

1998

Cloning, sequencing and partial characterization of the 5' flanking region of the mouse neutrophil collagenase gene : conservation with the human neutrophil collagenase regulatory region

Marcoli Cyrille
Yale University

Follow this and additional works at: <http://elischolar.library.yale.edu/ymtdl>

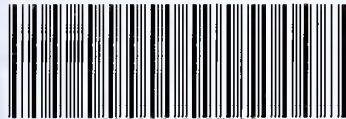
Recommended Citation

Cyrille, Marcoli, "Cloning, sequencing and partial characterization of the 5' flanking region of the mouse neutrophil collagenase gene : conservation with the human neutrophil collagenase regulatory region" (1998). *Yale Medicine Thesis Digital Library*. 2503.
<http://elischolar.library.yale.edu/ymtdl/2503>

This Open Access Thesis is brought to you for free and open access by the School of Medicine at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Medicine Thesis Digital Library by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

MED
T113
+Y12
6559

YALE UNIVERSITY LIBRARY



39002010487339

CLONING, SEQUENCING AND PARTIAL
CHARACTERIZATION OF THE 5' FLANKING REGION
OF THE MOUSE NEUTROPHIL COLLAGENASE GENE:
CONSERVATION WITH THE HUMAN NEUTROPHIL
COLLAGENASE REGULATORY REGION.

Marcoli Cyrille

YALE UNIVERSITY

1998

YALE
UNIVERSITY



CUSHING/WHITNEY
MEDICAL LIBRARY


Permission to photocopy or microfilm processing of this thesis for the purpose of individual scholarly consultation or reference is hereby granted by the author. This permission is not to be interpreted as affecting publication of this work or otherwise placing it in the public domain, and the author reserves all rights of ownership guaranteed under common law protection of unpublished manuscripts.

Marati Lejille

Signature of Author

5/11/98

Date



Digitized by the Internet Archive
in 2017 with funding from
The National Endowment for the Humanities and the Arcadia Fund

<https://archive.org/details/cloningsequencin00cyri>

**Cloning, Sequencing and Partial Characterization
of the 5' Flanking Region of the Mouse Neutrophil Collagenase Gene:
Conservation with the Human Neutrophil Collagenase Regulatory Region.**

**A Thesis Submitted to the
Yale University School of Medicine
in Partial Fullfillment of the Requirements for
the Degree of Doctor of Medicine**

Marcoli Cyrille

Yale University

1998

Med Lib

Thesis

+Y12

6559

YALE MEDICAL LIBRARY

AUG 13 1998

CLONING, SEQUENCING AND PARTIAL CHARACTERIZATION OF THE 5'
FLANKING REGION OF THE MOUSE NEUTROPHIL COLLAGENASE GENE:
CONSERVATION WITH THE HUMAN NEUTROPHIL COLLAGENASE
REGULATORY REGION

Marcoli Cyrille, Nathan Lawson, Theresa Zibello and Nancy Berliner. Department of Internal Medicine, Section of Hematology, Yale University School of Medicine, New Haven, CT.

Abstract

The expression of the secondary granule protein (SGP) genes lactoferrin (LF), transcobalamin (TBI), human neutrophil gelatinase (HNG) and collagenase (HNC) is coordinately regulated at the level of mRNA transcription. Our laboratory has previously identified a silencer element in the LF promoter which binds CCAAT displacement protein (CDP); when CDP is overexpressed in myeloid cells, it blocks the expression of all SGP genes. CDP is a ubiquitous repressor known to be implicated in the regulation of many myeloid-specific genes, and we wish to determine the means by which it acts as a stage specific repressor of multiple genes at different stages of differentiation. Here we describe the cloning, sequencing and partial characterization of the 5' flanking region of the Mouse Neutrophil Collagenase (MNC) gene. The MNC promoter shows a high degree of conservation over more than 200bp with the human homolog, with a 56bp stretch sharing 81.4% nucleotide identity. Within this region, we have identified well described control elements involved in myeloid gene regulation, including a canonical consensus sequence for the CCAAT enhancer binding protein alpha (C/EBP α). Functional assays will be necessary to establish the functional significance of these regions. CDP binding sites are highly promiscuous and can only be characterized functionally. The identification of the promoter will now allow such characterization.

TABLE OF CONTENTS

I. Overview of Transcriptional Regulation of Myeloid Differentiation.....	1
II. Secondary Granule Protein Gene Expression in Neutrophils.....	3
A.SGP gene transcription is coordinately regulated at the level of mRNA transcription.....	4
B. SGP gene expression is coordinately blocked in abnormal myeloid differentiation.....	6
III. Identification of a Shared Regulatory Pathway for Multiple Myeloid-Specific Genes.....	9
A. Identification of a silencer within the LF Promoter.....	9
B. Identification of CDP as binding protein for silencer element.....	9
IV. CDP and Positive Regulators in Myeloid Promoters.....	12
V. Statement of Purpose.....	14
VI. Materials and Methods.....	15
A. Cell lines and Cloning of the MNC cDNA.....	15
B. Generation of a Probe for Screening a Genomic Library.....	15
C. Titering of Library.....	16
D. Screening of 129SV genomic Library.....	16
E. Phage Amplification.....	17
F. Generation of Probes for isolation of 5' region of MNC gene.....	17
G. Isolation of the 5' Flanking Region of the MNC gene.....	17
H.Analysis of Sequences.....	18
VII. Results.....	19
A. Isolation of the MNC gene.....	19
B. Isolation of the 5' Flanking Region of the MNC gene.....	19
C. Sequence Analysis of the 5' Flanking Region of the MNC gene.....	23

VIII. Discussion.....	29
Acknowledgements.....	31
References.....	31

I. OVERVIEW OF TRANSCRIPTIONAL REGULATION OF MYELOID DIFFERENTIATION

Hematopoiesis is the process by which blood cells acquire defining phenotypes as the consequence of coordinated, cell-specific gene expression. The proliferation and differentiation of hematopoietic cells are orchestrated by a variety of extra and intra cellular signals. Extracellular signals, namely growth factors, signal cellular events by binding membrane and cytoplasmic receptors. They produce their physiologic effect by modulating the level of expression of key transcriptional regulators and directing downstream changes in proliferation and phenotype by regulating gene expression.

The regulation of gene expression is mediated through *cis*-acting regulatory sequences and the *trans*-acting regulatory factors which bind to them. The locus of biological activity of these regulatory elements is usually at the 5' flanking region of the transcription start site of the genes, namely the promoter.

The pathway to terminal neutrophil differentiation and maturation has not been completely outlined. However some of the necessary transcriptional regulators governing that process are emerging out of recent studies.

PU.1 appears to be required for myeloid differentiation. PU.1 is the product of Spi-1/Sfpi1/PU.1 protooncogene, a hematopoietic-specific member of the *ets* family expressed principally in monocytes, macrophages and B lymphocytes, but also in granulocytes and erythroid cells. Nearly all myeloid-specific genes contain biologically active PU.1 binding sites in their regulatory regions. In one study, mice lacking PU.1 function die 1 to 3 days before their expected birth and these PU.1-null mice show an absence of monocytes, T and B lymphocytes and granulocytes. Further, flow cytometry shows diagnostic Ig and T cell receptor gene rearrangements¹, as demonstrated by surface markers. Early myeloid-specific genes are expressed in embryonal stem cells from these mice. These cells in fact show a loss of all late markers of lymphoid and

myeloid lineage maturation^{2,3}. Another study of PU.1-null mice results in viable mice with an otherwise similar defect in myeloid cell development. For example, granulocyte-colony stimulating factor (G-CSF) and granulocyte macrophage-colony stimulating factor (GM-CSF) receptor mRNA can be detected in these cells. However these receptors are not detected at the protein level. The reasons for the discordant viability of the two models are unknown.

