uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Cathy Xin Yue Zhang

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Cellular and Molecular Medicine)

GRADE / DEGREE

Faculty of Medicine – Department of Cellular and Molecular Medicine

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Investigating the Function of Hemogen During Erythroid Differentiation

TITRE DE LA THÈSE / TITLE OF THESIS

Marjorie Brand

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

Laura Trinkle-Mulcahy

Marc Ekker

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Investigating the Function of Hemogen during Erythroid Differentiation

Cathy Xin Yue Zhang

This thesis is submitted as a partial fulfillment of the MSc degree in Cellular and Molecular Medicine.

February 16, 2010

Faculty of Medicine
Faculty of Graduate and Postdoctoral Studies
University of Ottawa
Ottawa, Ontario, Canada

Your file Votre référence
ISBN: 978-0-494-74154-2
Our file Notre référence
ISBN: 978-0-494-74154-2

Abstract

Hemogen is a murine protein expressed in active sites of hematopoiesis during embryogenesis and in the adult. Hemogen's function has not been clearly elucidated. Evidences suggest Hemogen's role in erythroid differentiation. Hemogen expression coincides with active sites of erythropoiesis. *β-globin* expression is a hallmark of erythroid differentiation. Proteomic data suggests that Hemogen associates with MafK which is part of the NF-E2 complex that acts as an activator of *β-globin* during erythroid differentiation. To gain insight into Hemogen's mechanistic function during erythroid differentiation, we attempted to decipher Hemogen's associating proteins during erythroid differentiation. We generated a Hemogen antibody functional in Hemogen immunoprecipitation (IP). We performed Hemogen IP and used mass spectrometry to identify putative Hemogen interacting proteins within the Hemogen IP elution. Our immuno-fluorescence microscopy (IFM) images of endogenous Hemogen revealed Hemogen as a nuclear protein within murine erythroleukemia cells before and after differentiation. In line with our IFM images, the majority of candidates identified within our putative list of Hemogen associating proteins are transcription regulators. Hemogen expression progressively increased with erythroid differentiation. Hence, Hemogen may function as a transcriptional regulator during erythroid differentiation.

List of Tables

List of Figures

List of Abbreviations

ATP	**Adenosine Triphosphate**
ATPase	**Adenosine Triphosphatase**
BRG-1	**Braham-Related Gene-1**
BAF	**BRG-1 Associating Factor**
BFU-E	**Blast Forming Unit-Erythroid**
BSA	**Bovine Serum Albumine**
bZIP	**Basic Leucine Zipper**
Chd4	**Chromodomain Helicase DNA Binding Protein 4**
CFU-E	**Colony Forming Unit-Erythroid**
CBP	**cAMP response element-binding (CREB) protein binding protein**
ChIP	**Chromatin Immunoprecipitation**
DAPI	**4', 6-diamidino-2-phenylindole**
DMSO	**Dimethyl Sulfoxide**
DNA	**Deoxyribonucleic Acid**
DNAse I	**Deoxyribonuclease I**
DNMT	**DNA methyltransferase**
DTT	**1, 4-Dithithreitol**
EDAG	**Erythroid differentiation-associated protein**
EDTA	**Ethyenediaminetetraacetic Acid**
EKLF	**Erythroid Krüppel-lke Factor**
FACS	**Fluorescence Activated Cell Sorting**

FOG-1	**Friend of GATA-1**
FPLC	**Fast Performing Liquid Chromatography**
GATA-1	**GATA-binding factor 1**
GATA-2	**GATA-binding factor 2**
HDAC	**Histone Deacetylase**
HPLC	**High-Performance Liquid Chromatography**
HS	**Hypersensitive Site**
HSC	**Hematopoietic Stem Cell**
IgG	**Immunoglobulin G**
ISWI	**Imitation Switch**
IP	**Immunoprecipitation**
IPI	**International Protein Index**
IPTG	**Isopropyl β-D-1-thopgalactopyranoside**
LB	**Luria Broth**
LC-ESI-MS-MS	**Liquid Chromatography-Electrospray Ionization-Tandem Mass Spectrometry**
LCR	**Locus Control Region**
LMO2	**LIM Domain Only 2**
Ldb1	**LIM Binding Domain 1**
MARE	**Maf Recognition Element**
MEL	**Murine Eythroleukemia**
mRNA	**Messenger Ribonucleic Acid**
NF-E2	**Nuclear Factor- Erythroid 2**

NuRD	**Nucleosome Remodelling and Histone Deacetylase Complex**
NURF	**Chromatin Remodeling Factor**
PBS	**Phosphate Buffered Saline**
PCNA	**Proliferation Cell Nuclear Antigen**
PCR	**Polymerase Chain Reaction**
PCV	**Pack Cell Volume**
PIC	**Protease Inhibitor Cocktail**
PMSF	**Phenylmethanesulfonyl Fluoride**
RBBP4	**Retinoblastoma-Binding Protein 4**
RBBP7	**Retinoblastoma-Binding Protein 7**
SDS	**Sodium Dodecyl Sulfate**
SDS-PAGE	**Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis**
SEC	**Size Exclusion Chromatography**
SMC	**Structural Maintenance of Chromosomes**
smcDH1	**Structural Maintenance of Chromosomes Hinge Domain 1**
SWI/SNF	**Switch/Sucrose Non-fermenting**
SWR1	**Switch2/Snf2 Related 1**
Sp1	**Specific protein 1**
TAL-1	**T-Cell Acute Lymphocytic Leukemia 1 (transcription factor)**
TCA	**Trichloric Acid**
USF	**Upstream Transcription Factor**

Table of Contents

Acknowledgements

This thesis cannot be completed without the support of generous people who have given me encouragement, advice and help along the way.

I owe my gratitude to my supervisor, Dr. Marjorie Brand who has always made her invaluable support available. I would like to sincerely thank her for her patient guidance in experimental techniques, her insightful advices and her encouragements along the way. I am also grateful to Dr. Jeffrey Dilworth for his generous advices and encouragements throughout this project.

I would like to thank my committee members: Dr. Marc Ekker and Dr. David Picketts for their insightful input throughout this project. I am indebted to Hong Ming, Dr. Yoichi Kawabe, Dr. Lawrence Puente and Feodor Price who aided me in countless ways. It is my pleasure to thank the talented post-doctoral fellows: Carmen Geogiana Palii, Chandra-Prakash Chaturvedi, Alison Hosey, Arif Aziz, Qi-Cai Liu and Kulwant Singh for their expert guidance and help. I am also grateful for the wonderful colleagues: Patricia Rakopoulos, LiFang Li and Jianguo Wu for their loving support.

Lastly, I would like to offer my deepest gratitude to my mom and dad for their unconditional support.

Chapter 1: Introduction

1.1 Overview

This thesis provides the initial description of Hemogen expression and possible function during erythroid differentiation in murine erythroleukemia (MEL) cells. Chapter 1 of this thesis introduces main themes of the thesis and rationale for the hypothesis. This chapter begins by introducing the process of erythrocyte differentiation and MEL cells used in this study as the model system of erythroid differentiation (1.2). The next sections introduce transcriptional regulation during erythroid differentiation (1.3) and structure (1.4) and regulation of the human *β-globin* locus (1.4). The previously identified associating Hemogen protein, MafK is introduced along with its role in *β-globin* regulation (1.5). Hemogen's expression dynamics and a summary of its possible role in erythroid differentiation are outlined (1.6). The final sections are statements of the objective (1.7) and the hypothesis (1.8).

1.2 Erythropoiesis

Mammalian erythrocytes or red blood cells are specialized cells that carry oxygen from the lungs and facilitate oxygen and carbon dioxide exchange within internal tissues of aerobic organisms (Hoffman et al. 2005). Erythrocytes are the most abundant cell in the human body with more than 25 trillion in the blood stream (Palis 2009). Erythrocyte lack organelles and are composed mainly of hemoglobin which constitutes more than 90% of the protein content in the mature erythrocyte (Palis 2009). Hemoglobin attributes to erythrocyte function (Hoffman et al. 2005). In adults, hemoglobin is a tetramer of two α- and two β-globin polypeptides. Each globin chain associated with an iron ion (Fe^{2+})-containing heme group to which oxygen reversibly binds. Mature mammalian red cells

are biconcave discs (Palis 2009). This shape allows for cellular deformability and facilitates navigation through capillary networks (Palis 2009). In adult, erythrocytes have a life-span of approximately 120 days in humans and 50 days in mice, therefore erythrocytes must be constantly replenished (Palis 2009 and Saxena et al. 2009). Hematopoietic stem cells (HSC) are self-renewing cells that give rise to all mature blood cells including erythrocytes (Hoffman et al. 2005) (figure 1).

Erythropoiesis is the differentiation program that results in the production of erythrocytes from hematopoietic stem cells. Erythropoiesis proceeds through cell intermediates and can be considered as three stages (Palis 2009). The first stage consists of early erythroid progenitors. Erythroid progenitors are sparse and difficult to isolate in sufficient numbers and purity. Their existence is therefore inferred in their capability to generate hemoglobinized progeny in *ex vivo* colonal erythroid cultures (Hoffman et al. 2005). The earliest erythroid-specific progenitor is the BFU-E (burst forming unit-erythroid) (Ingley et al. 2004; Hoffman et al. 2005 and Palis 2009). BFU-Es are not actively proliferating with most cells in the G_0/G_1 phase of cell cycle (Koury et al. 2002 and Hoffman et al. 2005). BFU-E generates the more mature erythroid-committed progenitors or CFU-E (colony forming units-erythroid) (Hoffman et al. 2005). CFU-Es are actively proliferating with most cells in S phase (Koury et al. 2002).

The second stage of erythroid differentiation consists of morphologically identifiable erythroid precursors. In mammals, this stage is characterized by progressive accumulation of hemoglobin, decrease in cell size and chromatin condensation (Palis

2009). As erythroid precursors differentiate, they cease to divide and accumulate at G_0 phase before enucleation (Koury et al. 2002). Proerythroblasts derived from CFU-E are the first morphologically recognizable erythroid cells (Ingley et. al 2002 and Hoffman et al. 2005). The cells are smaller than previous stages of differentiation (Hoffman et al. 2005). As proerythroblasts develop into basophilic erythroblasts, chromatin condensation occurs and an increased abundance of ribosomes causes the cytoplasm to become basophilic when stained by basic dye haematoxylin (Hoffman et al. 2008). As mitotic divisions continue, cells develop into polychromatophilic erythroblasts and produce hemoglobin making the cytoplasm both basophilic and eosinophilic (stained pink with acidic dye eosin Y) (Hoffman et al. 2008). Hemoglobin continues to accumulate in the cytoplasm making it more eosinophilic thus resulting in acidophilic orthochromatic erythroblasts (Hoffman et al. 2008). Orthochromatic erythroblasts eventually extrude their nuclei giving rise to enucleated reticulocytes (Palis 2009).

The third stage of erythroid differentiation consists of reticulocyte maturation. Transcription ceases with the degradation of ribosomal machinery as reticulocytes terminally differentiate into erythrocytes (Palis 2009). Active loss of all other organelles including endoplasmic reticulum and mitochondria occurs through internal degradation. Reticulocytes enter the bloodstream where the remaining ribosomes are lost and mature into erythrocytes (Palis 2008 and Palis 2009) (figure 1).

Murine erythroleukemia cells (MEL) are an established model system used to understand the mechanism of erythroid differentiation and leukemogenesis (Tsiftsoglou

et al. 2003a and Tsiftsoglou et al. 2009). MEL cells were created from murine tumors of Friend virus induced erythroleukemia (Friend et al. 1966). MEL cells are committed to the erythroid lineage and resemble proerythroblasts (Tsiftsoglou et al. 2009). Grown in suspension culture, MEL cells can proliferate indefinitely. However upon exposure to dimethyl sulfoxide (DMSO), MEL cells terminally differentiate into erythroid cells resembling orthochromatic erythroblasts (Tsiftsoglou et al. 2009). Over four days of DMSO treatment, cellular alterations of differentiated MEL cells include: decrease in cell size, activation of iron-transport, *de novo* heme biosynthesis, hemoglobin production, expression of erythrocytic antigens and spectrin, nuclear condensation and limitation in proliferative capacity leading to growth arrest at G_1 phase (Tsiftsoglou and Wong 1985 and Tsiftsoglou et al. 2003b). The emergence of erythroleukemias in Friend virus-infected mice is often associated with proviral insertion at the *Spi-1* locus which leads to deregulated expression of the PU.1 transcription factor (Paul et al. 1989). PU.1 inhibits GATA-1, a key stimulator of red blood cell differentiation, by direct interaction. Chemical induction of MEL differentiation results in early down-regulation of PU.1 levels thereby allowing GATA-1 to mediate erythroid differentiation (Delgado et al. 1994 and Rao et al. 1997).

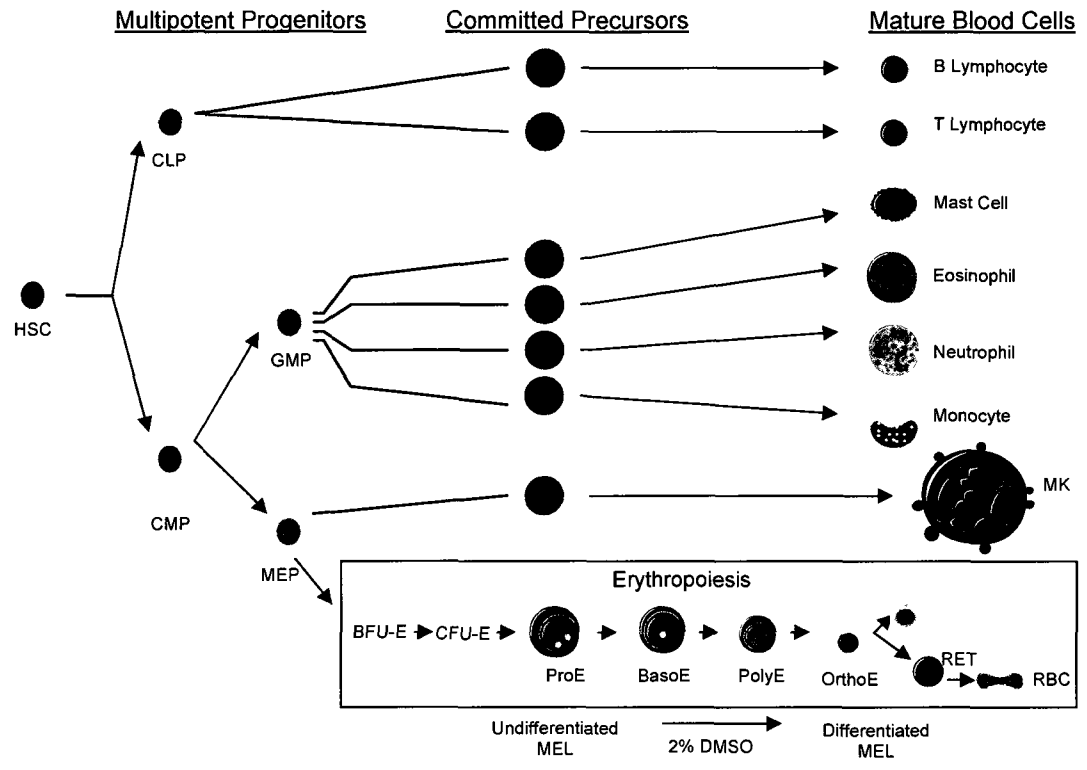


Figure 1: **Schematic of erythropoiesis in the context of hematopoiesis.** The pluripotent hematopoietic stem cell (HSC) can give rise to multipotent progenitors such as common myeloid progenitor (CMP) and common lymphoid progenitor (CLP). CMP can give rise to megakaryocyte/erythroid progenitors (MEP) and granulocyte/macrophage progenitors (GMP). Through a series of committed precursors, CLP can give rise to mature B lymphocytes or T lymphocytes, GMP can give rise to mature cells, eosinophil, neutrophil or monocyte and MEP can give rise to megakaryocytes (MK) or mature red blood cells (RBC). In erythropoiesis, erythroid progenitors such as BFU-E and CFU-E differentiate into pro-erythroblasts (ProE). Pro-erythroblasts give rise to basophilic erythroblasts (BasoE), polychromatophilic erythroblasts (PolyE) and orthochromatic erythroblasts (OrthoE). Enucleation of orthochromatic erythroblasts results in reticulocytes (RET). Reticulocytes mature in the blood-stream into red blood cells (RBC) or erythrocytes. Undifferentiated MEL cells resemble proerythroblasts. Differentiated MEL cells using 2% DMSO resemble orthochromatic erythroblasts.

1.3 Transcriptional Regulation of Erythroid Differentiation

Erythroid cells exert their function through the expression of erythroid specific genes that are regulated by lineage-specific transcription factors (Kim and Bresnick 2007). As intrinsic determinants of cellular phenotype, the study of transcription factor function provides insight into understanding the developmental program of erythroid differentiation (Perry and Soreq 2002 and Orkin and Zon 2008). Gene knockout experiments in transgenic mice and other model organisms as well as forced expression experiments of transcription factors revealed their importance in erythroid differentiation (Pevny et. al. 1991; Tsang et al. 1998; Iwasaki et al. 2003 and Mikkola et al. 2003). Moreover proper function and regulation of transcription factors is necessary to normal hematopoiesis as the aberrant function of many hematopoietic transcription factors are involved in human hematopoietic malignancies (Tenen et al. 2003; Shimizu et al. 2008 and Orkin and Zon 2008).

