The Role of Sp Family Members, Basic Kruppel-Like Factor, and E Box Factors in the Basal and IFN-γ Regulated Expression of the Human Complement C4 Promoter

Daniela Ulgiati, Lily S. Subrata and Lawrence J. Abraham

_J Immunol_ 2000; 164:300-307; 
doi: 10.4049/jimmunol.164.1.300
http://www.jimmunol.org/content/164/1/300

References
This article cites 40 articles, 20 of which you can access for free at: http://www.jimmunol.org/content/164/1/300.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

_The Journal of Immunology_ is published twice each month by The American Association of Immunologists, Inc., 1451 Rockville Pike, Suite 650, Rockville, MD 20852 Copyright © 2000 by The American Association of Immunologists All rights reserved.
Print ISSN: 0022-1767 Online ISSN: 1550-6606.
The Role of Sp Family Members, Basic Krupple-Like Factor, and E Box Factors in the Basal and IFN-γ Regulated Expression of the Human Complement C4 Promoter

Daniela Ulgiati, Lily S. Subrata, and Lawrence J. Abraham

The fourth component of human complement (C4) is a serum protein that is expressed in the liver and other organs. The promoter region of the C4 gene has been analyzed in reporter gene assays in two cell lines that represent hepatic (HepG2) and monocytic (U937) lineages. Analysis indicated that regions important for basal transcription in HepG2 cells included Sp1 and E box sites within the first 100 bp upstream of the transcription initiation site but not the nuclear factor-1 site important in the control of the mouse C4 gene. Also, a region encompassing −468 to −310 was able to repress activity 2-fold. However, when a CACCC or GT box sequence at −140 was mutated the repressive activity of the upstream region resulted in almost no activity. The −140 region consists of a series of four closely positioned GT boxes that were shown to bind Sp1, Sp3, and basic Krupple-like factor in EMSA. This novel two-part regulatory element may be involved in the regulated expression of C4. However, IFN-γ a major activator of C4 expression did not signal through this two-part regulatory element. We were able to map the position of an IFN-γ responsive element in U937. IFN-γ was able to increase transcription by up to 20-fold with mutations in the E box sequence at −78 to −73, thus completely abolishing induction. We conclude that the E box binding factors, which appear to be distinct from upstream stimulatory factors 1 and 2, are totally responsible for IFN-γ induction of C4.


Copyright © 2000 by The American Association of Immunologists 0022-1767/00/$02.00
Materials and Methods

Cell lines and culture conditions

Human hepatoma derived HepG2 (HB-8065) and myelo-monocytic U937 (CRL-1593) were obtained from the American Type Culture Collection (Manassas, VA) and maintained in culture medium at 37°C with 5% CO₂, consisting of either Eagles MEM with Earle’s salts and t-glutamine supplemented with 10% FBS, 0.1 mM nonessential amino acids, and 1 mM sodium pyruvate (HepG2) or RPMI 1640 with L-glutamine supplemented with 10% FBS (U937). Both media contained 100 IU/ml penicillin.

Construction of C4 promoter deletion and mutant luciferase fusion constructs

A SmaI/BamHI fragment of the C4 promoter containing −1137 to +44 was cloned into the luciferase reporter pGL3-basic vector (Clontech, Palo Alto, CA). Site-directed mutagenesis was performed using the Quick-Change Mutagenesis kit (Stratagene, La Jolla, CA), which enabled the incorporation of Nhel restriction sites extending 3’ from the positions −316, −141 (CACC), −101 (NF-1), −78 (E box), and −36. Deletion constructs were then prepared utilizing the newly incorporated Nhel site together with the Nhel site situated in the pGL3-basic vector. Restriction enzyme digestion of the mutant plasmids with Nhel resulted in varying lengths of upstream C4 promoter sequence being deleted from the full-length construct.