Also implicated in early myeloid differentiation is the family of CCAAT/Enhancer binding proteins (C/EBP)⁴. Shown to be a member of a family of differentially expressed genes, C/EBP was first isolated as a protein with a basic region/leucine zipper structure which binds a CCAAT site. It was originally purified and cloned from rat liver and appears to have a role in adipogenesis. C/EBP α -null mice have an impaired glucose metabolism and die within 24 hours of birth. Late myeloid differentiation is also impaired in these mice as they show absence of mature granulocytic cells, and very low expression of the G-CSF receptor^{5,6}.

Other transcription factors important for myeloid differentiation include MZF-1^{7,8}, c-myb⁹, and the retinoic acid receptors¹⁰. The evidence is growing that there are complexes of transcription factors which are necessary for regulated expression of many of the myeloid genes.

In order to understand leukemogenesis and normal myeloid differentiation it is important to identify and characterize *cis* and *trans* regulators that activate important genes in the myeloid lineage.

II. SECONDARY GRANULE PROTEIN GENE EXPRESSION IN NEUTROPHILS.

In the process of differentiation and maturation of neutrophils, a critical and pivotal step occurs at the transition from promyelocytes to myelocytes. Normal progenitor marrow cells begin to lose proliferative capacity¹¹ as they become committed to terminal neutrophil maturation¹². Leukemic cells from patients with AML are arrested in maturation at or before the promyelocyte stage, suggesting the critical importance of understanding the processes involved in regulating this commitment step.

Sequential phases of granulocyte differentiation show the acquisition of different types of granules. The promyelocyte stage is marked by primary ("nonspecific") granules which are attenuated in the later stages of both neutrophil and monocyte maturation¹³. As a promyelocyte advances to the myelocyte stage it acquires secondary ("specific") granules¹⁴. Specific granules carry four major identified proteins, namely, transcobalamin I (TCI), lactoferrin (LF), human neutrophil collagenase (HNC), and human neutrophil gelatinase (HNG). The rise of secondary granules provides a unique label of commitment to terminal neutrophil differentiation. Thus, the expression of genes responsible for the formation of secondary granule membranes and their contents should constitute a convenient probe for the mechanisms governing regular neutrophil differentiation.

HNG and HNC are believed to be responsible for the turnover of extracellular matrix and involved in the acquisition of metastatic potential in transformed cells. Our laboratory has previously cloned and sequenced approximately 1kb of the 5' flanking region of the human neutrophil collagenase gene (See Figure 1). This promoter has not been characterized functionally.

Under physiologic conditions, collagenases are enzymes that cleave peptide bonds in the triple-helical regions of collagen. Two sets of collagenases are known: the tissue collagenases and the collagenases implicated in gas gangrene.

The bacterium *Clostridium histolyticum* that causes gas gangrene secretes a battery of collagenases that contribute to their invasiveness by destroying the connective-tissue barriers of the host.

Neutrophil collagenase is a tissue collagenase and is thought to play a role in degrading collagen in inflammation. The action of the collagenases is potentially harmful to the host but activated collagenases can be inhibited by a family of specific tissue inhibitors produced by most mesenchymal cells. Neutrophil collagenase probably plays a critical role in the host response to infection as well as in conferring metastatic potential in malignancy.

A.)Secondary granule protein gene transcription is coordinately regulated at the level of mRNA transcription.

Studies in several systems have shown that secondary granule protein gene (SGP) expression is coordinately regulated at the level of mRNA transcription.

Upon induction with G-CSF the 32Dcl3 cell line (32D) undergo differentiation to mature neutrophils. The 32D cell line is a murine stem cell line that requires IL3. The Berliner laboratory proved that upon induction of neutrophil differentiation with G-CSF, the expression of a mouse neutrophil gelatinase gene showed a similar pattern to that of the LF gene. Nuclear run-on assays established that regulation of expression was transcriptional¹⁵. Our laboratory later isolated a cDNA clone for the mouse homolog of neutrophil collagenase, and showed comparable results in 32D cells¹⁶.

gaattcgatgccctccacaaatgaaacatgtgacttaaacaggtcatctaaatgtctctggaggtgatttcatctgtaa
 atggagatagtaatagcatctccctccaaggctatgtgaagatggattggggtaatatgtggaacacagtacaaa
 tgtctgggcaatcataatTTTTtagcaaatgttcattattatcatcatcagattgttctctttatctttccatggaaagtc
 taacttatagactagggattttctaataatggtgataccaaaccacatctccagaagcagctctaggccccatacaaag
 cctgctttgtaataacacagtcactcattctctccacatacaatgagggaggataagtacagagattctctagttgggct
 ccttgggtccaagatatctaccacaaacctcaaaatattccctgcataagttaattcaacctcaacatcttacatagtca
 aaaaatgtgatccacataacttcaggattagtccttgttcttttctcaaccaatTTTgtaaaaggaaataactaaaa
 catccaaaagcaaagagtagaaagttggagtttgaagactgtt**gaaagactccatttgggggtactaaatcacatt**
ataaaattgcaaaactgtcagagtgaattaacctatgttgcattatctctgatgcagcaattgtttaagacatggct
ttaaagagggactgcaagagagctgtgagtgaCACATGATGCTGTGAACCTCAGGGTGCTCGC
 CAGGGAAGGGCCCNNGGCGTCACCTTGCAACTCCCAAGAACGGACCTAT
 GGAGAAAGAAAGCCAGGAGGGGTCATGAGTTTGAAGAGAAGATCATGT
 TCTGCCTGAAGACGCTTCCATTTCTGCTCTTACTCCATGTGCAGATTTCCA
 AGGCCTTTCCTGTATCTTCTAAAGAGAAAAAATACAAAAATTGTTTCAGGT
 AAATGAACTATCAGTTAGGTGATGTTTGATTGCTGATTATTAGGAAATA
 GTGTGGGTTGTTTTCTCTTGTTAATTA AACAGCCTGAAAGAAATTATTTG
 GGTGATTAGTTTGAGATTGAACATCTCCTTTTAGTGATTGGAGGGTTCCA
 AGCTGGTAAGGACTAGGCCTTTTAAAAGACGGACATGAAAACAAGCTGC
 ATTC

Figure. 1 - DNA sequence of the 5' Flanking region of HNC gene. Capitalized sequence corresponds to transcribed region as intimated by cDNA data. Region of conservation with the mouse homolog is italicized, with highest conservation in bold faces. Translation start site is underlined.

MPRO cells are derived from murine bone marrow stem cells which have been retrovirally transduced with a dominant negative retinoic acid receptor (RAR α) by Tsai et al. who described two cell lines from such transductions: a stem cell factor-dependent cell line, EML, which fails to differentiate into early myeloid precursors until treated with all trans retinoic acid (ATRA), and a GM-CSF dependent line, MPRO, which is blocked in differentiation at the promyelocyte stage. Our laboratory has induced neutrophil differentiation in both EML and MPRO with appropriate growth factors in the presence of ATRA, and has shown phenotypic maturation to mature neutrophils associated with rapid increase in LF gene expression .