The eukaryotic genome along with chromosomal proteins is organized into chromatin. The nucleosome is the fundamental organizational unit of chromatin and consists of a core histone octamer (two copies of H2A, H2B, H3 and H4) encircled by 146 bp of DNA (Ruthenburg et al. 2007). Sequence specific hematopoietic transcription factors bind to DNA motif within chromatin and recruit co-factors via protein-protein interactions (Demers et al. 2007; Kim et al. 2007; Chaturvedi et al. 2009 and Kim, S-I. et al. 2009). Co-factors usually exist within multi-protein complexes. These protein complexes can have chromatin modifying and/or chromatin remodeling enzymatic activities that mediate transcriptional activation or repression (Kim and Bresnick 2007).

Chromatin modifying enzymes catalyze the addition and removal of post-translational modifications or epigenetic marks (such as acetylation, methylation ubiquitination and sumoylation) that commonly occur on N-terminal tails of core histones (Weaver 2005 and Kim and Bresnick 2007). Epigenetic marks on histone tails may affect DNA-histone interaction within a nucleosome and/or the folding of nucleosome arrays. The significance of most epigenetic marks are not fully understood, however lysine acetylation and methylation are recognized as modulator marks for transcriptional activation or repression (Kouzarides 2007). Acetylation of lysines on H3 and H4 usually correlates with chromatin accessibility and transcription. Acetylation of histone tails neutralizes positive charge of the lysines on histone tails thereby destabilizing DNA-histone and nucleosomal array interactions (Wolffe et al. 2000 and Weaver 2005). This results in an open chromatin environment for transcription factor access to nucleosomal DNA. Histone acetylation is a reversible process. Histone tail deacetylation catalyzed by histone deacetylases reverses these effects and results in a more repressive chromatin environment (Kim, J. et al. 2009). In addition, specific epigenetic marks such as acetyl-lysine can function as ligands to recruit additional regulatory factors (Fischle et al. 2003). Histone lysine methylation has varied effects depending on the number of methyl groups and position of lysine residues. For example, trimethylation of lysine 9 and 27 on H3 are localized to inactive genes while trimethylation of H3 lysine 4 is associated with active transcription. In some cases, H3 K9 tri-methylation has also been found in actively transcribed genes (Vakoc et al. 2006).

In addition to histone modifications, DNA methylation is another epigenetic mark (Kim, J. et al. 2009). The addition of methyl group to the fifth carbon position of cytosine bases CpG dinucleotides is associated with transcription repression. DNA methylation may directly interfere with transcription factor binding or through the recruitment of methyl cytosine-binding proteins that recruit repressive histone modifying co-repressor complexes (Fuks 2005 and Kouzarides 2007). DNA methylation can be inherited through generations of somatic cells and is mediated by DNA methyltransferases. Mammalian DNA methyltransferases (DNMT) can be grouped into two major classes. DNMT3A and DNMT3B are *de novo* methyltransferases that establish cytosine methylation patterns at unmethylated DNA during early embryogenesis while DNMT1 copies pre-existing methylation patterns onto newly synthesized DNA strands during replication (Jones and Liang 2009). In human, CpG dinucleotides are clustered into CpG islands. CpG methylation is generally associated with repetitive sequences that must remain silent. CpG dinucleotides within CpG island promoters of constitutively active genes, such as housekeeping genes, are typically unmethylated in normal tissues (Kim, J. et al. 2009).

Chromatin remodeling complexes alter histone-DNA interaction by ATP hydrolysis to regulate nucleosomal DNA accessibility. Nucleosomes may occupy key *cis*-regulatory binding sites of regulated genes. These *cis*-regulatory sites may be reliant on chromatin remodeling complexes to 'uncover' these sites (Cairns 2009). Transcriptional activation may occur by transcription factor binding to exposed binding sites in the linker DNA or nucleosome edge followed by the recruitment of remodeling

and modifying complexes to expose additional *cis*-regulatory sites. Alternatively, remodeling complexes may cause nucleosomes to mask binding sites of transcription factors (Cairns 2009). All chromatin remodeling complexes consist of a catalytic subunit with an ATPase domain (Saha et al. 2006). Several chromatin remodeling families have evolved distinct mechanisms of restructuring the nucleosome. Remodellers can organize nucleosomal arrays to promote repression (ISWI and NURF families), incorporate histone variants to promote activation (SWR1 family) or slide and eject nucleosome (SWI/SNF family) (Cairns 2009). Chromatin remodeling complexes also work in concert with chromatin modifying complexes. The NuRD complex consists of the chromatin remodeling CHD ATPase as well as histone deacetylases HDAC1/2. In some cases epigenetic marks also serve as ligands to attract additional regulatory factors to the chromatin. Bromodomains of the catalytic subunit of the SWI/SNF remodeling complex recognize acetylated histones while chromodomains of CHD recognize repressive methylated histones (Becker and Horz 2002 and Cairns 2009).

Erythropoiesis progresses through cellular hierarchies with differential gene expression. Hematopoietic stem cells are 'primed' to develop into blood cell types by expressing lineage restricted genes from multiple lineages. Gene expression during erythropoiesis is characterized by global gene repression and expression of erythroid specific genes (Ney 2006). Transcription regulatory motifs are constructed based on current data to understand how transcription regulators mediate erythroid lineage specifications. Lineage choices during hematopoiesis can be controlled by the levels of key transcription factors with a mutually antagonistic relationship. For instance, the

mutually antagonistic relationship between PU.1 and GATA-1 is hypothesized to be important in the choice between myeloid or megakaryotic/erythroid lineage. PU.1 and GATA-1 directly inhibit one another's activity (Swiers et al. 2006). Temporal control of gene expression can be mediated by feed-forward motifs. Feed-forward motifs arise when one transcription factor regulates an intermediate with either or both needed for regulation of the third target gene. For instance, c-myb is a transcription factor associated with erythroid progenitor proliferation and erythroid commitment during early stages of hematopoiesis but must downregulated during terminal erythroid differentiation (Greig et al. 2008). Expression of GATA-1 initally activates *c-myb* however as GATA-1 activates its co-factor FOG-1, *c-myb* is downregulated by GATA-1 and FOG-1 (Swier et al. 2006).

1.4 β-Globin Locus

The *β-globin* locus serves as a model for studying tissue-specific gene expression. Gene activity during erythroid maturation is dominated by globin expression (Hoffman 2005). The human *β-like globin* locus contains a group of genes, *ε-G γ-A γ- δ- β* ordered in sequence to their expression during development (figure 2) (Stamatoyannopoulos 2005). During human development, there are two major transitions in *β-globin* gene expression that occur: one from the embryonic *ε-globin* gene to fetal *γ-globin genes* and another from fetal *γ-globin genes* to adult *δ* and *β globin* genes. In primitive human embryonic erythropoiesis, *ε-globin* is expressed in large nucleated erythroid cells of the yolk sac blood islands. During fetal development, definitive erythropoiesis takes place in the fetal liver and produce circulating enucleated erythrocytes in which *γ-globin* expression predominates. Close to the time of birth and post-natally, the bone marrow

becomes the major site of erythropoiesis and *β-globin* gene becomes highly transcribed along with low levels of *δ-globin* expression (Kiefer et al. 2008). Epigenetic marks of the *β-globin* genes change during various developmental stages with the transcriptional activity of the gene. For instance, methylation of CpGs found in the 5' regions of *β-globin* genes at various stages of development is inversely correlated with the transcriptional activity of the gene. Positive histone modifications are acquired at active globin genes during differentiation while silencing histone modifications are associated with genes before their activation and after they are silenced (Kiefer et al. 2008).

High expression of the *β-globin* genes in erythoid cells depends upon an upstream enhancer region known as the locus control region (LCR) (Grosveld 1999). The LCR is approximately 6-25 kb upstream of the *ε-globin* gene (Tuan et al. 1985) and comprises of four erythroid-specific DNAse I hypersensitivity sites (HS) and one ubitquitous HS each 200-400 bp in size separated by 2-4 kbp of DNA (Liang et al. 2008). HSs contain binding sites for erythroid-specific transcription factors (such as EKLF, GATA-1 and NF-E2) and ubiquitous transcription factors (such as Sp1 and USF) (Liang et al. 2008). Its has been proposed that transcription factors bound at the LCR interact with each other to efficiently capture transcriptional complexes which are then recruited to basal globin promoters. This model suggests that the LCR and the globin promoter are brought into close proximity via looping of the intervening DNA to form an active chromatin hub (Tolhuis et al. 2002 and de Laat and Grosveld 2003). Recent assays that determine proximity between genes and regulatory elements demonstrate a close proximity between

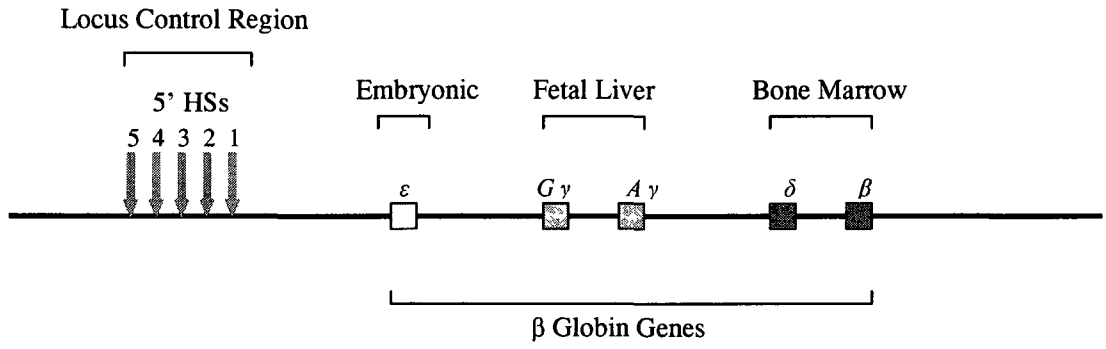


Figure 2: Diagram of the human *β-globin* locus. The five β-like globin genes are arranged in order of their expression during development. The locus control region is an upstream enhancer that is made of five DNAse I hypersensitive sites (grey arrows) containing many transcription factor binding sites including the Maf recognition element.

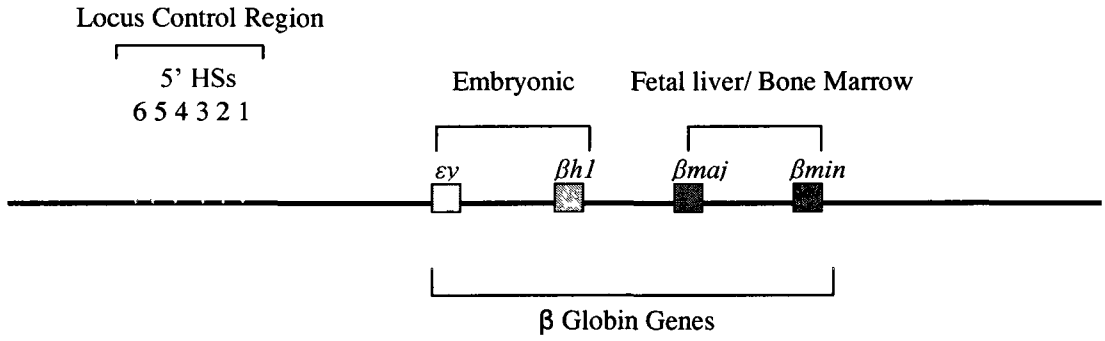


Figure 3: Diagram of the mouse *β-globin* locus. Embryonic globin genes are expressed in the yolk sac with the onset of erythropoiesis and silenced during definitive erythropoiesis in the fetal liver. The adult *βmaj* and *βmin* genes are expressed in the fetal liver and later in the bone marrow. . The locus control region is an upstream enhancer that is made of five DNAse I hypersensitive sites (grey arrows) containing many transcription factor binding sites including the Maf recognition element.

the LCR and actively transcribed globin genes at the active chromatin hub (Palstra et al. 2003).

1.5 MafK

The potential role of MafK in *β-globin* expression and erythroid differentiation was first noticed when 5'HS2 of the *β-globin* LCR was discovered to contain tandem Maf recognition elements (MARE) and conferred strong erythroid-specific enhancer activity (Mignotte et al. 1990; Moi and Kan 1990 and Ney et al. 1990). MafK is a basic leucine zipper (bZIP) transcription factor and belongs to the Maf family of transcription factors (Motohashi et al. 2002). bZIP proteins consist of a basic domain for DNA binding and an α-helix leucine zipper involved in dimer formation with other Mafs or bZIP proteins. Maf proteins are ubiquitously expressed (Motohashi et al. 2002).

Investigation of the protein that bound to the Maf recognition element (MARE) motif resulted in the identification of the NF-E2 complex (Andrews et al. 1993). NF-E2 is an obligate heterodimer composed of two subunits: MafK and a hematopoietic bZIP factor p45 (Andrews et al. 1993). NF-E2 is expressed in erythroid, megakaryocytic and mast cells (Andrews et al. 1993). During terminal erythroid differentiation, NF-E2 binds to its recognition element within the HS2 of the *β-globin* locus (Brand et al. 2004). Activity of the NF-E2 complex is necessary for *βmaj globin* expression within murine erythroleukemia (MEL) cells (Kotkow and Orkin 1995).

MafK's role in *β-globin* expression regulation was observed in a quantitative proteomic screen of MafK before and after erythroid differentiation in MEL cells (Brand et al. 2004). The study revealed MafK as a "dual function" protein that dynamically associates with co-repressive complexes before differentiation and co-activators after differentiation at the *β-globin* LCR and *βmaj* promoter (Brand et al. 2004). MafK preferentially associates with deacetylase complexes such as NuRD and SIN3 before differentiation which may actively repress *β-globin* activity (Brand et al. 2004). In response to DMSO erythroid differentiation, MafK preferentially associates with activating TAL-1 complexes and acetyltransferases such as TIP and CBP (Brand et al. 2004). Interestingly, the proteomic screen of MafK also identified a protein called Hemogen that preferentially associates with MafK after erythroid differentiation (Brand et al. 2004).

1.6 Hemogen

Hemogen is a murine protein with a theoretical molecular weight of 55.043 KDa (Yang et al. 2001) of unclear function. Analysis of Hemogen's primary sequence revealed an N-terminal basic domain and a C-terminal acidic domain (Yang et al. 2001). The functions of these domains have not been elucidated. Embedded within the N-terminal basic domain is a nuclear localization signal and a coiled-coil domain implicated in protein-protein interactions (figure 4). The human homolog of Hemogen known as erythroid differentiation-associated gene protein (EDAG), shares 43% identity with Hemogen while the rat homolog Hemogen, PR59 shares 70% identity with Hemogen (Yang et al. 2001) (figure 5).

