The Human gC1qR/p32 Gene, C1qBP
GENOMIC ORGANIZATION AND PROMOTER ANALYSIS*

Received for publication, October 4, 2000, and in revised form, February 5, 2001
Published, JBC Papers in Press, March 8, 2001, DOI 10.1074/jbc.M009064200

Angela J. Tye‡, Berhane Ghebrehiwet‡, Ning Guo§, Kedarnath N. Sastry¶, Billy K. C. Chow‡, Ellinor I. B. Peerschke†, and Boon-Leong Lim**

†Department of Zoology, University of Hong Kong, Pokfulam, Hong Kong, China, ‡Department of Medicine, State University of New York, Stony Brook, New York, 11794-8161, §Department of Pathology, Weill College of Medicine, Cornell University, New York, New York 10021, and ¶Department of Pathology and Laboratory Medicine, Boston University School of Medicine, Boston, Massachusetts 02118

The Human gC1qR/p32 Gene, C1qBP

gC1qR is an ubiquitously expressed cell protein that interacts with the globular heads of C1q (gC1q) and many other ligands. In this study, the 7.8-kilobase pair (kb) human gC1qR/p32 (C1qBP) gene was cloned and found to consist of 6 exons and 5 introns. Analysis of a 1.3-kb DNA fragment at the 5′-flanking region of this gene revealed the presence of multiple TATA, CCAAT, and Sp1 binding sites. Luciferase reporter assays performed in different human cell lines demonstrated that the reporter gene was ubiquitously driven by this 1.3-kb fragment. Subsequent 5′ and 3′ deletion of this fragment confined promoter elements to within 400 base pairs (bp) upstream of the translational start site. Because the removal of the 8-bp consensus TATATATA at -399 to -406 and CCAAT at -410 to -414 did not significantly affect the transcription efficiency of the promoter, GC-rich sequences between this TATA box and the translation start site may be very important for the promoter activity of the C1qBP gene. One of seven GC-rich sequences in this region binds specifically to PANC-1 nuclear extracts, and the transcription factor Sp1 was shown to bind to this GC-rich sequence by the supershift assay. Primer extension analysis mapped three major transcription start regions. The farthest transcription start site is 49 bp upstream of the ATG translation initiation codon and is in close proximity of the specific Sp1 binding site.

gC1qR is a biologically important, widely distributed, multiligand-binding and multifunctional protein (1). Numerous reports have claimed that gC1qR and its homologue could be isolated or identified in various cellular compartments, including plasma membrane, cytoplasm, mitochondria and nucleus. gC1qR isolated from the plasma membrane was originally characterized as a high affinity C1q-binding protein (2), and later many reports showed that gC1qR could interact with several proteins of the intrinsic coagulation/bradykinin-forming cascade, including high molecular weight kininogen (3), Factor XII (4), fibrinogen (5), and multimeric vitronectin (6). Conversely, intracellular gC1qR was shown to interact and subsequently down-regulate the surface expression of the α1b-adrenergic receptor (7) as well as bind to the kinase domain of protein kinase Cκ and thus prevent its substrate phosphorylation activity (8). In addition, gC1qR was also reported to bind to a nuclear splicing factor, SF2, and to many viral proteins, including HIV-1 Tat (9) and Rev (10), core protein V of adenovirus (11), EBNA-1 of the Epstein-Barr virus (12), and open reading frame P of the herpes simplex virus (13), implying that gC1qR may play a role in virus-host interaction.

The full-length cDNA of gC1qR encodes a pre-pro-protein of 282 residues from which a 73-residue-long N-terminal segment is removed by site-specific cleavage to generate the mature gC1qR (2). It was shown that the fusion of the residues 1–81 or 1–33 of the pre-pro-protein to the N terminus of the green fluorescent protein directed the fusion protein to mitochondria (14). However, the findings of Dedio et al. (14) do not exclude the possibility that gC1qR, like many other proteins, could be exported from the mitochondria by an unknown mechanism (15). This possibility is supported by a recent report showing that anti-gC1qR monoclonal antibody can reverse the anti-proliferation effects of hepatitis virus C core antigen on activated T cells (16).

