samples as in A using primer 2 corresponding to positions -318 to -294.
The TTR enhancer is bound by two C/EBP-like proteins in liver nuclei

A liver-specific enhancer found almost 2 kb upstream of the TTR cap site has been characterized in detail by transient transfections, gel shift analysis, and in vitro footprinting. A 110-bp Accl–MboII fragment (at −1.96 to −1.86) has full enhancer capacity and contains a binding site for API, two binding sites for C/EBP, and a weak HNF-4 site (Costa et al. 1988a; 1989). Particularly, the integrity of both C/EBP-binding sites is necessary for maximal enhancer effect in hepatoma cells. We analyzed this region for DNase I protection in liver and spleen nuclei. Genomic DNA samples were digested with SstI and RsaI, and the lower strand (noncoding) was purified with the complementary SP6 biotinylated RNA. The purified products were hybridized to a labeled 25-nucleotide-long primer and extended from a position 100 bp upstream of the Accl site. Only two regions were clearly protected in liver nuclei, at positions 30–40 and 75–89 downstream of the Accl site. The same sequences are precisely protected in vitro using liver nuclear extracts or recombinant C/EBP (Costa et al. 1988b). These sites are necessary for full enhancer function and in vitro bind specifically to the liver-enriched transcription factor C/EBP (Costa et al. 1988b). The first 20 nucleotides downstream of the Accl site can bind API activity but did not show a protected region in liver nuclei and had only a modest effect in transfected deletion constructs, despite the presence of API in liver extracts (Costa et al. 1988a). A low-affinity HNF-4-binding site known to exist at positions 53–64 did not show any protection against DNase I in nuclei and also had only a modest effect in transfection analysis (Costa et al. 1989).

To obtain similar band intensities for liver and spleen nuclei adjacent to the C/EBP sites, a 10× lower concentration of DNase I was used for liver nuclei (Fig. 4, cf. lanes 3 and 6 or 4 and 7). This region actually corresponds to the DNase I hypersensitive site 3 observed in liver nuclei (Fig. 1). However, the region downstream of the C/EBP-binding sites was underdigested in liver nuclei at the lower DNase concentrations that digested extensively within the enhancer. The region starting from around position 120 downstream of the Accl site gave no evidence of special accessibility or of specific protection in liver nuclei.

Figure 4. (A) Sensitivity to copper-phenanthroline of the TTR promoter in liver nuclei and naked liver genomic DNA. Genomic samples were processed to reveal the cleavage pattern on the upper strand of the promoter as in Fig. 3A. (B) Sensitivity to DNase I of the TTR enhancer in purified liver genomic DNA (lane 1), liver nuclei (lanes 2–4), and spleen nuclei (lanes 5–7). The lower strand was purified and extended with a primer-anchored 100 nucleotides upstream of the Accl site. All samples were digested at 25°C for 1 min with no DNase I [lanes 2 and 5], 1 μg/ml [lane 1], 10 μg/ml [lane 3], 20 μg/ml [lane 4], 80 μg/ml [lane 6], and 160 μg/ml [lane 7]. Protections in liver nuclei are shown by solid bars. Numbering is relative to the Accl site. The Accl–MboII fragment is a liver-specific enhancer (Costa et al. 1988a).

Discussion

One of the cornerstone techniques used in modern molecular genetics is footprinting, the detection of protein binding at specific DNA sites by the protection of DNA
Footprinting of TTR transcription control elements

**PROMOTER:**

-300 → -250 → -200 → -150 → -100 → -50 → +1 → +50

BgIII Haelll

HNF3 HNF3 INKl

TCCCA TATA

**ENHANCER:**

-2.15 Kb → -1.96 Kb → -1.85 Kb

NcoI Prmcr

Accl MboII

API C/EBP HNF4 C/EBP

**Figure 5.** Summary of DNase I sensitivity of the TTR promoter and enhancer in mouse liver nuclei. The BgIII–HaeIII promoter fragment from -330 to +88 nucleotides relatively to the start site (+1) and the Accl–MboII enhancer fragment were analyzed. Solid bars represent regions protected in liver nuclei. The sensitivities for the upper [mRNA-like] and lower strands are drawn above and beneath the line, respectively. The in vitro-binding sites of various proteins and the positions of the TGGCA and TATA sequences are shown by open bars (Costa et al. 1988a,b, 1989; K. Xanthopoulos, pers. comm.). Also shown are the positions of primers used to detect the purified genomic sequences.

from chemical or enzymatic attack (Schmitz and Galas 1979). In this paper we present a new simple and reliable method for obtaining information precise to the nucleotide about proteins that protect DNA sequences in the nuclei of cells. We applied this technique to the analysis of a mammalian gene whose regulatory regions have been previously mapped by the now traditional methods of transient transfection of normal and mutant recombinant DNA constructs. Regulatory regions are then analyzed for specific binding proteins.

