Expression of both the apolipoprotein (apo)E and apoC-I genes in the liver is specified by a 319-nucleotide hepatic control region (HCR-1) that is located 15 kilobase pairs downstream of the apoE gene and 5 kilobase pairs downstream of the apoC-I gene. In vivo footprint analysis of HCR-1 in intact nuclei revealed several liver-specific protein-binding sites that were not detectable by in vitro methods. In addition to three previously identified in vitro footprints, four in vivo footprints were identified in a region of HCR-1 that is required for directing gene expression to hepatocytes. Prominent liver-specific nuclear protein binding to these sites was associated with these footprints. Liver-specific nuclear protein binding to these sites was confirmed by oligonucleotide gel-retention assays. The in vivo analysis also identified a cluster of nuclear protein-binding sites in the Alu family repeat segment adjacent to the domain required for liver expression. Micrococcal nuclease digestion indicated the presence of a nucleosome in the central domain of HCR-1 in liver chromatin that was in phase with the nucleosome location in tissues that did not express the transgene. These results suggest that HCR-1 functions in a highly structured chromatin environment requiring a complex interaction of liver-enriched transcription factors.

The human apolipoprotein (apo)E gene spans 3.6 kb (1–3) and is located at the 5’ end of a 45-kb cluster of apolipoprotein genes on chromosome 19, all of which have the same transcriptional orientation (4). The apoC-I gene (4.7 kb) is located 5.3 kb downstream from the apoE gene, and a 4.4-kb apoC-I’ pseudogene is located 7.5 kb further downstream. The apoC-II gene (3.3 kb) is ~16 kb downstream from the apoC-I’ pseudogene. Recently, the apoC-IV gene (3.3 kb), located 555 bp upstream from the apoC-II gene, has been identified in this locus (5). Each of these genes contains four exons (except apoC-IV, which lacks the nontranslated first exon found in the other genes) with the introns located in similar intragenic positions, suggesting that this gene family evolved from a single ancestral gene (1, 3).

Human apoE, a 299-amino acid glycoprotein of $M_r = 35,000$ (3), is a major component of various plasma lipoprotein classes, including chylomicron remnants, very low density lipoproteins, and high density lipoproteins (6, 7). It is required for the receptor-mediated uptake of chylomicron remnants and facilitates the redistribution of cholesterol from peripheral tissues to the liver (6, 8). Although apoE is produced by specific cell types in many different tissues, more than 90% of the circulating apoE in human plasma comes from the liver (6, 9). Receptor binding-defective variants of apoE having the E2 phenotype are associated with type III hyperlipoproteinemia and premature atherosclerosis (6). A commonly occurring apoE variant, the E4 allele, has been linked to the development of Alzheimer’s disease (10–12). Apolipoprotein C-II is an essential cofactor for lipoprotein lipase, giving it an important role in the hydrolysis of lipoprotein triglycerides (13). The function of apoC-I may be to modulate or to inhibit the apoE-mediated cellular uptake of remnants (14, 15). The function of apoC-IV is unknown.

Simonet et al. (16, 17) demonstrated that expression of the apoE and apoC-I genes in the liver requires the presence of a distal downstream tissue-specific enhancer. Subsequent studies by Simonet et al. (18) and Shachter et al. (19) demonstrated that this hepatic control region (HCR) is located 19 kb downstream from the transcription start site of the apoE gene and 9 kb downstream from the transcription start site of the apoC-I gene. The HCR contains all sequences necessary to direct expression of both the apoE and apoC-I genes in hepatocytes (18): constructs that lacked the HCR were not expressed in the livers of transgenic mice, even at low levels. The presence of a previously characterized enhancer element, which lacks tissue specificity, in the promoter of the apoE gene (20, 21) was required for transcriptional activation. These results suggested that interaction of a unique hepatocyte-specific combination of distal elements in the HCR with a nonspecific activator sequence in the promoter directed the expression of the apoE/C-IV/C-II locus in the liver.

Recently, a second HCR sequence (denoted HCR-2), which shares 85% sequence identity with the initially identified HCR (henceforth referred to as HCR-1), was localized ~5.5 kb downstream of the apoC-I’ pseudogene and ~10 kb downstream of HCR-1 (22). A construct in which the HCR-2 was ligated to the human apoE gene directed high levels in the liver of the transgenic mouse; however, the function of HCR-2 in the apoE gene locus remains to be determined (22).