Transfections and detection of promoter activity

Before each transfection, HepG2 cells were grown to ~70–85% confluence and U937 were grown to 7 × 10⁶ cells/ml and then electroporated with plasmid DNA prepared using Qiagen Maxiprep-500 columns (Qiagen, Clifton Hill, Australia). To each cuvette (Bio-Rad, Hercules, CA) was added 15 µg full-length promoter construct or equimolar amounts of deletion constructs, 300 ng pRL-TK (Promega, Madison, WI) transfection control vector and 400 µl of cells (2.5 × 10⁶–2.5 × 10⁷ cells/ml). The cells were electroporated using a Bio-Rad gene pulse (240 V and 960 µF) and distributed into a 6-well culture tray containing 5 ml of culture medium together with the nuclear extract for 10 min on ice. The nuclear extract was then incubated with 80 fmol of 32P-labeled oligonucleotide for 30 min on ice before loading onto a 6% polyacrylamide gel. The gel was then electrophoresed at 150 V using 0.25× Tris-taurine-EDTA as the running buffer. EMSA gels were then dried under vacuum and exposed to x-ray film.

Results

Requirements for high level basal activity lie within the first 138 bp of the human C4 promoter

To determine promoter regions that contribute to expression of C4, a series of 5’ deletion constructs were generated by introducing mutations into a reporter construct that consisted of the −1137 to −44 region of the C4 gene (Fig. 1) linked to the luciferase coding region of pGL3-basic. Deletion of the regions between the mutations and the Nhel restriction site in the pGL3 vector created the 5’ deletion series. These constructs were used in transient transfection assays to determine the minimum promoter region required for basal expression of human C4 in the hepatic (HepG2) environment (Fig. 2). The full-length −1137 construct was able to mediate high level expression in HepG2. Truncation of the C4 promoter from −1137 to −310 (Fig. 2A) caused a 1.6-fold increase in promoter activity over the full-length construct. Further truncation to position −136 resulted in the maintenance of elevated basal transcription with a 2-fold increase observed when compared with the full-length −1137 construct. Recent studies by Vaishnaw et al. (13) have established the presence of negative regulatory elements in nuclear extracts (10–20 µg) protein were preincubated on ice for 10 min together with 1 µg poly(dl-dC) in a binding buffer consisting of 4% Ficoll, 20 mM HEPES (pH 7.9), 1 mM EDTA, 1 mM DTT, and 50 mM KCl. When required, competitor oligonucleotides or supershift basic Kruppel-like factor (BKL1) (20), or commercial Sp1, Sp3 or upstream stimulatory factor-1/2 Abs (Santa Cruz Biotechnology, Santa Cruz, CA) were then incubated with the nuclear extract for 30 min on ice. The nuclear extract was then incubated with 80 fmol of 32P-labeled oligonucleotide for 30 min on ice before loading onto a 6% polyacrylamide gel. The gel was then electrophoresed at 150 V using 0.25× Tris-taurine-EDTA as the running buffer. EMSA gels were then dried under vacuum and exposed to x-ray film.
the −1043 to −178 region in HepG2 cells. We have now defined the region to between −1043 and −310. Further truncation to −95 had a minor effect. However, truncations to position −72 resulted in a marked (2-fold) decrease in promoter activity (Fig. 2A) compared with the full-length construct. Further truncation of the full-length construct to position −30, leaving only the Inr and downstream elements, resulted in a 40-fold decrease in basal transcriptional activity. These later results are consistent with the requirements seen in human (13) and mouse hepatic cells (12), as the Inr and downstream elements are sufficient for start site selection and low level activity, with the immediate upstream region containing the Sp1 site at −57 to −49, conferring elevated basal transcription.

Mutation analysis demonstrates important roles for the −141 CACCC box and −78 E box motifs in maintaining basal expression of C4

The 6-bp NheI mutations used to create the 5′ deletion series were designed to introduce changes into putative transcription factor binding sites in the C4 promoter. These mutant constructs were used to identify critical elements involved in basal C4 transcription. Both the full-length promoter construct and the site-directed mutants were transiently transfected into HepG2 cells and assayed for transcriptional activity. The full-length C4 promoter and the construct containing a 6-bp linker mutation at location −316 (Fig. 2B) produced high levels of basal transcriptional activity. In contrast, a promoter construct carrying a 6-bp mutation at position −141 in a CACCC box (or GT box) regulatory element produced a 94% decrease in transcriptional activity compared with the full-length construct in HepG2 (Fig. 2B), suggesting an important role for this site in maintaining efficient levels of basal activity although the deletion results gave no indication of this. Mutation of the putative NF-1 site at position −101 did not significantly effect transcription, although this site has been shown to be protected by HepG2 nuclear extracts in DNase I footprinting (13). In contrast to our results, the equivalent NF-1 site in the mouse C4 promoter is required for full activity C4 promoter activity. Mutation of the E box motif (position −78) resulted in a 86% decrease in transcriptional activity when compared with the full-length construct. These results indicate a major role also for the E box motif in basal C4 activity. Mutation in a region starting at position −36 had the effect of decreasing transcriptional activity by 64%. This region may be protected by HepG2 extracts in DNase I footprinting (13) and serves a modulatory role in C4 promoter activity in both cell types.