In further studies, the laboratory examined the sequential expression of the primary and secondary granule protein genes in purified bone marrow progenitors induced with cytokines to undergo in vitro differentiation to terminal neutrophils. Over a period of 10-12 days, the cells progressively matured to the late metamyelocyte and band stage. Induction with G-CSF resulted in a stage-specific expression of primary and secondary granule proteins, with asynchronous expression of nonspecific granule proteins starting from days 1-5, and concomitant expression of lactoferrin and transcobalamin I (SGPs) from day 7-8¹⁷.

B.)SGP gene expression is coordinately blocked in abnormal myeloid differentiation

Leukemic cell lines proliferate independently from growth factors. The HL60 and NB4 cell lines were derived from patients with acute promyelocytic leukemia and are used in the study of neutrophil differentiation. Our laboratory has shown an irreversible block to SGP gene expression in inducible leukemic cells

HL60 cells are inducible to phenotypic maturation with 12-O-tetradecanoylphorbol-13-acetate (TPA) to monocytes and with dimethyl sulfoxide

(DMSO) to granulocytes. HL60 cells induced with DMSO undergo defective neutrophil maturation, manifested by a coordinate loss of secondary granule protein gene expression. In previous studies our laboratory has characterized the defect in secondary granule protein gene expression in uninduced and DMSO-induced HL60 cells. Nuclear run-on assays confirmed that the secondary granule protein mRNA are not transcribed. In addition, our laboratory has sequenced more than 600bp of the HL60 lactoferrin promoter region and shown that it is not defective structurally. Since the SGP genes are not linked these results suggest that the failure of gene expression in these cells reflects a regulatory defect in part via a *trans*-acting factor¹⁸.

Similar findings have been shown in the NB4 cell line which exhibits a defect in granulocytic maturation in ATRA-induced cells. They lack expression of all the secondary granule protein genes upon induction with ATRA¹⁹.

Further, specific granule deficiency (SGD), is a congenital disorder characterized by recurrent infections, the absence of neutrophil secondary granules, and marked deficiency or absence of their content proteins. We demonstrated insignificant mRNA levels for all of the secondary granule protein genes as well as for defensins from bone marrow cells obtained from a patient with SGD. These levels were comparable to the previous SGP levels measured in this patient's neutrophils. This is consistent with the notion that a common *trans*-acting factor governing secondary granule protein gene expression may be compromised in SGD^{20,21}.

Finally, studies of patients with newly-diagnosed acute leukemia have shown absence and/or abnormalities in secondary granule protein gene expression in apparently normal neutrophils which are actually part of the malignant clone.

In light of these studies our laboratory hypothesizes that one irregularity of the leukemic cell is a compromise of the differentiation program via common cis and trans acting factors

which interrupts later events in neutrophil maturation, as reflected in the coordinate loss of expression of secondary granule proteins.

III. IDENTIFICATION OF A SHARED REGULATORY PATHWAY FOR MULTIPLE MYELOID-SPECIFIC GENES.

A) Identification of a silencer within the LF promoter:.

The LF promoter has been extensively studied in our laboratory. LF promoter sequences spanning -384, -649, -726,-750, -800, and -916 bp upstream of the LF transcription start site were subcloned into the pGL2 promoterless vector. These constructs were assayed by transient transfection into many leukemic hematopoietic cell lines including K562, a cell line isolated from a patient with CML in blast crisis, and HEL, a human erythroleukemia cell line; these leukemic lines do not express LF. These plasmids were also assayed in the uninduced 32Dcl3, the IL-3 dependent murine myeloid progenitor cell line mentioned earlier, which expresses the LF gene only after G-CSF-induced differentiation. Measurable luciferase activity was detected for the -384, -649, -726,-750 and -800bp constructs. These results were duplicated in HeLa cells, a non-hematopoietic cell line. However, the -916bp construct only expressed background levels of activity suggesting the presence of a silencer between the 800bp and 916bp upstream of the LF gene transcription start site⁴⁰.

Our laboratory has subsequently shown that the silencing effect is mediated by a nucleotide sequence containing a triad of octamers reported below:

5' ATGTATTTACGAGATGTATTCTAGAAGCAGTATTCTAGCTTTTGAATTT 3'

Deletion plasmids where a 56bp sequence containing the triad of octamers was removed expressed levels of luciferase activity comparable to the -726bp suggesting that the 56bp deleted fragment is a cis-acting silencer of LF expression.

B) Identification of CDP as binding protein for silencer element:

Our laboratory has studied the identified silencer element for DNA-protein interactions by mobility shift assays and UV crosslinking. Data from these studies confirm the presence of a large protein-DNA complex (180kD protein) in LF non-expressing cells⁴⁰ which is absent from LF expressing cells. This has been shown to be bound by CCAAT displacement protein (CDP/cut) described more extensively below.

To further establish the role of CDP in LF gene expression Khanna-Gupta et al. isolated 32D cells stably transfected with a CDP expression vector. Tissue cultures of transfected cells, and single cell clones overexpressing CDP show a complete loss of LF mRNA upon induction with G-CSF, while these cells maintain the ability to undergo morphologic differentiation. Remarkably, CDP overexpression also prevents the expression of neutrophil gelatinase, neutrophil collagenase, and of the myeloid specific gene gp91 phox (Figure 2)^{39,40}.

The coordinate loss of secondary granule protein gene expression upon overexpression of CDP intimates that the coordinate loss of SGP expression in leukemia and SGD could reflect an irregularity of modulation of CDP binding and/or CDP expression. In light of the fact that CDP has been shown to interact with a variety of myeloid specific genes, including PU.1, our laboratory postulates that modifications in CDP binding may be crucial to stage-specific gene expression throughout the process of myeloid differentiation.

An understanding of how CDP interacts with different genes at different phases of the maturation program may therefore provide critical insights into the process of normal differentiation as well as the dysregulation in acute leukemia.