The human C1qBP gene was assigned to human chromosome 17q13.3 (17). A high degree of amino acid identity exists between the human, rat, and mouse gC1qR cDNA sequences (18). In this study, the full-length gene of human gC1qR was cloned, and its exon-intron boundaries were revealed. Furthermore, the transcription start site and its promoter elements were also mapped and characterized.

MATERIALS AND METHODS

Screening of the Human Genomic Library—A human genomic library in bacteriophage EMBL3 was purchased from CLONTECH (cat. no. HL1067J). The human cDNA of the C1qBP gene was used as the probe for library screening (2). The cDNA insert was released from its vector and gel-purified before radiolabeling with [32P]dATP by the random priming method. Positive plaques were picked, replated, and rescreened until single positive plaques were picked. The insert in a positive plaque was digested with various restriction enzymes, and the positive fragments were mapped by Southern blot analysis. Six overlapping subclones (Fig. 1) were obtained by inserting the positive fragments into the plasmid pBluescript. Each subclone was sequenced by the primer walking method using the ABI Prism 310 Genetic Analyzer (Applied Biosystems, Hong Kong Limited, Hong Kong).

Sequence Assembly and Analysis—Sequence data were assembled and analyzed by DNA processing software including MAC DNA/SIS (Hitachi, Japan) and DNA Strider (Christian Marek, Service de Biochimie, Department de Biologie, Institut de Recherche Fondamentale, CEA, Paris, France). Promoter analysis was performed using MatInspector software (version 2.2) (19) obtained from the German Research Center for Biotechnology, Braunschweig, D-38124, Germany. A transcription factor data base (TRANSFAC, version 4.0) was employed for

* This work was supported in part by Grants RGC HKU198/95 from the Hong Kong Research Grant Council (to B-L. L.) and RPG-95068-04-CIM from the American Cancer Society (to B. G.). 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.

** To whom correspondence should be addressed: Dept. of Zoology, the University of Hong Kong, Pokfulam, Hong Kong, China. Tel.: 852-22990826; Fax: 852-25599114; E-mail: blim@hkucc.hku.hk
**The Human gC1q-R/p32 Gene (C1qBP)**

**Table I**

| Primers designed for construction of recombinant plasmids and gel shift assays |
|---|
| **For recombinant plasmids**<sup>a</sup> | **Primer sequences** | **Position in 5'-flanking region** |
| Pro00Sac | 5′ attaGAGCTCT CGCCCTTTGAGGATG 3′ | −1314 to −1300 |
| Pro00Sac | 5′ attaGAGCTCT ACCAACTACACAGGAAGCTT 3′ | −1121 to −1101 |
| Pro00Sac | 5′ attaGAGCTCT GAGCCACCCGCTTCA 3′ | −916 to −901 |
| Pro00Sac | 5′ attaGAGCTCT GAGTACCCTACAG 3′ | −713 to −702 |
| Pro00Sac | 5′ attaGAGCTCT GTTTTTCCGAGGCTT 3′ | −506 to −489 |
| Pro95Sac | 5′ attaGAGCTCT TTGAAAGGTCCTGAGG 3′ | −364 to −350 |
| G1A2Bgl | 5′ attaGAGCTCT CCGGAAAGCCTGCGGAA 3′ | −18 to −1 (inverted) |
| Pro5Bgl | 5′ attaGAGCTCT GCGCCCTTTGAGGATG 3′ | −1314 to −1300 |
| G1A2Sac | 5′ attaGAGCTCT CCGGAAAGCCTGCGGAA 3′ | −18 to −1 (inverted) |

<sup>a</sup> Four bases and a restriction site were added at the 5′ end.

<sup>b</sup> All primers were designed from their GC with respect to their location in the 5′ flanking region of the human gC1qBP gene. “S” indicates sense primers, and “A” is the antisense version of its sense primer. L8S and L8A, which do not contain GC-rich sequences, were designed as negative controls.