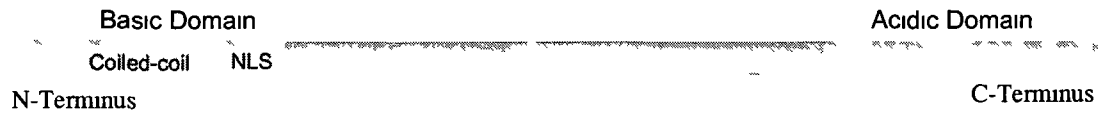


Figure 4: Schematic of Hemogen primary sequence features. N-terminal Hemogen has a cluster of basic residues. Embedded within the basic domain are a coil-coiled domain and a nuclear localization sequence (NLS). C-terminal Hemogen has a cluster of acidic residues.

```
Hemogen     MDMGKGRPRLKLPQMPEAHPQKSCAPDIIGSWSLRNREQLRKRKAEAQGRQTSQWLLGEQ 60
RP59        MDVRKDQSHLTLNQTPEAHVEKSHAPDIIGSWSLRNREQLKKRKAEAQGRQTSQWQLGEQ 60
EDAG        MDLGKDQSHLKHHQTPDPHQEENHSPEVIGTWSLRNRELLRKRKAEVHEKETSQWLFGEQ 60

Hemogen     KKRKYQRTGKGNKRGRKRQGNVEQKAEPWSQTERERVQEVLVSAEEETEHPGNSATEALP 120
RP59        KKRKYQRTGKGNRRGRKRQGNVEQKAESWSQTENERVQEVLVPAEEETEYSGNTATQALP 120
EDAG        KKRKQQRTGKGNRRGRKRQQNTELKVEPQPQIEKEIVEKALAPIEKKTEPPG-SITKVFP 119

Hemogen     LVPSPTKAVPADQCSEAHQESIQCQERAIQNHSQTH--------LSPTTCQGIAVLQHSP 172
RP59        LRASPTKAVPTEHCSEVPQESLQCQEITIQNHSQTHQRRDKAEALSPTTCQEIGVLQYSP 180
EDAG        SVASPQKVVPEEHFSEICQESNIYQEN-------------------FSEYQEIAVQNHSS 160

Hemogen     KMCQDMAEPEVFSPNMCQETAVPQTYPPKALEEMAAAEPLSPKMC--------------- 217
RP59        KMCQDMAEPEARSPKMCQETAVPQTYSPKAHEDMAEYEALSPKMCRETPVPQYHSSKIPQ 240
EDAG        ETCQHVSEPEDLSPKMYQEISVLQ----------------------------------- 184

Hemogen     -------------QETTVSPNHSSKVPQDMAGPEALSPNMCQEPTVPQEHTLKMCHDVAR 264
RP59        DMTGPEALAPDMCQETTVSQNHSSKVPQDMAGPEALSLKMCQEPTVLQEHTLKICQDVAR 300
EDAG        --------------------DNSSKICQDMKEPEDNSPNTCQVISVIQDHPFKMYQDMAK 224

Hemogen     PEVLSPKTHQEMAVPKAFPCVTPGDAAGLEGCAPKALPQSDVAEGCPLDTTPTSVTPEQT 324
RP59        PDVLSSKTHQEMTVPKALPCITPGDAAGPEGCSPKTLPQS--------DVTPMSVTPGKN 352
EDAG        REDLAPKMCQEAAVPKILPCPTSEDTADLAGCSLQAYPKPDVPKGYILDTDQNPAEPEEY 284

Hemogen     TSDPDLGMAVTEGFPSEARECTVSEGVSTKTHQEAVEPEFISHETYKEFTVPIVSSQKTI 384
RP59        TSHPDLGTAVAEGCPSEAGECIVSEDISTKTHQEAVEPKFLSHKTYKEFTVPIVSSHKTI 412
EDAG        N-FTDQGIAETEGLFPKIQEIAEPKDLSTKTHQESAEPKYLPHKTCNEIIVPKAPSHKTI 343

Hemogen     QES----------------------------PEPEQYSPETCQPIPGPENYSLETCHEMS 416
RP59        QES----------------------------PGPEEYSPG--------------SRHEMS 430
EDAG        QETPHSEDYSIEINQETPGSEKYSPETYQEIPGLEEYSPEIYQETSQLEEYSPEIYQETP 403

Hemogen     GPEDLSIKTCQDREEPKHSLPEGAQKVGGAQGQDADAQ--DSENAGAFSQDFTEMEEENK 474
RP59        EPEDLSIKTCKNRDGPKHSLPEGAQEVGGAQRQDPDAQ--DIEDAGAFSQDLEE------ 482
EDAG        GPEDLSTETYKNKDVPKECFPEPHQETGGPQGQDPKAHQEDAKDAYTFPQEMKEKPKEE- 462

Hemogen     ADQDPEAPASPQGSQETCPENGIYSSALF 503
RP59        ----------------------------
EDAG        -------PGIPAILNESHPENDVYSYVLF 484
```

Figure 5: Amino sequence aligment of Hemogen (mouse), RP59 (rat) and EDAG (human) by ClusterW (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Hemogen (gi number 16716441) shares 70 and 43% amino acid sequence identity with RP59 (gi number 11140172) and EDAG (gi number 13445577) respectively. Conserved residues between Hemogen, RP59 and EDAG are hightlighted.

Hemogen expression is present within active hematopoietic sites during murine embryonic ontogeny (Yang et al. 2001). At E8.5 of mouse embryogenesis, *in situ* hybridization experiments revealed *Hemogen* transcripts within the blood islands of the yolk sac and primitive circulating blood cells (Yang et al. 2001). *Hemogen* transcripts were found at E9.5 and E10 in circulating blood cells (Yang et al. 2001). Around the same time, primitive nucleated erythrobasts appear within the vasculature of the yolk sac between E7.25-E9.0 (Palis 2008) and enter circulation at E8.25 (Palis 2008). At E9.5, the fetal mouse liver is colonized by external hematopoietic elements soon after it begins to form as an organ (Palis 2008). Later at E11.5, *Hemogen* is expressed exclusively in the fetal liver while its expression within circulating blood cells is down-regulated to undetectable levels (Yang et al. 2001).

In adult mice, *Hemogen* is expressed within adult hematopoietic compartments (Yang et al. 2001). Northern blot analysis of Hemogen tissue expression revealed *Hemogen* expression within the bone marrow and spleen. Hemogen is weakly expressed in peripheral blood (Yang et al. 2001). To determine which hematopoietic cells express *Hemogen*, adult mouse bone marrow cells were FACS sorted into multipotential HSCs and mature blood cells such as CD3+T cells, B220+B cells, Ter-119+erythrocytes and granulocytes (Yang et al. 2001). *Hemogen* is primarily expressed in HSCs and not expressed in mature blood cell lineages investigated (Yang et al. 2001). *Hemogen* transcripts were also found within the red pulp of the spleen where erythropoiesis predominates (Yang et al. 2001).

Interestingly, *Hemogen* deregulation is associated with hematological neoplasm. Levels of *EDAG* mRNA found in peripheral blood cells of acute myeloid leukemia (AML) patients correlated to pathological processes (An et al. 2005). High levels of *EDAG* mRNA were detected in peripheral blood cells of *de novo* acute AML patients and patients who were refractory to treatment while lower levels of EDAG mRNA were observed in AML patients experiencing complete remission (An et al. 2005). High *EDAG* transcripts were also found in the peripheral blood mononuclear cells of other leukemias (Li et al. 2004). The over-expression EDAG within HSCs in transgenic mice was insufficient to induce leukemia but led to a myeloproliferative disorder resembling human chronic myeloid leukemia (Li et al. 2007). Transgenic mice over-expressing Hemogen within HSCs resulted in disrupted hematopoietic homeostasis and led to enhanced extramedullary myelopoiesis in the spleen and reduced lymphopoiesis (Li et al. 2007). Over-expression of Hemogen in hematopoietic cell line studies implicated Hemogen's ability to promote proliferation (Li et al. 2004).

The function of Hemogen has not been clearly elucidated at the molecular level. However evidences suggest that Hemogen is a transcription factor. Primary sequence analysis of Hemogen suggests a putative nuclear localization sequence at the N-terminal region (Yang et al. 2001). Immuno-staining of exogenous Flag-tagged Hemogen expressed in COS-7 cells revealed Hemogen's nuclear localization (Yang et al. 2001). Moreover, a luciferase reporter assay revealed that the Gal4 fusion human homolog of Hemogen, EDAG is a transcriptional activator in 293 cells (Li et al. 2007).

Several lines of evidence implicate Hemogen's involvement in erythroid differentiation. *Hemogen* transcripts were found within active sites of erythropoiesis. *In situ* hybridization of mRNA during embryogenesis revealed Hemogen's expression in murine yolk sac (Yang et al. 2001), where primitive erythropoiesis predominates (Orkins and Zon 2008 and Palis 2008), as well as in circulating primitive erythrocytes (Yang et al. 2001). Moreover, Hemogen transcripts are exclusively expressed within the red pulp of the adult mouse spleen where erythropoiesis predominates (Tsiftsoglou et al. 2009). In addition, the rat homolog of Hemogen, RP59 is expressed within erythroid precursors of fetal liver and circulating blood during late prenatal stages. RP59 was also observed in erythroid precursors within the red pulp of the adult spleen (Kruger et al. 2002).

Moreover, evidences suggest that the *Hemogen* gene is a direct target of GATA-1 (Yang et al. 2006 and Li, et al. 2008), an important transcriptional stimulator of erythropoiesis (Fujiwara, Y. et al. 1996). GATA-1 binding sites were found in the *Hemogen* promoter (Yang et al. 2006). Electrophoresis mobility shift assay and chromatin-immunoprecipitation experiments demonstrate the association of GATA-1 to the *Hemogen* promoter (Yang et al. 2006). Luciferase assays and gene-knock down experiments suggest that GATA-1 is an activator of human Hemogen's (EDAG) expression in erythroid cells (Li et al. 2008). Furthermore, *EDAG* transcript levels correlate to GATA-1 levels within normal and leukemic peripheral mono-nuclear blood cells (Yang et al. 2006).

1.7 Statement of Intent

The mechanistic function of Hemogen has not been elucidated. Interestingly, several reports implicated Hemogen's involvement in erythroid differentiation. *Hemogen* transcripts were found in active sites of erythropoiesis and the rat homolog of Hemogen, RP59 was found within erythroid precursors. *Hemogen* expression may be activated by GATA-1; the key stimulator of erythropoiesis. Moreover, Hemogen may associate with the NF-E2 complex which activates *β-globin* expression, a hallmark of erythroid differentiation. Elucidating Hemogen's associating partners during erythroid differentiation will provide insight into the mechanistic function of Hemogen during erythroid differentiation.

1.8 Hypothesis

We hypothesize that Hemogen functions as a transcriptional regulator. Hemogen is localized to the nucleus and has transactivation activity. Moreover, a proteomic screen of MafK indicated that Hemogen associates with transcription regulators during erythroid differentiation.

Chapter 2: Materials and Methods

2.1 Vector Design and Cloning

Using an antigenicity plot (http://www.bioinformatics.org/JaMBW/3/1/7/), the primary sequence of Hemogen was assessed for regions of high antigenicity. The N-terminal segment of Hemogen was selected and cloned into a bacterial expression vector that conferred a histidine tag. Primers (5'ATGGATCCTATGGACATGGGGAAGGG (*BamHI*), 5'ACGCGTCGACTTATCCTTGGCATGTGGTAGGAGAGAGGT(*SalI*)) were designed to amplify the N-terminal Hemogen from a plasmid containing the mouse full-length Hemogen cDNA (pET28b+ Hemogen full-length). After PCR amplification, the amplified product was digested with *BamHI*/*SalI* restriction enzymes and sub-cloned into pET28b+ vector to generate the recombinant vector (pET28b+ N-terminal Hemogen) using standard cloning techniques (figure 6).

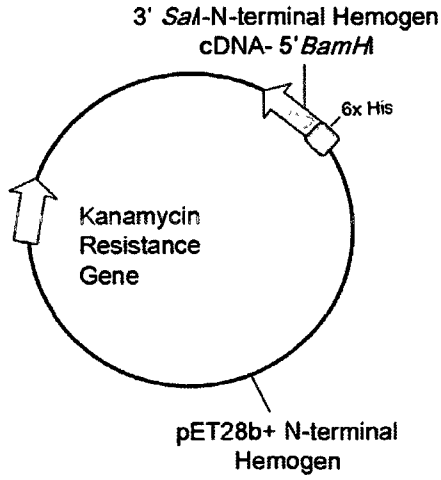


Figure 6: Vector map of pET28b+ N-terminal Hemogen. Vector map of pET28b+ N-terminal Hemogen vector used for protein expression in bacteria. The pET28b+ N-terminal Hemogen vector encodes a hexa-histidine tag N-terminal to the N-terminal *Hemogen* insert. Kanamcyin resistance gene allowed for the use of drug selection during cloning in DH5α *E.coli*. Vector map is not to scale.

The full-length Hemogen was expressed as a triple Flag -fusion protein in MEL cells. Primers (5'CCCAAGCTTATGGACATGGGGAAGGGC (*HindIII*), 5'CTAGTCTAGAAAATAGAGCAGAGCTGTAGATGCCA (*Xba*I)) were designed to amplify the full-length Hemogen from pET28b+ Hemogen full-length. After PCR amplification, the amplified product was digested with *HindIII*/*Xba*I restriction enzymes and sub-cloned into p3xFlag-CMV-14 vector to generate a recombinant vector (p3xFlag-CMV-14-Hemogen full-length) using standard cloning techniques (figure 7).

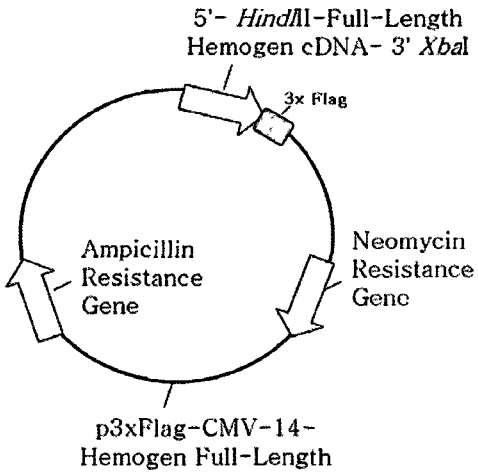


Figure 7: Vector map of p3xFlag-CMV-14-Hemogen-Full-Length. The p3xFlag-CMV-14-Hemogen-Full-Length vector encodes a triple Flag tag C-terminal to the *Hemogen* insert. Ampicillin resistance gene allowed for the use of drug selection during cloning in DH5α *E.coli*. Neomycin resistance gene allowed for the use of drug selection during MEL cell culture. Vector map is not to scale.

2.2 N-terminal Hemogen Expression and Purification

E.coli strain HD5α was transformed with the ligation product after sub-cloning. Positive clones were sequenced to ensure no sequence mutations occurred during PCR. One positive clone was transformed into *E. coli* strain BL21. The transformed BL21 *E. coli* was used to inoculate a 1L of LB medium. This culture was incubated at 37 °C until an OD600 of 0.5. Protein expression was induced with 1mM IPTG for four hours at 37 °C. Following induction, the bacteria was pelleted by centrifugation (10,000xg for 15 minutes at 4 °C). The supernatant were removed and the pellet was stored overnight at – 80 °C. The next day, the pellet was resuspended in 60ml of sonication buffer (50mM NaH_2PO_4 pH7.6, 300mM KCl, 10mM Imidazole and Complete Mini EDTA-free Protease Inhibitor Cocktail tablets (Roche)). The resuspended pellet was sonicated at 4 °C at 3x 20%, 2x 25%, 1x 30% and 1x 35% with 30 second pulses and 30 second wait-time in between on ice. The sonicated bacteria solution was centrifuged (10,000xg, 30 minutes at 4 °C) to remove insoluble cell debris. The supernatant after centrifugation was collected and incubated with inversion (2 hours at 4 °C) in 2ml suspension volume of Talon Metal Affinity Resin (Clontech) equilibrated in sonication buffer. The protein-bounded resin suspension was centrifuged (200xg, 5 minutes) to pellet the beads and washed (2x for 10 minutes at 4 °C) in 6 bead suspension volumes of wash buffer (50mM NaH_2PO_4 pH7.6, 5mM Imidazole, 500mM KCl, 8 mM β-mercaptoenthanol (Fisher) and Complete Mini EDTA-free Protease Inhibitor Cocktail tablets (Roche)). The protein-bound resin was subsequently rinsed 2x with 10 suspension volumes of wash buffer. The resin was loaded into a gravity flow column (Biorad) and the protein was eluted in six 500 μl fractions using (50mM NaH_2PO_4 pH7.6, 400mM Imidazole, 500mM KCl, 8 mM

β-mercaptoenthanol (Fisher), 10% glycerol and Complete Mini EDTA-free Protease Inhibitor Cocktail tablets (Roche)).

2.3 Cell Culture

Murine erythroleukemia (MEL) cells, strain 745A, were cultured in RPMI-1640 medium (Sigma R8758 containing L-glutamine) supplemented with 10% fetal bovine serum (Sigma) and 1% penicillin/streptomycin (Hyclone). MEL cells were cultured in suspension using a spinner flask (Bellco) and placed in a 37 °C humidified incubator with 5% CO_2 in air. MEL cells arrested at the proerythroblast stage were induced to differentiation using 2% DMSO (Sigma) at 0.25 x 10^6 cells/ml.

2.4 Immunofluorescence Microscopy

Undifferentiated and MEL cells differentiated for three days were removed from culture medium and resuspended in sterile PBS1x at 10^6 cells/ml. Coverslips (22x22 Fisher Scientific) were rinsed 3x with 70% ethanol, 3x with double distilled water and washed for 5 minutes in 0.1M HCl. After coverslips were dried under UV, they were coated with 0.01 mg/ml poly-L-ornithine. The cell suspension was placed on poly-L-ornithine coated coverslips for 20 minutes for cell attachment. Coverslips were rinsed with PBS1x to remove unattached cells. Cells were fixed with 4% paraformaldehyde (diluted from 16% EM grade paraformaldehyde, Electron Miscroscopy Science) in PBS1x for 10 minutes. Coverslips were washed with 50 mM NH_4Cl for 5 minutes and rinsed 2x with PBS1x. Cells were permeablized with 1% Triton X-100 in PBS1x for 5 minutes and washed (2x 5 minutes) with PBS1x. Coverslips were incubated with 3% bovine serum albumine (BSA)

(Hyclone) in PBS1x for 2 hours at room temperature. Incubation with Hemogen antibody (1:300) (sc-68360; Santa Cruz Biotechnology) in 3% BSA was carried out overnight at 4 °C in a humidified chamber. After incubation, cells were washed (2x 10 minutes) in PBS1x. Incubation with anti-rabbit secondary antibody conjugated to FITC (1:300) (Jackson ImmunoResearch Laboratories Inc.) was carried out for 2 hours at room temperature. Coverslips were washed (2x 10 minutes) in PBS1x before they were mounted onto slides using hardset antifade media with DAPI (Vector Laboratories, Inc.). Data were collected from confocal microscope (Zeiss LSM 510) using LSM 510 digital image processing software.