The method described here overcomes the major obstacle to in vivo footprinting, that is, the huge excess of irrelevant DNA present in the genome of higher organisms that makes the detection of single-copy sequences by hybridization-based probes very difficult. The use of a biotinylated riboprobe complementary to sequences of interest both provides a simple selection technique and a ready means (alkali treatment) of removing the selective agent. By reduction of the diluting DNA to approximately the complexity of a bacterial genome, primer extension of the selected fragments from a radiolabeled probe then allows detection of all of a collection of DNA fragments that are complementary to the probe.

The gene that we then analyzed in detail, TTR, encodes a serum protein that binds thyroxine. In adults it is transcribed mainly in the liver and in the choroid plexus of the brain (Dickson et al. 1985; Soprano et al. 1985). The gene is transcribed early in embryogenesis, the mRNA being detectable by 13 days and the transcription rate in fetal liver nuclei being at least 50% of the adult level by 17–19 days of fetal life (Powell et al. 1984). The sequences defined by transfection into hepatoma cells [-1.96 to -1.86 and -202 to +50], as enhancer and promoter sequences have been shown to suffice in transgenic animals for extensive and mainly liver-specific transcription of transgenes (Yan et al. 1990). The transgenic experiments do not allow the conclusion that quantitatively equal transcription occurs in transgenes containing only the described enhancer and promoter but highlight the importance of these sequences in hepatocytes. These regions also happened to be the strongest DNase I hypersensitive sites in liver nuclei, strongly suggesting their involvement in hepatocyte expression of TTR. Other hypersensitive sites may be involved in the precise tissue distribution of TTR expression.

Much effort in the last several years has been devoted to enumerating DNA-binding proteins that activate transcription from binding sites that lie several to many thousands of nucleotides away from the RNA start site. Some of these proteins are tissue-specific in the sense that they are present in liver but not in most other tissues. Analysis of the TTR promoter and enhancer identified four such proteins: HNF-3, which is present mainly in liver and lung (Lai et al. 1990); HNF-4, which is present in liver, kidney, and intestine (Sladek et al. 1990); C/EBP, which is found in liver, lung, intestine, fat cells, and some brain cells (Birkenmeier et al. 1989; Xanthopoulos et al. 1989; Kuo et al. 1990); and HNF-1, which is present in liver, kidney, and intestine (Frain et al. 1989; Baumhueter et al. 1990). The results presented here show that some, but not all, of the expected DNA-binding sites [based on experiments on nuclear extracts] are occupied in chromatin. Two HNF-3 sites in the promoter and two C/EBP sites of the enhancer are bound by...
proteins and are presumably part of a stable chromatin conformation associated with liver-specific TTR transcription. In addition to these tissue-limited factors, the sites binding NF-1-like proteins and TFIID are also part of this stable complex. Interestingly, the promoter region around −130 that can bind both HNF-1 and C/EBP in vitro is clearly accessible to DNase I. This may be due to the presence of the two occupied flanking HNF-3 sites that could prevent proteins from binding in between, although DNase I can reach the template to cut it. In vitro binding experiments could presumably test if the presence of one factor would prevent access for other proteins in the −90 to −150 region. Although the C/EBP sites in the enhancer appeared to be occupied, the two binding sites for this factor in the promoter were not. However, a large region around the C/EBP site at around −185 is clearly accessible to DNase I and shows no specific protection. In transient transfections, the deletion of this site lowers TTR expression by a factor of two but has no effect if the enhancer is present. Therefore, among all of these redundant C/EBP sites, perhaps the only ones involved in ongoing transcription in vivo are in the enhancer.