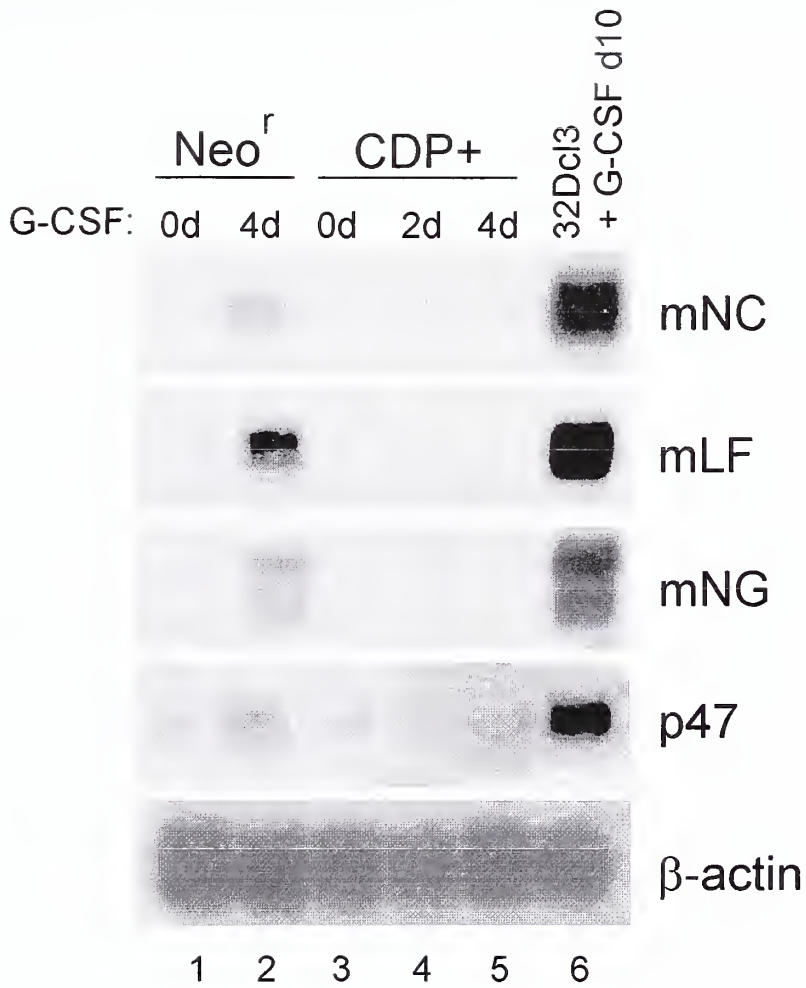


Figure 2.- Northern Blot showing suppression of SGP gene expression in 32D cells. (Courtesy of Nathan Lawson. See reference 38.)

IV. CDP AND POSITIVE REGULATORS IN MYELOID PROMOTERS.

Tissue-specific control of gene expression appears to be partly mediated by negative regulatory elements as exemplified in the growth hormone²², MHC²³, and lysozyme²⁴ gene systems. These systems uniformly exhibit tightly regulated expression that is tissue and stage specific, as with the repression of the lysozyme gene in the stages prior to myeloid differentiation²⁵, and the suppression of immunoglobulin gene expression in immature pre-B cells²⁶.

Originally described as a repressor of the sea urchin testis-specific H2B gene promoter²⁷, CDP is constitutively expressed in most cells. CDP is a member of a homeodomain family involved in lineage specification. This family includes the *Drosophila cut* protein and CDP is related to the canine *clox*²⁸ (Cut like homeobox) and murine *cux-1*^{29,30,31}. CDP/cut proteins contain four possible DNA-binding motifs with overlapping but distinct sequence specificities³², but often binds specifically to disparate sequences.

The role of CDP has been demonstrated in the repression of a wide variety of genes including the human myeloid-specific cytochrome b558 heavy chain gene, gp91 phox, the human *c-myc* gene³³, human γ -globin gene, mouse NCAM gene, and mouse myosin heavy-chain gene. Recent studies of the interaction of CDP with the histone gene locus would lead one to infer that it is a part of a complex with Rb, and may have a critical role in modulating cell cycle events³⁴.

The control of gene expression is usually dependent on a complex interplay of positive and negative regulatory factors. For example, although constitutively expressed in myeloid cells, CDP shows decreased binding for preferential binding of CCAAT enhancer protein, and resultant gp91-phox expression in maturing neutrophils^{35,36}. There is growing evidence that such a complex interaction is central to

the regulation of myeloid specific genes, especially of the myeloid cytokine receptors. As mentioned earlier binding of PU.1 appear to be required for myeloid gene activation^{4,5}. Another group of positive regulators, namely the C/EBP factors also appear to be major modulators of myeloid gene expression with binding sites often proximal to PU.1 binding elements. All C/EBP genes are activated in myeloid differentiation with relative abundance of the different proteins. In addition to PU.1 and C/EBP, the AML1 transcription factor seems necessary for optimal myeloid promoter function. Mutation of AML1 binding site has been shown to dramatically decrease myeloid promoter function^{4,5,37}.

Tissue-specific gene expression is likely to be highly complex and tightly orchestrated by the positive and negative arms of regulation. We hypothesize that CDP plays an integral role in this complex interaction.

V. STATEMENT OF PURPOSE

With the aim of further characterizing the spectrum of CDP-regulated genes and the means by which it acts as a stage specific repressor of multiple genes we have endeavored to identify and sequence the mouse neutrophil collagenase promoter and compare it to the LF and HNC promoters.

We propose that one of the ways CDP may mediate its effects is through common cis-acting elements in the promoters of its target genes.

VI. MATERIALS AND METHOD

A. Cell lines and Cloning of the full length MNC cDNA. (done by Nathan Lawson)

The EML line, a murine stem cell line, is stem cell factor (SCF)-dependent, multipotential, and displays a specific block to differentiation along the myeloid lineage with SCF and IL3. When induced with IL3 and SCF in the presence of high doses of all-*trans* retinoic acid (ATRA), EML cells differentiate to promyelocytes. Transfer to GM-CSF in the presence of ATRA will promote full differentiation to neutrophils. MPRO cells were selected for growth in GM-CSF, and are blocked in differentiation at the promyelocyte stage. When induced with ATRA, they too will mature to neutrophils. Induction of EML with IL3, SCF, and ATRA for three days, followed by culture in GM-CSF in the absence of ATRA results in a cell line blocked at the promyelocyte stage. This line, termed EPRO, is similar in phenotype to the MPRO cell line. Using cDNA representational difference analysis to compare the EPRO cell line to the multipotent stem cell line from which it was derived, EML, our laboratory had previously isolated a full length 1.9 kb MNC cDNA. That fragment was subsequently cloned in pBluescript (Stratagene, LaJolla,CA) vector.

B. Generation of a probe for screening a genomic library.

The above construct was used to transform DH 10B electromax cells according to standard protocol. The transformed cells were lysed by alkali and the amplified clone was isolated by ultracentrifugation, CsCl/EtBr gradients according to standard protocol.. The product was digested with appropriate restriction enzymes and purified

on an agarose gel. The 1.9 kb MNC cDNA was radiolabelled by nick translation and used as a probe to screen the 129 SV mouse genomic library which was kindly provided by Dr. Archibald Perkins.

C. Titering of library.

Phage were diluted in SM (NaCl 5.8g/L; MgSO₄/7H₂O 2.0 g/L; 1M Tris, pH 7.5 50ml/L; 2% Gelatin 5ml/L) to estimated concentrations of 10³, 10⁴, and 10⁵/ml and plated on small NZY agar plates with 1-2µl/plate. 100 µl of overnight host bacteria, LE 392, were adsorbed to phage at 37°C for 15 minutes, diluted in 3ml of molten agarose and poured onto warmed plates which were subsequently incubated at 37°C overnight. Plaques were subsequently counted to determine phage titer.

D. Screening of 129SV genomic library.

Phages were plated at 10⁶ plaque forming units (pfu)/plate and 30 plates were used to screen approximately 1.5x10⁶ recombinants. Host bacteria were adsorbed to phages at 300µl/plate at 37°C for 30 minutes prior to overnight incubation. Duplicate BioTraceNT nitrocellulose membranes (Gelman Sciences, Ann Arbor, MI) were lifted and hybridized at 65°C overnight to the nick-translated MNC cDNA probes in hybridization solution (4xSSC; 5x Denhardt's; 1% SDS; salmon sperm DNA; 0.1% Na pyrophosphate) and washed with 2xSSC (300mmol/L NaCl, 60 mmol/L Na Citrate)/0.5% SDS twice at room temperature for 5 minutes followed by 15 min. of washing in 2xSSC/0.1% SDS at room temperature and finally one high stringency wash in 0.5xSSC/0.1% SDS at 65°C for 30 min.. Films were exposed for 48 hrs. for visualization of positive plaques. Positive plaques were purified by secondary and tertiary screenings using the above method.

E. Phage amplification.

The tertiary core phage was titered as described above and the phage were plated on NZY agarose plates at 1×10^5 pfu/plate and grown overnight at 37°C for confluent lysis. The small plates were then overlaid with 9ml of SM and gently shaken at room temperature overnight. The SM was then collected spun and the supernatant stored. Tertiary phages were thus amplified to 10^9 pfu/plate. The phagemids were subsequently excised according to standard protocol and digested by several restriction enzymes. Genomic fragments were subcloned into the pbluescript plasmid vector.

F. Generation of probes for isolation of 5' region of the MNC gene..

The 1.9 kb MNC cDNA was digested with appropriate enzymes to generate a 500bp probe for the 5' portion of the MNC cDNA. That fragment was subsequently radiolabelled by Nick translation as described above. That probe was used for initial screening of genomic clones and preliminary sequences obtained were used to generate upstream genomic oligonucleotides. These ^{32}P labeled 22-24bp single stranded oligonucleotides, including one corresponding to the 5' end of the MNC cDNA were used for further screening of genomic clones.

G. Isolation of the 5' flanking region of the MNC gene.

The genomic constructs were used to transform DH 10B electromax cells by electroporation. Transformed cells were plated and incubated overnight at 37°C. Nitrocellulose filters (Gelman Sciences, Ann Arbor,MI) were lifted and hybridized at 37°C overnight to the probes described above. Positive colonies were grown and lysed

according to Qiagen Maxi protocols and plasmids were isolated and submitted to the Yale sequencing facility for sequencing by the dideoxy chain termination method.

H. Analysis of sequences.

Sequences were analyzed using the databank of the NIH and DNA star.

VII. RESULTS.

A. Isolation of the MNC gene.

Using the 129 SV genomic library, approximately 1.5×10^6 phages were screened with the 1.9kb MNC cDNA resulting in 17 positive clones. From these 5 clones were purified by secondary and tertiary screens.

B. Isolation of the 5' flanking region of the MNC gene.

Two of the 5 positive clones were digested and subcloned in pBluescript and screened with the 500bp MNC cDNA corresponding to the 5' region and sequenced. These sequences were used to design single stranded oligonucleotides, including one corresponding to the 5' end of the MNC cDNA, with which to obtain further upstream sequences.

5 subclones were isolated, representing three independent DNA fragments. Two of these clones were identical, 5kb in size, and contained the first exon of the MNC gene and 4kb of the 5' flanking region. Two other identical clones (1.8kb) contained most of the first intron and part of the second exon. We have subsequently established the first intron to be 1,839bp in length. The last positive clone was found to be a subclone of the latter two clones (Figures 3 & 4).

Previous studies of the MNC cDNA have established the translation start site and suggested a relatively short 5' untranslated (UTR) region of 73 nucleotides. It remains to be determined by S1 nuclease and primer extension if there is additional 5' untranslated sequences although this is unlikely since the HNC 5' UTR is approximately 95 n.t.

The clone containing the 5' flanking region was digested with several restriction enzymes and a map of that region was generated (Figures 6,7 & 8).



Figure 3.- Pst digests of two positive subclones. The subclones are identical in size approximately 5kb in size.

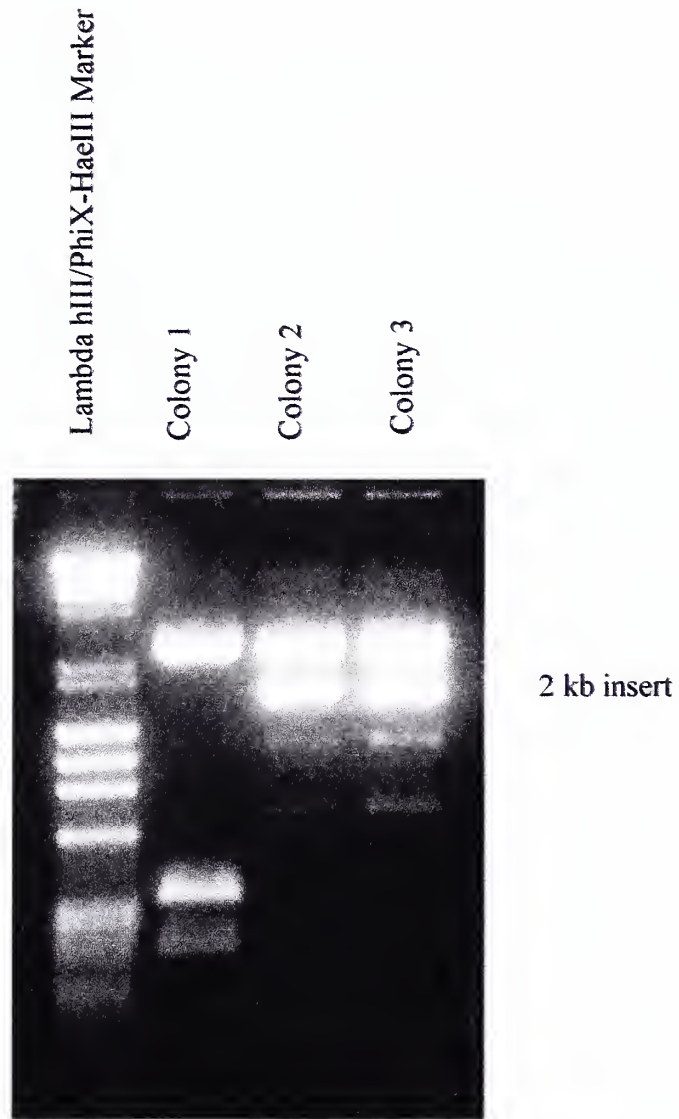


Figure 4.- EcoRI/HindIII digests of three positive subclones. Two are identical in size and contain the genomic subfragment with the first intron of the MNC gene.

C. Sequence analysis of the 5' flanking region of the MNC gene.

Sequence of 1314bp 5' of the putative transcription start site as well as 1839bp of the first intron was obtained (Figure 5). Analysis of the sequence at the 5' end of the MNC gene shows several well-described control sequences. A CCAAT element is found at -6 as well as an AP.1 recognition site at that position. There is a c-myb binding element at -111, a CACCC motif at -200 and a PU.1 binding site at -253. A C/EBP α binding site is found in the reverse direction at -530 and further upstream a C/EBP β element is found at -772. Multiple binding sites for the zinc-finger transcription factor GATA-1 are at -796, -1038, and -1169 as well as a TGGCA element at -1000. Finally we found at -1269 and at -1221 consensus sequences to the c-ets-2 and c-ets-1 binding factors.

The sequence between -125 and -69 shows a high degree of sequence conservation with the human homolog of neutrophil collagenase. That region shares a 81.4% similarity index with a region in the 5' flanking region of the human neutrophil collagenase (Figure 9). In the human counterpart that region contains a canonical consensus binding site for C/EBP α whereas the sequence in the mouse genome is imperfect. Further the region spanning 254bp upstream of the 5' untranslated region maintains a high degree of conservation, sharing a 71.1% similarity index with the corresponding region in the promoter of the human neutrophil collagenase. The role of this region in regulating MNC expression is not yet established but the high conservation in regions that do not encode protein suggests that it may be important.

No comparative octamer repeats, namely a negative regulatory element, to the LF promoter were found in our preliminary analysis and sequence. This is probably not relevant since CDP binding is promiscuous and can only be defined functionally.

agttcattttagggctgagantaaaaacttagtctgtgaactacttctttcaggattgaggaacaggagtagattcgtaagacaggcaggcaggaagaga
c-ets-2 PU.1
 gcccagtgcncgtgtcaccacatcatctggaagccctcactcacctctagcactctggggatgtctgagacatctatctctctccttgcctttgtcct
GATA-1 Ap.2
TGCCA motif GATA-1/Octamer motif
 ggatgtaggccaagtttccactgtctgtgagccacagtactgatggccagcctgattcaaatagcaggttgatgaactagagcctggtataggatggg

 ctctgttctctcaccatactcatgggctgagctctatgcaaaagtgagcatgtgccttcttagtctccagcctctgctctctgcaaaattatagcaatagtgtg
GATA-1
 tgctctgaaagctgtatgagaataggcatttagtgtcattaacatgtttgaaacattgtaagtacaacagatgtagtttctcttaactatcactatctttgtct
C/EBPβ
 ttctctgaaaaaccaccttatagtgaatgttttagtacatctatagcagtcattgccacacagaagctgtgttacgtatgtataggacgtgtgattttaa

 aagtgaatgttcatgtttgtgtgtgctgaagtgggattgtttgtttgtttgtgacagggctcttaactagcctggaactaccaagtagactaggtgactggc
C/EBPα (reversed)
 caataagctgcaaacacccccctgctggcacagaccagaccagattatactcatgtaccactaagttcagtttttacctgcatctgagaatcaaatcagg
AP-2
 ccctcatgcttgtgtgacaagtactttcctaacaagccatcatctctagatgctgttttggccccacacaagcctgcttggcacagctatcaggcagcagt
PU.1
 ggtggcttagaaacagaagtcttacaggtgtgccaccgcttccaagatgatcgctgaacattctaaaatattgccccgattaactagcacaacttctcattac
CACCC AP-1
 acacagacagaaactggaagactgtagccccgtgccagccactcacctgacagtttacaagtagcgatccatctcttctgttgaatgactccccca
c-myb;C/EBPα (in human homolog)
 ctctaggctcaaatcacattctaaccattgaagaactgtcagagtgcggttagctatgttgcctcactgcttaggcagaagttgctcaagatgtagctttataag
CCAAT Sp.1
 ggaccagacaagagcagtgattggcagGGGATGCTGTGAGCTTGAGGGCACTTGCCAGAGGGGCAGAGGAAGGC

Figure 5.- Sequence of 5' Flanking region of MNC gene. Well described control elements are illustrated. Exon is in capitals. Region of high conservation is italicized with highest conservation in bold faces.

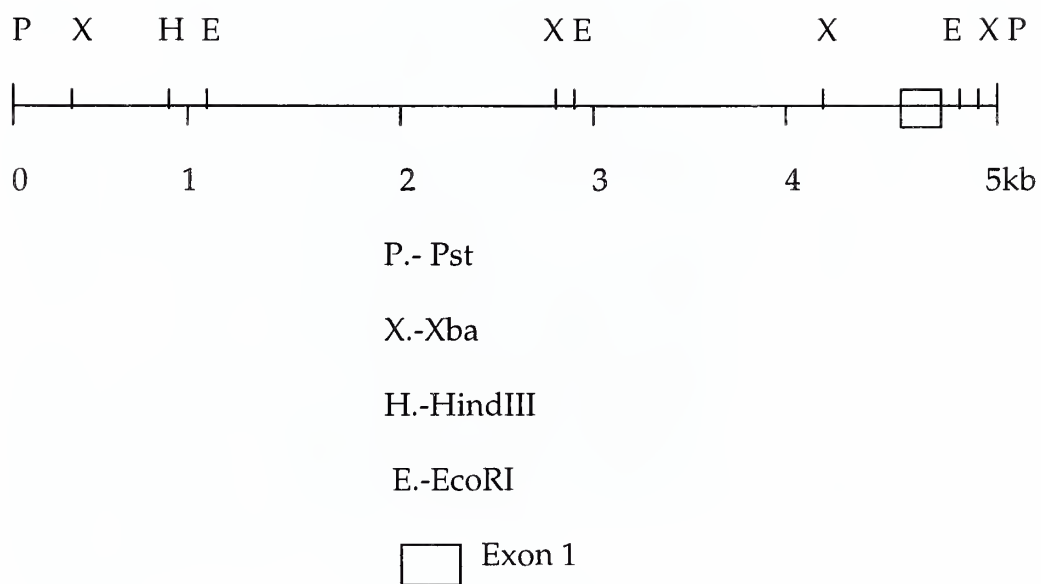


Figure 6.- Restriction map of 5' flanking region of MNC gene.

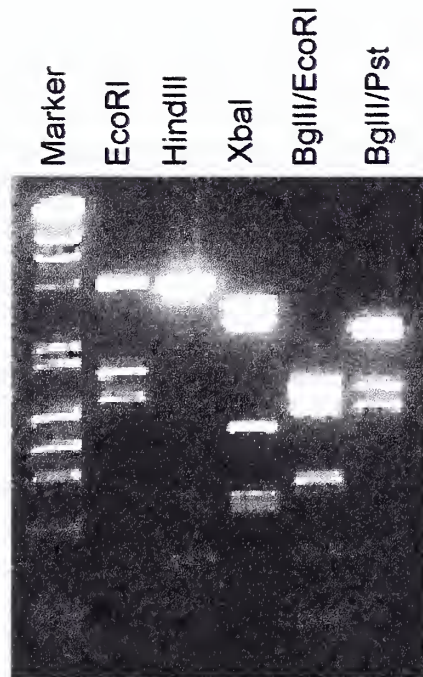


Figure 7.- Restriction digests of plasmid construct containing the 5kb insert. The cloning site is Pst and our molecular weight standard was Lambda DNA cut with HindIII and PhiX174 with HaeIII.

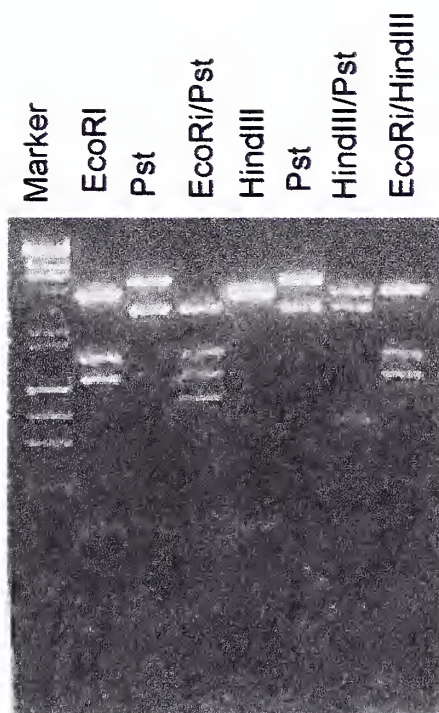


Figure 8.- Restriction digests of plasmid construct containing the 5kb insert. The cloning site is Pst and our molecular weight standard was Lambda DNA digested with HindIII and PhiX174 cut with HaeIII.

c-myb;C/EBP α binding element

5'	_____	Human	3'
AAATCACATTAT	AAAATTGCAAAACT	GTCAGAGTGTAATTAACCTATGTTGCTTCATA	
IIIIIIIIII	III IIII I	II IIIIIIIII IIIII	
AAATCACATTCT	AACATTGAAGAACT	GTCAGAGTGCGGTTAG-CTATGTTGCCTCATA	
5'		Mouse	3'

Figure 9.- Region of 81.4% identity in human and mouse neutrophil collagenase promoters. This region is italicized and in bold faces in the complete sequences reported in previous figures.

