Human Eosinophil Charcot-Leyden Crystal Protein: Cloning and Characterization of a Lysophospholipase Gene Promoter

By Hilary I. Gomolin, Yuji Yamaguchi, Angela V. Paulpillai, Laura A. Dvorak, Steven J. Ackerman, and Daniel G. Tenen

The Charcot-Leyden crystal (CLC) protein is a lysophospholipase expressed exclusively by eosinophils and basophils. During eosinophilic differentiation of eosinophil-committed cell lines, CLC steady state mRNA levels increase significantly. This increased expression is transcriptionally regulated during butyrate induction of an eosinophilic subtype (C15) of the promyelocytic leukemia cell line HL-60, as shown by nuclear run-on assays. The transcriptional start site of the CLC gene was identified 43 bp upstream of the 5' end of the longest available cDNA sequence. The gene encoding CLC protein was cloned from a chromosome 19-specific library and a fragment overlapping the transcriptional start site was isolated and sequenced. Plasmid constructs (in the pXP2 luciferase expression vector) containing 411 and 292 bp of genomic sequence upstream of the CLC transcriptional start site directed reporter gene expression in transient transfections of HL-60-C15 cells, as well as other myeloid (U937) and nonmyeloid (HeLa and RPMI 8402) cell lines. However, the differential expression of the two CLC promoter constructs in these cell lines suggests that the −292 to −411 bp region of the promoter may confer some specificity for expression in the eosinophil lineage. The CLC promoter sequence contains two consensus GATA binding sites, a purine-rich sequence that presents potential binding sites for PU.1, a member of the ets family of genes, as well as sequences described in other myeloid-specific promoters. This is the first demonstration of a functional eosinophil promoter that could serve as a model for identifying DNA elements and trans-activating factors that regulate gene expression during the commitment and differentiation of the eosinophil lineage. © 1993 by The American Society of Hematology.

THE CHARCOT-LEYDEN crystal (CLC) protein is a unique constituent of human eosinophils and basophils and identification of the characteristic hexagonal bi-pyramidal crystals formed by this protein is a hallmark of eosinophil-associated inflammation.1,2 Although CLC protein possesses lysophospholipase activity3 (lysolecithin acylhydrolase E.C.3.1.1.9), its role in eosinophil or basophil function and associated inflammatory responses remains obscure.4 CLC protein is one of the major constituents of the eosinophil, comprising an estimated 7% to 10% of total cellular protein.4

The recently isolated cDNA for human CLC protein4 encodes an amino acid sequence that shows no similarities to available sequences of prokaryotic or eukaryotic lysophospholipases. In contrast, there is a 25% to 30% sequence homology over a 100 to 120 amino acid region of the carboxyl-terminal domains of four IgE-binding proteins.5-7 One of these, Mac-2,8 is the major integrin laminin-binding protein of macrophages. These proteins are members of a superfamily of β-galactoside binding S-type animal lectins.9 Whether CLC protein is bifunctional and also exhibits lectin carbohydrate-binding activity remains to be determined.

CLC mRNA and protein is expressed exclusively by eosinophils and basophils, and its expression increases during butyrate or interleukin-5 (IL-5) induction in two eosinophil-committed sublines of HL-60, clones C15 and 3+C-5, respectively.4,10 To characterize the cis-acting elements that control CLC gene expression during eosinophil differentiation, we sought to clone and analyze the CLC promoter.

In this report, we show that CLC expression is transcriptionally upregulated in HL-60-C15 cells during butyrate induction. We have cloned the CLC gene, identified and sequenced its promoter, and shown that the CLC promoter is functional after transfection into HL-60-C15 cells. Consensus sequences that correspond to GATA and PU.1 binding sites are present in the 411 bp upstream of the transcriptional start site. Study of the CLC promoter will hopefully elucidate unique transcriptional features of the differentiation and maturation of the eosinophil lineage.