Characterization of factors binding to the E box motif

The transfection results from both the deletion and mutation constructs identified the E box as a major transcriptional element.
The E box sequences in human and mouse are identical and in mouse it has been shown that this site is a major element driving basal expression of the C4 gene (12, 13). The murine results also suggest that members of the USF family distinct from USF-1 or USF-2 bind to the rodent sequence (21). To determine the nature of the transcription factors interacting with this region, EMSA were conducted using extracts derived from HepG2 cells and the −89 to −64 sequence containing the E box motif as a probe (Fig. 3). Two major complexes (A and B) were seen (Fig. 3A), both of which could be substantially competed away using 10-fold molar excess of unlabeled E box probe. When a consensus USF binding site was used as a competitor, the lower complex B could be competed away using 10-fold excess, indicating that it contained factors that recognized the USF binding site. However, the major complex, A, was not substantially removed by USF site competition at 250-fold molar excess, indicating that this complex was not recognizing the USF binding site but may recognize a site that overlaps with the E box site in the C4 promoter. To determine whether complex B was USF-1 or USF-2, supershift assays were done with a USF-1/2 Ab. Neither pre- or postincubation of the extract with Ab caused any discernible change in the gel retardation profile seen in the absence of Ab (Fig. 3B), whereas control reactions using a consensus USF binding site showed substantial reactivity with bound complexes (data not shown), suggesting that complex B did not contain proteins antigenically related to USF-1 or USF-2.

**Factors binding to the −141 CACCC box region include Sp1, Sp3, and BKLF**

The reporter gene analysis of mutation within the −141 CACCC box indicated that it was very important in regulating expression from the C4 promoter. To determine the nature of the transcription factors that interacted with this region, EMSA were conducted using double-stranded oligonucleotides that corresponded to the −152 to −113 region of the promoter. Inspection of this 40-bp region indicated that a total of four GT boxes were present, two in the reverse (CACCC box) orientation (Fig. 4A). The region has the potential to form a variety of palindromic structures involving adjacent GT box and CACCC box sequences. Complexes binding to this region were identified in EMSA using oligonucleotides representing either the entire 40-bp region (4XGT), a truncated region including the upstream 33 bp but missing the proximal 3′ GT box (3XGT), and two mutant sequences that had either −141 to −137 (Umut) or −128 to −123 (Dmut) changed to a NheI restriction site GCTAGC (Fig. 4A). Using HepG2 nuclear extracts, 4XGT bound four major complexes A, B, D, and E (Fig. 4B). Self competition indicated that complex B could be removed with a 10-fold excess and complex A with a 50-fold excess of unlabeled 4XGT. However removal of complexes D and E required a 250-fold excess and may indicate that very high concentrations of these factors are present in the extract and/or represents nonspecific interactions. Use of the shortened 3XGT probe in EMSA (Fig. 4C), indicated that only complex B was able to bind to the upstream region. Introduction of the mutation at −128 to −123 (probe Dmut) to the 4XGT sequence had the effect of abolishing binding of complex E and substantially reducing formation of complex B (Fig. 4C).

The upstream GT box mutation (Umut) abolished formation of complexes B, D, and E. Taken together, the results indicated that complex A was interacting with the −118 to −112 region and complex B was interacting with the −141 to −127 region but was influenced by downstream sequences. Surprisingly, complex D was interacting with the −141 to −127 region but also required the downstream region −118 to −112 to bind. Complex E appeared to require the entire 40-bp region for binding, interacting with the −141 to −127 region, the −127 to −122 region and the −118 to −112 region deleted in the 3XGT oligonucleotide.

To establish whether Kruppel-like Sp (22) or EKLF (20) family members were involved in complex formation within the −140 region, competition EMSA were conducted using cognate GC and GT box sequences as well as the GC-rich AP2 consensus sequence (Fig. 4D). Both the GC and GT box Sp1 consensus competitors were able to specifically remove complex A only; complex B, D and E were unaffected. The consensus AP2 competitor had the effect of removing complex A but also reduced the binding of complexes B and E. Considering the fact that 250-fold molar excess of the AP2 consensus could not completely prevent formation of these complexes indicated that complexes B and E may be interacting nonspecifically with the AP2 oligonucleotide (Fig. 4D).

The identity of complex A as containing Sp family members was confirmed using Abs specific for either Sp1 or Sp3 (Fig. 4E). Both Abs were able to interfere specifically with complex A formation indicating that complex A contained both Sp1 and Sp3; furthermore, both proteins were required for complex formation as either Ab almost completely abolished complex A binding. As complex B was interacting specifically with the CACCC box sequence at −141 to −127, supershift analyses were conducted using an anti-BKLF Ab. BKLF has been shown to be a major CACCC box binding activity in liver (20). The BKLF-specific Ab was able to interfere with the formation of complex B (Fig. 4E), suggesting that complex B was BKLF.

**Functional interaction between the −141 CACCC box region and upstream sequences**

Results from the 5′ deletion experiments indicated that the GT box sequences at −141 were not required for basal expression of C4, because deletion of the region from −137 to −141 including the reverse strand GT box sequence CACCC, had no significant effect on promoter activity. However, mutation of the CACCC site to a

![FIGURE 3. EMSA analysis of the −89 to −64 E box region with the 26-bp probe, Ebox, using HepG2 nuclear extracts. The two major complexes interacting with the E box recognition site, A and B are indicated by the arrows. A, Competition EMSA using either the Ebox probe or a consensus USF binding site. Fold molar excess (×) of unlabeled competitor is as indicated. B, Supershift EMSA using USF-specific Ab. Ab was added either before or after incubation of the probe with the nuclear extract as indicated.](http://www.jimmunol.org/)

The Journal of Immunology
NheI site in construct mut-141 almost completely abolished activity. These results taken together indicated that sequences upstream of the GT box were important and had a strongly negative effect on promoter activity but only in the absence of the GT element. A possible explanation is that deletion of both these elements nullifies the balancing effect of the factors binding at the two sites and results in activated transcription generated by sequences downstream including the Sp1 and E box motifs. To further delineate the upstream region responsible for the negative effect, a series of deletions were introduced into the full-length construct containing the mutation of the GT box CACCC site at -141 (Fig. 2, construct mut-141). When the constructs were tested by transfecting HepG2 cells, it was determined that the region responsible for the repressive effect was located between -468 and -310, as deletion to -468 had no effect on the very low activity of the parent construct, mut-141, but deletion to -310 restored high level expression to above wild-type levels (Fig. 5). Analysis of the region between -472 and -310 using TFSEARCH version 1.3 (40) revealed the presence of potential transcription factor binding sites for GATA and Nkx2 (see Fig. 1); cognate binding activities for both GATA-6 (23) and Nkx2.8 (24) are expressed in liver. Members of the Nkx and GATA families of transcription factors have been shown to interact in other tissue types (25) and may potentially interact in hepatocytes.

IFN-γ responsiveness of the C4 promoter

Previous data have indicated that expression of C4 is increased in response to IFN-γ in cells of both hepatic and myeloid origin. In the hepatic cell line HepG2, C4 expression is expressed constitutively at high levels. Mitchell et al. (14) found that C4 expression
was increased in response to IFN-γ primarily via posttranscriptional events. In myeloid cells such as peripheral blood monocytes and the U937 cell line, however, expression is undetectable but is strongly increased in response to IFN-γ. Part of this increase presumably involves transcriptional activation of the C4 promoter (18). To determine whether the C4 promoter played a role in the increased C4 synthesis seen upon induction with IFN-γ, the responsiveness of the various C4 promoter constructs were tested in U937 cells.

Treatment of cells transfected with the full-length, −1137 WT construct with varying concentrations of IFN-γ produced a dose- and time-dependent increase in luciferase production when compared with the transcriptional activity obtained for uninduced samples with maximal induction (20- to 25-fold) occurring at a concentration of 100 U/ml of IFN-γ (data not shown). Induction using 100 U/ml IFN-γ was half-maximal at 5 h (Fig. 6A).

To localize the region responsible for the IFN-γ up-regulation observed, both the deletion and mutant constructs (see Fig. 2) were tested in U937 cells. Of significant interest were deletion constructs and site-directed mutants that eliminated putative IFN-γ responsive elements of the type identified in the TRANSFAC search. These elements were located at −327 to −320, −316 to −309, and −39 to −32 (see Fig. 1) and were of the type CTG/TGN/TANNY found in HLA class II (26–28) and FcR (29) genes to mediate IFN-γ responsiveness. When the deletion constructs were transiently transfected into U937 cells, the data revealed that only 95 bp of upstream promoter sequence was required for full inducibility by IFN-γ (Fig. 6B). Deletion of two of the putative γ response elements at positions −327 to −320 and −316 to −309 (Fig. 6B, construct −136) did not affect induction significantly. Truncation of the promoter to position −72 (Fig. 6B) resulted in the elimination of IFN-γ induction, indicating that an element required for IFN-γ inducibility was positioned between −95 and −72.

Results obtained from the mutant constructs confirmed that the −95 to −72 region and specifically the E box was the major element controlling IFN-γ responsiveness (Fig. 6C). All of the other mutations retained significant responsiveness to IFN-γ, although expression levels differed. Interestingly, the promoter construct containing a mutation in the CACCC box at position −141, which was shown to cause a marked decrease in basal transcriptional activity, was still inducible by IFN-γ (Fig. 6C, construct mut-141). These results indicate that, as well as being an important element contributing to basal C4 transcription, the E box was the major element mediating IFN-γ responsiveness.

To determine whether IFN-γ induction was correlated with changes in the nature of the nuclear proteins interacting with the E box region, EMSA analysis was undertaken. Using the E box probe and nuclear extracts derived from U937 cells, EMSA results indicated that similar complexes to those seen in HepG2 extracts were formed but complex A was less abundant than that seen in the HepG2 extract (Fig. 7). U937 cells also were incubated with 100 units IFN-γ for 0, 0.5, 3, or 6 h before preparation of extracts. The EMSA results showed that both complexes A and B increased following induction with IFN-γ (Fig. 7), and an additional complex B' appeared at 3 h. Following 6 h of IFN-γ treatment, all three complexes decreased when compared with the 3 h treatment.

FIGURE 6. Effect of IFN-γ on the expression of C4 in U937 cells. A, U937 cells were transfected with construct −1137 WT and then treated with 100 U/ml of recombinant human IFN-γ for 3, 6, 12, or 24 h before assaying for luciferase activity. Shown are mean induction ± SE from three experiments. B and C, Activities of deletion and mutant constructs were compared with the parent −1137 WT construct with or without IFN-γ treatment for 24 h. The results represent the mean luciferase activity data ± SE for three experiments.

FIGURE 7. EMSA analysis of the E box region following IFN-γ induction. E box probe binding activity from nuclear extracts of untreated or 0.5, 3, or 6 h IFN-γ treated U937 cells were analyzed as indicated. The major complexes A and B (indicated by arrowheads) that were seen in HepG2 extracts (Hep) were also present in the U937 extracts. An additional IFN-γ inducible complex B' is indicated by the arrowhead.
Discussion

Investigation of the requirements of the human C4 promoter for expression revealed that high levels of basal transcription were attained with a region comprising 141 bp of upstream promoter sequence. Promoter constructs containing only the initiator and downstream elements resulted in a decrease in activity to a very low but measurable level. These results are comparable to data from mouse studies which found that initiator and downstream elements were sufficient for minimal transcription and start site selection; however, promoter activity was enhanced with the presence of both the NF1 and E box motifs (12). Within the immediate upstream promoter region of the human gene, two elements have been identified recently as being important in driving expression, an Sp1 site (−57 to −49) and an E box motif (−78 to −73) (13). Our data confirm these results and, in addition, identify an additional regulatory site downstream of the Sp1 site in the −36 to −31 region. In contrast to the mouse studies, we have found evidence for only a minor involvement of the NF-1 site in driving transcription of the human gene. This is despite the fact that DNase I footprint analysis indicates that proteins are interacting with the putative NF-1 site of the human C4 gene (13). It is possible that, if NF-1 is involved in transcriptional control, then it may have a role other than in basal transcription in hepatic cells and may be involved in cell-type or activation-specific activities.