VIII. DISCUSSION

We report here the cloning of the 5' flanking region of the MNC gene, the first exon and the first intron (data not shown) as well as mapping studies of 4kb of the promoter of the MNC gene. We further report preliminary characterization of cis-acting 5' promoter sequences.

Our studies have been aimed at understanding the role of CDP in regulating the MNC gene within the context of the neutrophil secondary granule proteins. It has already been shown that CDP overexpression blocks the expression of all SGP genes including the MNC gene³⁸. It is postulated that a disruption in early myeloid differentiation mediates continued binding of CDP and thereby blocks SGP gene expression in abnormal myeloid differentiation.

The silencer element identified in the LF promoter is highly conserved between the mouse and human LF. The overall homology at the nucleotide level between the human and mouse LF promoters in this region is 82.5% with conservation of the octamer repeats suggesting a high functional importance in regulating gene expression³⁹. Our preliminary results suggest that CDP recognizes a different sequence in the MNC promoter, pending further upstream sequencing. This is not surprising since CDP lacks a well defined consensus binding sequence. Studies by Liu and Skalnik^{35,36,40} have shown differential affinity for the different DNA binding domains of CDP for different sites within the gp91 phox promoter. Furthermore, association of recombinant proteins through their GST portions increased binding affinity, suggesting that DNA binding activity is enhanced by potential interactions of different binding domains binding to different regions of the promoter⁴⁰. A putative silencer within the MNC promoter should be sought.

The role of the *cis*-acting sequence elements in the regulation of MNC remains to be established with physiologic studies. However several are of potential interest.

The transcription start site has not yet been mapped but perhaps the most important *cis* regulating elements lie within the 254 bp upstream of the untranslated region which share a high degree of conservation with the HNC promoter. As mentioned earlier that region corresponds to a C/EBP α canonical consensus sequence in the HNC promoter. Many studies suggest that in the hematopoietic system, C/EBP factors may be specifically expressed in myeloid cells, as indicated earlier. C/EBP α , C/EBP β and C/EBP δ have been detected in all myeloid cell lines. C/EBP α has been identified as the major form of these proteins that binds to the promoters of the M-CSF receptor and GM-CSF receptor at the α sites and trans-activate them in myeloid cell lines^{4,5}.

It has been shown that C/EBP β regulates the IL1 β promoter consistent with the role of this member of the family in activation by cytokines like IL6 and LPS. Targeted disruption of C/EBP β leads to defects in killing of bacteria and tumor cells by macrophages^{4,5}. It is not surprising to find a binding site for this transcription factor as one would expect an increase in MNC synthesis as part of the broader expansion of the granulocyte mass in the inflammatory response.

The transcription factor PU.1 has been previously indicated to be crucial for terminal myeloid gene expression and differentiation^{4,5}. Most of the myeloid promoters lack a TATA box. They are often dependent on an SP.1 site which confers specificity, activity and inducibility. Many of these promoters, including the MNC 5' flanking region, have a PU.1 binding site upstream of the transcription start site. It is thought that PU.1 activates these promoters by recruiting TATA binding protein^{4,5}. Functionality of the site remains to be shown.

Of note, many GATA-1 canonical consensus sequences as well as a c-myb binding site are reported. c-myb is involved in early hematopoiesis whereas GATA-1

has a predominant role in erythropoiesis. Interestingly we did not identify an MZF-1 element.

In summary, we have isolated two genomic clones which together include 4kb of the 5' flanking region of the MNC gene, the first exon and part of the second exon. The 5' flanking DNA contains several potential cis-acting regulatory elements whose physiologic role remains to be more fully defined. We have found the region immediately upstream of the translated sequence highly conserved between the human and mouse homologs over more than 200bp. Confirmation of the role of these sequences in mediating MNC expression, as well as identification of potential CDP binding sites, awaits functional analysis of the promoter.

ACKNOWLEDGEMENTS

This work has been funded by the Howard Hughes Medical Institute to which I am forever grateful. I wish to express my appreciation and thanks to Theresa Zibello who has been supportive in so many ways, Nathan Lawson for his technical expertise and Nancy Berliner for her guidance and inspiration. A.K. Gupta and N. Maun I couldn't have done it without you.

REFERENCES

-
- ¹Scott EW, Simon MC, Anastasi J, Singh H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* **265**: 1573-1577, 1994.
- ²Olson MC, Scott EW, Hack AA, Su GH, Tenen DG, Singe H, Simon MC. PU.1 is not essential for early myeloid gene expression but is required for terminal myeloid differentiation. *Immunity* **3**: 703-714, 1995.

-
- ³Torbet BE, Anderson KL, Connors K, Smith KA, Henkel G, Sinh CD, McKurcher SC, Maki RA. The disruption of the PU.1 gene results in the loss of GM-CSF and G-CSF receptors. *Blood* **88**: 629a, 1996.
- ⁴Zhang DE, Hohaus S, Voso MT, Chen H-M, Smith LT, Hetherington CJ, and Tenen DG. Function of PU.1, C/EBP, and AML1 in early myeloipoiesis: Regulation of multiple myeloid CFS receptor promoters. *Current Topics in Microbiology and Immunology*, L Wolff and AS Perkins, eds. Springer-Verlag, Berlin, p. 137, 1996.
- ⁵Zhang P, Zhang D-E, Wang N, Hetherington CJ, Darlington GJ, and Tenen DG. Absence of G-CSF signaling and neutrophil development in C/EBP α deficient mice. *Blood* **88**: 630a, 1996.
- ⁷Hromas R, Collins SJ, Hickstein D, et al. A retinoic acid-responsive human zinc finger gene, MAF1, preferentially expressed in myeloid cells. *J Biol Chem* **266**: 14183-14187, 1991.
- ⁸Bavisotto L, Kaushansky K, Lin N, Hromas R. Antisense oligonucleotides from the stage-specific myeloid zinc finger gene MZF-a inhibit granulopoiesis in vitro. *J Exp Med* **174**: 1097-101, 1991.
- ⁹Mucenski ML, McClain K, Kier AB, Swerdlow SH, Schreiner CM, Miller TA, Pietryga DW, Scott WJ, Potter SS. A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. *Cell* **65**: 677, 1991.
- ¹⁰Sporn MB, Roberts AB. Role of retinoids in differentiation and carcinogenesis. *Cancer Res* **43**: 3034-3040, 1983.
- ¹¹Janoschowitz H, Moore MAS, Metcalf D. Density gradient segregation of bone marrow cells with the capacity to form granulocytic and macrophage colonies in vitro. *Exp Cell Res* **68**: 220, 1971.
- ¹²Warner HR, Athens JW. Analysis of granulocyte kinetics in blood and bone marrow. *Ann NY Acad Sci* **113**: 523, 1964.
- ¹³Bainton DF. Neutrophil granules. *Br. J Hematol* **29**: 19, 1975.
- ¹⁴Bainton DF, Ulliyot JL, Farquhar MG. The development of the neutrophil secondary granule. *J Exp Med* **134**: 907, 1971.
- ¹⁵Graubert T, Johnston J, Berliner N. Cloning and expression of the cDNA encoding mouse neutrophil gelatinase: Demonstration of coordinate secondary granule protein gene expression during terminal neutrophil maturation. *Blood* **82**: 3192, 1993.
- ¹⁶Lawson N, Khanna-Gupta A, Berliner N. (1998). Isolation and Characterization of the cDNA for mouse neutrophil collagenase. Demonstration of shared negative regulatory pathways for neutrophil secondary granule protein gene expression. *Blood* **91** (7), (April 1, 1998)