2.5 Nuclear Extraction

MEL cells differentiated for three days were suspended at 10^8 cells/ml and stored in MS-30 buffer (50 mM Tris-HCl, pH 7.9, 0.625 mM EDTA, 27.6% Glycerol). The cell suspension was thawed on ice and washed once with five pack cell volume (PCV) of PBS1x buffer. Cells were incubated with five PCV of buffer A (10mM Kepes K+, pH7.9, 1.5 mM $MgCl_2$, 10mM KCl, 0.5 mM dithiothreitol and Complete Mini EDTA-free Protease Inhibitor Cocktail tablets (Roche)) for 10 minutes and resuspended in two PCV of buffer. A Kontez homogenizer (type B pestle) was used to disrupt cell membranes. The pellet recovered following centrifugation corresponded to crude nuclei. Nuclei were resuspended in one PCV of buffer C (20mM Kepes K+, pH7.9, 1.5 mM $MgCl_2$, 600mM KCl, 25% Glycerol, 0.2 mM EDTA and 0.5 mM dithiothreitol and Complete Mini EDTA-free Protease Inhibitor Cocktail tablets (Roche)) was homogenized by Kontez homogenizer (type B pestle) and was incubated on rotator for 30 minutes at 4 °C. Soluble

nuclear proteins are recovered after centrifugation. The nuclear extract is dialyzed against 50 times its original volume using buffer D (20mM Kepes K+, pH7.9, 5 mM $MgCl_2$, 100mM KCl, 20% Glycerol, and 0.5 mM dithiothreitol and 1mM PMSF).

2.6 Antibody Cross-linking and Immunoprecipitation

The anti-Hemogen serum was cross-linked to sepharose protein-A resin (Invitrogen). 100 µl of Protein-A sepharose resin were rinsed twice with 500 µl of 0.1 M phosphate buffer, pH 8.2 with 0.1% (w/v) ovalbumine. 100 µl of antibody solution (33% (v/v) anti-Hemogen serum, 0.1 M phosphate buffer, pH 8.2) was incubated with 100 ul Protein-A sepharose resin for 1 hour. The supernatant was removed. The resin was washed once with 500 µl of 0.1 M phosphate buffer, pH 8.2 containing 0.1% (w/v) ovalbumine and twice with 500 µl 0.2 M triethanolamine (Fluka), pH 8.2. The resin was washed with 5mg/ml (w/v) dimethyl pimelimidate (Pierce, Rockford, USA) in 0.2 M triethanolamine, pH 8.2 for 30 minutes to cross-link the antibody to the resin. The cross-linking reaction was stopped by washing the resin once with 500 µl of 50 mM Tris, pH 7.4. The resin was washed 3x with 500 µl of PBS1x with 0.1% (w/v) ovalbumine and 3 minutes in 100 mM glycine, pH 3. Resin was washed 3x with 500 µl of IP100 buffer (25mM Tris, pH 7.9, 5mM $MgCl_2$, 100 mM KCl, 10% glycerol, 0.1% (v/v) NP-40, 0.3 mM DTT and Complete Mini EDTA-free Protease Inhibitor Cocktail tablets (Roche)). MEL nuclear extract (NE) was incubated with antibody bound resin for 15 hours at 4 °C. The NE was removed. Proteins bound beads were washed with IP300 buffer (25mM Tris, pH 7.9, 5mM $MgCl_2$, 300 mM KCl, 10% glycerol, 0.1% (v/v) NP-40, 0.3 mM DTT and Complete Mini EDTA-free Protease Inhibitor Cocktail tablets (Roche)) and equilibrated

with same IP buffer containing 100 mM KCl. Bound proteins for western blot analysis were eluted with loading buffer (143mM Tris, pH 6.8, 28.6% Glycerol, 5.7% SDS, 286 mM DTT) for five minutes at 72 °C. Bound proteins for mass spectrometry (MS) analysis were eluted with loading buffer (50 mM Tris pH 6.8, 20% Glycerol, 4% SDS, 1mM DTT) at 50 °C. MS elution samples volumes (400 μl) were reduced by two fold using a speed vacuum centrifuge. Mock IP using pre-immune serum was done in parallel as the specific Hemogen IP. MafK IP was performed similarly to Hemogen IP.

2.7 Silver Staining for Mass Spectrometry

IP protein samples in loading buffer (50 mM Tris pH 6.8, 20% Glycerol, 4% SDS, 1mM DTT) were denatured at 95°C for 5 minutes and resolved on 10% SDS-polyacrylamide gel (375mM Tris-HCl, pH 8.8, 0.1% (v/v) SDS, 0.05% (v/v) APS, 0.05% (v/v) TEMED) at a constant 30 mA. High molecular weight protein molecular standard (Bio-Rad Latoratories Inc.) was used to denote molecular weight. The resolving gel was fixed overnight in 50% ethanol and 10% acetic acid and placed in 30% ethanol for 15 minutes. The gel was washed (3x 5 minutes) in ddH_2O and placed in 0.2g/L $Na_2S_2O_3$ solution. The sensitized gel was washed (3x 30 seconds) in ddH_2O and incubated in 2.0g/L $AgNO_3$ for 25 minutes. The gel was washed (3x 30 seconds) in ddH_2O to remove $AgNO_3$ solution. The gel was placed in developing solution (60g/L Na_2CO_3, 0.2g/L $Na_2S_2O_3$ and 500 μl/L 37% formaldehyde) until the desired level of staining was achieved. The development was stopped by discarding the developing solution and adding 6% acetic acid for 10 minutes. The gel was washed (4x 15 minutes) in ddH_2O.

2.8 Reduction, Alkylation and Trypsin Digestion of Silver Stained Gel Bands for Mass Spectrometry

Gel bands were cut from sliver stained gel compatible to mass spectrometry and individually cut into pieces by individual surgical razor blades (Feather). Gel pieces were washed (1x 10 minutes) in ddH_2O and destained in destain solution (15 mM potassium ferricyanide and 50 mM sodium thiosulfate) with gentle vortexing for 15 minutes. The destain solution was made fresh by mixing equal volumes of 30 mM potassium ferricyanide and 100 mM sodium thiosulfate. The gel pieces were then washed (1x 10 minutes) in ddH_2O with gentle vortexing. Gel pieces were washed in destain solution and subsequently in ddH_2O for a total of 5 times. Gel pieces were dehydrated by adding 100% acetonitrile for 5 minutes at room temperature and dried in a vacuum centrifuge for 10 minutes. For reduction, 10 mM DTT was added to gel pieces for 30 minutes at 37 °C. Gel pieces were cooled to room temperature and alkylated using 100 mM iodoacetamide for 30 minutes at room temperature in the dark. 100% acetonitrile was added (1x 5 minutes) at room temperature to dehydrate the gel pieces. The gel pieces were rehydrated (1x 10 minutes) with 50 mM ammonium bicarbonate. One hundred percent acetonitrile was added (1x 5 minutes) at room temperature to dehydrate the gel pieces followed by completely drying the gel pieces (1x 10 minutes) at room temperature in a vacuum centrifuge. Gel pieces were rehydrated on ice for 20 minutes using 30 μl of trypsin solution. Stock trypsin solution was prepared by suspending 20 μg of porcine trypsin (Promega) in 100 μl of the accompanying cold resuspension buffer. The working trypsin solution was made by diluting the stock trypsin solution 10x using 50 mM ammonium bicarbonate. Sufficient 50 mM ammonium bicarbonate was added to

submerge all gel pieces. Gel pieces were gently mixed (1x 30 seconds) and incubated at 37 °C overnight. 30 μl of ddH_2O was added to each digested sample and incubated (1x 10 minutes) with gentle vortexing. Samples were sonicated (1x 10 minutes) and centrifuged (1x 30 seconds). The supernatant was collected. 30 μl of extraction buffer (50% (v/v) acetonitrile and 5% (v/v) formic acid) was added (1x 10 minutes) with gentle vortexing. The samples were sonicated (1x 10 minutes) and centrifuged (1x 30 seconds). The supernatant was collected. The procedure of adding sonication buffer, sonication and collection of supernatant was performed for a total of two times. The supernatants containing protein extracts were reduced to approximately 5 μl in a vacuum centrifuge at room temperature. The final volume of the sample was adjusted to 10 μl using 0.22 μm filtered ddH_2O (by using Corning 0.22 μm 150 ml filters).

2.9 Zip Tip™ Cleanup

10 μl protein extract samples were acidified with 1 μl of 1% trifluoroacetic acid and mixed by pipetting. 10 ul Zip tip™ (Millipore) contains a C_{18} resin that binds to acidified peptides. Zip tips™ were equilibrated by pipetting (2x 10 μl of 50% acetonitrile) followed by pipetting 3x using 10 μl of 0.1% trifluoroacetic acid. Acidified peptides were bound to zip tip™ resin by pipetting 10 times using the acidified extracts. Zip tips™ resin was washed 5x using 10 μl of 0.1% % trifluoroacetic acid. Peptides were eluted from the zip tips™ by pipetting 5 times 5 μl of 50% acetonitrile. Peptide elution was repeated for a total of two trials.

2.10 μLC-ESI-MS-MS

Each peptide sample treated with zip-tips was diluted in 0.1% trifluoroacetic acid to 45 μl and loaded into a 96-well plate. The Finnigan Surveyor Autosampler Plus and Finnigan Surveyor MS Pump Plus HPLC system was used to automatically inject each sample into a trap column (Michrom C_{18} CapTrap) at 20 μl/minute. Peptides retained within the trap column were eluted over a 32 minutes gradient from 98% (v/v) HPLC Buffer A (0.2% (v/v) formic acid in water) and 2% (v/v) HPLC Buffer B (0.2% formic acid in acetonitrile) to 60% (v/v) HPLC Buffer A and 40% (v/v) HPLC Buffer B. Peptides that eluted from the trap enter an analytical column consisting of a Picofrit (New Objectives) column packed with Zorbax 5-micron C_{18} chromatographic packing material (Agilent). The Picofrit column is constructed with an integrated nanospray emitter that sprayed the chromatographically separated peptides into a nanospray ionization source. A split flow system was used to produce a column flow of approximately 250-300 nanolitres/minute through the Picofrit column. Charged peptides were analyzed by a Finnigan LTQ ion trap mass spectrometer. The mass spectrometry was set to scan from 400-1500 Da followed by data dependent MS/MS scans of the four most abundant ions. The software system used to control the HPLC and mass spectrometer was Xcalibur 2.0 SUR 1 (Thermo). Dynamic exclusion was used to exclude repetitive scanning of the same ions. The dynamic exclusion duration was 30 seconds. A permanent exclusion list was also used to reject background ions commonly seen in nanospray ionization. Sequest (Eng 1994) and Mascot (Perkins et al. 1999) algorithms were used to correlate the observed mass spectra with the theoretic mass spectra of peptides from a murine database to determine the

peptide sequence. The identification probability of each peptide identified from Sequest was determined by proteinprophet (Keller et al. 2002).

2.11 Silver Stain Analysis for Small-scale IP Elution

IP protein samples in loading buffer (143mM Tris, pH 6.8, 28.6% Glycerol, 5.7% SDS, 286 mM DTT) were denatured at 95 °C for 5 minutes and resolved on 10% SDS-polyacrylamide gel (375mM Tris-HCl, pH 8.8, 0.1% (v/v) SDS, 0.05% (v/v) APS, 0.05% (v/v) TEMED) at a constant 30 mA. High molecular weight protein molecular standards (Bio-Rad Latoratories Inc.) were used to denote molecular weight. The resolving gel was fixed in 50% ethanol and 10% acetic acid overnight. The gel was placed in staining solution (0.8% (w/v) silver nitrate, 0.08% (v/v) sodium hydroxide, 1.6% (v/v) ammonia) for 15 minutes and then washed (1x 5 minutes) with ddH_2O. The gel was gently washed in developing solution (0.25% (v/v) citric acid, 0.6% (v/v) formaldyhyde) to reveal the proteins. Once the desired level of staining was achieved, the developing solution was removed and the stopping solution (50% (v/v) ethanol and 10% (v/v) acetic acid) was added. The gel was washed (1x 30 minutes) in stop solution and (1x 5 minutes) in gel drying solution (25% (v/v) ethanol and 5% (v/v) glycerol). The gel was stored by drying between cellophane sheets using the DryEase® Mini-Gel Drying System (Invitrogen) as per the manufacturers' instructions.

2.12 Western Blot

Protein extracts and immunoprecipitation elutions were heated at 95 °C for 5 minutes in loading buffer and separated by SDS-PAGE as previously described. All Blue Precision-

Blue protein standards (Bio-Rad Laboratories Inc.) were used as molecular weight markers. Briefly, proteins from SDS-PAGE were transferred to a nitrocellulose membrane (Bio-Rad Latoratories Inc.) and blocked with 5% (w/v) skim milk powder in PBS1x for 30 minutes on a rocking platform. The protein blot was incubated in primary antibody in PBS1x for 1 hour at room temperature or overnight at 4 °C on a rocking platform. The blot was washed (3x 10 minutes) with 0.05% (v/v) Tween-20 in PBS 1x and blotted with an appropriate secondary antibody conjugated to horseradish peroxidase (BioRad) for 1 hour at room temperature in 3% (w/v) skim milk powder in PBS1x. The blot was washed (3x 5 minutes) in 0.05% (v/v) Tween-20 in PBS1x. Proteins within the blot were visualized by horse radish peroxidase chemiluminescent reaction using Pierce ECL Western Blot Substrate (Rockford, IL, USA) and exposed on autoradiographic film (Kodak). Anti-Hemogen serum and pre-immuno serum were dilute at 1:5000 for western blot. Commercial antibodies against MafK (1:1000) (sc-16), p45 (1:500) (sc-19), Hemogen (1:1000) (sc-68360), (1:500) TFIIH (sc-20697) and Dnmt-1 (1:500) (sc-20701) were used for western blot.

2.13 Size Exclusion Chromatography

Proteins from differentiated MEL nuclear extract were separated by size using a 6 10/300 GL column (GE Healthcare) on the FPLC AKTA system. Collected fractions were 500 μl each. Fractions were collected starting at elution volume of 5ml. The void volume was 7.87 ml as approximated by the elution volume of Blue Dextran. To concentrate the fractionated proteins for western blot, proteins within each fraction were precipitated overnight using 10% trichloroacetic acid (TCA). Protein pellets were recovered by

centrifugation. Subsequent acetone wash removed residual TCA. Precipitated protein pellets were dissolved in 100mM KCl containing loading buffer (143mM Tris, pH 6.8, 28.6% Glycerol, 5.7% SDS, 286 mM DTT) for western blot. Antibodies against MafK (1:1000) (sc-16), p45 (1:500) (sc-19), Hemogen (1:1000) (sc-68360) and Dnmt-1 (1:500) (sc-20701) were used.

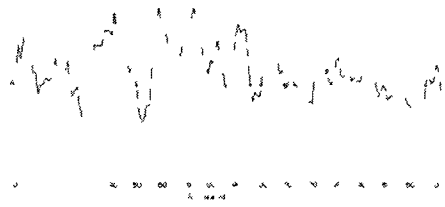
Chapter 3: Results

3.1 Hemogen Antibody Production

At the start of the project, there was no commercially available Hemogen antibody. To study Hemogen by western blot and immuno-precipitation, we decided to produce a rabbit polyclonal Hemogen antibody. Firstly, we assessed the primary sequence of Hemogen to select a region of high antigenicity (figure 8). The N-terminal region consists predominantly of hydrophilic residues that could serve as epitopes for antibodies. Therefore, we used PCR to amplify the N-terminal region of Hemogen from a plasmid containing full length Hemogen cDNA. Full length Hemogen cDNA was previously amplified by reverse-transcription PCR from Hemogen mRNA in MEL cells and cloned into a pET28b+vector (Novagen). We cloned the N-terminal Hemogen in the pET28b+bacterial expression vector that conferred a N-terminal his-tag (figure 9). The recombinant plasmid was submitted for sequencing and no mutations were found within the protein coding regions of the plasmid (data not shown). We therefore transfected the N-terminal Hemogen expression vector into BL21 *E.coli* (Stratagene), induced protein-production using IPTG and purified the synthesized his-tagged N-terminal Hemogen using metal affinity chromatography.

N-Terminal Segment

1 mdmgkgrprl klpqmpeahp qkscapdiig swslrnreql rkrkaeaqgr qtsqwllgeq
61 kkrkyqrtgk gnkrgrkrqg nveqkaepws qterervqev lvsaeeeteh pgnsatealp
121 lvpsptkavp adqcseahqe siqcqeraiq nhsqthlspt tcqgiavlqh spkmcqdmae



Middle Segment

181 pevfspnmcq etavpqtypp kaleemaaae plspkmcqet tvspnhsskv pqdmagpeal
241 spnmcqeptv pqehtlkmch dvarpevlsp kthqemavpk afpcvtpgda aglegcapka
301 lpqsdvaegc pldttptsvt peqttsdpdl gmavtegffs earectvseg vstkthqeav

C-Terminal Segment

361 epefishety keftvpivss qktiqespep eqyspetcqp ipgpenysle tchemsgped
421 lsiktcqdre epkhslpega qkvggaqgqd adaqdsenag afsqdfteme eenkadqdpe
481 apaspqgsqe tcpengiyss alf

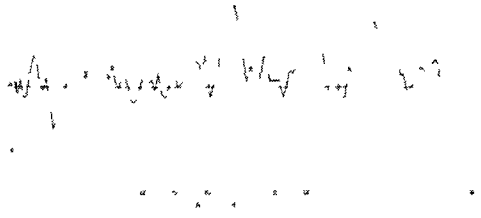


Figure 8: Antigenicity plot of three Hemogen segments. Each amino acid within the primary sequence was assessed for its surface accessibility and hydrophilicity. High scores on the y-axis indicate high antigenicity. The linear sequence of each domain is above its respective antigenicity plot. Highlighted Hemogen residues correspond to sequence used for our Hemogen antibody. Underlined Hemogen residues correspond to sequence used for the commercial sc-68360 Hemogen antibody. The antigenicity plot was calculated from (Hopp and Woods 1981). More information can be found at: http://www.bioinformatics.org/JaMBW/3/1/7/.