MATERIALS AND METHODS

Cell culture. An eosinophil-committed subline, HL-60-C15 (a gift of Dr S. Fischkoff, University of Pennsylvania School of Medicine, Oncology Division, Philadelphia, PA) derived from the HL-60 promyelocytic leukemia line was maintained in RPMI 1640 supplemented with 10% fetal bovine serum (FBS; Hyclone, Logan, UT) and passed twice weekly in an atmosphere of 5% CO2. Cells (2 × 10^6/mL) were induced with n-butyrate (0.5 mmol/L; Sigma, St Louis) for 48 hours, as previously described.11

Northern blot analysis. Total RNA was prepared from induced and noninduced HL-60-C15 cells by the guanidium isothiocyanate method.12 RNA (15 to 20 μg/lane) was denatured in formamide/formaldehyde followed by electrophoresis in 1% agarose/formaldehyde gels. RNA was then transferred to Hybond-N nylon membranes (Amersham, Arlington Heights, IL), and the blots probed with a 598-bp cDNA encoding CLC protein4 and a 285 RNA oligonucleotide probe (5'-AACGATCAGAGTAGTGG-3') for internal control of RNA loading. Hybridizations with the random-primed cDNA probe for CLC were performed at 43°C in 50% formamide,
labeled full-length CLC cDNA and a clone with a 20-kb insert was isolated. A 1.5-kb EcoRI fragment of the phage clone, overlapping the 5' end of the CLC cDNA, was identified by Southern blotting, subcloned into PUC18 (Promega, Madison, WI), and sequenced on both strands using the dyeodeoxy method with Sequenase (US Biochemical, Cleveland, OH). Oligonucleotide primers for sequencing were synthesized by phosphoramidite chemistry on an Applied Biosystems Model 381A DNA Synthesizer (Applied Biosystems, Foster City, CA). Regulatory consensus sequences within the CLC promoter were identified by searching the transcription factors database.15

**RNase protection assay.** A DNA fragment encompassing bp −292 to +77 of the CLC gene was subcloned using polymerase chain reaction (PCR) into HindIII/Bgl II cut pGEM11Zf (Promega) and linearized with HindIII (all PCR constructions were sequenced). A [32P]CTP-labeled antisense transcript was prepared with T7 Polymerase (Promega) as described, with the following changes: 1 μg of template DNA was used, and 50 μCi of [α-32P]CTP and unlabeled CTP were used to a final concentration of 12 μmol/L. In vitro transcript (5 × 105 cpm) was heated to 85°C for 5 minutes with 10 μg of yeast tRNA, U937 mRNA, or HL-60−C15 mRNA (from cells induced for 2 days with 0.5 mM butyrate) and annealed overnight at 45°C. Samples were digested with RNase A and RNase T1 as described and separated on a 6% polyacrylamide, 8.3 mol/L urea gel. M13mp18 sequence obtained with a −40 primer (5'-GTTTTCCCAGTCACGAC-3'; US Biochemical) was used as a size marker. Autoradiographs were performed with Kodak XAR-5 film at −80°C for 12 hours.

**Transfections.** The CLC (−292)/luciferase (luc) construct contained genomic material from a PCR-generated DNA fragment from bp −292 to bp +77, at the start ATG of CLC cDNA, inserted into the HindIII and Bgl II sites of pXP2. A larger construct, CLC (−411)/luc, was likewise inserted into the pXP2 vector with an additional 119 bp of upstream DNA. All PCR-generated constructs in pXP2 were sequenced. Transfections were performed as described, with minor modifications. Logarithmically growing HL-60−C15 cells (1 to 2 × 106) were washed once and resuspended in RPMI 1640 without biotin at 2.8 × 106 cells/mL. The cells were incubated with 20 μg cesium chloride-purified DNA and 2 μg CMV-GH (a human growth hormone expression construct, kindly provided by Dr L. Zon, Division of Hematology, Children’s Hospital, Boston, MA) for 15 minutes at room temperature in BioRad 0.4-cm gap electroporation cuvettes (BioRad, Richmond, CA). The CMV-GH cotransfection provided for standardization among different cell lines and separate transfection experiments. Cells were then electroporated at 280 V, 960 μF (BioRad Gene Pulser; BioRad), incubated for 15 minutes on ice, and transferred to 10 mL prewarmed RPMI 1640/10% FBS. The HeLa, U937, and RPMI 8402 cell lines were electroporated at 960 μF and 150, 300, or 270 V, respectively, conditions previously optimized for these lines. Four to 6 hours after electroporation, cell extracts were prepared in 300 μL 1% Triton X-100, 25 mM/L Gly-gly, 15 mM/L MgSO4, 4 mM/L EGTA, 1 mM/L dithiothreitol (DTT), and 100 μL of the extract (equivalent to ~5 × 104 cells) was analyzed for luciferase activity in a monolight 2010 Luminometer (Analytical Luminence Laboratory, San Diego, CA). Culture supernatants were analyzed for growth hormone levels by radioimmunoassay (Nichols Institute Diagnostics, San Juan Capistrano, CA) as described by the manufacturer. Relative light units of luminescence are reported per nanogram per milliliter of growth hormone.

**RESULTS**

CLC expression is transcriptionally upregulated during butyrate-induced eosinophilic differentiation of HL-60−
Fig 2. CLC protein gene promoter sequence. Consensus sequences are boxed and shaded for two GATA-binding sites, two PU.1 sites, the CCCCACCC element found in other myeloid promoters, and a putative TATA box. The transcriptional start site is designated by a forward pointing arrow and lower case letters represent the overlap with the published CLC cDNA sequence. The position of a splice donor exon is indicated starting at bp +29. The first exon is underlined. The ATG marking the translational start site, and the 5 amino acids encoded in the first exon are indicated underneath the DNA sequence. The vertical arrows at bp +292 and +77 mark the boundaries of the genomic sequence used for RNAse protection, as well as in the −292/luc construct used in transfection studies. The cDNA and genomic sequences have been deposited in the Genbank data base (accession nos. L01664 and L01665, respectively).

Identification of the CLC transcriptional start site. If transcription of the CLC gene began at the 5' end of the known cDNA sequence, a 34-bp band should be protected in an RNAse protection assay using the probe depicted in Fig 3A. However, a major band comigrating with an 85-bp DNA fragment was protected in HL-60-C15 mRNA, but not U937 mRNA or yeast tRNA control samples (Fig 3B). In these gel systems, the correct size of the RNA species is 80 bp in length.21 A consensus CAP site,22 CANYYY, was identified 43 bp upstream of the 5' end of the previously isolated CLC cDNA, and a consensus TATAA sequence,23 AATAA, 31 bp upstream of the CAP site. The adenosine designated +1 at the transcriptional start site (Fig 2) corresponds to a protected RNA fragment of 77 bp. Additional minor protected species were also observed that extended up to 10 bp 5' of the major transcriptional start site (Fig 3), suggesting additional starts; the presence of several weaker start sites may reflect the lack of a perfect TATA box in the promoter as shown in Fig 2.

Sequence analysis. The 411 bp of sequence upstream from the transcriptional start site was evaluated for potential regulatory consensus sequences (Fig 2). As in promoters of other leukocyte-specific genes such as TdT,24 PIM-1,25 CD11b,14 and CD18,26 no CCAAT box was identified. However, sequences were present in the upstream 411 bp that correspond to binding sites of a number of known transcription factors (Fig 2). Two GATA binding sites27 are present at −11 (on the sense strand) and −207 (on the antisense strand); GATA-binding proteins have previously been shown to be important in the regulation of hematopoietic

---

The gene encoding the CLC protein has been localized to chromosome 19 by somatic cell hybrid analysis20 and more recently sublocalized to 19q13.1 by fluorescence in situ hybridization using cosmid clones (B. Trask and H. Mohrenweiser, unpublished observations). A 20-kb genomic clone encoding CLC protein was isolated by screening a human chromosome 19 Charon 40 phage library with the CLC cDNA. A 1.5-kb EcoRI fragment that hybridized to an oligonucleotide constructed from the 5'-most end of the longest sequenced cDNA4 was isolated. As shown in Fig 2, a portion of the EcoRI fragment was identical to the first 48 bp of the CLC cDNA sequence and contained 454 bp of additional 5' sequence.

Cloning of the CLC gene. The gene encoding the CLC protein has been isolated to chromosome 19 by a yeast artificial chromosome library and more recently sublocalized to 19q13.1 by fluorescence in situ hybridization using cosmid clones (B. Trask and H. Mohrenweiser, unpublished observations). A 20-kb genomic clone encoding CLC protein was isolated by screening a human chromosome 19 Charon 40 phage library with the CLC cDNA. A 1.5-kb EcoRI fragment that hybridized to an oligonucleotide constructed from the 5'-most end of the longest sequenced cDNA was isolated. As shown in Fig 2, a portion of the EcoRI fragment was identical to the first 48 bp of the CLC cDNA sequence and contained 454 bp of additional 5' sequence.
specific transcription and recently were shown to be expressed in eosinophils and basophils. A purine-rich element is present between bp –180 and –175 on the antisense strand and bp –65 and –60 on the sense strand that is identical to the binding site of the myeloid- and B-cell-specific, ets-related transcriptional activator PU.1. In addition, there is a sequence, CCCCACCC (between bp –86 and –79), which has been identified in other developmentally expressed myeloid-specific genes, such as myeloperoxidase, cathepsin G, human neutrophil elastase, and eosinophil peroxidase (Yamaguchi, Tenen, and Ackerman, manuscript in preparation). The genomic sequence diverges from that of the cDNA 92 bp downstream of the transcriptional start site; the presence of a perfect splice donor consensus sequence at this point strongly suggests that this represents the start of the first intron (Fig 2).

**Transfection analyses.** Two constructs, containing 411 and 292 bp of genomic DNA upstream of the transcriptional start site and extending 3' just proximal to the start ATG of the CLC cDNA, were linked to the luciferase reporter gene in pXP2. These constructs directed reporter gene activity an average of 70-fold and 30-fold above background (promoterless luciferase vector), respectively, when transfected into the eosinophil-committed HL-60–C15 cells (Fig 4). In three separate experiments, the larger construct (–411/luc) stimulated twofold to threefold greater expression of the luciferase reporter than the smaller construct (–292/luc). Thus, the sequences that direct the basal expression of the CLC gene are contained within at least the first 292 bp upstream of the transcriptional start site. To assess the presence of eosinophil-lineage–specific regulatory elements in the CLC promoter, both the –411/luc and –292/luc CLC promoter constructs were transfected into other myeloid (U937, monocytic) and nonmyeloid (HeLa, cervical carcinoma; RPMI8402, T-lymphoblastic) cell lines (Fig 5); as shown above, deletion of bp –292 to –411 of the CLC promoter resulted in decreased expression in the eosinophil-committed HL-60–C15 line. In contrast, this deletion resulted in approximately twofold to threefold increased expression of luciferase activity in all the other cell lines tested, suggesting that this region of the promoter may confer some cell specificity to the expression of the CLC gene in the eosinophil-committed HL-60–C15 line.

Because we have shown that CLC expression is in part transcriptionally upregulated by butyrate induction of the HL-60–C15 cells, transfections were performed with these constructs in the induced cells as well; no consistent increase in activity was obtained with butyrate induction (data not shown), suggesting that the elements necessary for butyrate responsiveness are not contained within the first 411 bp upstream of the transcriptional start site.

**DISCUSSION**

Isolation of the CLC gene provides an opportunity to elucidate the cis promoter elements and trans-acting factors that regulate tissue- and differentiation-specific transcription in myeloid cells in general and eosinophils in particular. This effort represents the first report of a functionally active eosinophil-associated gene promoter.