**Cell Culture, Transient Transfection, and Promoter Assays**—The human cell line PANC-1 was selected for gel shift assays. Cells were cultured to exponential phase as mentioned above and harvested. End-labeling of double-stranded oligos (Table I) was done using the Ready-To-Go T4 polynucleotide kinase labeling kit and [γ-32P]ATP (5000 Ci/mmol) (Amersham Pharmacia Biotech). Nuclear extract preparation and gel shift assays were performed as described previously (21, 22).

**RESULTS**

**Genomic Organization of the Human C1qBP Gene**—Fig. 1 is a schematic representation of the genomic organization of the gene, which, including its 5′- and 3′-flanking regions, spans about 7.8 kb. From the first codon of the initiation methionine to the stop codon of the gene, the gene spans 6055 bp (Fig. 2). By alignment of the cDNA and the genomic sequences, intron-exon boundaries were defined. There are 6 exons and 5 introns in the C1qBP gene. The size of the exons range from 94 (exon 3) to 232 bp (exon 1), and that of the introns range from 128 (intron 1) to 232 bp (exon 1), and that of the introns range from 128 (intron 1) to 3156 bp (intron 2). Amino acid codons are split by introns 1 and 2 at the junctions of their adjacent exons. Table III A poly(A) signal is located 369 bp from the stop codon. The entire genomic sequence has been deposited in the GenBank<sup>TM</sup> under accession no. AF338439.

**Characterization of the 5′-Flanking Region of the Human C1qBP Gene**—A 1.3-kb nucleotide sequence located upstream of the ATG initiation codon of the gene was analyzed (Fig. 3) using the TRANSFAC transcription factor data base. Putative

---

1 The abbreviations used are: kb, kilobase pair(s); AP, transcription factor activator protein; bp, base pair(s).
TABLE II
Expression plasmids for luciferase promoter assay

| Constructs | Primers employed | Restriction sites at 5’-end | Restriction sites at 3’-end | Fragment size | Deletion from |
|------------|------------------|-----------------------------|-----------------------------|---------------|---------------|
| ERES 1     | Pro0Bgl + G1A2Sac | BglII                       | SacI                        | 1.3           | Nil (reverse) |
| ERES 2     | Pro0Sac + G1A2Bgl | SacI                        | BglII                       | 1.1           | 5’            |
| ERES 3     | Pro200Sac + G1A2Bgl | SacI                    | BglII                       | 0.9           | 5’            |
| ERES 4     | Pro400Sac + G1A2Bgl | SacI                        | BglII                       | 0.7           | 5’            |
| ERES 5     | Pro600Sac + G1A2Bgl | SacI                        | BglII                       | 0.5           | 5’            |
| ERES 6     | Pro800Sac + G1A2Bgl | SacI                        | BglII                       | 0.4           | 5’            |
| ERES 7     | Pro954Sac + G1A2Bgl | SacI                        | BglII                       | 0.4           | 5’            |
| ERES 2H    | Pro0Sac + G1A2Bgl  | SacI                        | HindIII                     | 0.67          | 3’            |
| ERES 2P    | Pro0Sac + G1A2Bgl  | SacI                        | PsPAI                       | 1.1           | 3’            |

* Nil, no deletion.

Fig. 1. The genomic organization of the human C1qBP gene. Exons are highlighted as dark boxes. Dark, bold lines below the gene indicate the overlapping subclones. The arrows indicate the positions of sequencing primers and the approximate length sequenced by each primer. Restriction enzyme sites are shown at locations above the gene.

Fig. 2. DNA sequences of the human C1qBP gene. The first nucleotide of the translation start codon ATG is designated as 1. Exon-intron boundaries are shown. All intronic sequences, the 5’-flanking region of the gene, and the 3’-flanking region as defined by the cDNA amino acid coding sequences are shown as lowercase letters. A poly(A) signal located 369 bp from the stop codon is underlined.