-
- ¹⁷Berliner N, Hsing A, Sigurdsson F, Zain M, Graubert T, Bruno E, Hoffman R. G-CSF induction of normal human bone marrow progenitors results in neutrophil-specific gene expression. *Blood*, **85**:799, 1995.
- ¹⁸Johnston J, Rintels P, Chung J, Sather J, Benz EJ, Berliner N. The lactoferrin gene promoter: Structural integrity and non-expression in HL60 cells. *Blood* **79**: 2998, 1992.
- ¹⁹Khanna-Gupta A, Kolibaba K, Berliner N. NB4 cells exhibit bilineage potential and an aberrant pattern of neutrophil secondary granule protein gene expression. *Blood* **84**: 294, 1994.
- ²⁰Johnston J, Boxer LA, Berliner N. Correlation of RNA levels with protein deficiencies in specific granule deficiency. *Blood* **80**: 2088, 1992.
- ²¹Lomax KJ, Gallin JI, Benz EJ, Rado TA, Boxer LA, Malech HL. Selective defect in myeloid cell lactoferrin gene expression in neutrophil specific granule deficiency. *J Clin Invest* **83**: 514, 1989.
- ²²McCormick A, Wu D, Castrillo J-L, Dana S, Strobl J, Thompson EB, Karin M. Extinction of growth hormone expression in somatic cell hybrids involves repression of the specific *trans*-activator GHF-1. *Cell* **55**: 379, 1988.
- ²³Miyazaki J, Appella E, Ozato K. Negative regulation of the major histocompatibility class I gene in undifferentiated embryonal carcinoma cells. *Proc Natl Acad Sci USA* **83**: 9537, 1986.
- ²⁴Blau H. Differentiation requires continuous active control. *Ann Rev Biochem* **61**: 1213, 1992.
- ²⁵Baniahmad A, Muller M, Steiner C, Renkawitz R. Activity of two different silencer elements of the chicken lysozyme gene can be compensated by enhancer elements. *EMBO J.* **6**: 2297, 1987.
- ²⁶Park K, Atchison ML. Isolation of a candidate repressor/activator, NF-E1 (YY-1, delta) that binds to the immunoglobulin kappa 3' enhancer and the immunoglobulin heavy chain mu E₁ site. *Proc Natl Acad Sci USA* **88**: 9804, 1991.
- ²⁷Barberis A, Superti-Furga G, Busslinger M. Mutually exclusive interaction of the CCAAT-binding factor and of displacement protein with overlapping sequences of a histone gene. *Cell* **50**: 347-359, 1987.
- ²⁸Andres V, Nadal-Ginard B, Mahdavi V. *Clox*, a mammalian homeobox gene related to *Drosophila cut* encodes DNA-binding regulatory proteins differentially expressed during development. *Development* **116**: 324-334, 1992.
- ²⁹Valarche I, Tissier-Seta JP, Hirsch MR, Martinez S, Gordis C, Brunet J-F. The mouse homeodomain protein Phox-2 regulates NCAM promoter activity in concert with Cux/CDP/cut and is a putative determinant of neurotransmitter phenotype. *Development* **119**: 881-896, 1993.

-
- ³⁰Andres V, Nadal-Ginard B, Mahdavi V. *Clox*, a mammalian homeobox gene related to *Drosophila cut* encodes DNA-binding regulatory proteins differentially expressed during development. *Development* **116**: 324-334, 1992.
- ³¹Blochinger K, Bodmer R, Jack J, Lan LY, Jan YN. Primary structure and expression of a product from *cut*, a locus involved in specifying sensory organ identity in *Drosophila*. *Nature (London)* **333**: 629-635, 1988.
- ³²Aufiero B, Neufeld EJ, Orkin SH. Sequence specific DNA binding of individual *cut* repeats of the human CCAAT displacement/*cut* homeodomain protein. *Proc Natl Acad Sci USA* **91**: 7757-7761, 1994.
- ³³Cowell IG. Repression versus activation in control of gene expression. *Trend Biochem Sci* **19**: 38-42, 1994.
- ³⁴van Wijnen AJ, van Gurp MF, de Ridder MC, Tufarelli C, Lat TJ, Birnbaum M, Vaughan PS, Giordano A, Krek W, Neufeld EJ, Stein JL, Stein GS. CDP/*cut* is the DNA-binding subunit of histone gene transcription factor HiNF-D: A mechanism for gene regulation at the G1/S phase cell cycle transition point independent of transcription factor E2F. *Proc Natl Acad Sci USA* **93**: 11516-11521, 1996.
- ³⁵Skalnik DG, Strauss EC, Orkin SH. CCAAT displacement protein as a repressor of the myelomonocytic-specific gp91-phox gene promoter. *J Biol Chem* **266**: 16736, 1991.
- ³⁶Neufeld EJ, Skalnik DG, Lievens PM, Orkin SH. Human CCAAT displacement protein is homologous to the *Drosophila* homeoprotein, *cut*. *Nature Genetics* **1**: 50, 1992.
- ³⁷Ford AM, Bennett CA, Healy LE, Towatari M, Greaves MF, Enver T. Regulation of the myeloperoxidase enhancer binding proteins Pu1, C-EBP α , β , and γ during granulocyte-lineage specification. *Proc Natl Acad Sci USA* **93**: 10838-10843, 1996.
- ³⁸Lawson N., Khanna-Gupta A., and Berliner N. Isolation and characterization of the cDNA for the mouse neutrophil collagenase. Demonstration of shared negative regulatory pathways for neutrophil secondary granule protein gene expression. *Blood* **91** (7), (April 1 1998)
- ³⁹Khanna-Gupta A., Zibello T., Kolla S., Neufeld E., and Berliner N.. CCAAT Displacement Protein (CDP/*cut*) recognizes a silencer element within the lactoferrin gene promoter. *Blood* **90** (7):2784-2795, 1997
- ⁴⁰Catt DM, Luo W, Skalnik DG. Study of the novel DNA-binding domains of CCAAT displacement protein, a repressor of the gp91-phox promoter. *Blood* **88**: 47a, 1996.

HARVEY CUSHING / JOHN HAY WHITNEY
MEDICAL LIBRARY

MANUSCRIPT THESES

Unpublished theses submitted for the Master's and Doctor's degrees and deposited in the Medical Library are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but passages must not be copied without permission of the authors, and without proper credit being given in subsequent written or published work.

This thesis by _____ has been
used by the following persons, whose signatures attest their acceptance of the
above restrictions.

NAME AND ADDRESS

DATE

YALE MEDICAL LIBRARY



3 9002 01048 7339