Further analysis of the regulatory sequences of the HCR-1 region demonstrated that full liver-specific activity is contained within a 319-nucleotide domain, as assayed in transgenic mice (23). In addition, HCR-1 has a nuclear scaffold attachment capability that may contribute to an apparent position inde-
dependence in directing liver-specific transgene expression. In vitro footprinting detected three liver-specific protein footprints in the purified HCR-1 DNA fragment (23). In this assay, isolated DNA fragments were mixed with mouse liver nuclear extracts to identify protein-binding sites. Unexpectedly, a critical 5′ region required for the activity of the HCR in transgenic mice showed no protein-binding sites, suggesting that this approach may not reveal all of the functional elements in HCR-1. This limitation may be due to the loss or inactivation of nuclear factors during nuclear extract preparation and to the absence of native chromatin structure.

In the current study, we employed in vivo footprinting in intact nuclei from human apoE transgenic mice to analyze nuclear protein binding to HCR-1. We used mice expressing the HEG.LE1 construct (18) in which the complete apoE gene, together with 5 kb of 5′-flanking sequence and 1.7 kb of 3′-flanking sequence, is ligated to a 3.8-kb downstream fragment containing the HCR-1 domain. The results indicate a complex linear-specific nuclear protein-binding pattern that clarifies the regulatory elements of HCR-1.

### MATERIALS AND METHODS

#### In Vivo Footprinting Analysis

Intact nuclei were isolated from 3–6-month-old, hemizygous transgenic ICR mice bearing 70 copies of the HEG.LE1 construct (18), illustrated in Fig. 1A. Nuclei were partially digested with DNase I or micrococcal nuclease (MNsase). Then, the DNA was extracted and analyzed by ligation-mediated polymerase chain reaction (LMPCR) as described below.

#### Digestion of Nuclei by DNase I—Nuclei were isolated, as described by Dang et al. (23), from either the livers or kidneys of transgenic mice bearing the HEG.LE1 transgene construct (18). The nuclear pellet was resuspended in Buffer A without EDTA (15 mM Tris-HCl (pH 7.4), 0.15 mM spermine, 0.5 mM spermidine, 0.5 mM phenylmethylsulfonyl fluoride) containing 5% glycerol and 5 mM MgCl₂. The nuclei were diluted to 10 A₂₆₀ₐₐₘₐₐₐₐ₁ with the same buffer, and 0.5-ml aliquots were incubated on ice for 30 min with or without different amounts of DNase I (2–40 units). The reactions were quenched by addition of 25 μl of 0.5× EDTA, 25 μl of 5× NaCl, 5 μl of 10% SDS, 12.5 μl of 1× Triton-X (pH 8), and 10 μl of 10 mg/ml proteinase K, and then incubated at 50°C for 1 h. Aliquots were extracted with phenol. DNA was precipitated with ethanol, then resuspended in 0.1× TE buffer. DNA sequencing analysis have been described (26–29). In the present study, illustrated in Fig. 1B and C, sequence-specific oligonucleotide primers were used first to generate double-stranded copies of the region of interest (using 1.5–5 μg of genomic DNA from DNase I-digested nuclei or Mnase-digested nuclei) in the HCR or to the sequences adjacent to the start site of transgene transcription. The primer locations on the HCR or near the first exon are shown in Fig. 1C. A common linker was ligated to the blunt end of the double-stranded fragments. To increase specificity, a second primer that is internal to the first primer and a primer hybridized to the linker primer were used for subsequent PCR amplification. For maximum resolution, DNA amplification was carried out by either of two reaction systems: (a) Sequenase (U.S. Biochemical Corp.) for first strand synthesis and Taq DNA polymerase (Perkin-Elmer) for PCR (26), or (b) Vent DNA polymerase (New England Biolabs) for both reactions (27). Amplified fragments were detected by a radioactively labeled third sequence-specific primer that is internal to the second primer. DNA was extracted with phenol/chloroform, and the footprint pattern was revealed by gel electrophoresis on a 6% denaturing sequencing gel (30) followed by autoradiography.

#### Oligonucleotides for in vivo footprinting analysis were synthesized via a nucleic acid synthesis system (Millipore, Bedford, MA), or ordered from Oligo Etc. (Wilsonville, OR). Primers for the common linker were as follows: 1, GGGGTACCCCGGGAGATTCGTAATC; and 2, GAATTCCAGATCC. Specific primers complementary to HCR-1 or the human apoE gene with their last nucleotide positions indicated (Fig. 1C) were those listed below.