Our results indicate a very important function for the E box element located at −78 to −72. The E box motif forms part of a growing number of proteins, termed basic helix-loop-helix (bHLH) transcription factors which bind the consensus CANNTG. The E box element located within the C4 promoter is characterized as class B due to the cognate binding sequence, CACGGT. Included within this family are c-MYC (30) and USF (31). Galibert et al. (21) have characterized the binding of a transcription factor as similar, albeit distinct, from USF interacting with the mouse C4 E box motif. Our EMSA data show that two nuclear protein complexes interact with the region; however, only one of these (complex B, Fig. 3) recognizes the consensus USF binding site, suggesting that it is a member of the USF family. However, supershift analysis indicated that it was not antigenically related to USF-1 or USF-2 and may be another member of the class B bHLH family.

The experiments performed to investigate IFN-γ inducibility of the human C4 promoter indicated a dose-dependent increase in transcriptional activity of the full-length construct when treated with IFN-γ. The kinetics of induction showed enhanced transcriptional activity reminiscent of that previously seen in MHC class II genes (32). An increased response was observed at 3 h, with maximal activity shown between 6 and 12 h. After an induction time of 24 h, the transcriptional activity was decreasing but still elevated over basal levels. Even taking into account the 1-h lag time in the synthesis of luciferase protein (33, 34), these data suggest that the IFN-γ response element functional in the C4 promoter would be similar to a IFN-γ regulatory element (γ-IRE) (32) rather than an immediate response IFN-γ activation site (GAS) element, where induction with IFN-γ is abolished after 3 h (35, 36). However, investigation of the putative γ-IREs within the C4 promoter demonstrated that none of the classical IFN-γ elements were functional. Instead the E box at −78 was totally responsible for induction.

The identity of the transcription factors binding to the E box remain unknown but, as outlined above, EMSA indicates that one of them (complex B) recognizes consensus USF binding sites and the other (complex A) does not. However complex A does recognize at least part of the E box sequence as mutations at this site abolish binding. Both of these complexes increase dramatically upon IFN-γ stimulation and may compete for binding at overlapping cognate sites. In addition, a new complex (B’) appears following IFN-γ treatment. Further work is required to determine whether increased transcriptional activity is due to an increase in the concentration of the proteins or to phosphorylation or other modifications of the proteins involved. However, given that the kinetics of induction of the binding activities precedes transcriptional activity, de novo synthesis of components of the EMSA complexes is likely, as in the case of IFN-γ activation of the MHC class II genes (32).

A novel two-part regulatory element was discovered within the human C4 promoter which appears critical for basal activity in hepatic cells. The reporter gene analysis results indicated the presence of repressor sequences between −468 and −310 (which contain putative binding sites for GATA and Nkx2) that had the effect of decreasing promoter activity by almost 50%. In addition, these distal elements appeared to be acting in concert with a complex of Sp1/3 and BKLF-binding GT box elements around −140. This interaction has the effect of masking the very strong negative effects due to the distal region. The mechanism for this masking effect is currently unknown, but our hypothesis is that interaction with the −140 region prevents interaction of the upstream element with the proximal basal elements. Abolition of the −140 sequences allows interaction of the repressor region with the proximal elements that have the effect of preventing transcriptional activity. It is of interest to determine whether extracellular inflammatory or other signals are able to impinge upon the factors interacting at the −140 and/or upstream repressor sites and regulate expression of C4 in a similar way. Also, given the role of BKLF and other EKLF family members in the cell type-specific regulation of a number of genes (20, 37–39), it is likely that this complex regulatory element may play a role in the tissue-specific expression of the C4 gene.

Acknowledgments

We thank Joanna Guy for expert technical assistance and Dr. Merlin Crossley for provision of the BKLF Ab and helpful advice.

References