Somewhat surprisingly, the HNF-1 and HNF-4 sites in the promoter were not occupied in liver chromatin. Either of these promoter sites might be required transiently during development or cell division for the assembly of the active transcription template. The HNF-4 site was shown to be essential for expression in transient transfections in hepatoma cells but may not be necessary for ongoing transcription in the liver. We must now examine the footprints of this gene in fetal cells as early as possible to see which protein-binding sites are occupied earlier in development. How practical is the described technique for examining fetal tissue? There are two problems in this connection: sufficient cell numbers and sufficient cell purity. From the point of view of cell number, the technique is very promising. In the present experiments, samples of 3 × 10⁷ cells were used, but the exposure times to observe the footprint were only 24–36 hr. It is therefore possible to obtain signals from less cells simply by using higher specific activity primers and longer autoradiographic exposure times. The detection of protected regions presumably necessitates that most templates be occupied by proteins. Adult liver has ~60% hepatocyte nuclei, most of which are tetraploid in adults, so as much as 80–90% of the liver DNA is from hepatocyte origin. This number is perhaps only 50% for fetal liver because of the presence of blood cell precursors. However, the fetal liver cells can be at least partially purified from these contaminants and should therefore be accessible for study back to at least 14–15 days of fetal life.

In vivo footprinting can distinguish which sites are stably involved in ongoing transcription. However, the precise nature of the bound proteins cannot be assessed. Many DNA regulatory sites can bind specifically to a number of different proteins, and in particular, we recently showed that the HNF-3-binding site defines a family of transcription factors (Lai et al. 1990; V.R. Prezioso, pers. comm.). The precise identification of proteins occupying genomic template awaits the development of sensitive cross-linking methods. Another limitation of our results is that we are describing the occupancy of regulatory elements in isolated nuclei, not in whole cells. The reagent DMS can penetrate whole cells and methylates purines (Ephrussi et al. 1985). Protection from DMS methylation is therefore the method of choice to detect in vivo footprints but is difficult to carry out on whole organs such as liver and spleen. However, we consider that a major redistribution of factors during nuclei isolation is unlikely. The nuclear proteins binding to the regulatory sequences can be extracted from nuclei only with high salt concentrations, suggesting that transcription factors are stably bound during isolation. This is further confirmed by the observation that when multiple binding sites are present for a specific protein (C/EBP), only a subset of these sites are actually occupied.

Although this method will be useful in studying proteins bound in vivo to regulate RNA synthesis, it obviously can be extended to other very important processes: Do putative origins of DNA replication give distinct evidence of protein interactions before and/or during DNA replication? Are sites of frequent recombination occupied by proteins in meiotic cells? Are proteins thought to be important in folding chromosomes in place at particular sites? Many of these problems should benefit from a simple genomic sequencing technique.

Materials and methods

Isolation and digestion of nuclei

Nuclei were isolated at 4°C from 10-week-old C57Bl female mice as described (Gorski et al. 1986), with the following modifications. Tissues were homogenized in 9 volumes of buffer A (15 mM Tris-HCl at pH 7.4, 0.2 mM spermine, 0.5 mM spermidine, 80 mM KCl, 2 mM K-EDTA, 1 mM DTT, 0.5 mM PMSF, 3 mg/ml aprotinin, 0.5 mg/ml leupeptin, 1 mg/ml pepstatin) containing 2.2 M sucrose. Lysates were passed through cheesecloth, homogenized with four strokes of an A-type pestle in a Dounce homogenizer, and layered over a cushion of buffer A containing 2 M sucrose and 10% glycerol. Gradients were spun at 24,000 rpm in a SW27 rotor for 45 min. Nuclei were resuspended in buffer A with 5% glycerol (3 ml per gradient), spun at 2000 g for 5 min, and resuspended in buffer A at ~10⁶ nuclei per ml. For the mapping of DNase I hypersensitive sites, nuclei were digested for 10 min on ice with DNase I in buffer A (without EDTA) with 5% glycerol and 5 mM MgCl₂. For genomic sequencing, nuclei were prewarmed for 2 min at 25°C in buffer A with 5% glycerol and an equal volume of buffer A with 20 mM MgCl₂ and DNase I was added for 1 min. Reactions were quenched with EDTA (to 25 mM), followed immediately by the addition of 250 mM NaCl, 0.2% SDS, 25 mM Tris-HCl (pH 8.0), and 0.2 mg/ml proteinase K. For copper–phenanthroline digestion, nuclei at 10⁸/ml in buffer A with 0.2 M sucrose were preincubated at 25°C for 2 min, a half-volume of H₂O₂ at 4 mM in buffer A was added, and another half-volume of copper–phenanthroline freshly prepared in buffer A was added immediately (final concentration of 0.18 mM CuSO₄ with 0.80 mM H₂O₂).
Footprinting of TTR transcription control elements

1,10-phenanthroline; Kuwabara and Sigmen 1987). After 30 sec, the reaction was quenched with 100 mM thiourea and 12 mM 2,9-dimethylphenanthroline, followed by the addition of proteinase K–SDS mix as for DNase I samples. DNA was purified by two phenol/chloroform/isoamyl alcohol (25 : 24 : 1) extractions and ethanol precipitation.

**DNase I hypersensitive site mapping**

About 20 µg of genomic DNA was cleaved with BamHI, electrophoresed in a 0.8% agarose gel, and transferred to Zetaprobe (Bio-Rad) in alkali according to the manufacturer’s recommendations. The membranes were further hybridized to random primer-labeled fragment in 0.5 M Na2HPO4 (pH 7.2), 7% SDS, and 1 mM NaEDTA overnight at 65°C. Filters were washed to a final stringency of 0.1 x SSC and 0.1% SDS at 65°C for 30 min.

**Isolation of specific genomic sequences**

Purified genomic DNA was digested with restriction enzymes BglII and HaelII for the TTR promoter or with SsrI and Rsdl for the enhancer. Digestion with at least one 4-bp cutter seems to be necessary to avoid the formation of gelatinous mass during hybridization, due to the formation of long fibers annealed by repetitive sequences. For a final gel lane, 200 µg of each sample was used. After further purification by phenol extraction, DNA was denatured in 100 µl of 0.2 M NaOH at 65°C for 5 min. The solution was neutralized with 100 µl of 0.2 M HCl and 15 µl of 1 M HEPES/NaOH (pH 7.4) added simultaneously. Then 35 µl of 100 mM HEPES/NaOH (pH 7.4), 3 mM NaCl, 5 mM EDTA, 0.25% Tween-20, and 0.5 mg/ml yeast RNA were added, with SP6 biotinylated RNA to a final concentration of 0.3 µg/ml.

SP6 RNAs were synthesized [Krieg and Melton 1987] using a 3 : 1 ratio of UTP to biotin-11-UTP [BRL and Enzo Biochemicals] and represented the entire restriction fragment to be analyzed. In a final volume of 50 µl, 10 µg of linearized plasmid DNA was mixed in 20 mM Tris-HCl (pH 7.4), 10 mM NaCl, 6 mM MgCl₂, 1 mM spermidine, 1 mM each ATP, GTP, and CTP, 0.75 mM UTP, 0.25 mM biotin–UTP, 200 units RNasin, and 75 units SP6 polymerase. The reaction was left for 1 hr at 40°C, after which fresh SP6 polymerase and RNasin were added at the same concentrations as before and the reaction proceeded for another hour. Two hundred units of DNase I [labeled RNase free] was added with 200 units of RNasin, and incubated at 37°C for 1 hr. The mixture was extracted with phenol, recovering the interphase, and precipitated with ethanol. The RNA was resuspended in 50 µl of 10 mM Tris-HCl, 50 mM NaCl, 10 mM MgCl₂, 1 mM DTT, 200 units RNasin, and 200 units of DNase I and incubated at 37°C for 1 hr. The SP6 RNA preparation had to be digested extensively with DNase I to eliminate all of the DNA template that would otherwise create a background sequence in the primer extension. After phenol extraction, the SP6 RNA was precipitated twice at room temperature with 0.1 volume of 7.5 M NH₄ acetate and 2.5 volumes of ethanol to eliminate free nucleotides. Twenty micrograms or more of full-length RNA was routinely obtained.

Hybridizations were performed at 65°C for 12–16 hr, after which 5 µl of packed streptavidin–agarose gel (Sigma) pre-washed in hybridization buffer was added for 45 min at 37°C. Agarose beads were recovered by spinning in a microfuge for 5 sec and washed once with 10 mM HEPES/NaOH (pH 7.4) and 1 mM EDTA. Purified genomic DNA was eluted in 0.2 M NaOH, agarose beads were pelleted, and the supernatant was incubated at 70°C for 20 min to hydrolyze the RNA. After addition of 100 mm Tris-HCl (pH 7.4), 0.2 M HCl, and 0.5 µg of yeast RNA, the samples were phenol-extracted and precipitated in siliconized tubes with three volumes of ethanol.

**Primer extension of purified genomic products**

The highly labeled primer was synthesized essentially as described (Saluz and Jost 1989), using a 29-mer template and a 9-mer annealed primer. The 9-mer anneals 4 residues from 3’ of the 29-mer, and the 25-mer extended product is synthesized with Klenow in the presence of a mix of labeled and unlabeled nucleotides to incorporate ~10 labeled nucleotides in the 25-mer, but it can be synthesized with only labeled nucleotides. Five picomoles of 29-mer annealed to 25 pmoles of 9-mer were extended in a final volume of 4 µl, for 30 min at 23°C containing 2 units of Klenow in 10 mM Tris-HCl (pH 7.4), 80 mM NaCl, 10 mM MgCl₂, 1 mM DTT, 200 µM of cold nucleotides of one kind, and 8 µM of labeled nucleotides of the other kind. This highly labeled primer was purified from an acrylamide–urea gel and used in the next 5 days.

The purified cleaved DNA obtained from 200 µg of genomic DNA was resuspended in 5 µl of 20 mM NaOH in siliconized microtubes in the presence of 2 x 10⁵ cpm of primer at ~30,000 Ci/m mole, incubated at 60°C for 3 min, and neutralized on ice with 1 µl of HCl (0.1 M) and 1.5 µl 10 x extension buffer [1 x extension buffer is 10 mM Tris-HCl (pH 8.5), 25 mM KCl, and 3.5 mM MgCl₂]. Annealing was done at 60°C for 10 min, after which the samples were put on ice. Deoxynucleotides (final 0.1 mM each), extension buffer [1 x final], and Taq DNA polymerase [AmpliTag from Cetus; 10 U/ml final] were added to a final volume of 20 µl. Extension was done at 70°C for 10 min (60°C for the enhancer primer that has a higher content of AT residues). Then 170 µl of 25 mM Tris-HCl (pH 7.4), 1 mM EDTA, 0.2 % SDS, and 2.5 µg/ml yeast RNA were added, and samples were heated at 80°C for 30 min. Ten microliters of 5 M NaCl was added to the samples, which were then extracted with phenol/CHCl₃/isoamyl alcohol (25 : 24 : 1), discarding any interphase, precipitated with 3 volumes of ethanol on dry ice, and run overnight at 1000 V on 80-cm-long 6% urea sequencing gels. Electrophoresis was stopped after the free primer came out of the gel. After fixation in 10% methanol and 10% acetic acid, the gels were dried on cellulose sheets and exposed to Kodak XAR-5 films at ~80°C between intensifying screens for 12 hr to 3 days.

**Acknowledgments**

We thank K. G. Xanthopoulos, T. Deckter, and D. J. Lew for critical reading of the manuscript, and A. English and L. Cousseau for preparation of the manuscript. J. M. was supported by a postdoctoral fellowship of the American Cancer Society and by the Swiss National Science Foundation. This work was supported by grants from the National Institutes of Health [CA16006] and the American Cancer Society [MV424R] to J.E.D.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

**References**

Axelrod, J.D. and J. Majors. 1989. An improved method for phototoolfootprinting yeast genes in vivo using Taq polymerase. *Nucleic Acids Res.* 17: 171–183.
Baumhueter, S., D.B. Mendel, P.B. Conley, C.J. Kuo, C. Turk, M.K. Graves, C.A. Edwards, G. Courtois, and G.R. Crabtree. 1990. HNF-1 shares three sequence motifs with the POU domain proteins and is identical to LF-B1 and APF. *Genes Dev.* 4: 372–379.

Becker, P.B., S. Ruppert, and G. Schutz. 1987. Genomic footprinting reveals cell type-specific DNA binding of ubiquitous factors. *Cell* 54: 435–443.

Becker, M.M., Z. Wang, G. Grossman, and K.A. Becherer. 1989. Genomic footprinting in mammalian cells with ultraviolet light. *Proc. Natl. Acad. Sci.* 86: 5315–5319.

Birkenmeier, E.H., B. Gwynn, S. Howard, J. Jerry, J.I. Gordon, M.M. Becker, M. Wang, G. Grossman, and K.A. Becherer. 1989. Genomic footprinting in mammalian cells with ultraviolet light. *Proc. Natl. Acad. Sci.* 86: 5315–5319.

Horikoshi, M., T. Hai, Y.-S. Lin, M.R. Green, and R.G. Roeder. 1989. Transcriptional regulation of the gene encoding C/EBP by sequence-specific DNA binding proteins. *Genes Dev.* 3: 1146–1156.

Clayton, D.F., M. Weiss, and J.E. Darnell Jr. 1985. Liver-specific RNA metabolism in hepatoma cells: Variations in transcription rates and mRNA levels. *Mol. Cell. Biol.* 5: 2633–2641.

Costa, R.H., E. Lai, and J.E. Darnell Jr. 1986. Transcriptional control of the mouse prealbumin (transthyretin) gene: Both promoter sequences and a distinct enhancer are cell specific. *Mol. Cell. Biol.* 6: 4697–4708.

Costa, R.H., E. Lai, D.R. Grayson, and J.E. Darnell Jr. 1988a. The cell-specific enhancer of the mouse transthyretin (prealbumin) gene binds a common factor at one site and a liver-specific factor(s) at two other sites. *Mol. Cell. Biol.* 8: 81–90.

Costa, R.H., D.R. Grayson, K.G. Xanthopoulos, and J.E. Darnell Jr. 1988b. A liver-specific DNA-binding protein recognizes multiple nucleotide sites in regulatory regions of transthyretin, a hepatic protein, albumin, and simian virus 40 genes. *Proc. Natl. Acad. Sci.* 85: 3840–3844.

Costa, R.H., D.R. Grayson, and J.E. Darnell Jr. 1989. Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and α1-antitrypsin genes. *Mol. Cell. Biol.* 9: 1415–1423.

Courtois, G., J.G. Morgan, L.A. Campbell, G. Fournel, and G.R. Crabtree. 1987. Interaction of a liver-specific nuclear factor with the fibrinogen and α1-antitrypsin promoters. *Science* 238: 688–692.

Courtois, G., S. Baumhueter, and G.R. Crabtree. 1988. Purified hepatocyte nuclear factor I interacts with a family of hepatocyte-specific promoters. *Proc. Natl. Acad. Sci.* 85: 7937–7941.

Dickson, P.W., G.J. Howlett, and G. Schreiber. 1985. Rat transthyretin (prealbumin): Molecular cloning, nucleotide sequence, and gene expression in liver and brain. *J. Biol. Chem.* 260: 8214–8219.

Ephrussi, A., G.M. Church, S. Tonegawa, and W. Gilbert. 1985. B lineage-specific interactions of an immunoglobulin enhancer with cellular factors in vivo. *Science* 227: 134–140.

Frain, M., G. Swart, P. Monaci, A. Nicosia, S. Stampfl, R. Frank, and R. Cortese. 1989. The liver-specific transcription factor LF-B1 contains a highly diverged homeobox DNA binding domain. *Cell* 59: 145–157.

Gorski, K., M. Carneiro, and U. Schibler. 1986. Tissue-specific in vitro transcription from the mouse albumin promoter. *Cell* 47: 767–776.

Gross, D.S. and W.T. Garrard. 1988. Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* 57: 159–197.

Horiokoshi, M., T. Hai, Y.-S. Lin, M.R. Green, and R.G. Roeder. 1988. Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex. *Cell* 54: 1033–1042.

Huibregtse, J.M. and D.R. Engelke. 1989. Genomic footprinting of a yeast tRNA gene reveals stable complexes over the 5′-flanking region. *Mol. Cell. Biol.* 9: 3244–3252.

Jackson, P.D. and G. Felsenfeld. 1985. A method for mapping intranuclear protein-DNA interactions and its application to a nucleosome hypersensitive site. *Proc. Natl. Acad. Sci.* 82: 2296–2300.

Johnson, P.F. and S.L. McKnight. 1989. Eukaryotic transcriptional regulatory proteins. *Annu. Rev. Biochem.* 58: 799–839.

Krieg, P.A. and D.A. Melton. 1987. In vitro RNA synthesis with SP6 RNA polymerase. *Methods Enzymol.* 115: 397–415.