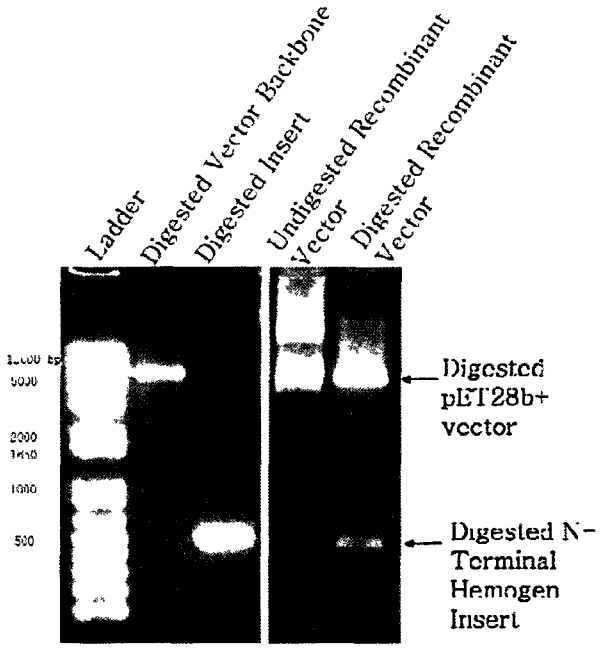


Figure 9: Cloning N-terminal of *Hemogen* cDNA into pET28b+ bacterial expression vector.

2% DNA agarose gel of digested pET28b+ vector (Novagen) and N-terminal *Hemogen* cDNA insert. N-terminal Hemogen cDNA was PCR amplified from a pET28b+ vector containing full length *Hemogen.* PCR products were 493 base pairs in size and had 5' *BamHI* and 3' *SalI* restriction enzyme recognition sites. PCR products (insert) and pET28b+ back bone (5369 base pairs) were digested with *BamHI* and *SalI* and ligated to generate the pET28b+ N-terminal Hemogen recombinant vector.

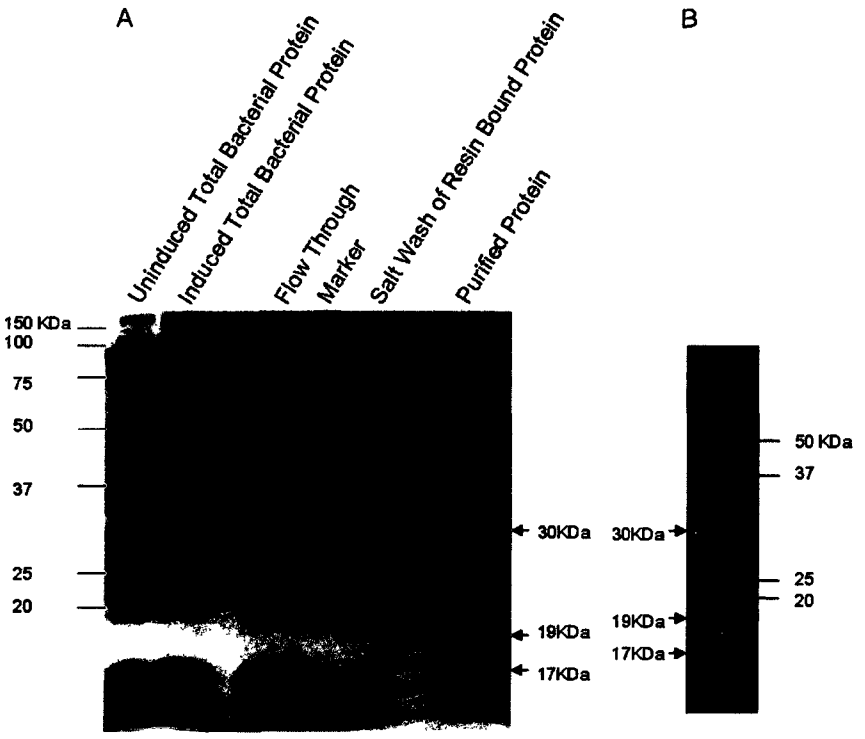


Figure 10: Purification of His N-terminal Hemogen from *E.coli* extracts.

Black arrows indicate N- terminal Hemogen.

(A) Coomassie Blue gel of His-tagged N-terminal Hemogen purification. BL21 *E.coli* was transformed with pET28b+ N-terminal Hemogen vector. After transformation, *E.coli* culture was grown to 0.5 O.D. and sampled as uninduced total bacterial proteins. N-terminal Hemogen production was induced by 1mM IPTG for four hours. N-terminal Hemogen was released from bacteria by sonification. After centrifugation, soluble bacterial proteins were incubated with TALON metal affinity resin (Hyclone). Unbound proteins were sampled as flow-through. Salt wash of 0.5mM KCl was used to remove non-specific proteins on the resin. His-tagged proteins were eluted from the resin using 0.5 M imidazole. The molecular weight is indicated on the left side. See materials and methods for details.

(B) His-tag western blot of purified N-terminal Hemogen protein. Proteins were separated by 12% SDS-PAGE, transferred to nitrocellulose membrane and revealed using His-tag antibody. The molecular weight is indicated on the right side.

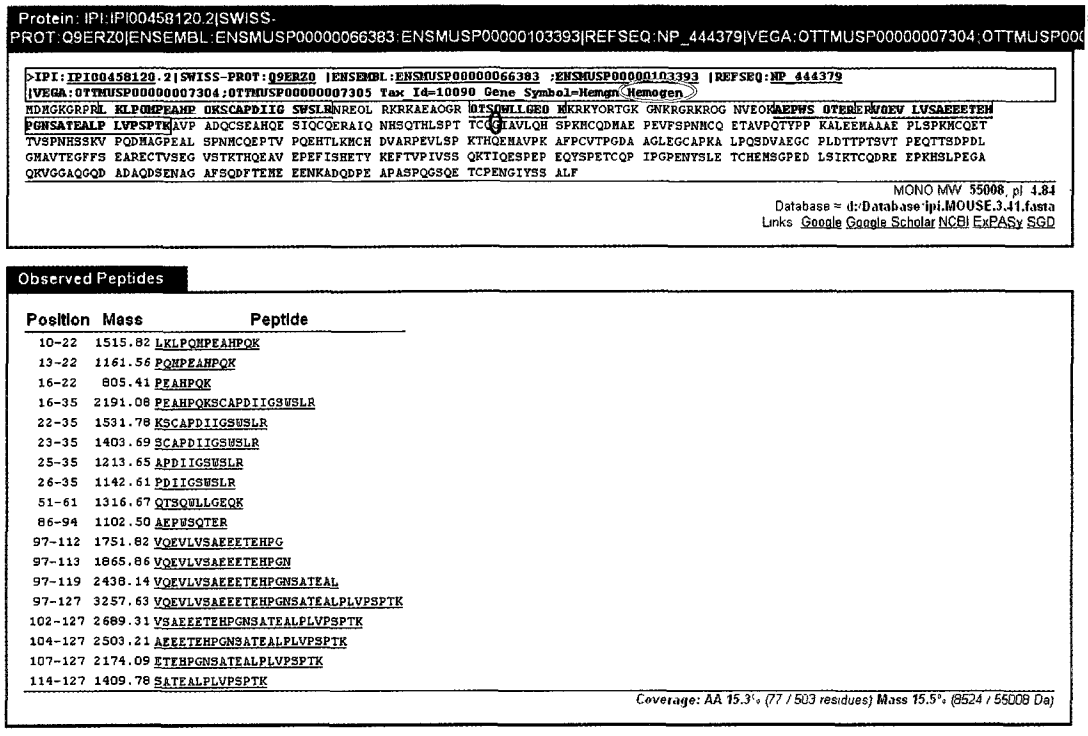


Figure 11: Sequest mass spectrometry identification of the ~30 KDa band from purified N-terminal Hemogen.

A sample of purified N-terminal bacterial Hemogen was loaded into 12% SDS-PAGE gel stained with Coomassie Blue. The ~30 KDa band was cut, treated with trypsin and loaded into the auto-sampler for subsequent mass spectrometry analysis. The red oval outline indicates the identification of Hemogen. The green oval outline indicates where the N-terminal Hemogen terminates. There is 43% coverage of the N-terminal Hemogen primary sequence.

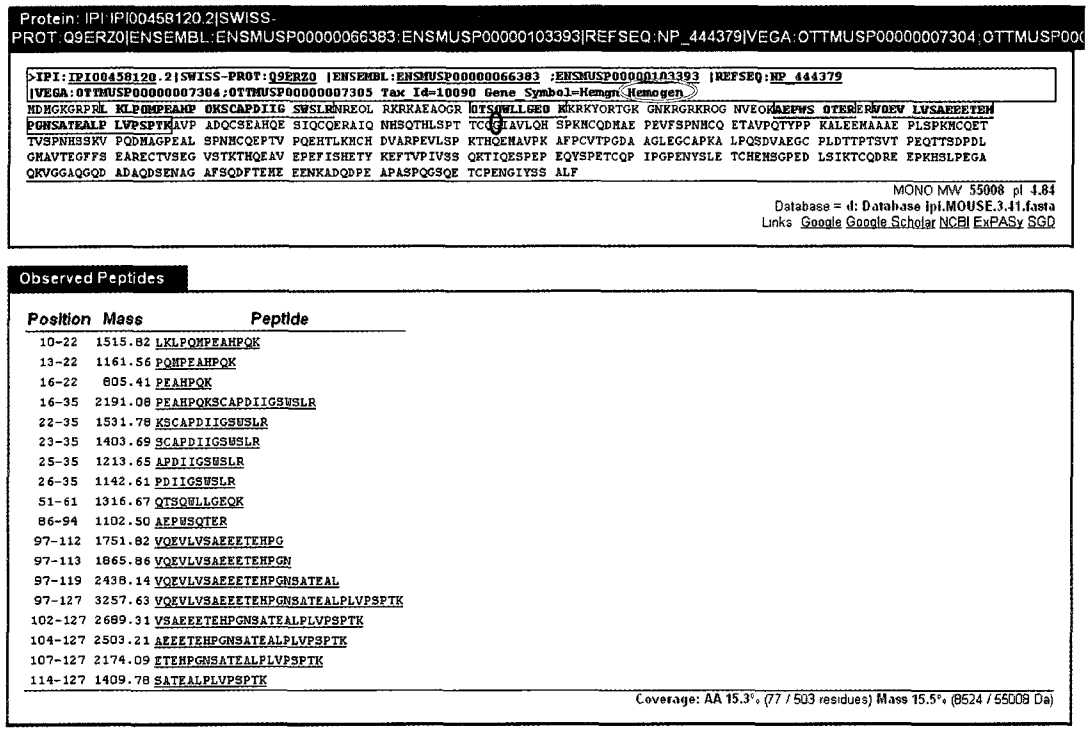


Figure 12: Sequest mass spectrometry identification of the ~19 KDa band from purified N-terminal Hemogen.

A sample of purified N-terminal bacterial Hemogen was loaded into 12% SDS-PAGE gel stained with Coomassie Blue. The ~19 KDa band was cut, treated with trypsin and loaded into the auto-sampler for subsequent mass spectrometry analysis. The red oval outline indicates the identification of Hemogen. The green oval outline indicates where the N-terminal Hemogen terminates. There is 49% coverage of the N-terminal Hemogen primary sequence.

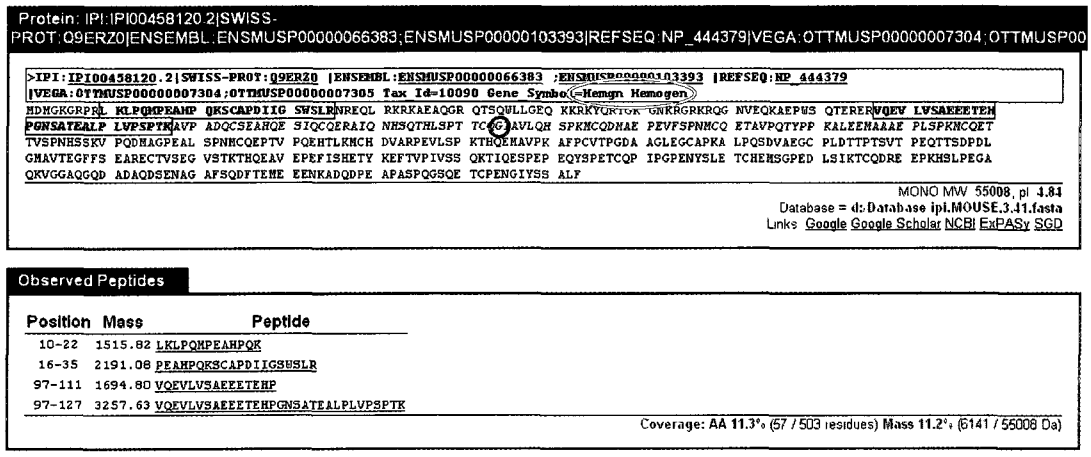


Figure 13: Sequest mass spectrometry identification of the ~17 KDa band from purified N-terminal Hemogen.

A sample of purified N-terminal bacterial Hemogen was loaded into 12% SDS-PAGE gel stained with Coomassie Blue. The ~17 KDa band was cut, treated with trypsin and loaded into the auto-sampler for subsequent mass spectrometry analysis. The red oval outline indicates the identification of Hemogen. The green oval outline indicates where the N-terminal Hemogen terminates. There is 35% coverage of the N-terminal Hemogen primary sequence.

We observed by Coomassie Blue staining, three species with molecular weights: 30 KDa, 19 KDa and 17 KDa within the purified Hemogen fraction (figure 10). To assess whether each species was derived from the vector encoding his-tagged N-terminal Hemogen, we performed a his-tag western blot of the purified fraction. All three species had a his-tag suggesting that all three species represented recombinant N-terminal Hemogen (panel B of figure 10). To verify this, we cut out the gel bands, digested and extracted the peptides and submitted our samples for mass spectrometry analysis. Each species was identified by mass spectrometry as Hemogen (figures 11-13).

We sent the purified protein to a commercial antibody production company (R and R Research, Seattle, Washington, USA) for antibody production in two rabbits. The immunization procedure involved collection of pre-immune serum, initial antigen exposure injection, followed by five immune-response boosting injections. The entire procedure took four months. We monitored the efficacy of the anti-Hemogen serum throughout the immunization procedure by performing western blot of differentiated MEL nuclear extract using test sera (figure 14). Western blot using pre-immune (pre-bleed) sera from both rabbits 9474P and 9520P, which served as our negative control, revealed no signal (figure 14). Western blot signal increased from first test serum to second test serum of both rabbits indicative of an increase in immune-response toward Hemogen.

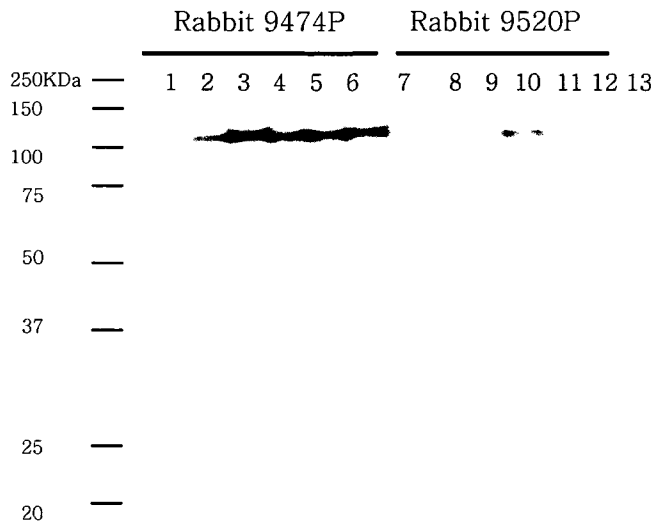


Figure 14: Western blot of differentiated MEL nuclear extracts using anti-Hemogen sera. Anti-Hemogen sera from two rabbits were received at scheduled time-points from R and R Research, Seattle, Washington, USA. Differentiated MEL nuclear extracts were loaded into a single large well of a SDS-PAGE gel. Proteins separated by electrophoresis were transferred onto nitrocellulose membrane. Membranes were cut into 0.4 cm wide stripes. Anti-Hemogen sera (1:5000) were individually applied to membrane stripes. See Materials and Methods for western blot details.

Table 1: Description of Figure 14. Collection time point is with respect to initial antigen exposure. Highlighted rows correspond to rabbit 9474P sera western blot. Rows not highlighted correspond to rabbit 9520P sera western blot. Commercial Hemogen antibody (sc-68360) was used in lane 13.

Lane Number	Test Serum Order	Collection Time Point
1	pre-bleed	1 week before
2	1st	1 month after
3	2nd	2.5 months after
4	3rd	3 months after
5	4th	4 months after
6	last	4.5 months after
7	pre-bleed	1 week before
8	1st	1 month after
9	2nd	2.5 months after
10	3rd	3 months after
11	4th	4 months after
12	last	4.5 months after

The serum from rabbit 9474P produced a more intense signal in the Hemogen western blot (figure 14). Rabbit 9474P may have developed more Hemogen antibodies or antibodies that more efficiently recognized Hemogen.

From a Hemogen western blot of differentiated MEL nuclear extract, we observed a predominant band at 100 KDa (figure 14). However, the theoretical size of Hemogen is approximately 55 KDa. We performed several experiments to verify that our antibody is specific to Hemogen. Primarily, we performed a N-terminal Hemogen western blot using our Hemogen antibody. The antibody recognized the recombinant Hemogen protein previously identified by mass spectrometry (figure 15). To verify that the 100 KDa band corresponded to Hemogen within MEL nuclear extracts, we created Flag-tagged full-length Hemogen expression vector (figure 16) and developed stable MEL cell clones. A Flag western blot of nuclear extract containing Flag-tagged Hemogen revealed a 100 KDa band (figure 17). Flag was not present in untransfected MEL cells (figure 17). Since the known cDNA sequence of Hemogen was transfected into MEL cells, the 100 KDa band observed in our Flag western blot should correspond to Hemogen. During the course of this experiment, an anti-Hemogen antibody was commercialized by Santa-Cruz (sc-68360). In contrast to our antibody, sc-68360 was raised against the mid-section of Hemogen (residues 181-360). We used the commercial Hemogen antibody as a control in our experiments (figure 14). We performed a Hemogen western blot of MEL nuclear extracts using the commercial Hemogen antibody (figure 14) in parallel with our Hemogen antibody (figure 14). Each antibody was raised to recognize non-overlapping

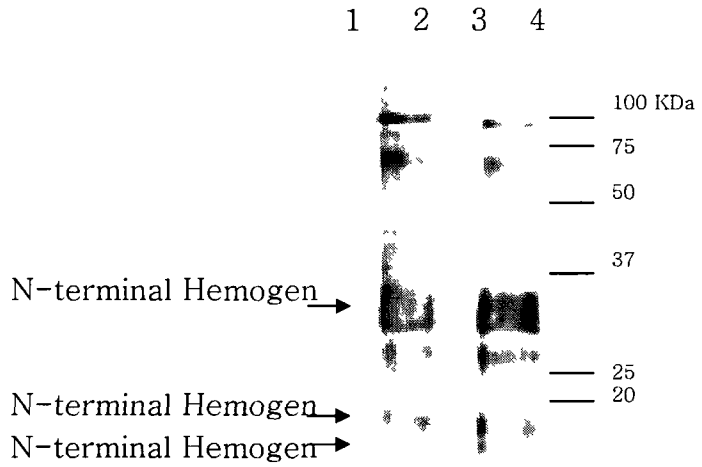


Figure 15: Western blot of His N-terminal Hemogen using anti-Hemogen sera.

N-terminal Hemogen western blot using anti-Hemogen sera. Purified his N-terminal Hemogen (from figure 10) was loaded into a single large well of a SDS-PAGE gel. Proteins separated by electrophoresis were transferred onto nitrocellulose membrane. Membranes were cut into 0.4 cm wide stripes. Each membrane stripe contains approximately 0.89 μg of protein. The anti-Hemogen and pre-bleed sera from two rabbits were individually applied to membrane stripes. See Materials and Methods for western blot details.

(1) Pre-bleed serum from rabbit 9474P (Serum collected 1 week before initial antigen exposure).

(2) 1st test bleed serum from rabbit 9474P (Serum collected 1 month after initial antigen exposure).

(3) Pre-bleed serum from rabbit 9520P (Serum collected 1 week before initial antigen exposure).

(4) 1st test bleed serum from rabbit 9520P (Serum collected 1 month after initial antigen exposure).

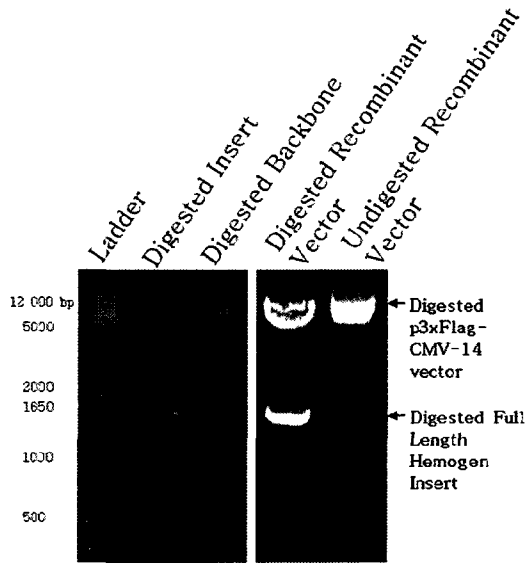


Figure 16: Cloning full-length *Hemogen* cDNA into p3xFlag-CMV-14 mammalian expression vector.

1% DNA agarose gel of digested p3xFlag-CMV-14 (Invitrogen) vector and *Hemogen* cDNA insert. Hemogen cDNA was PCR amplified from a pET28b+ vector containing full length *Hemogen.* PCR products (inserts) were approximately 1.5 kilobase pairs in size and had 5' *HindIII* and 3' *XbaI* restriction enzyme recognition sites. PCR products and p3xFlag-CMV-14 back bone (6.3 kilobase pairs) were digested with *BamHI* and *XbaI* and ligated to generate the p3xFlag-CMV-14-Hemogen-Full-Length recombinant vector.

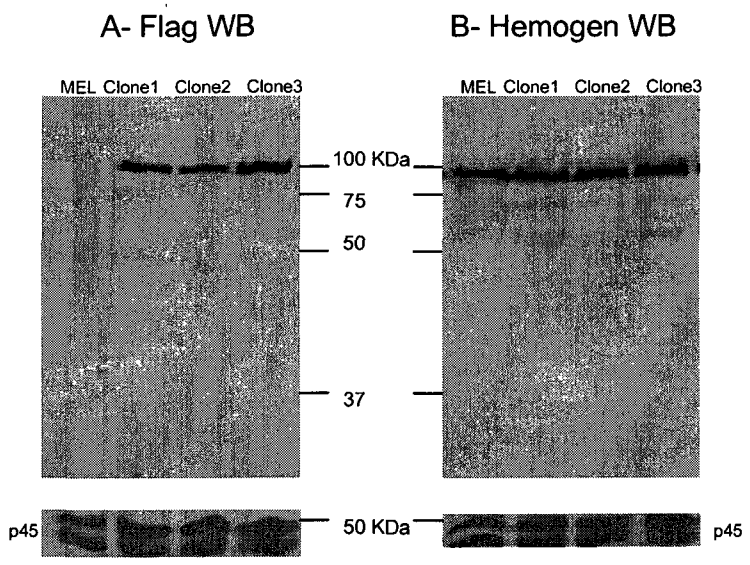


Figure 17: Western blots of stable MEL Clones expressing Flag tagged full-length Hemogen.

(A) Flag western blot of nuclear extracts from untransfected MEL cells or Flag-Hemogen expressing clones (clone 1, clone 2 and clone 3). Western blot of the hematopoietic specific transcription factor, p45 was performed on the same blot as a loading control.

(B) Hemogen western blot (sc-68360) of nuclear extracts from untransfected. MEL cells or Flag-Hemogen expressing MEL clones (clone 1, clone 2 and clone 3). Western blot of the hematopoietic specific transcription factor, p45 was performed on the same blot as a loading control.

primary sequences (residues 181-360) of Hemogen (figure 8). The two Hemogen antibodies both recognized the same 100 KDa protein within differentiated MEL nuclear extracts (figure 14). This and other results indicate that our Hemogen antibody is specific to Hemogen. As with the recombinant N-terminal Hemogen, we performed our SDS-PAGE under denaturing and reducing conditions.

Our polyclonal Hemogen antibody is functional in Hemogen immuno-precipitation (IP) (figure 18). As a negative control for our IP, we performed a mock IP using pre-immune serum in parallel with Hemogen specific IP. Hemogen was absent in our mock IP and present only within Hemogen specific IP. Moreover, Hemogen is immuno-depleted in specific IP flow-through (figure 18) and is enriched in specific IP elution with respect to the input (figure 18). These observations indicate the efficacy of our antibody for Hemogen IP. Therefore, the Hemogen antibody provided us the means to purify Hemogen complexes for the identification of Hemogen associating proteins.

3.2 Characterization of Hemogen Expression within MEL Cells

To investigate the cellular localization of endogenous Hemogen within MEL cells, we performed Hemogen immuno-staining in both undifferentiated (figure 19) and differentiated (figure 20) MEL cells using commercial Hemogen antibody. Secondary antibody alone was used as a negative control (figures 19 and 20). Our immuno-fluorescent microscopy (IFM) results revealed that Hemogen is localized to the nucleus.

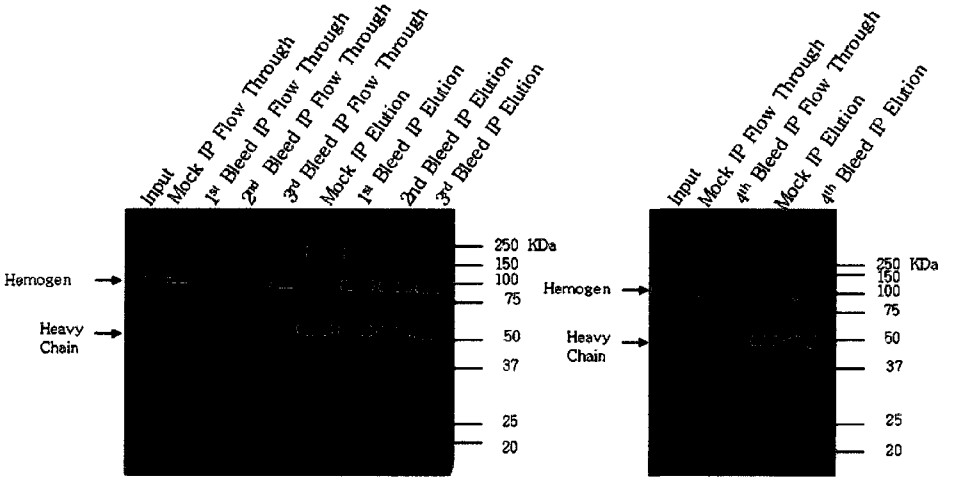


Figure 18: Hemogen IP of differentiated MEL nuclear extract using anti-Hemogen sera.

MEL nuclear extracts made from cell differentiated for three day were incubated with Hemogen anti-serum or pre-immuned conjugated protein-A sepharose resin. Captured proteins were eluted by boiling in SDS buffer and separated by 12% SDS-PAGE. Individual immunoprecipitations were performed with the Hemogen serum (9474P) received at various time points or with pre-immune serum as negative control (mock). Western blots were revealed with the commercial Hemogen antibody (sc-68360). Equal sample volumes were loaded.

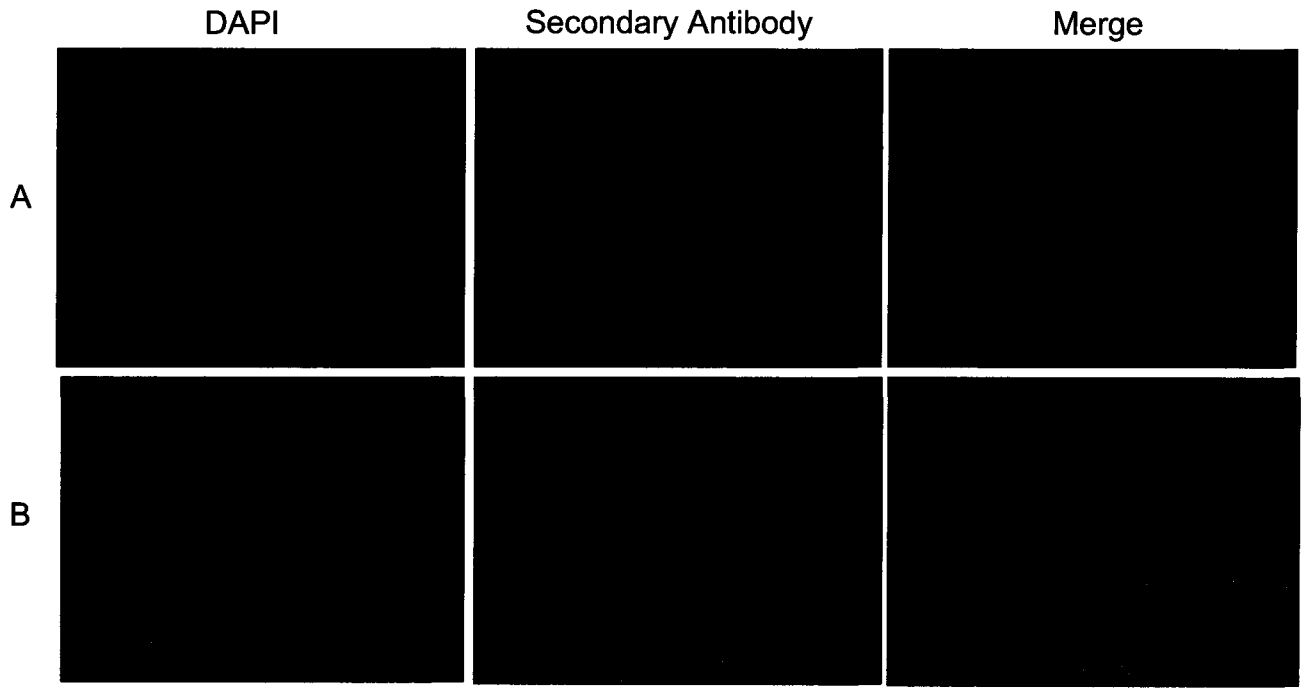


Figure 19: Immuno-fluorescence staining of endogenous Hemogen within undifferentiated MEL cells.

Cellular localization of endogenous Hemogen in undifferentiated MEL cells. Cells were affixed to poly-lysine coated slides and immunostained with commercial Hemogen primary antibody sc-68360 (1:300). FITC- conjugated anti-rabbit secondary antibody (1:300) was visualized in green. DAPI in blue was used to visualize DNA. First column exhibits DAPI staining of cellular DNA. Regions intensely stained by DAPI are highly condensed chromatin. Third column is a merge of both DAPI and FITC signals.

(A) Cells were treated only with FITC- conjugated secondary antibody. The absence of FITC emitted signal serves a negative control for the secondary antibody.

(B) Cells were treated with both primary Hemogen and anti-rabbit FITC- conjugated antibodies.

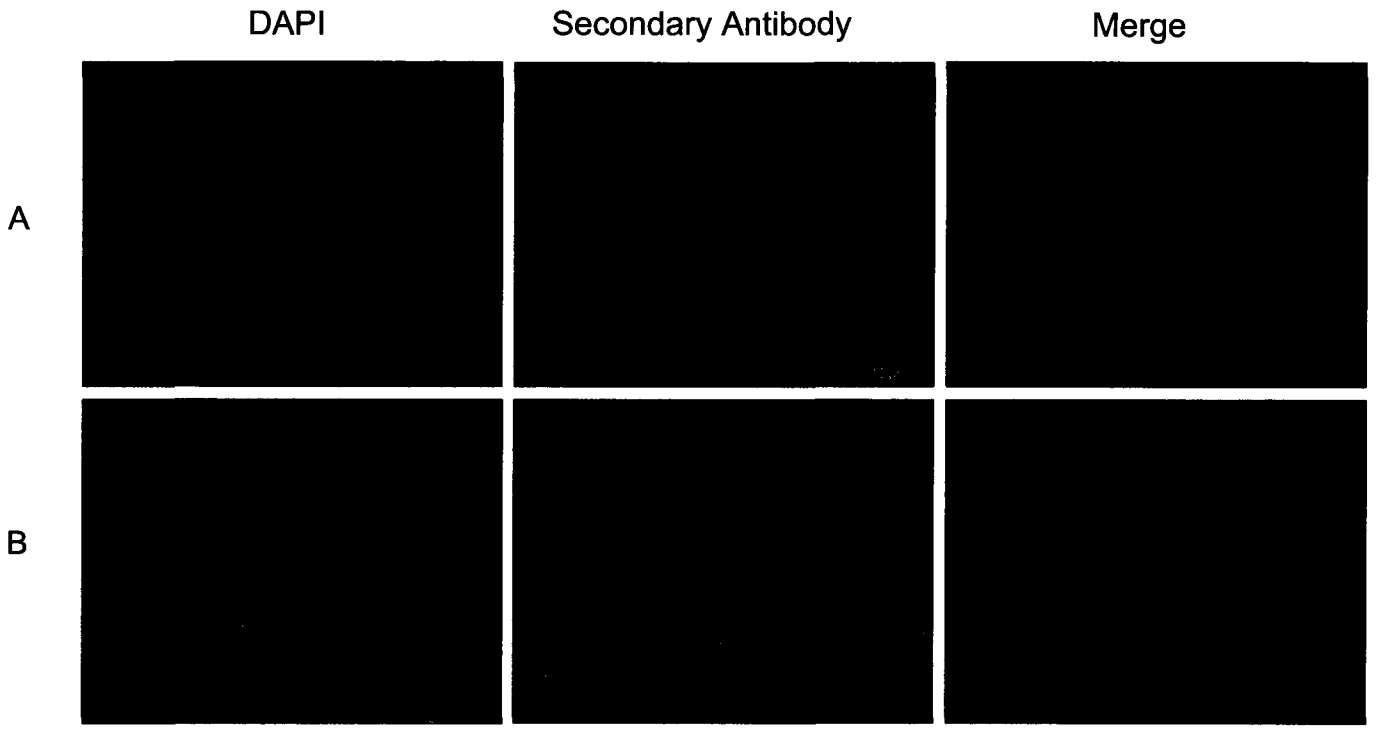


Figure 20: Immuno-fluorescence staining of endogenous Hemogen within differentiated MEL cells.

Cellular localization of endogenous Hemogen in differentiated MEL cells. Cells were affixed to poly-lysine coated slides and immunostained with commercial Hemogen primary antibody sc-68360 (1:300). FITC- conjugated anti-rabbit secondary antibody (1:300) was visualized in green. DAPI in blue was used to visualize DNA. First column exhibits DAPI staining of cellular DNA. Regions intensely stained by DAPI are highly condensed chromatin. Third column is a merge of both DAPI and FITC signals.

(A) Cells were treated only with FITC- conjugated secondary antibody. The absence of FITC emitted signal serves a negative control for the secondary antibody.

(B) Cells were treated with both primary Hemogen and anti-rabbit FITC- conjugated antibodies.

To determine the expression dynamics of Hemogen during erythroid differentiation of MEL cells, we performed a western blot of endogenous Hemogen from nuclear extracts of undifferentiated and 24 hours, 48 hours and 72 hours differentiated MEL cells (figure 21). Hemogen protein levels progressively increased with differentiation. As a positive control for MEL differentiation, we detected a progressive up-regulation of p45 after DMSO-induced differentiation (figure 21) as previously reported (Nagai, Igarashi et al. 1998 and Johnson, Christensen et al. 2001). Levels of the p89 subunit of the general transcription factor complex TFIIH were constant indicating similar protein amounts between samples (figure 21).

3.3 Investigation of Hemogen Associating Proteins

Hemogen was previously identified by a proteomic screen of MafK to preferentially interact with MafK after MEL differentiation (Brand et al. 2004). To confirm the interaction between MafK and Hemogen, we performed a MafK IP using differentiated MEL nuclear extracts. We performed MafK IP in parallel with mock IP using iso-type matched IgG as our negative control. We observed MafK solely within MafK IP (figure 22). The absence of MafK in mock IP confirmed the specificity of our IP. Moreover, we observed Hemogen solely within specific MafK IP (figures 22). The observation of MafK within Hemogen IP elution provided confirmation of the MafK and Hemogen interaction.

The observed MafK and Hemogen interaction and the fact that MafK is part of large multi-protein complexes in both undifferentiated and differentiated MEL cells

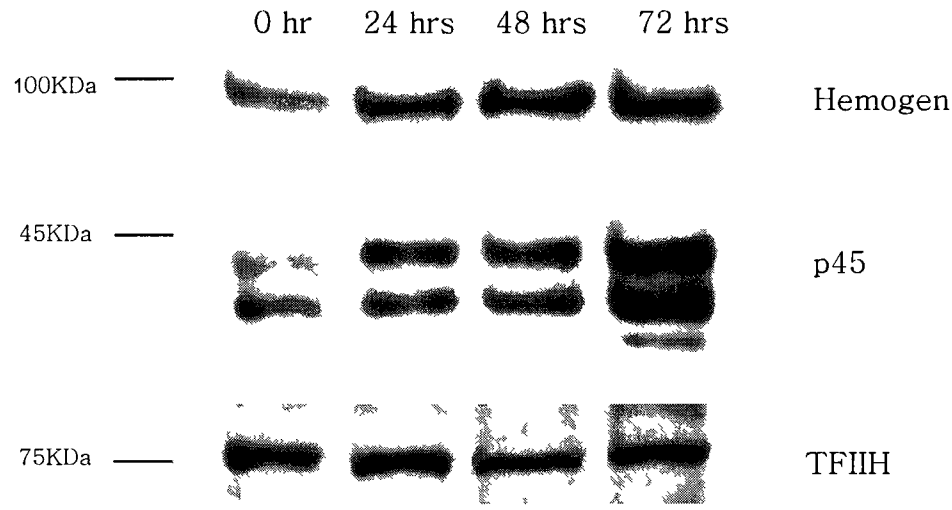


Figure 21: Western blot of MEL differentiation time-course. Western blot of Hemogen (sc-68360), TFIIH (sc-20697) and p45 (sc-291) of MEL nuclear extracts before DMSO-induced differentiation (0 hr) and 24, 48 and 72 hours after 2% DMSO-induced differentiation in MEL cells.

Total protein amounts between samples were equilibrated using Bradford assay. TFIIH was used as a loading control. The increase in p45 levels during MEL differentiation was used as a positive control for MEL differentiation (Nagai, Igarashi et al. 1998; Johnson, Christensen et al. 2001).

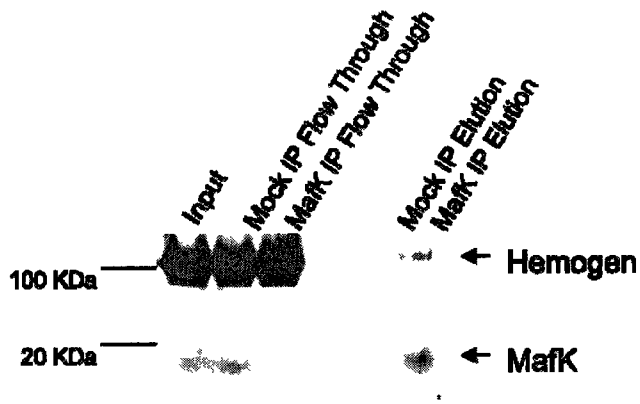


Figure 22: MafK IP of differentiated MEL nuclear extract.

MafK and Hemogen western blots of MafK IP indicated MafK and Hemogen interaction within differentiated MEL nuclear extract.

Three days differentiated MEL nuclear extracts were incubated with commercial Hemogen antibody or isotype-matched IgG conjugated protein-A sepharose resin. Captured proteins were eluted by boiling in SDS buffer and separated by 12% SDS-PAGE. MafK IP was performed with the MafK antibody (sc-C16) in parallel with mock IP using IgG. Western blots were revealed with the commercial Hemogen antibody (sc-68360) or MafK antibody (sc-C16). Equal sample volumes were loaded.

implicated that Hemogen may exist in protein complexes. We performed Hemogen IP using our polyclonal Hemogen antibody (9474P) in parallel with mock IP using cognate pre-immunized rabbit serum. In preparation for large-scale Hemogen IP for mass spectrometry, we optimized the endogenous Hemogen IP by minimizing the amount of non-specific proteins in the mock IP visualized by silver staining (figure 23) by optimizing the elution temperature and buffer.

For mass spectrometry analysis of Hemogen associating proteins, we performed a large scale IP of differentiated MEL nuclear extract. We used our polyclonal Hemogen antibody for Hemogen specific IP. As a negative control, we used the pre-immune serum. We separated proteins within the IP elution by SDS-PAGE and silver stained the SDS-PAGE gel using a protocol compatible for subsequent mass spectrometry analysis. Gel sections were cut from the specific IP lane and equivalent sections were cut from the mock IP lane (figure 24). Peptides were extracted from the gel pieces and subjected to mass spectrometry analysis.

By mass spectrometry, we observed Hemogen solely in our specific IP. Proteins identified within mock IP elution are considered non-specific and were excluded from our list of candidate Hemogen interactors (tables 2 and 3). Our list of putative Hemogen interactors after erythroid differentiation included members of the chromatin remodeling and modifying complexes, sequence-specific DNA-binding transcription factors, basal transcription factors, members of the mediator complex, anaphase promoting complex and architectural proteins.

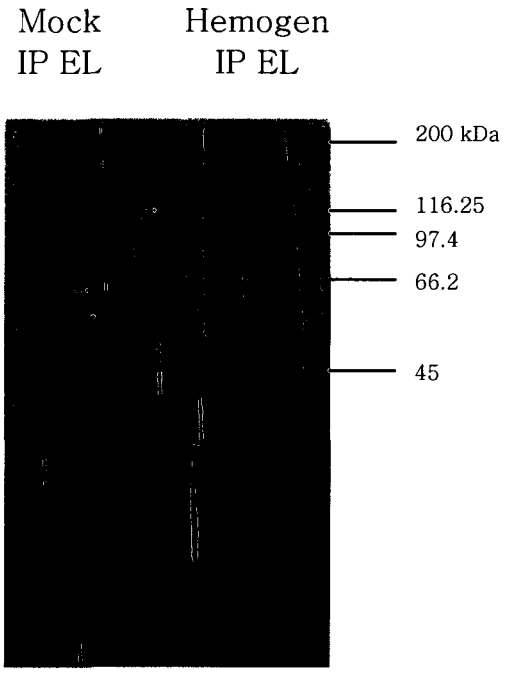


Figure 23: Silver stain of Hemogen IP elutions of differentiated MEL nuclear extracts.

Silver stained SDS-PAGE gel of small-scale mock IP elution (mock IP EL) and Hemogen IP elution (Hemogen IP EL). Three days differentiated MEL nuclear extracts were incubated with Hemogen anti-serum or pre-immuned serum conjugated protein-A sepharose resin. Captured proteins were eluted by boiling in SDS buffer and separated by 12% SDS-PAGE. Immunoprecipitation (IP) was performed with the Hemogen serum (9474P) or pre-immune serum (mock). Molecular weights are indicated to the right.

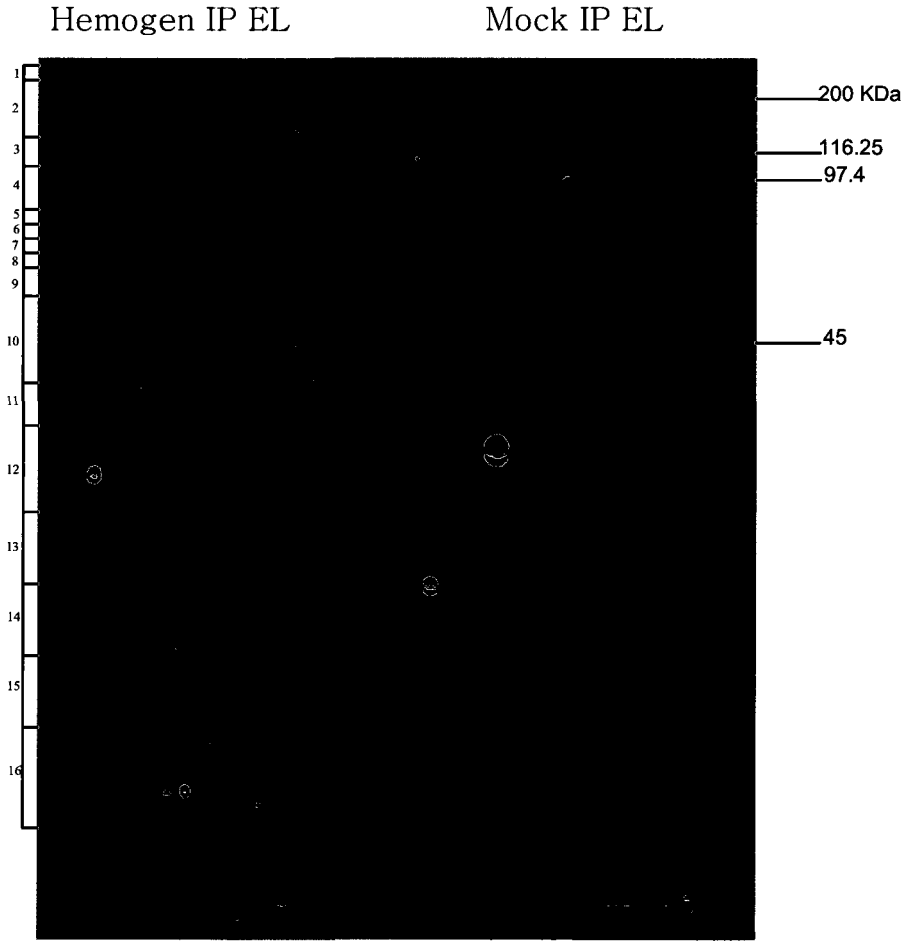


Figure 24: Silver stain of Hemogen IP elutions of differentiated MEL nuclear extracts for mass spectrometry.

Silver stained SDS-PAGE gel of mock IP elution (mock IP EL) and Hemogen IP elution (Hemogen IP EL) for mass spectrometry analysis. Silver stained SDS-PAGE gel of small-scale mock IP elution (mock IP EL) and Hemogen IP elution (Hemogen IP EL). Immunoprecipitation (IP) was performed with the Hemogen serum (9474P) or pre-immune serum (mock). Molecular weights are indicated to the right. Gel sections were cut and numbered as indicated on the left. Gel sections 2 to 7 contain Hemogen (see table 3).

Table 2: Tandem mass spectrometry identification of specific Hemogen associated proteins by IP after MEL erythroid differention classified by function. Sequest and Mascot algorithms were used to correlate observed mass spectrum with therotical mass spectrum from a database to determine peptide sequence. IPI- international protein index number. Table does not include Hemogen.

Name	IPI	MW (KDa)	Number of Unique Peptide (Sequest)	Coverage (%) (Sequest)	Identification Probability (%) (Sequest)	Mascot Score	Gel Band # (fig 24)
TRANSCRIPTION							
Sequence-specific DNA-binding Proteins							
CEB/Pβ	116613.1	31.2	1	9.8	52.8	n/a	12
E12 of E2-α	122714.2	67.7	1	2.2	98.77	n/a	8
POU2F2	230116.3	60.1	1	2.9	79.7	n/a	14
TAL-1/SCL	123645.1	34.3	1	2.2	78.3	n/a	10
Basal Machinery							
TAF5	223695.1	87.2	2	5	n/a	47	5
TAF10	321331.3	21.8	2	17	n/a	76	13
Mediator Complex							
MED12, Isoform 2	620781.4	242.2	1	0.6	99.62	60	2
MED17	459787.2	72.4	1	3.7	99.2	n/a	7
Cofactors							
Bcor	229530.4	111.33	1	1.5	99.6	71	2
CHD5, Isoform 2	608001.5	222.52	1	0.6	n/a	62	2
DNMT-1	469323.1	183.1	5	3.9	100	95	2
FOG-1	132648.1	105.9	1	2.2	95.12	52	3
INO80	378561.2	176.4	2	2.7	100	n/a	2
RBBP6	551082.3	195.2	1	0.7	99.4	58	2
SAF-B	378561.2	116.85	3	2.7	100	53	2
TRRAP	330902.4	291.5	1	0.4	99.81	n/a	1
NuRD/Mi-2 Complex							
CHD4	396802.1	217.6	3	2.8	100	67	1,2
MTA2	128230.1	75	3	2.4	99.56	75	8
p66 α	229784.2	67.3	1	1.9	96.4	n/a	8
p66 β	128615.1	65.4	1	5.2	99.6	78	8
RBBP4	122696.5	47.6	1	5.5	99.74	n/a	9
RBBP7	552530.2	47.8	3	11	n/a	84	9
BAF Complex (SWI/SNF type)							
BAF53a	323660	47.4	1	3	n/a	81	9
BAF60b	115461.4	55.4	1	10	83.73	n/a	9
BRG-1	460668.2	181.7	3	2	100	n/a	2
Transcription Cyclins							
Cyclin-K	153134.1	61.3	1	4.7	99.1	n/a	5
Cyclin-T1	316214.3	80.5	2	4	100	n/a	5,6

name	IPI	MW (KDa)	Number of Unique Peptide (Sequest)	% coverage (Sequest)	Identification Probability (%) (Sequest)	Mascot Score	Gel Band # (fig. 24)
CELL CYCLE MODULATOR							
APC complex							
APC1	137228 1	218 1	4	3 3	100	80	2
APC5	380355 1	83	2	3 7	100	75	7
CHROMOSOME CONDENSATION/SEGREGATION							
smcHD1	137433 3	225 5	6	4 8	100	130	2
NSBP1	311626 5	45 3	3	6 9	99 5	63	8
PROTEINS OF UNKNOWN/UNCLEAR FUNCTION							
WDC146	170242 1	145 8	5	4 4	100	101	2
ZC3H18	648513 1	105 6	1	2 6	99 6	n/a	2
OTHERS							
Cbll1	310594 2	54 3	1	6	99 7	n/a	9
Nucleoporin	330624 1	45 3	1	6 9	99 5	n/a	8
Microspherule protein 1	222228 5	58 2	1	2 2	86 2	n/a	9

Table 3: Tandem mass spectrometry identification of specific Hemogen associated proteins by IP after MEL erythroid differention within each gel band. Hemogen identification is highlighted. Sequest and Mascot algorithms were used to correlate observed mass spectrum with therotical mass spectrum from a murine database to determine peptide sequence. IPI-international protein index number.

Gel Band 1

Name	IPI	MW (KDa)	Number of Unique Peptide (Sequest)	Coverage (Sequest) (%)	Identification Probability (%) (Sequest)	Mascot Score	Number of Unique Peptide (Mascot)	Coverage (%) (Mascot)
CHD4	396802.1	217.8	1	0.7	98.3	n/a	n/a	n/a
TRRAP	330902.4	291.5	1	0.4	99.8	n/a	n/a	n/a

Gel Band 2

Name	IPI	MW (KDa)	Number of Unique Peptide (Sequest)	Coverage (Sequest) (%)	Identification Probability (%) (Sequest)	Mascot Score	Number of Unique Peptide (Mascot)	Coverage (%) (Mascot)
APC1	137228.1	218.1	4	3.3	100	80	3	2
BCOR	229530.4	111.33	1	1.5	99.6	71	2	3
BRG-1	460668.2	181.7	3	2	100	n/a	n/a	n/a
CHD4	396802.1	217.6	3	2.8	100	67	6	5
CHD5, Isoform 2	608001.5	222.52	n/a	n/a	n/a	62	3	1
DNMT-1	469323.1	183.1	5	3.9	100	95	3	2
Hemogen	458120.2	55	4	10.6	100	346	11	30
INO80	378561.2	176.4	2	2.7	100	n/a	n/a	n/a
MED 12, Isoform 2	620781.4	242.2	1	0.6	99.62	60	5	4
Nucleoporin	330624.1	45.3	1	6.9	99.5	n/a	3	3
RBBP6	551082.3	195.2	1	0.7	99.4	58	2	2
SAF-B	378561.2	116.85	3	2.7	100	53	3	2
smcHD1	137433.3	225.5	6	4.8	100	130	12	8
WDC146	170242.1	145.8	5	4.4	100	101	3	4
ZC3H18	648513.1	105.6	1	2.6	99.6	n/a	n/a	n/a

Gel Band 3

Name	IPI	MW (KDa)	Number of Unique Peptide (Sequest)	Coverage (Sequest) (%)	Identification Probability (%) (Sequest)	Mascot Score	Number of Unique Peptide (Mascot)	Coverage (%) (Mascot)
FOG-1	132648.1	105.9	1	2.2	95.12	52	2	3
Hemogen	458120.2	55	2	6.2	100	126	4	13

Gel Band 4

Name	IPI	MW (KDa)	Number of Peptide Unique (Sequest)	Coverage (Sequest) (%)	Identification Probability (%) (Sequest)	Mascot Score	Number of Unique Peptide (Mascot)	Coverage (%) (Mascot)
Hemogen	458120.2	55	1	8.1	99.8	284	6	19

Gel Band 5

Name	IPI	MW (KDa)	Number of Peptide Unique (Sequest)	Coverage (Sequest) (%)	Identification Probability (%) (Sequest)	Mascot Score	Number of Unique Peptide (Mascot)	Coverage (%) (Mascot)
Cyclin-T1	316214.3	80.5	2	4	100	n/a	n/a	n/a
Hemogen	458120.2	55	4	8.3	100	912	14	42
TAF5	223695.1	87.2	n/a	n/a	n/a	47	2	5

Gel Band 6

Name	IPI	MW (KDa)	Number of Peptide Unique (Sequest)	Coverage (Sequest) (%)	Identification Probability (%) (Sequest)	Mascot Score	Number of Unique Peptide (Mascot)	Coverage (%) (Mascot)
Cyclin-T1	316214.3	80.5	2	4	100	62	3	5
Hemogen	458120.2	55	6	8.3	100	n/a	n/a	n/a

Gel Band 7

Name	IPI	MW (KDa)	Number of Peptide Unique (Sequest)	Coverage (Sequest) (%)	Identification Probability (%) (Sequest)	Mascot Score	Number of Unique Peptide (Mascot)	Coverage (%) (Mascot)
APC5	380355.1	83	2	3.7	100	75	2	3
Cyclin-K	153134.1	61.3	1	4.7	99.1	n/a	n/a	n/a
Hemogen	458120.2	55	2	6.2	100	174	7	22
MED 17	459787.2	72.4	1	3.7	99.2	n/a	n/a	n/a

Gel Band 8

Name	IPI	MW (KDa)	Number of Peptide Unique (Sequest)	Coverage (Sequest) (%)	Identification Probability (%) (Sequest)	Mascot Score	Number of Unique Peptide (Mascot)	Coverage (%) (Mascot)
E12 of E2α	122714.2	67.7	1	2.2	98.77	n/a	n/a	n/a
p66 α	229784.2	67.3	1	1.9	96.4	n/a	n/a	n/a
p66 β	128615.1	65.4	1	5.2	99.6	78	1	5
MTA1/2	128230.1	75	3	2.4	99.56	75	3	5
NSBP1	311626.5	45.3	3	6.9	99.5	63	2	9

Gel Band 9

Name	IPI	MW (KDa)	Number of Peptide Unique (Sequest)	Coverage (Sequest) (%)	Identification Probability (%) (Sequest)	Mascot Score	Number of Unique Peptide (Mascot)	Coverage (%) (Mascot)
CBLL1	310594 2	54 3	1	6	99 7	n/a	n/a	n/a
BAF53a	323660	47 4	n/a	n/a	n/a	81	3	3
BAF60b	115461 4	55 4	1	10	83 73	n/a	n/a	n/a
RBBP4	122696 5	47 6	1	5 5	99 74	n/a	n/a	n/a
RBBP7	552530 2	47 8	n/a	n/a	n/a	84	3	7
Microspherule protein 1	222228 5	58 2	1	2 2	86 2	n/a	n/a	n/a

Gel Band 10

Name	IPI	MW (KDa)	Number of Peptide Unique (Sequest)	Coverage (Sequest) (%)	Identification Probability (%) (Sequest)	Mascot Score	Number of Unique Peptide (Mascot)	Coverage (%) (Mascot)
TAL-1/SCL	123645 1	34 3	1	2 2	78 3	n/a	n/a	n/a

Gel Band 11

No unique specific IP protein identified

Gel Band 12

Name	IPI	MW (KDa)	Number of Peptide Unique (Sequest)	Coverage (Sequest) (%)	Identification Probability (%) (Sequest)	Mascot Score	Number of Unique Peptide (Mascot)	Coverage (%) (Mascot)
CEB/Pβ	116613 1	31 2	1	9 8	52 8	n/a	n/a	n/a

Gel Band 13

Name	IPI	MW (KDa)	Number of Peptide Unique (Sequest)	Coverage (Sequest) (%)	Identification Probability (%) (Sequest)	Mascot Score	Number of Unique Peptide (Mascot)	Coverage (%) (Mascot)
TAF10	321331 3	21 8	n/a	n/a	n/a	76	2	17

Gel Band 14

Name	IPI	MW (KDa)	Number of Peptide Unique (Sequest)	Coverage (Sequest) (%)	Identification Probability (%) (Sequest)	Mascot Score	Number of Unique Peptide (Mascot)	Coverage (%) (Mascot)
PO2F2	230116 3	60 1	1	2 9	79 7	n/a	n/a	n/a

Gel Band 15

No unique specific IP protein identified

To approximate the size of the Hemogen complexes and to investigate whether there are multiple Hemogen complexes within the nuclear extract, we fractionated protein complexes within differentiated MEL nuclear extracts by size using size exclusion chromatography (SEC). We used superose 6 10/300 GL column with an optimal separation range of 500 KDa to 5 KDa. We performed Hemogen western blot on every second SEC fraction to determine the fractions most abundant in Hemogen (figure 25). Once we determined the elution profile of Hemogen by SEC, we compared Hemogen's elution profile with the elution profiles of purified globular proteins of known molecular weights to approximate the size of the Hemogen complex (figure 26). The majority of Hemogen within differentiated nuclear extract was approximately between 382 KDa to 117.46 KDa (figure 26). We observe only one peak in which Hemogen eluted out of the SEC column (figure 25). Therefore Hemogen may belong to one complex or within several Hemogen complexes similar in size.

Western blot of SEC fractions can provide an indication of protein-protein interactions if two proteins are detected within the same fraction. Since we observed an interaction between MafK and Hemogen by IP, we assessed whether MafK cofractionated with Hemogen. Hemogen eluted out of the size exclusion column slightly before MafK. However there was an overlap in the elution patterns of Hemogen and MafK in which both proteins are co-eluted (figure 25).

Since five peptides of Dnmt-1 were identified within Hemogen specific IP by mass spectrometry analysis, we investigated whether Dnmt-1 co-fractionated with

Hemogen and observed the presence of Dnmt-1 in all fractions containing Hemogen (figure 25).

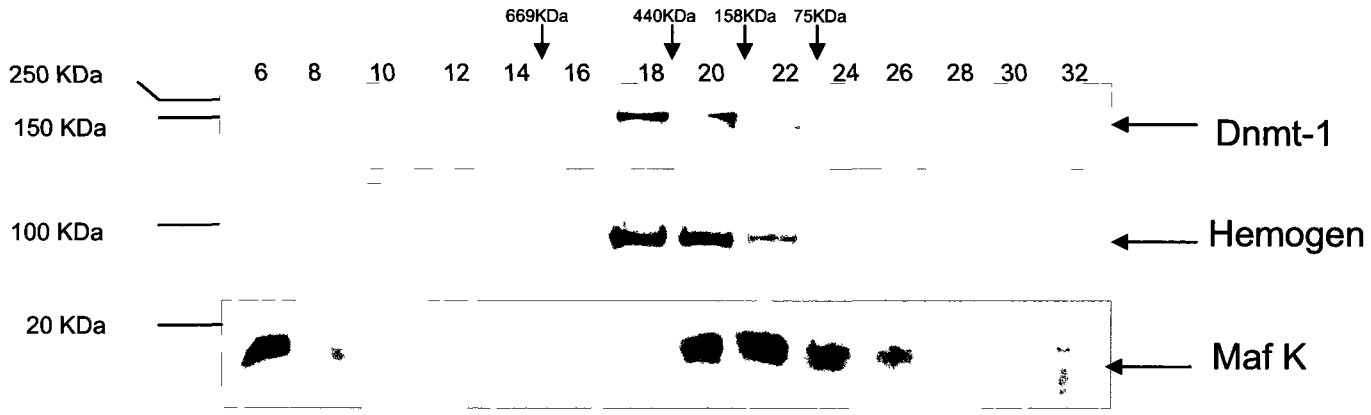


Figure 25: Western blot of size exclusion chromatography (SEC) fractions using differentiated MEL nuclear extracts.

Differentiated MEL nuclear extract were loaded into Superose 6 10/300 column. 500 µl fractions were initially collected at elution volume of 5 ml. The fraction number is labeled across the top. The void volume, approximated by Dextran blue, is within the 5th fraction collected. Approximate sizes of fractions are labeled in kilodaltons across the top according to globular protein standards of known molecular weight (table 3).

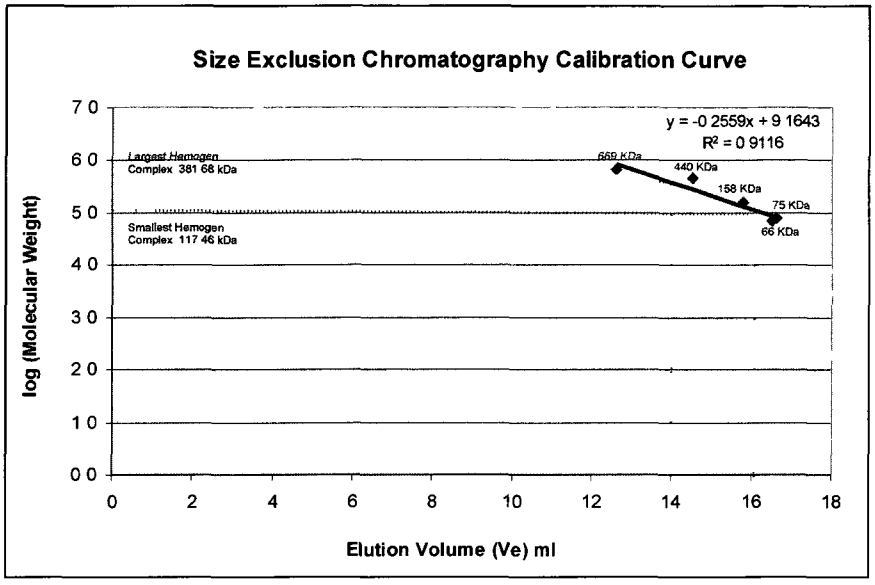


Figure 26: Size exclusion chromatography (SEC) calibration curve.

Size exclusion chromatography calibration curve using protein standards (table 3). Hemogen complex elution volumes were obtained from Hemogen western blot of SEC fractions. Size of Hemogen complexes were calculated from trend-line equation of the calibration curve.

Table 4- Data Points for SEC Calibration Curve

Protein Standards	Molecular Weight (Da)	log(Molecular Weight)	Elution Volume (Ve, ml)
thyroglobulin	669000	5 8	12.6
ferritin	440000	5.6	14.52
aldolase	158000	5.2	15.78
conalbumin	75000	4.9	16.62
albumin	66000	4.8	16.51

Chapter 4: Discussion

Hemogen is a murine protein expressed in active sites of hematopoiesis during embryogenesis and adulthood (Li et al. 2001). The mechanistic function of Hemogen has not been currently elucidated. However, several reports indicate that *Hemogen* deregulation is associated with neoplasm. *Hemogen* is over-expressed in the peripheral blood cells of several types of leukemias (Li et al. 2004) and is associated with poor prognosis in patients suffering from acute myeloid leukemia (An et al. 2005). Over-expression of *Hemogen* in HSCs disrupted hematopoietic homeostasis and resulted in extramedullary myelopoiesis in the spleen and reduced lymphopoiesis (Li et al. 2007). Moreover, evidence also support Hemogen's role in erythroid differentiation. Hemogen or its homologs are expressed in primitive erythrocytes, erythroid precursors and tissues of active erythropoiesis (Kruger et al. 2002 and Yang L.V. et al. 2001). One of the hallmarks of erythroid differentiation is the expression of hemoglobin. Hemogen associates with MafK, a subunit of the NF-E2 complex that activates *β-globin* expression during erythroid differentiation (Johnson et al. 2001). Moreover, the key stimulator of erythroid differentiaton, GATA-1 activates the expression of human Hemogen homolog, EDAG (Li et al. 2008). Hemogen's association to leukemogenesis and the erythroid differentiation perhaps warrants the study of Hemogen's role in erythroid differentiation. Investigating the associating partners of Hemogen is a step into understanding Hemogen's mechanistic function.

Previous evidence suggested that Hemogen is a nuclear protein. Immuno-staining of exogenous Hemogen within COS-7 cells revealed Hemogen's nuclear localization (Li

et al. 2001). However the cellular localization of endogenous Hemogen within erythroid cells has never been investigated. Therefore to investigate Hemogen's cellular localization within MEL cells, we performed immunostaining of endogenous Hemogen in both undifferentiated and differentiated MEL cells. We used DAPI, which stained DNA, to demark the nuclear region. Hemogen was found within the nucleus of undifferentiated and differentiated MEL cells (figures 19 and 20).

To investigate whether Hemogen protein levels are influenced by erythroid differentiation, we performed western blot of MEL nuclear extracts at different time points of differentiation. As terminal erythroid differentiation progressed in MEL cells, we observed a progressive up-regulation of Hemogen protein (figure 21). Since GATA-1 is an activator of Hemogen (Li et al. 2008) and GATA-1 activity increases during MEL erythroid differentiation (Choe et al. 2001), the increase in Hemogen levels may be associated with increases in GATA-1 activity.

Previous MafK proteomic screen implicates an interaction between MafK and Hemogen in MEL cells. We performed a co-IP experiment and confirmed this interaction (figure 22). The observed MafK and Hemogen interaction and the fact that MafK is part of large multi-protein complexes in both undifferentiated and differentiated MEL cells (Brand et al. 2004) implicated that Hemogen may exist in protein complexes.

Hence we created Hemogen antibodies that were applicable in Hemogen IP to identify its associating partners. N-terminal Hemogen bacterial protein was used as

antigen for Hemogen antibody production in two rabbits. We observed three proteins species within our purified fraction (figure 10). Each species was identified as Hemogen by mass spectrometry (figures 11-13). The approximate theoretical size of N-terminal Hemogen is 17 KDa. Since we used denaturing conditions to reduce secondary and tertiary protein structures, size discrepancy between the three species observed may be due to post-translation modification and the inherent character of the primary sequence. A report demonstrated that SDS has different affinity for free amino acids. Specially, acidic amino acids do not associate with SDS (Maley et al. 1977). Proline minimally associates with SDS (Maley et al. 1977). Glutamic acid attributes to 11.9% of Hemogen's total residue content while proline contributes to 10.3% (Li et al. 2001). Both the prevalence of glutamic acid and proline in Hemogen are higher than average levels (Brendel et al. 1992). Therefore it is possible that SDS did not efficiently confer negative charge to the protein thereby reducing the speed of migration during electrophoresis. