Lactoferrin Gene Promoter: Structural Integrity and Nonexpression in HL60 Cells

By Jennifer J. Johnston, Peter Rintels, Josan Chung, John Sather, Edward J. Benz, Jr, and Nancy Berliner

Lactoferrin is a member of the transferrin family of iron-binding proteins. It is found in several glandular epithelial tissues and human neutrophils, where it is localized to secondary granules. To examine the mechanisms controlling lactoferrin gene expression in neutrophils and defects in its expression in acute leukemia, we have cloned a lactoferrin cDNA from a chronic myelogenous leukemia library, and used it to obtain genomic clones representing the chromosomal lactoferrin gene. Using polymerase chain reaction, primer extension, and S1 analysis, we have identified the 5' end of the lactoferrin mRNA. We have defined a putative promoter region for the gene, and characterized its first two exons. In addition, we have examined the structure of these regions in DNA from HL60 cells. HL60 is a leukemic cell line that undergoes phenotypic neutrophil maturation on exposure to dimethyl sulfoxide (DMSO). However, the cells cannot be induced to express any secondary granule protein genes. We have shown that the 5' end of the lactoferrin gene, including the putative promoter region, is entirely normal in HL60. By Northern analysis, nuclear run-on studies, and primer extension assays we have shown that the gene is not transcribed in DMSO-induced HL60 cells. This supports the hypothesis that the defect in HL60 is an abnormality in the production or activity of a transacting regulator of lactoferrin gene expression.

© 1992 by The American Society of Hematology.

ACTOFERRIN (LF) is an approximately 80,000-Kd protein of known amino acid sequence1 found in several glandular epithelial tissues2 and human neutrophils, where it is localized to secondary granules.2,3 It is a member of the family of iron-binding proteins that includes transferrin, with which it shares almost 60% homology at the amino acid level.1 Although its exact biologic role remains uncertain, LF has been implicated in host microbial defense both directly4-9 and indirectly through modulation of neutrophil function.6,8 A role in the feedback regulation of neutrophil mass has also been suggested.7,10 Recently, multiple isoforms of lactoferrin have been identified in both human milk and neutrophils that vary in their iron-binding and RNase activity, most likely because of differences in posttranslational glycosylation.11

The LF gene has been localized to human chromosome 3, band 3q21-q23,12 a region that also contains the human transferrin gene family cluster.13 Current evidence suggests that this family of genes evolved first through an intragenic duplication,14 followed by further duplication events13 that have resulted in a family of genes encoding proteins with divergent tissue expression and functions.

Neutrophil LF expression has been examined in detail. Previous studies performed by this laboratory demonstrated that LF mRNA expression and protein synthesis begin at the myelocyte stage, in concert with the appearance of secondary granules.15 These studies were further confirmed by in situ hybridization16 using an LF partial cDNA clone.17 Thus, LF appears to be a useful marker of terminal differentiation in myeloid cells. Therefore, not surprisingly, conditions characterized by disordered myeloid differentiation, ie, acute leukemia and the myelodysplastic syndromes, are also notable for disordered or absent LF gene expression, despite an otherwise morphologically mature appearance of the neutrophils.18,19

Thus, the details of LF genetic control regions should provide important insights into the processes governing tissue and stage-specific gene expression, as well as yielding important information about terminal granulocyte maturation. A partial LF cDNA clone has previously been isolated in our laboratory.17 A full-length cDNA from mammary tissue, along with partial genomic 5' data, has recently been reported.20,21 We have derived an LF cDNA from a human chronic myelogenous leukemia (CML) library. This has been used to identify genomic clones spanning the full length of the LF gene. We report the characterization of the promoter region and potential cis-acting regulatory elements of the gene. We have mapped the first two exons and shown that they are homologous to their transferrin counterparts.

We have used this information to examine the LF gene in HL60 and confirmed that it is structurally intact, and that the sequence of the 5' region of the gene is completely normal. This supports the hypothesis that the failure of the expression of the gene in that line is not a result of a structural defect at the LF locus. We have also shown that the LF gene is not transcribed in HL60. These data strongly support the hypothesis that the inability to express LF in that cell line reflects the inability to induce mRNA transcription, perhaps as a result of the absence or dysfunction of a necessary transacting factor.

