Regulation of Transcription of the TATA-less Human Complement Component C4 Gene


*J Immunol* 1998; 160:4353-4360; ;
http://www.jimmunol.org/content/160/9/4353

References

This article cites 49 articles, 27 of which you can access for free at:
http://www.jimmunol.org/content/160/9/4353.full#ref-list-1

Subscription

Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

Permissions

Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

Email Alerts

Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
Regulation of Transcription of the TATA-less Human Complement Component C4 Gene


The 5'-sequences flanking the human complement component C4 genes (C4A and C4B) have been analyzed for their ability to direct expression of a reporter gene in cell lines that constitutively express or do not express C4. No difference in the level of reporter gene expression was detected in cells transfected with C4A- or C4B-specific constructs. A series of reporter constructs containing progressively truncated C4 promoter fragments transfected into the hepatocyte Hep G2 cell line, identified the sequence contained within the region −178 to −39 as that associated with maximal reporter gene expression. This region contains consensus binding motifs for nuclear factor 1 (−110 to −97), Sp1 (−57 to −49), and three basic helix-loop-helix (−137 to −132, −98 to −93, and −78 to −73)-like transcription factors. Electromobility shift assays and DNase I footprinting analysis showed specific DNA-protein interactions of the C4 promoter at the nuclear factor 1, two E box (−98 to −93 and −78 to −73), and Sp1 binding domains. Site-directed mutagenesis of the Sp1 binding site resulted in total abrogation of reporter gene expression and mutation of the E box (−78 to −73) resulted in a 8-fold reduction in expression. We conclude that the Sp1 binding site at position −57 to −49 is critical for accurately initiated, basal transcription of C4.


Received for publication October 27, 1997. Accepted for publication January 7, 1998.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

5 These authors contributed equally to this work.

Address correspondence and reprint requests to Dr. Bernard J. Morley, Rheumatology Section, Division of Medicine, Imperial College of Science, Technology and Medicine, Hammersmith Hospital, Du Cane Road, London, W12 0NN, U.K.

Copyright © 1998 by The American Association of Immunologists

0022-1767/98/$20.00

Regulation of Transcription of the TATA-less Human Complement Component C4 Gene


The 5'-sequences flanking the human complement component C4 genes (C4A and C4B) have been analyzed for their ability to direct expression of a reporter gene in cell lines that constitutively express or do not express C4. No difference in the level of reporter gene expression was detected in cells transfected with C4A- or C4B-specific constructs. A series of reporter constructs containing progressively truncated C4 promoter fragments transfected into the hepatocyte Hep G2 cell line, identified the sequence contained within the region −178 to −39 as that associated with maximal reporter gene expression. This region contains consensus binding motifs for nuclear factor 1 (−110 to −97), Sp1 (−57 to −49), and three basic helix-loop-helix (−137 to −132, −98 to −93, and −78 to −73)-like transcription factors. Electromobility shift assays and DNase I footprinting analysis showed specific DNA-protein interactions of the C4 promoter at the nuclear factor 1, two E box (−98 to −93 and −78 to −73), and Sp1 binding domains. Site-directed mutagenesis of the Sp1 binding site resulted in total abrogation of reporter gene expression and mutation of the E box (−78 to −73) resulted in a 8-fold reduction in expression. We conclude that the Sp1 binding site at position −57 to −49 is critical for accurately initiated, basal transcription of C4. The Journal of Immunology, 1998, 160: 4353–4360.

The liver is the main source of serum C4, and this is reflected in the high level of C4 mRNA and protein expression by the human hepatoma cell line Hep G2 (15, 16). Additional cell types that synthesize and secrete C4 include PBMC (17), glomerular epithelial cells (18), proximal tubular cells of the kidney (19), fibroblasts (20), and synoviocytes (21). During the acute phase response, serum concentration of C4 increases by ~50%. This is predominantly mediated by IFN-γ, which is the only cytokine to induce C4 expression in Hep G2 cells (16), II-1, II-6, and TNF-α having no effect (16, 22, 23). We have previously shown that in Hep G2 cells the major effector mechanism by which IFN-γ up-regulates C4 expression is through stabilization of C4 mRNA (24).

To define the cis-acting elements necessary for high level tissue-specific expression of human C4, we have determined the capacity of progressively truncated C4 promoter fragments to drive expression of a reporter gene in two cell lines, Hep G2 (hepatocyte) and HeLa (epithelial) cells. In vitro DNA-protein interactions were studied using DNase I footprinting and electromobility shift assays.
(EMSA). We now report that maximal promoter activity is asso-
ciated with a compact fragment (−126 to +62) that contains con-
sensus binding motifs for NF-1 (−110 to −93), two E boxes (−98 to −93 and −78 to −73), and Sp1 (−57 to −49). Moreover, our data also demonstrate that the presence of the Sp1 binding site is critical for C4 expression.

Materials and Methods

Cell culture

Cultured cell lines were obtained from the European Collection of Animal Cell Cultures (Wiltshire, U.K.). All reagents were supplied by Life Tech-
nologies (Paisley, Scotland) unless otherwise noted. Hep G2 cells were cultured in DME, supplemented with 15% FCS, 2 mM l-glutamine, 1 mM sodium pyruvate, 100 U/ml penicillin, and 100 μg/ml streptomycin. K562 and HeLa cells were grown in RPMI supplemented with 10% FCS, 2 mM l-glutamine, 100 U/ml penicillin, and 100 μg/ml streptomycin.

C4-β globin reporter constructs

A promoterless β globin fragment was amplified using the PCR from a plasmid, pNb1.1 (25), containing the entire β-globin gene. Primers Y1 (ggtgaaggtcgttcacagctacccc, +22 to +39) and Y2 (cactatgaagactctgcaggtgttcactagcaacctc, +1670 to +1653) were used to ampli-
fy a 1650-bp amplicon that spanned the β-globin gene (primer sequence complementary to the β-globin gene is in upper case), including 29 bp of the 5′-untranslated region and terminating 64 bp downstream of the poly-
adenylation site. This fragment was cloned via NotI restriction sites in the primer sequence into the plasmid pBlueScript SK+ (Stratagene, Cam-
bridge, U.K.). This reporter construct was termed pβ8.

C4 5′-flanking regions were generated using the PCR from cosmid DNA containing the C4 genes, as described previously (13). A series of seven 5′ primers, progressively closer to the C4 cap site and a common 3′ primer termed C4.25 (AGCACGCTTCATGCTGAGGAG, +1805 to +1784), pC4.12 (G1I.2, TGGCTACTTGTGGCTGAGG, −1558 to −1539), pC4.8 (C4.8, CTTACCTGGACCTCCTGGCG, −1043 to −1024), pC4.4 (C4.4, GTATTATTTCTGGGCAAGATTAC, −178 to −155), pC4.3 (C4.3, GGTGCCCCACACTCTGGG, −126 to −106), pC4.2 (C4.2, TGT CAGGTGGTTCCCAGGC, −81 to −63), and pC4.1 (C4.1, AGGTCCA GAGTCACTGTGGC, −99 to −91), respectively, and were amplified using the primers described above. Primers were end-labeled with polynucleotide kinase with [γ-32P]ATP before use in the PCR. After phenol/chloroform extraction of the PCR amplicons, probe concentrations were determined by agarose gel electrophoresis.

Results

In a previous study, we mapped DNase I hypersensitive sites around the C4 gene loci including regions proximal and distal to the C4 promoters (13). A single DNase I hypersensitive site was
identified, for each of the C4 genes, located within 500 bp upstream of the cap sites, only in a cell line expressing C4 (Hep G2). No other hypersensitive sites were identified downstream or within the body of the gene. The investigation reported here has therefore concentrated on the region containing this hypersensitive site and extended the original findings by functionally characterizing the cis-acting sequences regulating the C4 promoter.

**Promoter efficiency of the C4 5′-flanking region**

A series of fusion genes consisting of progressively truncated C4 5′-flanking sequences linked to a β-globin reporter gene (see Materials and Methods) were constructed to identify the functional consensus binding motifs regulating transcription of the C4 gene in the hepatocyte cell line, Hep G2 (Fig. 1). The 3′ end of the C4 sequence in all the constructs was +62, and the 5′ end was variable. The longest constructs, specific for C4B and C4A, extended −1805 and −1668 upstream of the C4 cap site, respectively. Each pβ4 test construct was cotransfected with the control plasmid Hβ (see Materials and Methods) by calcium phosphate precipitation. Transcription of Hβ and pβ4 constructs resulted in cDNA extension products of 62 and 112 nucleotides, respectively. These data confirmed that the majority of the transcripts for both test and control constructs were being initiated from the previously documented endogenous cap sites for C4 (26) and β-globin (32). Transient transfections and primer extension analysis were performed in duplicate in four independent experiments.

Levels of transcription from reporter and control constructs were evaluated by scanning densitometry of the 112- and 62-bp bands. Densitometric values of the control plasmid (Hβ) were used as an internal control to normalize for variations in transfection efficiency. The effect of C4 5′ truncation on reporter gene expression, relative to the level of transcription determined from the full-length C4 fragment (−1558 to +62) in construct pβC4A, is summarized in Table I. No significant difference in transcriptional efficiencies of the C4A (pβC4A) and C4B (pβC4B) promoter regions was observed. Maximal promoter efficiency was associated with two constructs pβ4.4 (−178 to +62) and pβ4.3 (−126 to +62) with mean relative promoter efficiencies (± SE) of 1.9 ± 0.09 and 1.9 ± 0.12, respectively. These data are also indicative of the presence of a negative regulatory element within the region −178 to −1043, as the mean promoter efficiency of construct pβ4.8 was 0.9 ± 0.04 compared with 1.9 ± 0.09 for pβ4.4. Deletion of sequence from −126 to −81 (pβ4.2) reduced transcription of β-globin (0.8 ± 0.04). Total abrogation of gene transcription occurred with the construct pβ4.1 (−39 to +62), demonstrating that the sequence from −81 to −39 is essential for promoter function and directs 42% of maximal transcriptional activity.

**DNA-protein interactions of the C4 promoter**

To complement the promoter function analyses described above, we determined nuclear DNA-protein interactions of the C4 5′-flanking regions using EMSAs (see Materials and Methods).

---

**Table I. Mean promoter efficiencies of the C4 reporter constructs**

<table>
<thead>
<tr>
<th>pβC4 Reporter Construct</th>
<th>Mean Promoter Efficiency (± SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pβC4B (−1805 to +62)</td>
<td>0.9 ± 0.1</td>
</tr>
<tr>
<td>pβC4A (−1558 to +62)</td>
<td>1.0</td>
</tr>
<tr>
<td>pβ4.8 (−1043 to +62)</td>
<td>0.9 ± 0.04</td>
</tr>
<tr>
<td>pβ4.4 (−178 to +62)</td>
<td>1.9 ± 0.09</td>
</tr>
<tr>
<td>pβ4.3 (−126 to +62)</td>
<td>1.9 ± 0.12</td>
</tr>
<tr>
<td>pβ4.2 (−81 to +62)</td>
<td>0.8 ± 0.04</td>
</tr>
<tr>
<td>pβ4.1 (−39 to +62)</td>
<td>0.0</td>
</tr>
<tr>
<td>pβSp1mut (−81 to +62)</td>
<td>0.0</td>
</tr>
<tr>
<td>pβEmut (−81 to +62)</td>
<td>0.1 ± 0.04</td>
</tr>
<tr>
<td>pβEGmut (−81 to +62)</td>
<td>0.0</td>
</tr>
</tbody>
</table>

* Densitometric analysis of the cDNA transcripts from the primer extension analysis of the RNA derived from transiently transfected Hep G2 cells. An arbitrary value of 1.00 was assigned to the level of transcription directed by the construct pβC4A. Correction for transfection efficiency and RNA loading was made by normalization of values against the control construct Hβ.
Three probes were generated, termed P4 (-178 to +62), P2 (-81 to +62), and P1 (-39 to +62) as described (see Materials and Methods). A SIGNAL SCAN search for transcription factor consensus motifs (33) was conducted on the C4 5' flanking sequence associated with maximal expression (up to position -178). This identified putative binding domains for several factors (illustrated in Fig. 2A) including NF-1 (-110 to -97), Sp1 (-57 to -49), and three E box binding sites termed 5'E box (-137 to -132), middle E box (-98 to -93), and 3'E box (-78 to -73).

In agreement with the data from the transient transfection analysis of construct pB4.1 (see Table I), probe P1 containing the same C4 promoter sequence (-39 to +62) did not form any DNA-protein complexes in EMSAs (data not shown). When probe P2 was used in EMSAs with nuclear protein extracts from Hep G2...
cells, and in the absence of any competitor, three distinct complexes were observed (shown in Fig. 2B, lanes 2, 5, and 9). These binding complexes were termed P2A, P2B, and P2C. The addition of excess unlabeled probe P2 to the binding reactions effectively competed all DNA-protein interactions, whereas a nonspecific competitor in the form of EGmut (P2 with mutated Sp1 and E box binding sites) had no effect on complex formation (data not shown). Using a combination of specific competitors and the mutant P2 probes (i.e., Sp1mut and Emut), it was possible to demonstrate that the three electromobility shifts with probe P2 were due to specific interactions with factors binding to the Sp1 (−57 to −49) and 3’ E box (−78 to −73) consensus binding sites (Fig. 2B, lanes 3 to 11). The addition of the excess unlabeled Sp1 oligonucleotide abrogated the binding of P2 with factors corresponding to P2A and P2B, whereas P2C was altered but not abolished (Fig. 2B, lanes 4 and 8). These data were supported using Sp1mut as probe (Fig. 2B, lanes 3 and 10). This suggested that P2C represented more than one DNA-protein complex. The complex remaining after Sp1 competition was termed P2C2. When the Emut probe was used, there was no obvious alteration in complex formation (Fig. 2B, lane 6). However, when the E box and Sp1 competitors were used with Sp1mut and Emut probes, respectively, all DNA-protein complexes were effectively competed (Fig. 2B, lanes 7 and 11). It is possible that Sp1 binding masks the effect of E box competition or there is competition for access to binding sites. DNase I footprinting analysis (see below) was used to address this issue.

Probe P4 bound factors present in crude nuclear protein extracts from Hep G2 cells, in addition to Sp1 and the E box binding factor. These were specifically competed away using a 50-fold molar excess of unlabeled double-stranded consensus oligonucleotide specific for NF-1 (data not shown). These data suggest a positive role for the NF-1 site at positions −110 to −97, which is supported by the enhanced levels of transcription directed by the reporter construct (pB4.4) containing this site (see Table I).

DNase I footprinting analysis (see Materials and Methods) was used to confirm the data obtained from transient transfection analyses and EMSAs and to determine whether binding to the 3’ E box and Sp1 sites occurred simultaneously. End-labeled probes corresponding to pB4.4 (P4), pB4.2 (P2), pBSp1mut, pBEmut, and pBEGmut were used to define the position of the binding sites occupied by nuclear protein extracts from Hep G2 cells. P4 contains the consensus binding motifs for NF-1 (−110 to −97), Sp1 (−57 to −49), and three E box binding sites: the 5’ E box (−137 to −132), middle E box (−98 to −93), and 3’ E box (−78 to −73). Footprint analysis showed strong binding at the NF-1, middle E box, 3’ E box, and Sp1 binding sites (see Fig. 3). The 5’ E box binding motif was not associated with a footprint. The promoter region associated with pB4.2 and containing the Sp1 and 3’ E box consensus binding motifs confirmed that both these sites were associated with strong footprints (shown in Fig. 3) and occupancy of the sites occurred simultaneously. In addition, a strong footprint was present at the extreme 3’ end of the probe in the region +35 and beyond. Since the transient transfection analyses demonstrated that this region was not functional, further characterization of this footprint was not pursued.

An Sp1 binding site at position −57 to −49 is essential for minimal C4 promoter activity

The data from the transient transfection analyses and the EMSAs led us to suggest that C4 promoter activity for the region −81 to +62 was achieved entirely through binding to the Sp1 and 3’ E box consensus motifs. To test this hypothesis, we prepared C4/β-globin reporter constructs containing mutagenized Sp1 and 3’ E box consensus binding domains (see Materials and Methods).
clearly show that both Sp1 and 3 binds simultaneously to provide maximum transcrip-

tional activity.

The −81 to +62 promoter region contains the elements necessary for tissue-specific expression of C4

We next attempted to address the issue of tissue-specific expres-
sion of C4. Two additional cell lines, HeLa and K562, were used for this analysis. Northern blot analysis of total RNA derived from these cell lines showed that Hep G2 cells express abundant C4 mRNA, but no mRNA was detected in either HeLa or K562 cells (data not shown). Reporter constructs and control plasmid Hβ were transiently transfected into HeLa cells by elec-
troporation (see Materials and Methods). These experiments were performed in duplicate on three occasions. In contrast to the data obtained when the same experiments were performed in Hep G2 cells, reporter gene expression from the test constructs was not detected in the HeLa cell line, while levels of expression of the control plasmid Hβ in HeLa cells were similar to those observed in Hep G2 cells (data not shown). These data therefore suggest that the C4 reporter construct, pβ4.2, which in Hep G2 cells confers promoter activity, also contains the sequence(s) necessary to direct tissue-specific expression of C4. This led us to examine the pos-
sibility that the specificity of DNA-protein interactions with probe P2 may be different when EMSAs and DNase I footprinting assays were performed using crude nuclear protein extracts prepared from HeLa and K562 cells. Surprisingly, the complexes formed with probe P2 and the mutant probes using extracts from these nonex-
pressing cell lines appeared to be identical with those seen with Hep G2 cell extracts (Fig. 5). This was confirmed by the DNase I footprint analysis (data not shown). The binding sites associated with Sp1 and 3’ E box using Hep G2 nuclear extracts were very similar when nuclear extracts from K562 and HeLa cell lines were used. In summary, despite observing a tissue-specific pattern of expression of the C4 reporter constructs in transient transfection analysis, we were unable to demonstrate any differences in actual binding of nuclear proteins derived from those cell lines that did or did not express C4.

Discussion

We have used a combination of transient transfection studies and DNA-protein binding assays to define the promoter elements that regulate transcription of the human C4 gene. The hepatocyte cell
These experiments led us to define the region containing the Sp1 binding site (up to position −57) as the minimal human C4 promoter. This region lacks orthodox TATA and CCAAT boxes. This raises important questions regarding C4 transcriptional start site selection and promoter activation because for the majority of tissue-specific eukaryotic genes, this follows nucleation of the TATA box with basal transcriptional machinery (reviewed in Ref. 37). Certain genes appear to lack a TATA box in the orthodox −30 position but either possess cryptic sequences (38), which may localize TATA box binding proteins, or have displaced TATA boxes. In the case of C4, a cryptic TATA box immediately upstream of the cap site seems unlikely, since the construct pβ4.1 was unable to direct any reporter gene expression. The evidence presented here identifies the Sp1 binding site at position −57 to −49 as the critical element for initiating gene transcription which is in common with a number of other recently studied TATA-less tissue-specific promoters (39–42). Recognition of such TATA-less promoters has led to the investigation and delineation of novel promoter mechanisms (reviewed in Ref. 37). Smale and Baltimore (43) have explored the hypothesis that Sp1 binding sites may function as surrogate TATA motifs where the initial localization of Sp1 leads to recruitment of the TATA box-binding proteins (TBP) through Sp1-TBP interactions (44). This suggestion has been supported by functional studies of tissue-specific TATA-less promoters (39–42). In almost all cases, an Sp1 consensus binding motif, located ∼50 bp upstream of the cap site, was critical to promoter function with respect to both level and accuracy of transcription. Indeed a survey of all known eukaryotic promoters determined that the peak location of Sp1 binding sites was at −50 (45) as for C4.

Our data also suggest that the C4 promoter fragment from position −81 to +62 contains the necessary elements to confer tissue-specific C4 expression. Transfection of HeLa cells, which do not express C4, with the C4 reporter constructs failed to demonstrate any detectable reporter gene expression. These data suggest that either the Sp1 or the 3′ E box, or indeed an association between the factors binding these elements, may confer tissue specificity. There is well-documented evidence illustrating the existence of a family of proteins able to bind the Sp1 consensus binding motif (46–48), and it has been suggested that a complex interaction of several factors at this site may be involved in tissue-specific gene expression. Similar interactions between ubiquitously expressed class A bHLH proteins and tissue-specific class B bHLH factors have also been reported (49–51) to confer tissue-specific gene expression. Although the promoter fragment (−81 to +62) confers tissue-specific expression, functional analysis using EMSAs and DNase I footprinting demonstrated that the Sp1 and 3′ E box binding sites bound specific Sp1 and E box binding proteins in nuclear extracts derived from both HeLa and K562 cell lines (Fig. 5). This is perhaps not surprising because factors capable of binding these consensus motifs are ubiquitously expressed. In supershift assays using three different Abs to USF-1 and USF-2, all were found to interact with E box binding proteins bound to probe P2, which were present in all three nuclear extracts (HeLa, K562, and Hep G2) examined (T. J. Mitchell, S. J. Rose, and B. J. Morley, unpublished observations). Findings similar to those reported here have been observed in the promoter of the MUC1 gene which codes for mucin and which is expressed mainly at the apical surface of glandular epithelial cells. Analysis of the 5′ sequences flanking the human MUC1 gene has demonstrated both Sp1 and E box binding factors to be involved in positive regulation of MUC1 in cell types that express mucin and in transcriptional repression of the gene in cell lines that do not express mucin (52, 53). Further work will be required to elucidate whether tissue specificity of human C4 transcription is due to restricted expression of specific...
E box- and/or Sp1-regulatory proteins or whether other proteins are involved.

References