**Set A**

- p229–1, TTCCACCCGAAAATTTCCAG
- p229–2, GGGTCGAGTGGATGGGTCAGGGTG
- p229–3, TGGATGTTGAGGTGCGATGCCATGGCCCG

**Set B**

- p348–1, GAAACAGGCTTCGGGTC
- p348–2, GGGGATACCCACTACTGACCAT
- p348–3, GGCACCATGACTGACTCCTCA

**Set C**

- p36–1, AGACACAGGAAGTTTACG
- p36–2, AGGGAGAGGGAGGATGGAGGACTC
- p36–3, GGGCGTACGGAGTCCCGAAGG

**Set D**

- p60–1, CTTCGAGGCTCAGAGGCCAC
- p60–2, ACAAGGATTTTCG GCC TACCTC
- p60–3, CACCCCTGCCCCCTTCACCCCCTC

**Set E**

- p268–1, AGGAGACCTTCCTGCGGCCC
- p268–2, ATGCCACCTCCACACTCCACTG
- p268–3, CTCCAACATCCACTGACCCCCC

**Set F**

- p372–1, GGTCGGCGGTTTCAAAACC
- p372–2, CT TT GCTGTTGAGGGTATCGAG
- p372–3, GG GGCGGAGTCGTCGA T

**Set P**

- p98–1, CGCAGTACCTCCTCTC
- p98–2, ACTACCTGGAGCCGAGGCGAAGTG
- p98–3, CAGGCGCAGGCTCGAAGGTC

### Sequencing 1–21

#### Gel Retardation Assays

The assays were carried out essentially as described (20) with minor modifications. In brief, 1.2 μg of mouse liver or kidney nuclear extract was incubated with 2 μg of poly(dI-dC) at 20°C for 5 min in 20 μl of nuclear dialysis buffer. Where appropriate, specific oligonucleotide competitors were added to each sample and incubated at 20°C for 10 min. Then, 2.5 ng of end-labeled oligonucleotide was added to the mixture and incubated at 20°C for 15 min. Samples were resolved by electrophoresis in 5% polyacrylamide gels in 0.1× TBE (45 mM Tris borate, 1 mM EDTA) buffer, and then the gel was dried and examined by autoradiography. For antibody supershift assays, 2 μl of an antibody to a specific transcription factor was added to the reaction mixture either before or after the addition of poly(dI-dC) and oligonucleotide probes. Then, the reaction was incubated for an additional 20 min.

#### Preparation of Nuclear Extracts—Nuclear extracts from mouse liver
RESULTS

In Vivo Footprinting—In vivo footprinting of HCR-1 reveals a dense concentration of protein-binding sites in the minimal domain required for full liver-specific expression, shown in Figs. 2 and 3. In this assay, DNase I penetrates the membranes of intact nuclei and cleaves accessible DNA mainly in those chromatin regions that are decondensed and transcriptionally active. Bound nuclear factors protect the underlying nucleotides from being digested; these sites are identified by subsequent LMPCR, and the exact location of protected sites is determined by analyzing the sequence of guanine nucleotides downstream of the apoE gene and generated by SpH1/BamHI digestion. It is ligated to the 5’ end of human apoE gene together with 5 kb of 5’-flanking sequence and 1.7 kb of the 3’-flanking sequence (18). B, the scheme for LMPCR. Gene-specific primer 1 is hybridized to the region of interest in the appropriately cleaved genomic DNA and extended with a DNA polymerase to generate a blunt end at a site randomly cleaved with DNase I. This blunt end is ligated to a unidirectional common linker. The product is subjected to a PCR reaction using gene-specific primer 2, which is internal of primer 1; and the linker primer, which hybridizes with the linker sequence. An end-labeled third primer, which is internal to primer 2, is used to label the PCR product. C, oligonucleotide primer sets used for LMPCR aided in vivo analysis of the HCR and the human apoE gene promoter region. Solid arrows show the position and orientation of the primers.

Similar to a cluster of five footprints was detected in the Alu family sequence adjacent to HCR-1 (Fig. 2, E and F), although these footprints tended to be less distinct than those located within the liver-specific HCR-1 domain. Since functional tests in transgenic mice demonstrated that the Alu sequence region was not required for full HCR-1 activity (23), the potential role (if any) of this repeated sequence in HCR-1 function remains unclear. Previously Detected in Vitro Footprints Are a Subset of in Vivo Footprints—Previous in vitro analysis had revealed a limited number of protein-binding sites in the HCR domain when nuclear extracts were incubated with isolated DNA (23). In this approach, some footprints could have been missed or artificial footprints might have been observed for various reasons. For example, the appropriate structure for specific protein binding may not be presented by purified, short, linear DNA fragments in vitro. In addition, some transcription factors may have been inactivated or lost during the preparation of nuclear extracts. Nevertheless, all six of the protein-binding sites detected previously by the in vitro method were identified in the in vivo assay in a 774-bp HCR-containing fragment were observed by in vivo footprinting, as summarized in Fig. 3. Three of these footprints (Footprints 4–6; Figs. 2 and 3) were in the minimum domain required for high level liver-expressing activity (23), two of which (Footprints 5 and 6) also had been found by an independent laboratory (19). In addition, three footprints (Footprints 8, 9, and 12; Figs. 2 and 3) downstream from the minimum HCR domain were detected by both in vitro and in vivo methods.

Most of the protein-binding sites identified in intact nuclei are slightly larger than when detected using purified DNA and nuclear extracts (23). Most notably, Footprint 4 is 20 bp longer in the in vivo experiments. This extended protein-binding se-
sequence includes two tandemly repeated TGTTTGC motifs in addition to the two tandem copies of this motif located within the in vitro footprint region. Only Footprint 5 had essentially the same length when determined by both in vitro and in vivo methods.

**New HCR Footprints in the Liver Expression Domain Detected by in Vivo Methods**—Previous studies showed that nucleotides 6–72 were involved in liver expression activity, with nucleotides 72–122 being absolutely required for any HCR activity (23). However, no footprints were detected in this region by in vitro analysis with liver nuclear extracts (19, 23). In contrast, the in vivo method revealed three protein-binding sites (Footprints 1b, 2, and 3; Fig. 2, panels A–D) in this essential regulatory sequence. Footprints 2 and 3, covering nucleotides 61–128 in the required liver expression domain, each contained a sequence that was closely related to the TGTTTGC motif of Footprint 4, differing by only 1 nucleotide at both sites (Fig. 3). The 5′ end of Footprint 1b was marked by a
These radioactive bands were substantially reduced by incubation with specific factors to the new footprint sequences observed in the present study. Therefore, these footprint sequences were not identified for Footprint 8 although protein binding had been confirmed. No transcription factor family was identified that bound to these specific footprint-binding sites. Only Footprint 10, which covered part of the A-rich 3' tail of the Alu family sequence, showed nearly equivalent levels of oligonucleotide retardation with both liver and kidney nuclear extracts.

Examination of the newly detected in vivo footprint sequences for motifs that correspond to high affinity binding sites for known transcription factors revealed potential homologies in the Footprint 1b segment of HCR-1. A sequence of 7 nucleotides that corresponds to the 3' portion of the TF-LF2 high affinity sequence (33) is located at nucleotide 31 in reverse orientation. The TF-LF2 high affinity oligonucleotide was a strong competitor for one of the retarded bands of the Footprint 1b oligonucleotide in the gel retardation assay (Fig. 5, lanes B–D). Footprint 1b also contains a CACCC erythroid Krüppel-like factor (EKLF) binding sequence at nucleotide 36 (Fig. 3). An oligonucleotide consisting of a high affinity EKLF binding sequence (32) did not compete with the labeled Footprint 1b probe for the binding of nuclear proteins, even when added to nuclear extracts at 80-fold excess (Fig. 5, lanes E–G). However, a combination of the TF-LF2 and EKLF oligonucleotides at lower concentrations (20-fold) competitively bound each major retardation band of the Footprint 1b probe (Fig. 5, lanes H–J). These preliminary studies suggest that the cooperative interaction between at least two different factor families may be required for Footprint 1b function.

The finding that an EKLF-binding motif is located in the HCR is surprising, since this factor is associated essentially with erythroid cell-specific expression (36). However, another CACCC-binding protein, basic Krüppel-like factor (BKLF) has been identified in the adult mouse liver (37). The possibility that either of these factors might interact with the CACCC motif in the HCR was examined by gel supershift assay using specific antibodies to EKLF and BKLF. However, neither antibody caused a supershift of the Footprint 1b probe or affected its mobility in any assay (data not shown). Thus, the CACCC motif in the HCR is likely to be recognized by other factors.

As discussed under “New HCR Footprints in the Liver Ex-
The samples were analyzed by polyacrylamide gel electrophoresis, and autoradiograms of dried gels are shown. L = mouse liver; K = mouse kidney. Dots indicate major retarded bands.

**Fig. 4.** Gel retardation analysis. Mouse liver and kidney nuclear extracts were incubated with \(^{32}P\)-end-labeled oligonucleotides that were synthesized according to \textit{in vivo} footprint sequences in the HCR-1 (A) and its adjacent \textit{Alu} sequence (B). Unlabeled oligonucleotides were used as specific competitors. The samples were analyzed by polyacrylamide gel electrophoresis, and autoradiograms of dried gels are shown. L = mouse liver; K = mouse kidney. Dots indicate major retarded bands.

**Fig. 5.** Competition gel-retention analysis. Mouse liver nuclear extracts were incubated with 2.5 ng of end-labeled oligonucleotides synthesized according to the \textit{in vivo} Footprint 1b sequence. Unlabeled DNA-binding consensus sequences of nuclear factors EKLF (32) and TF-LF2 (33) were used as specific competitors of the amounts indicated. The samples were analyzed by polyacrylamide gel electrophoresis, and autoradiograms of dried gels are shown. Dots indicate major retarded bands.

expression Domain Detected by \textit{in Vivo} Methods," the Footprint 2 and Footprint 3 sites each contain a sequence that is similar to the tandemly repeated TGGTTC motif of Footprint 4, differing by only 1 nucleotide (Fig. 3). Although the Footprint 4 sequence was shown previously to bind HNF3\(\alpha\) (prepared by cell-free translation of an \textit{in vitro} expression vector) (23), this factor did not bind to Footprint 2 or Footprint 3 oligonucleotides (data not shown). Similarly, HNF3\(\alpha\) did not bind to the Footprint 8 oligonucleotide, which also contains a TGGTTC-like motif (23) (data not shown). Thus, either additional factors are required for HNF3\(\alpha\) binding to these sites, or different transcription factors may mediate the contribution of these footprint sequences to HCR-1 activity. Gel-retention assays using \textit{in vitro} prepared HNF-1, HNF-4, or C/EBP as described (23) showed that these liver-enriched transcription factor families, commonly found to be active in liver-specific gene expression, also did not bind to the HCR Footprint 2 or 3 sequences (data not shown).

**DNase I Hypersensitivity of HCR-1—Footprints 2–7** were distinguished from each other by DNase I-sensitive sites. These sites were especially prominent between Footprints 2 and 3 at nucleotide 103 (Fig. 2, C and D) and between Footprints 6 and 7 at nucleotide 268 (Fig. 2B). The strong site that separated Footprints 2 and 3 (Fig. 2, C and D) corresponded to the approximate location of a liver-specific nuclease-hypersensitive site previously found by an independent \textit{in vivo} method (23). Notable hypersensitive sites also were detected between Footprints 8 and 12 in the \textit{Alu} family segment. Distinct hypersensitive sites were not observed at either end of this \textit{Alu} family cluster of footprints, nor at either end of the HCR-1 cluster covering Footprints 1a–7.

The relative sensitivity of nuclease-hypersensitive sites in the liver-specific domain was examined further by digestion of nuclei with different amounts of DNase I (Fig. 6). Three levels of sensitivity were observed as a consequence of using successively greater nuclease concentrations, suggesting different degrees of DNA exposure or distortion at these sites. The two sites that were prominent at low levels of DNase I, located at nucleotides 103 and 274/268 (equivalent sites detected on either strand of duplex DNA), were not detected when increasing amounts of the enzyme were used. Two additional hypersensitive sites that were prominent at intermediate levels of nuclease were detected near nucleotides 183 and 255. These sites were located on either side of the Footprint 5 and 6 sequences, the most DNase I-resistant domain observed in the central liver-specific portion of the HCR (Fig. 6).

Many additional hypersensitive sites were detected (Fig. 6) when greater amounts of DNase I were incubated with nuclei that were either much less prominent at low enzyme concentrations (i.e. nucleotides 58 and 135) or not detected at the lowest enzyme concentration used (i.e. nucleotides 78 and 114). Many of these latter sites requiring extensive DNase I digestion for detection were spaced 10–12 nucleotides apart. This spacing of weakly hypersensitive sites has been reported previously to be characteristic of rotationally phased DNA that is wrapped around nucleosome cores or transcription factor complexes (38). The possibility that nucleosomes might be associated in a specific arrangement within the HCR was investigated.

**Micrococcal Nuclease Digestion of HCR-1—**Nuclei were partially digested with MNase, an enzyme that preferentially digests the DNA between nucleosomes (reviewed in Refs. 39–41). The partial digestion products were subjected to LMPCR to identify the boundaries of nucleosome-like structures. As shown in Fig. 7, prominent bands were detected having the nucleotide spacing of weakly hypersensitive sites that were not detected at either end of this HCR-1 cluster covering Footprints 1a–7.

DNase I digestion of nuclei with different amounts of DNase I was examined further by digestion of nuclei with different amounts of DNase I (Fig. 6). Three levels of sensitivity were observed as a consequence of using successively greater nuclease concentrations, suggesting different degrees of DNA exposure or distortion at these sites. The two sites that were prominent at low levels of DNase I, located at nucleotides 103 and 274/268 (equivalent sites detected on either strand of duplex DNA), were not detected when increasing amounts of the enzyme were used. Two additional hypersensitive sites that were prominent at intermediate levels of nuclease were detected near nucleotides 183 and 255. These sites were located on either side of the Footprint 5 and 6 sequences, the most DNase I-resistant domain observed in the central liver-specific portion of the HCR (Fig. 6).

Many additional hypersensitive sites were detected (Fig. 6) when greater amounts of DNase I were incubated with nuclei that were either much less prominent at low enzyme concentrations (i.e. nucleotides 58 and 135) or not detected at the lowest enzyme concentration used (i.e. nucleotides 78 and 114). Many of these latter sites requiring extensive DNase I digestion for detection were spaced 10–12 nucleotides apart. This spacing of weakly hypersensitive sites has been reported previously to be characteristic of rotationally phased DNA that is wrapped around nucleosome cores or transcription factor complexes (38). The possibility that nucleosomes might be associated in a specific arrangement within the HCR was investigated.
with the site between Footprints 6 and 7 reflecting a staggered protection (hypersensitivity at nucleotides 268 and 274 on opposite DNA strands) by nuclear proteins (Figs. 2B and 6). The distances between the 3 and 131 bands, as well as between the 268 and 347 bands (128 and 79 nucleotides, respectively), are consistent with the range of lengths observed for internucleosome linker fragments (41).

The band at nucleotide 131 was detected in the liver and the kidney, but not in the brain (Fig. 7). Previous studies showed that the HEG.LE1 construct is expressed at high levels in only the liver, with no detectable expression in other tissues (18). However, the LE1 fragment was shown to contain two discrete silencer activities that specifically inhibited transgene kidney expression as directed by kidney-specific elements near the transcription start site (18). In the absence of these silencers, the apoE gene construct containing the liver-specific HCR domain was expressed at high levels in the kidney as well as the liver, but with no expression in the brain (18). These studies also indicated an interaction between the downstream HCR domain and the promoter region in directing expression of the apoE gene. Thus, the presence of the 131 nucleotide MNase band in the brain may be associated with an open chromatin environment that permits (liver) or can be induced to permit (kidney) expression. Similarly, the absence of the 131-nucleotide MNase band may be a consequence of a more condensed or blocked chromatin environment that is consistent with a lack of gene expression in this tissue.

The finding of MNase bands at nucleotides 3, 268, and 347 in all three tissues suggests that nucleosome or chromatin structure is ordered or phased in some way by the HCR sequence or by neighboring sequences. In this regard, studies by others have demonstrated that nucleosome positioning can be directed by the sequence of the bound DNA (42, 43). A previous report (44) has shown that Alu family members can confer rotational positioning on nucleosomes, and that Alu sequences can influence patterns of nucleosome formation on adjacent sequences. It is noteworthy that the 3' end of an Alu family sequence is located relatively close to HCR-1 (23). Thus, our results are consistent with an ordered positioning of a nucleosome over the central segment of the HCR (nucleotides 131–268), independent of its enhancer activity in different tissues. This positioning may be facilitated by the nearby Alu family member.

At the higher concentration of micrococcal nuclease used in these studies (Fig. 7, lanes 4, 6, and 8), HCR-1 DNA was more susceptible to digestion in the liver than in the brain or kidney. This increased susceptibility may be due to a partial accessibility of the DNA to micrococcal nuclease as a result of the sequence-specific binding of transcription factors to the footprint sites described above.

In Vivo Footprints in the Promoter of the ApoE Gene—Since the HEG.LE1 construct is expressed only in the livers of transgenic mice (18), we might expect to find evidence of transcription factor binding on the transgene promoter in hepatic nuclei. Therefore, we examined the sequences around the transcription start site of the transgene, comparing in vivo nuclear protein binding in the liver to that of the kidney. In this construct, expression of the HEG.LE1 transgene in the kidney appears to be blocked specifically by silencer activity in the

Fig. 6. DNase I digestion of HCR-1 chromatin. Liver nuclei were examined as indicated in Fig. 2. Primer set D was used in the experiment. From lanes 3–7, increasing concentrations of DNase I were used. N = naked genomic DNA digested with DNase I; G = G-specific cleavage of genomic DNA isolated from nuclei that were not treated with DNase I. Arrows indicate prominent DNase I-hypersensitive sites between footprints, and the dots indicate additional DNase I-hypersensitive sites. The asterisks indicate sites of greatest DNase I hypersensitivity. Footprint numbers are indicated.

Fig. 7. MNase digestion of HCR-1. Nuclei isolated from the liver, brain, and kidney of transgenic mice bearing the HEG.LE1 construct were digested with MNase. The DNA was extracted and analyzed by LMPCR. Primer set B was used in the experiment. MNase concentrations in lanes 4, 6, and 8 are higher than those in lanes 3, 5, and 7. N = naked genomic DNA digested with MNase; G = G-specific cleavage of genomic DNA isolated from nuclei of HEG.LE1 transgenic mouse liver that were not treated with MNase.
plexity for HCR-1 that was not possible to detect by This approach allowed us to determine a protein-binding com-

gent and cultured cells (45), the ability to analyze the chroma-

tisin environment of human HCR-1 in transgene mouse tissues (16, 18).

The nucleotide sequence of the human apoE gene promoter shows a typical TATA box motif at nucleotides −28 to −33 (3). This motif is covered by a prominent protein footprint liver nuclei, but not in kidney nuclei (Footprint 2, Fig. 8). A prominent footprint also is found over the start site of transcription in liver nuclei, but not in kidney nuclei (Footprint 3, Fig. 8).

Downstream of this region, the first exon also is protected by nuclear factors in only the liver nuclei (protected segment 4, Fig. 8). Footprints 2, 3, and 4 are almost contiguous, separated only by DNase I-sensitive sites, consistent with the formation of a transcription complex. This pattern of nuclear protein binding would be expected from a transgene that is actively transcribed in the liver, but not in the kidney.

On the other hand, Footprint 1 was observed in both liver and kidney (Fig. 8), indicating the presence of a transcription factor abundant in both tissues. This sequence contains the GGGCGGG Sp1 binding motif. Our previous in vitro studies showed that it was detected as a prominent footprint with nuclear extracts from several different cultured cells (20).

**DISCUSSION**

In these studies, we have characterized the nuclear protein binding of HCR-1 in a chromatin environment that represents its in vivo active state. We demonstrated that transgenic ani-

mal tissues in which a single cell type is dominant (e.g. the hepatocyte in the liver) can be used for in vivo footprinting. This approach allowed us to determine a protein-binding com-

plexity for HCR-1 that was not possible to detect by in vitro methods. The results showed that nearly the entire length of the minimum fragment required for high levels of liver-specific expression was covered with prominent footprints that were delineated by DNase I-hypersensitive sites. In addition, evidence for nucleosome association with actively functioning HCR-1 was obtained. Since human apoE gene constructs ap-

pear to be expressed without much specificity in widely diver-

gent and cultured cells (45), the ability to analyze the chroma-

tin environment of human HCR-1 in transgene mouse tissues has provided an important advance in understanding its mech-

anism of action.

The current results are consistent with our earlier report in which we examined the functional domain of the HCR in trans-

genec mice (23). In those studies, the intact human apoE gene with −5 kb of 5′-flanking and 1.7 kb of 3′-flanking sequence were ligated to different fragment lengths of the HCR, the expression of each construct was examined in transgenic mouse tissues. The minimal fragment needed to direct full levels of expression in the liver consisted of nucleotides 6–325. Shorter HCR fragments resulted in a loss or an attenuation of marker gene (the apoE transgene) transcription in the liver. The current finding that nucleotides 5–302 are fully occupied by nuclear factors in the liver but not other tissues, as determined by in vivo footprinting, fully supports the earlier results (23).

We used a line of transgenic mice bearing the HEG.LE1 construct (18) having about 70 copies of the transgene inte-

grated into the host genome (23). There does not appear to be a deleterious effect of multiple gene copies on the expression of the transgene. For example, there was relatively little nonspe-

ific background in the DNase I footprints that were detected (i.e. Fig. 2, A–D), and MNase digestion (Fig. 7) yielded a con-

sistent pattern in a tissue where expression was high (liver) and in tissues that lacked expression (kidney and brain). Fur-

thermore, the footprint patterns detected in the promoter and HCR domains were consistent with the pattern of transgene expression found for liver and kidney. It seems apparent that regulatory factors were not limiting and that the HCR and promoter have essentially the same chromatin environment in all transgene copies.

The results of in vivo footprinting, summarized in Fig. 9, extend the regulatory picture of the HCR. While previously identified in vitro footprints (19, 23) were confirmed, four new nuclear protein-binding sites were identified in the region re-

quired for liver-specific expression. The location of these new sites was consistent with the previously reported expression pattern of several subfragments of the HCR in which nucleo-

tides 6–325 were determined to provide full enhancer activity in the liver (23). The HCR contains all necessary regulatory sequences for liver expression; the apoE promoter does not direct any hepatocyte expression in the absence of the HCR (18). However, the HCR appears to interact with the proximal promoter, since a nonspecific enhancer element at nucleotides −161 to −140 is required for expression in almost any tissue (18). Consistent with this mechanism, simultaneous activation of both liver-specific promoter sequences and a far distal liver enhancer element to direct tissue-specific expression has been demonstrated for the albumin gene (46).

A second cluster of footprints was found close downstream of the HCR in the Alu family sequence. While this sequence does not appear to contribute liver specificity to the HCR, it may mediate HCR function or play a role in other transcriptional events. Alu family members have been associated with tran-

scriptional control in other genes (47, 48), and they may play a role in chromatin structure by ordering nucleosome positioning in adjacent sequences (44). In the expression of the HEG.LE1 transgene, the Alu sequence appeared to be in a more open chromatin configuration in the liver than in the kidney, as sug-

gested by the presence of in vivo footprints. However, a potential functional for the Alu family sequence awaits further study.

The possibility that a nucleosome is associated with the central portion of the liver-specific regulatory region of the HCR is suggested by the results of MNase digestion (Fig. 7). While the 138-nucleotide segment of protected DNA in the HCR is slightly less than the typical 142–146-nucleotide region that is characteristically protected by nucleosome cores (41),
This motif has been shown to bind HNF3. The nucleosome-protected region contains four TGTTTGC motifs. In addition, the Footprint 4 sequence found at the 5' edge of the nucleosome-protected region contains four TGTTTGC motifs. This motif has been shown to bind HNF3α, a member of a transcription factor family with a histone H5-like DNA recognition motif, that introduces a pronounced bend into DNA upon binding (49). The TTTG motif also is characteristic of the core DNA-binding sequence in high mobility group proteins (50). These proteins also act by introducing prounced bending in DNA at their binding site (51), thereby bringing neighboring DNA sequences into closer proximity. The DNA-bending action would be expected to result in enhanced nucleosome sensitivity. The binding of a HNF3 family member to the Footprint 4 sequence may help to position a nucleosome, as well as to influence DNA structure and nuclear protein binding in adjacent sequences, similar to that found for transcription factors in the albumin gene distal enhancer (52). Alternatively, the DNA sequence of HCR-1 itself may be sufficient to direct nucleosome association; evidence has been presented elsewhere that DNA/histone interactions alone are sufficient to position a single nucleosome in regulatory domains (42, 43).

The finding that nucleosome phasing in the liver was the same as in the kidney, and similarly in the brain, was surprising since the HEG.LE1 construct is expressed only in the liver of transgenic mice (18). Since the DNase I footprints of the HCR in the liver showed essentially no background (i.e. Fig. 2, A–D), it seems likely that all copies of the transgenic bound transcription factors. Likewise, nuclear protein binding in the promoter at the transcription start site (Fig. 8) showed no background. These data suggest that the HCR sequence or neighboring sequences play a major role in nucleosome positioning. Furthermore, it is likely that transcription factors can bind to the HCR nucleosome complex. There is precedence for this latter possibility; the glucocorticoid receptor binds to a nucleosome formed by the mouse mammary tumor virus long terminal repeat, and the binding correlates with the induction of gene expression (53). The possibility of transcription factor binding to the HCR nucleosome complex will be investigated in future studies.

Footprint 1b at the 5' end of the HCR contains a CACCC core sequence characteristic of the binding site of the EKLF, a recently identified erythroid-specific transcription factor that activates transcription of both murine and human β-globin genes (54). The EKLF element is essential for appropriate and optimal promoter function of both globin (55–60) and non-globin erythroid genes (61–63). In the current study, gel retardation and competition studies suggested that a member of the EKLF family may bind to the Footprint 1b sequence via the cooperative action of the TF-LF2-binding protein, a factor that may also bind to Footprint 5 (Figs. 4 and 5). This possibility was tested by doing gel supershift assays with antibodies to EKLF and to BKLF (37), a related factor that also binds to the CACCC motif, but is relatively abundant in the liver. However, neither antibody influenced liver nuclear protein binding to the HCR Footprint 1b oligonucleotide in any combination of assays (data not shown). Thus, the potential role of the EKLF family in the HCR and apoE gene transcriptional activation remains uncertain.

It is noteworthy that a GATA-1-binding site is located in the Footprint 9 sequence in the second cluster of HCR-1 region footprints (23). Studies by others have shown that GATA-1 is the prototypic member of a family of zinc finger proteins that recognize the GATA consensus sequence (64, 65). Potential GATA-binding sites are found in the regulatory elements of all the erythroid expressed genes (66) and in the globin locus control region (67–70). The expression of GATA-1 coincides with the onset of erythropoiesis in the yolk sac (71). Interestingly, the yolk sac is the major expression site of the apoE gene during fetal development (72). Further studies are required to determine if the GATA-1-binding sites of Footprint 9, and possibly the CACCC EKLF-binding site of Footprint 1b, play a role in the developmental expression of the apoE gene.

The application of LMPCR to nuclei from transgenic mouse tissue following the partial digestion of chromatin by DNase I or MNase has provided fresh insights into the control of apoE expression in the liver. Protein binding has now been demonstrated in liver nuclei for regions of HCR-1 shown previously to be required for its activity in transgenic mice. An association of the central liver-specific domain of HCR-1 with a nucleosome is also indicated. These findings constitute an essential extension beyond standard in vitro approaches for understanding the unique properties of HCR-1. The in vivo results indicate the apparent importance of the entire 6–325-nucleotide length of the HCR-1 domain for its function, and they provide a guide for future studies of its role in directing the expression of multiple genes in the apoE locus to the liver.

Acknowledgments—We thank Angela Chen for help with manuscript preparation, Gary Howard for editorial review, and John Carroll, Steven Gonzales, and Amy Corder for graphics.