MATERIALS AND METHODS

Cloning of the LF cDNA. The construction and screening of the CML cDNA library has been previously reported.22 This library was used to generate a probe for the 5' portion of the LF cDNA. A sense oligonucleotide derived from published data and an antisense oligonucleotide based on sequence from the partial LF cDNA were used in the polymerase chain reaction (PCR) to generate a DNA product corresponding to the remaining coding sequence of the LF cDNA.23 The product was purified on an

From the Section of Hematology, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT.

Supported in part by grants from the National Institutes of Health (P.R., E.J.B., N.B.), and by a Swebilius Research grant. N.B. is also supported by a Junior Faculty Research Award from the American Cancer Society.

Address reprint requests to Nancy Berliner, MD, Department of Internal Medicine, Section of Hematology WWW423a, Yale University School of Medicine, 333 Cedar St, New Haven, CT 06510.

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

© 1992 by The American Society of Hematology.

0006-4971/92/7911-0001$3.00/0

agrose gel, sequenced to confirm its identity, and used to screen the CML cDNA library. A 2.4-kb full-length LF cDNA was obtained. The cDNA was fully sequenced using the dye deoxy chain termination method.24

Cloning of genomic LF. cDNA probes corresponding to the 3' and 5' portions of the LF cDNA were used to screen a Charon 4A human genomic library. Phages were plated at 10^5 plaque-forming units (pfu)/plate. Duplicate nitrocellulose filters were lifted and hybridized at 42°C overnight to nick-translated cDNA probes. Positive plaques were purified by secondary and tertiary screening. Genomic fragments were subcloned into plasmid vectors and analyzed by restriction enzyme mapping and DNA sequencing.

Southern blot analysis. High molecular weight DNA was prepared from CML, HL60, and normal control cells by previously described methods.25 DNA was digested with appropriate restriction enzymes, size-fractionated by agarose gel electrophoresis, and transferred to nitrocellulose filters by the Southern method.6 An EcoRI probe was prepared from the LF cDNA, which contained the full coding sequence but lacked the 3' untranslated region. Filters were hybridized overnight to a 32P-labeled, nick-translated probe of the LF gene, and washed at 55°C in 0.1% sodium dodecyl sulfate (SDS) and 15 mmol/L NaCl/1.5 mmol/L Na citrate (SSC) before autoradiography.

Primer extension. Simultaneous reactions were performed using two antisense oligonucleotides corresponding to sequences in the 5' end of the cDNA, and separated by approximately 50 bp. Oligonucleotides were labeled with adenosine triphosphate (γ-ATP) using polynucleotide kinase. Labeled oligonucleotides were heated to 100°C for 5 minutes and cooled on ice. Annealing was performed in 50 mmol/L Tris, 6 mmol/L MgCl2, 40 mmol/L KCl, and 10 mmol/L dithiothreitol at 65°C for 15 minutes, followed by slow cooling to room temperature over 1 to 2 hours. dNTPs were added to 2 mmol/L, and primer extension was performed with AMV reverse transcriptase for 10 minutes at room temperature, 10 minutes at 37°C, and 60 minutes at 42°C. Reaction was stopped with 1 μL 0.5 mol/L EDTA, and the reaction products were EtOH precipitated. The pellet was resuspended in 10 μL TE, with 1 ng/mL RNAase, and digested at 37°C for 5 minutes. Loading dye was added, consisting of 95% formamide and 20 mM EDTA; samples were heated to 95°C, and analyzed on 8% polyacrylamide gels.

SI analysis. SI analysis was performed using modifications of the method of Berk and Sharp.26 A DNA probe was prepared from a genomic subclone containing the first exon of the LF gene. To prepare the probe, a PCR reaction was performed using a sense primer to the polylinker region of the genomic subclone and the more 5' of the two antisense oligonucleotides used in the primer extension. The PCR product was digested with BamHI to yield a product of approximately 250 bp. The probe was treated with calf intestinal phosphatase and labeled with γ-ATP using polynucleotide kinase. Counts of end-labeled probe, 5x10^6 cpm of labeled RNA, was electrophoresed through a 6% polyacrylamide gel at 4°C at 100 V for approximately 6 hours and then exposed to film overnight at −70°C.