The CLC protein, a lysophospholipase, is uniquely expressed in both eosinophils and basophils in comparable amounts. We have previously performed Northern blotting of RNA from several sources to identify the CLC transcript; these analyses showed a 0.9-kb mRNA species in...
various preparations of peripheral blood eosinophils and basophils, an eosinophilic subline of HL-60 (3+C−5) induced towards eosinophilic differentiation with BCGF-II (as a source of IL-5), as well as the parental HL-60 line induced with dimethylsulfoxide. In contrast, the CLC mRNA was absent in peripheral blood neutrophils, monocytes, T-cell and B-cell lines, or HL-60 cells induced towards monocytic differentiation with vitamin D₃. Subclones of the HL-60 promyelocytic cell line have recently served as models for eosinophil and basophil differentiation. For example, we have shown that BCGF-II treatment of the eosinophilic HL-60 subline 3+C−5 increased the expression of mRNA for the eosinophil granule cationic proteins, eosinophil cationic protein (ECP), and eosinophil-derived neurotoxin (EDN), in addition to the CLC protein. By comparison, treatment of the HL-60−C15 subline with n-butyrate similarly induced the increased expression of mRNA for all the eosinophil-associated granule proteins concomitantly with a pronounced maturation of these cells, as evidenced by the development of large numbers of cytoplasmic granules. In this regard, the present analyses show that CLC gene expression is transcriptionally upregulated during butyrate-induced differentiation of the eosinophilic HL-60−C15 cell line. However, no increase in luciferase activity was observed using any of these constructs in butyrate-treated HL-60−C15 cells, indicating that the element(s) necessary for butyrate responsiveness is not contained within the first 411 bp upstream of the transcriptional start site. Responsiveness of other genes to butyrate has been demonstrated to be mediated by scaffold/matrix-attached regions (SAR/MAR elements), suggesting that butyrate exerts its effects at the level of chromatin structure. These elements generally have been isolated as large fragments localized several kilobases upstream of the transcriptional start site; further experiments will be aimed at attempting to localize such elements using larger CLC promoter constructs.

Experiments showing the expression of both the −411/−292 CLC promoter constructs in other myeloid (U937) and nonmyeloid (HeLa and RPMI 8402) cell lines suggest the lack of eosinophil-lineage-specific regulatory elements in the CLC promoter sequence analyzed thus far. However, deletion of bp −292 to −411 of the CLC promoter resulted in decreased expression in the eosinophilic HL-60−C15 subline, but increased expression in all the other lines tested. This observation suggests the presence of
a cell-specific regulatory element(s) in this region that may bind factors uniquely expressed in HL-60–C15 cells and/or a possible repressor-binding element that might inhibit CLC gene expression in other myeloid or nonmyeloid lineages.

We have identified the CLC transcriptional start site by RNAse protection. The predominant transcription initiation site lies 43 bp upstream of the 5' end of the previously characterized cDNA sequence, 31 bp downstream from the sequence TATAA, which could serve as a TATAA box.23 However, the presence of additional weaker start sites extending up to 10 bp upstream of the principal start may reflect the lack of a perfect TATAA box. Like the promoters of other leukocyte-specific genes such as TdT,24 PIM-1,25 CD11b,18 and CD18,26 the CLC promoter also lacks a potential binding site for PU.1.

Interestingly, the sequences between -180 and -175 on the antisense strand and between -65 and -60 on the sense strand contain the purine-rich sequence that represents a potential binding site for PU.1,31 a member of the ets superfamily of genes previously described as being expressed exclusively in monocytes and B cells. We have recently shown that a PU.1 binding site is important for myeloid-specific expression of the CD11b gene,55 which is also expressed by eosinophils.56 Finally, an additional sequence (CCCCA-CCC), found in other myeloid-specific gene promoters such as myeloperoxidase,32 cathepsin G,33 human neutrophil elastase,34 and eosinophil peroxidase,35 was identified between bp -86 and -79. The functional significance of these putative regulatory elements and the role of their cognate DNA binding factors in eosinophil development awaits detailed deletion and mutagenesis studies of the CLC and other eosinophil-specific genes.

ACKNOWLEDGMENT

The authors thank Timothy Bum, Heike Pahl, Helene Rosenberg, and Anne Satterthwaite for helpful discussions; Edward Albee for excellent technical assistance; and Mary Singleton for manuscript preparation.

REFERENCES

5. Albrandt K, Orida NK, Liu FT: An IgE-binding protein with a distinctive repetitive sequence and homology with an IgE receptor. Proc Natl Acad Sci USA 84:6859, 1987
30. Yamaguchi Y, Albee E, Steinberg HN, Nagata S, Zon LI, Tenen DG, Ackerman SJ: Expression of eosinophil granule protein genes and GATA transcription factors during IL-5 induced eosinophil differentiation. (manuscript in preparation)
Human eosinophil Charcot-Leyden crystal protein: cloning and characterization of a lysophospholipase gene promoter

HI Gomolin, Y Yamaguchi, AV Paulpillai, LA Dvorak, SJ Ackerman and DG Tenen