Kuo, C.F., K.G. Xanthopoulos, and J.E. Darnell Jr. 1990. Fetal and adult localization of C/EBP: Evidence for combinatorial action of transcription factors in cell-specific gene expression. *Development* 109: 473–481.

Kuwabara, M.D. and D.S. Sigman. 1987. Footprinting DNA-protein complexes in situ following gel retardation assays using 1,10-phenanthroline-copper ion: *Escherichia coli* RNA polymerase-lac promoter complexes. *Biochemistry* 26: 7234–7238.

Lai, E., V.R. Prezioso, E. Smith, O. Litvin, R.H. Costa, and J.E. Darnell Jr. 1990. HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally. *Genes Dev.* 4: 1427–1436.

Landschulz, W.H., P.F. Johnson, E.Y. Adashi, B.J. Graves, and S.L. McKnight. 1989. Isolation of a recombinant copy of the gene encoding C/EBP. *Genes Dev.* 2: 786–800.

Mitchell, P.J. and R. Tian. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* 245: 371–378.

Mueller, P.R. and B. Wold. 1989. In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* 246: 780–786.

Nakajima, N., M. Horikoshi, and R.G. Roeder. 1988. Factors involved in specific transcription by mammalian RNA polymerase II: Purification, genetic specificity, and TATA box-promoter interactions of TFIID. *Mol. Cell. Biol.* 8: 4028–4040.

Powell, D.J., J.M. Friedman, A.J. Oulette, K.S. Krauter, and J.E. Darnell Jr. 1984. Transcriptional and post-transcriptional control of specific messenger RNAs in adult and embryonic liver. *J. Mol. Biol.* 179: 21–35.

Rupp, R.A.W., U. Kruse, G. Muthaup, U. Grobel, K. Beyer, and J.E. Darnell Jr. 1984. Transcriptional and post-transcriptional control of specific messenger RNAs in adult and embryonic liver. *J. Mol. Biol.* 179: 21–35.

Salt, J.-P. Jost. 1981. A simple high-resolution procedure to study DNA methylation and in vivo DNA-protein interactions on a single-copy gene level in higher eukaryotes. *Proc. Natl. Acad. Sci.* 86: 2602–2606.

Sasse-Dwight, S. and D.J. Gralla. 1990. Role of eukaryotic-type functional domains found in the prokaryotic enhancer repressor factor Crp. *Cell* 62: 945–954.

Schmitz, A. and D.J. Galas. 1979. The interaction of RNA polymerase and lac repressor with the lac control region. *Nucleic Acids Res.* 6: 2388–2397.

Selsle, S.B. and J.E. Majors. 1987. In vivo DNA-binding properties of a yeast transcription activating protein. *Mol. Cell. Biol.* 7: 3260–3267.

Sladek, F.M., W. Zhong, E. Lai, and J.E. Darnell Jr. 1990. Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes Dev.* 2353–2365.

Soprano, D.R., J. Herbert, K.J. Soprano, E.A. Schon, and D.S.
Goodman. 1985. Demonstration of transthyretin mRNA in the brain and other extra hepatic tissues in the rat. J. Biol. Chem. 260: 11793–11798.

Xanthopoulos, K.G., J. Mirkovitch, T. Decker, C.F. Kuo, and J.E. Darnell Jr. 1989. Cell-specific transcriptional control of the mouse DNA-binding protein mC/EBP. Proc. Natl. Acad. Sci. 86: 4117–4121.

Yan, C., R.H. Costa, J.E. Darnell Jr., J. Chen, and T.A. Van Dyke. 1990. Distinct positive and negative elements control the limited hepatocyte and choroid plexus expression of transthyretin in transgenic mice. EMBO J. 9: 869–878.

Zhang, L. and J.D. Gralla. 1989. In situ nucleoprotein structure at the SV40 major late promoter: Melted and wrapped DNA flank the start site. Genes Dev. 3: 1814–1822.
Rapid in vivo footprinting technique identifies proteins bound to the TTR gene in the mouse liver.

J Mirkovitch and J E Darnell

*Genes Dev.* 1991, 5:
Access the most recent version at doi:10.1101/gad.5.1.83

References

This article cites 39 articles, 25 of which can be accessed free at:
http://genesdev.cshlp.org/content/5/1/83.full.html#ref-list-1

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.