The major transcription start sites. The farthest extension band was observed at 49 bp upstream of the ATG translation initiation codon and is an adenine residue. This nucleotide is positioned just downstream to an SP1 binding site found in the promoter region. All of these regions lie within 50 bp upstream of the major transcription start sites. The farthest extension band was observed at 49 bp upstream of the ATG translation initiation codon and is an adenine residue. This nucleotide is positioned just downstream to an SP1 binding site found in the promoter region. All of these regions lie within 50 bp upstream of the translation initiation codon.
tion clones were constructed by cloning various restriction fragments produced by polymerase chain reaction into a reporter vector, pGL3-Basic. Because these fragments were cloned upstream to a luciferase reporter gene, transcriptional activities of the promoter-luciferase cartridge could be studied by transfecting various clones to the human cell line PANC-1 (Fig. 5). It was found that the essential region for promoting transcription is very close to the start site of translation. By comparing luciferase activities between constructs ERES2 to ERES7 and that of constructs ERES7 with ERES-2H and -2P, the essential promoter region could be mapped to a region spanning the translation start site to +2364 bp. This observation is supported by the fact that no significant difference in the luciferase activities was observed upon deletion of the 5′-9 sequence from +2364 to +21319 (ERES7). However, in the two constructs in which the 3′ region was deleted, the luciferase activities were completely abolished. Because no 5′ deletion studies were carried out beyond +2364, we cannot conclude and finely map the essential promoter elements in the +21 to +2365 region upstream of the translation start site. Experimental data also showed that inverting the 1.3-kb 5′ flanking region (ERES1) greatly decreased the transcription activities.

Deletion Analysis on the Other Human Cell Lines—Three constructs, ERES2, ERES7, and ERES-2P were transfected to the other human cell lines to test the cell specificity of the 

---

**TABLE III**

Exon-intron boundaries of the human C1qBP gene

| Exon | Location on cDNA | Exon size | Exon 5′ sequence | Exon 3′ sequence | Intron | Intron 5′ sequence | Intron 3′ sequence | Intron size | Amino acid interrupted |
|------|------------------|-----------|------------------|------------------|--------|--------------------|--------------------|------------|-----------------------|
| 1    | 1–232            | 232       | CACACCGACG       | CAGCAGGAGAAG     | 1      | GT GAGGTTTA         | AGAGAGGAGGAGG     | 568        | Gly-78                |
| 2    | 233–383          | 151       | GAGACAAAGCC      | CAGGGGAGAAGG     | 2      | GT AGATACCTG        | AGAGAGGAGGAGG     | 3156       | Lys-128               |
| 3    | 384–477          | 94        | AATCACCGGTC      | AGAGACGGAGAAG     | 3      | GT AAGCCTAAT        | GGTGGGAGGAGG     | 1104       | Nil                   |
| 4    | 478–576          | 99        | CCGAGACCTG       | AGAGACGGAGAAG     | 4      | GT ACCAGAGGA        | TCTTCTTCTCAG     | 253        | Nil                   |
| 5    | 577–689          | 123       | CTTGGACAGAG      | AGAGACGGAGAAG     | 5      | GT AGATCGTCTT      | TCCTCCTCAGG     | 128        | Nil                   |
| 6    | 700–845          | 516       | GCTCTATATG       | AGAGACGGAGAAG     |        |                    |                    |            |                       |

---

**FIG. 3.** Analysis of the 5′-flanking region of the human C1qBP gene. The first nucleotide of the translation start codon ATG (bolded) is designated as +1 and the preceding nucleotide as –1. All potential transcription factor binding sites, promoters, and elements that may be responsible for transcription control are underlined and labeled below the line.
The Human gC1q-R/p32 Gene (C1qBP)
The Human gC1q-R/p32 Gene (C1qBP)

FIG. 6. Promoter assays in different human cell lines. The relative promoter activities of all cell lines were normalized by the pGL3-Basic activity. The figure represents the mean of two individual transfection experiments.

gC1qBP gene. With regard to the number of various transcription control elements in the 5' flanking region, it may be possible that different transcription control elements are essential for the gene expression in different human cells.

The three-dimensional structure of gC1qR was revealed by x-ray crystallography. The mature protein molecule has one N-terminal α-helix followed by seven consecutive antiparallel β-strands and two C-terminal α-helices (29). Three molecules form a doughnut-shaped quaternary structure with an internal channel 10 Å in diameter. By aligning the exon boundaries and the three-dimensional structure, it is observed that: exon 1 encodes the first 77 amino acid residues containing the mitochondria-targeting sequence (14); exon 2 encodes the N-terminal α-helix and the first two β-strands; exon 3 encodes β-strand 3; exon 4 encodes β-strands 4 and 5; exon 5 encodes β-strands 6 and 7; and the last exon encodes for the C-terminal α-helices.

As gC1qR was shown to be a multiligand-binding protein, it would be interesting to map the binding sites on gC1qR to various ligands.

More recently, gC1qR was found to function as a receptor for the internalin B (InlB) invasion protein of Listeria monocytogenes (30) and to bind to protein A of Staphylococcus aureus (31). Another report showed that the hepatitis virus C core antigen can interact with gC1qR, and the anti-proliferation effects of hepatitis virus C core antigen on activated T cells can be reversed by anti-gC1qR monoclonal antibodies (16). All of these reports further suggest that gC1qR is a biologically important, widely distributed, multiligand-binding and multifunctional protein (1).

REFERENCES
1. Ghebrehiwet, B., and Peerschke, E. I. (1998) Immunobiology 199, 225–238
2. Ghebrehiwet, B., Lim, B. L., Peerschke, E. I., Willis, A. C., and Reid, K. B.

FIG. 7. Gel-shift assays. 8 μg of PANC-1 nuclear extracts were used for each individual reaction. A, gel shift was observed only in the L6 DNA probe, corresponding to the sequence –96 to –76 of the human C1qBP gene. B, a competition assay was performed using 8 μg of the PANC-1 nuclear extracts in the presence of an increasing concentration of unlabeled DNA probe (0-, 10-, and 100-fold). The formation of DNA-protein complexes was gradually inhibited only by the specific oligo but not by an unrelated oligo. In the negative control experiments (−ve) using an unrelated oligo, only the formation of DNA-protein complex 1 was inhibited, but the formation of the other three complexes was not disturbed; this showed that DNA-protein complex 1 was an artifact.

Fig. 8. Supershift assay was performed using the L6 oligo as a probe. When 4 μg of anti-Sp1 antibodies were added to the mixture, the band of DNA-protein complex 4 shifted upward from its original position (as shown when no antibody was added), whereas the other three bands were unaffected. DNA-protein complex 4 should be a complex of oligo L6 and the transcription factor Sp1. Bovine serum albumin was used as the negative control in this assay, which was unable to display DNA-protein complex 4.
17075

The Human gC1q-R/p32 Gene (C1qBP)

3. Joseph, K., Ghebrehiwet, B., and Kaplan, A. P. (1999) *Clin. Immunol.* **92**, 246–255
4. Joseph, K., Ghebrehiwet, B., Peerschke, E. I., Reid, K. B., and Kaplan, A. P. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 8552–8557
5. Lu, P. D., Galarakis, D. K., Ghebrehiwet, B., and Peerschke, E. I. (1999) *Clin. Immunol.* **90**, 360–367
6. Lim, B. L., Reid, K. B., Ghebrehiwet, B., Peerschke, E. I., Leigh, L. A., and Preisser, K. T. (1996) *J. Biol. Chem.* **271**, 26739–26744
7. Xu, Z., Hirasawa, A., Shinoura, H., and Tsujimoto, G. (1999) *J. Biol. Chem.* **274**, 21149–21154
8. Storz, P., Hauser, A., Link, G., Dedio, J., Ghebrehiwet, B., Pinzenmaier, K., and Johannsen, F. (2000) *J. Biol. Chem.* **275**, 24601–24607
9. Yu, L., Loewenstein, P. M., Zhang, Z., and Green, M. (1995) *J. Virol.* **69**, 3017–3023
10. Luo, Y., Yu, H., and Peterlin, B. M. (1994) *J. Virol.* **68**, 3850–3856
11. Mathews, D. A., and Russell, W. C. (1998) *J. Gen. Virol.* **79**, 1677–1685
12. Wang, Y., Finnan, J. E., Middeldorp, J. M., and Hayward, S. D. (1997) *Virology* **236**, 18–29
13. Bruni, R., and Roizman, B. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 10423–10427
14. Dedio, J., Jahnen-Dechent, W., Bachmann, M., and Muller-Esterl, W. (1998) *J. Immunol.* **160**, 3534–3542
15. Soltys, B. J., and Gupta, R. S. (1999) *Trends Biochem. Sci.* **24**, 174–177
16. Kittlesen, D. J., Chianese-Bullock, K. A., Yao, Z. Q., Braciale, T. J., and Hahn, Y. S. (2000) *J. Clin. Invest.* **106**, 1239–1249
17. Guo, N., Weremowicz, S., Lynch, N., Lim, B. L., Schwaeble, W., Peerschke, E. I., Morton, C. C., Reid, K. B., Ghebrehiwet, B., and Sastry, K. N. (1997) *Cytogenet. Cell Genet.* **77**, 283–284
18. Lynch, N. J., Reid, K. B., van-den-Berg, R. H., Daha, M. R., Leigh, L. A., Ghebrehiwet, B., Lim, W. B., and Schwaeble, W. J. (1997) *FEBS Lett.* **418**, 111–114
19. Quandt, K., Frech, K., Karas, H., Wingender, E., and Werner, T. (1995) *Nucleic Acids Res.* **23**, 4878–4884
20. Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A. E., Kel, O. V., Ignatieva, E. V., Ananko, E. A., Podkolodnaya, O. A., Kolpakov, F. A., Podkolodny, N. L., and Kelchanov, N. A. (1998) *Nucleic Acids Res.* **26**, 364–370
21. Lee, K. A. W., Binderleif, A., and Green, M. R. (1988) *Gene Anal. Tech.* **5**, 22–31
22. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1987) *Current Protocols in Molecular Biology*, Wiley, New York
23. Hugnon, I. C., Davies, A., Gove, C., Moscoso, G., Moniz, C., Fossa, Y., Farzanbeh, F., and Towner, P. (1997) *Biochim. Biophys. Acta* **1353**, 98–102
24. Kiang, J. G., Gist, I. D., and Tsokos, G. C. (2000) *Mol. Cell Biochem.* **204**, 169–178
25. Ranger, A. M., Gerstenfeld, L. C., Wang, J., Kon, T., Bae, H., Gravallese, E. M., Glimcher, M. J., and Glimcher, L. H. (2000) *J. Exp. Med.* **191**, 9–22
26. Lutterbach, B., and Hiebert, S. W. (2000) *Gene* **245**, 223–235
27. Macias, D., Perez, P. J., Garcia, G. L., Carmona, R., and Munoz, C. R. (1998) *Anat. Embryol.* **198**, 307–315
28. Lim, B. L., White, R. A., Hummel, G. S., Schwaeble, W., Lynch, N. J., Peerschke, E. I., Reid, K. B., and Ghebrehiwet, B. (1998) *Gene* **229**, 229–237
29. Jiang, J., Zhang, Y., Krainer, A. R., and Xu, R. M. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 3572–3577
30. Braun, L., Ghebrehiwet, B., and Cossart, P. (2000) *EMBO J.* **19**, 1458–1466
31. Nguyen, T., Ghebrehiwet, B., and Peerschke, E. I. B. (2000) *Infect. Immun.* **68**, 2061–2068
The Human gC1qR/p32 Gene, C1qBP: GENOMIC ORGANIZATION AND PROMOTER ANALYSIS
Angela J. Tye, Berhane Ghebrehiwet, Ning Guo, Kedarnath N. Sastry, Billy K. C. Chow, Ellinor I. B. Peerschke and Boon-Leong Lim

J. Biol. Chem. 2001, 276:17069-17075.
doi: 10.1074/jbc.M009064200 originally published online March 8, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M009064200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 30 references, 12 of which can be accessed free at http://www.jbc.org/content/276/20/17069.full.html#ref-list-1