Northern blot analysis. RNA was extracted from 32D cells at various timepoints after induction with granulocyte colony-stimulating factor (G-CSF) conditioned medium (kindly provided by Giovanni Rovera, Wistar Institute), as well as from MEL and WEHI-3 cells using the method by Chirgwin et al.27 Total RNA, 10 ng, was electrophoresed in a 0.9% agarose/formaldehyde denaturing gel, blotted to nitrocellulose, and hybridized overnight in 50% formamide at 42°C to a full-length mouse GATA-1 cDNA probe labeled using nick-translation with 32P dCTP. The blot was washed for 30 minutes at 55°C in 0.1X SSC, 0.1% SDS.

RESULTS

Structure of LF cDNA. A 2.3-kb cDNA clone was obtained from the CML library. This clone contains the entire coding sequence of the LF propeptide, including a short stretch of 5' untranslated sequence (Fig. 1). The obtained sequence differs by only 2 bp from the recently reported cDNA isolated from mammary gland tissue,20 and predicts an identical protein sequence. This confirms that
were probed with the LF cDNA probe. Six bands were identified and were identical for both, indicating that the gene is grossly intact in HL60 cells (Fig 1). These results suggest an approximate size for the LF gene of 35 kb (Fig 2B). Thus, LF approaches the reported transferrin gene size of 33 kb. Chicken ovotransferrin, closely related to both TF and LF1,14 is, by contrast, only 10.3 kb.15

The first two exons were identified through hybridization of labeled oligonucleotides to genomic subclones, and the size of the intervening intron determined from alignment of overlapping phage clones. The 5' end of the first exon was mapped using S1 nuclease protection and primer extension (see below). This predicts a length of the first exon and the five prime untranslated to be 82 bp (v transferrin 91 bp). The length of the second exon is 167 bp (v transferrin 173).
The intervening intron is approximately 3 kb for LF, versus 1.7 for transferrin.

*S1 and primer extension analysis.* The transcription start site was mapped using both primer extension and S1 analysis, using probes as described in Materials and Methods. Primer extension using two different antisense oligos 50 bp apart, as well as S1 analysis using a PCR-generated probe primed with the more 5' of the two oligonucleotides, established the transcription start site 27 bp 5' of the end of our cDNA (Fig 3). This predicted a 5' untranslated region of 39 bp. There was no primer extension or S1 product detected using RNA from HL60 cells as template. Repeat primer extension using RNA from DMSO-induced HL60 cells also showed no product (data not shown). This rules out the possibility of the presence of prematurely terminated mRNA products in the latter cells.

*Nuclear run-on assay of LF in HL60 cells.* To establish that the lack of LF mRNA production in HL60 cells was caused by a failure of mRNA transcription, nuclear run-on assays were performed on uninduced and DMSO-induced HL60 cells and CML peripheral blood leukocytes (Fig 4). No LF message was detectible before or after DMSO induction. Control probes showed that uninduced cells expressed both c-myc and myeloperoxidase; as previously reported, expression of both genes was suppressed on induction. CML was used as a positive control for LF expression.

*Sequence analysis of the 5' regions of the genomic LF gene.* Sequence of 950 bp 5' of the transcription start site, as well as 700 bp of the first intron, was obtained (Fig 5). Analysis of the sequence of the 5' end of the LF gene shows several well-described control sequences. A TATA element is found at −30, a CCAAT element at −59, a CACCC element at −68, and the ubiquitous SP1 binding site at −24 (Fig 4). A hexamer, CTGGGA, that has been found in several proteins involved in the acute phase response (fibrinogen, haptoglobin, α1 antitrypsin) is found at −280.32 A binding site for the zinc finger transcription factor GATA-1, currently felt to be a predominantly erythroid-specific factor 33−35 is located at −367 bp. Finally, a region at −350 bp is highly homologous to the consensus sequence (5'-GGTCANNTGACC-3')36 of the estrogen response element, from which it differs by only 2 of 13 bp (GGT-CAGGCCGTAC).
Fig 3. Analysis of LF mRNA transcript initiation site by primer extension and S\(_1\) nuclease mapping. See Materials and Methods for details. (A) Results of primer extension using two oligonucleotide primers 60 bp apart. Lane M contains markers with size in nucleotides at the left. Lane 1 shows primer extension using the more 3’ oligonucleotide. Lane 2 shows primer extension product using the more 5’ oligonucleotide, which was also used to define the 3’ end of the S\(_1\) probe. (B) Results of S\(_1\) analysis. Probe was a 250-bp PCR product primed at the 3’ end with the antisense oligonucleotide used as a primer in lane 2 of the primer extension. Lane M contains size markers. Lane 1 shows protected fragment from RNA from CML; lane 2 shows the same result using RNA from HL60; lane 3 contains tRNA.

Northern analysis. The finding of a GATA-1 site in the LF promoter prompted us to look for GATA-1 message in a cell line that is capable of expressing LF message. The only cell line available for this purpose is the 32D mouse progenitor cell line. Total RNA was isolated from this cell line both before and after induction of a mature neutrophil phenotype with G-CSF. These RNAs were then analyzed by Northern blotting. MEL cell RNA was used as a positive control for GATA-1 expression, and WEHI RNA as a negative control. Figure 6 shows GATA-1 expression in 32D cells at all stages of induction. As expected, there is abundant GATA-1 expression in MEL, and absent expression in WEHI cells. Hybridization to the upper ribosomal band is seen in all lanes.

Gel shift assays. To determine if the GATA-1 site in the LF promoter could serve as a binding sequence for GATA-1 protein, mobility shift assays were performed. A double-stranded oligonucleotide probe of the GATA-1 binding sequence from the LF gene promoter was bound to nuclear extracts from 32D cells, MEL cells, and HL60 cells (Fig 7). Binding of the LF GATA-1 site was seen in both MEL and 32D extracts. Specificity of binding was confirmed by the finding that the binding could be competed out by a fragment containing a GATA-1 site from the \(\gamma\)globin promoter, and by the unlabeled LF fragment. However, the binding could not be competed by a fragment identical to the LF fragment but which contained two mutations in the GATA-1 site. A previously described breakdown product from the GATA-1 protein is seen in the 32D lanes resulting

Fig 4. Analysis of transcription of LF gene in HL60 cells by nuclear run-on assay. See Materials and Methods for details. \(^{32}\)P-UTP labeled RNA synthesized by nuclear run-on assay of nuclei from uninduced and DMSO-induced HL60, and peripheral blood leukocytes from CML was hybridized to slot-blotted DNA probes for Bluescript plasmid, actin, c-myc, myeloperoxidase (MPO), and LF.
CHARACTERIZATION OF THE LACTOFERRIN PROMOTER

Fig 5. Sequence of the LF promoter region. Uppercase letters with protein sequence denote exons. Dashed lines above the sequence indicate locations of TATA, SP1, and CCAAT elements, the first of which is in reverse orientation. Dotted line above sequence indicates acute-phase reaction element. Carets indicate estrogen response element; straight line marks location of the GATA-1 binding site.

Sequence analysis of HL60 LF gene promoter. Oligonucleotide primers were synthesized corresponding to 5' sequences, and PCR was used to amplify 600 bases of the LF promoter region from HL60 genomic DNA (Fig 5). These fragments were sequenced in their entirety, and found to be identical to their counterparts in the transferrin gene. These data add to the evidence that LF is a relatively late by-product of a series of gene duplications originating with an ancient

We report here the cloning of the genomic LF gene, mapping studies of the transcription start site and preliminary characterization of cis-acting 5' promoter sequences. As expected, there is striking homology of size and sequence between the first two LF exons identified here and their counterparts in the transferrin gene. These data add to the evidence that LF is a relatively late by-product of a series of gene duplications originating with an ancient
iron-binding protein. Serial duplications have given rise both to internal sequence homologies within LF and transferrin as well as to a transferrin "supergene" family. This family currently includes the human melanoma-associated antigen p97 ("melanotransferrin"), a membrane-associated iron-binding protein found primarily in melanomas, nevi, small bowel, and certain fetal tissues. p97 is located on the same region of chromosome 3 that includes transferrin, LF, and the transferrin receptor.

Our further studies have been aimed at understanding the expression of LF within the context of the neutrophil secondary granule. Two findings in pathologic conditions support the theory that the coordinate appearance of secondary granule content proteins is regulated at the level of mRNA accumulation. In specific granule deficiency, a rare disorder characterized by absence of neutrophil secondary granules and their content proteins, the deficiency of LF has been correlated with almost complete absence of LF mRNA. In this patient, LF was detectable in glandular epithelial tissues, indicating that neutrophil LF expression is under a completely separate regulatory control intimately linked to secondary granule expression, but not necessarily to morphologic differentiation. This suggests that this disease may be caused by a defect in RNA accumulation common to all of the secondary granule proteins. This could perhaps be explained by the absence of a shared transcriptional regulator directing coordinate expression of the secondary granule protein genes.

Further evidence for this comes from studies of HL60 cells, which can be induced with a number of agents toward morphologic maturation to segmented neutrophils, but which do not express LF or any of the other secondary granule proteins. Our studies show that the LF gene is grossly intact in the HL 60 cell line. Although we cannot exclude the possibility that small, critical regulatory regions have been lost, these data support the hypothesis that a critical trans-acting promoter is absent, or that a silencer has been left unperturbed by currently available inducers.

This situation may in fact reflect a common in vivo behavior of leukemic cells. Fearon et al., using molecular
probes of X-linked polymorphisms and aneuploid chromosomal abnormalities, demonstrated that terminal differentiation of the leukemic clone to mature polymorphonuclear cells occurs and may even be detectable in patients who appear in morphologic remission. Several investigators have documented cytochemical abnormalities in these mature neutrophils, including, but not limited to, LF deficiency.

Schofield et al.\(^9\) found LF completely absent in the mature neutrophils of 4 of 14 patients with acute myelogenous leukemia and markedly diminished in two others. In all patients studied, the LF content of mature neutrophils decreased between remission and relapse, as did myeloperoxidase and nonspecific esterase content, but complete absence was uncommon. Miyauchi et al.\(^{10}\) likewise detected LF deficiency in some, but not all, of the neutrophils of 14 patients studied. In two patients in whom all neutrophils were LF deficiency, ultrastructural study showed complete absence of secondary granules, as had been described previously by Bainton et al.\(^{11}\) These abnormalities typically are not seen in remission neutrophils, strongly implying that the anomalies are by-products of disordered differentiation of the leukemic clone. Thus, like HL60 cells in vitro, many in vivo leukemic cells lack the ability to express LF in the backdrop of a global defect in the expression of secondary granules.

The role of the cis-acting sequence elements in the regulation of LF remains to be established with physiologic studies. However, several are of potential interest. The presence of a sequence 85% homologous with the consensus sequence for the estrogen response element is consistent with the estrogen inducibility of LF in reproductive epithelium.\(^{46,47}\) Martinez et al.\(^{48}\) studied several mutants of this consensus sequence for activity in a reporter gene assay, but did not test the particular configuration we have identified.

A binding site for a transcription factor apparently common to acute-phase reactants is also present, but its significance has not yet been defined.\(^3\) While one would expect increased LF synthesis under conditions of "stress" as part of the broader expansion of granulocyte mass, it is not clear that individual neutrophils will increase their LF synthesis under such conditions, even though the presence of such an activator implies this response.

The presence of a binding site for the erythroid growth factor GATA-1 (also known as NF-E1, GF-1, Eryf-1) is a surprise. This transcription factor has been cloned\(^33,34\) and its expression has been reported to be restricted to erythroid cells, megakaryocytes, and mast cells.\(^33,35\) However, the only myeloid cell line studied has been HL60, which fails to express LF. We have studied the expression of GATA-1 in 32D cells differentiated toward the myeloid lineage with G-CSF and found that in this cell line GATA-1 is expressed even after differentiation. Based on this study, mobility shift assays for GATA-1 binding to the consensus site in the LF promoter were performed using extracts from the 32D cells. Extracts from MEL cells, known to produce abundant GATA-1 protein, and HL60 cells, known not to express GATA-1, were used as controls. Our assays show specific binding to the GATA-1 consensus site in the LF promoter. These studies show that the GATA-1 site in LF can be bound by GATA-1 protein, and furthermore that at least one myeloid cell line expresses GATA-1. These studies suggest a possible role of GATA-1 in the control of LF expression during hematopoiesis, but proof of this awaits functional studies currently being undertaken.

In summary, we have examined the structure and expression of the LF gene in neutrophil precursors and compared it with HL60 cells. We have cloned a cDNA of neutrophil LF and established its identity with the cDNA cloned from mammary gland. We have isolated six genomic clones which together include the full length of the LF gene. Its first two exons bear striking homology to the transferrin gene consistent with its origin from a gene duplication. We have identified the 5' regulatory region and several potential cis-acting elements whose physiologic role remains to be more fully defined. We have found that the gene is grossly intact in the HL60 myeloid leukemic cell line, despite its inability to express it following induction toward myeloid differentiation. Nuclear run-on assays have established that the gene is not transcribed in those cells. This strongly suggests that the defect in these cells is an abnormality in the production or function of a transacting regulator of LF gene expression within neutrophils.

**REFERENCES**


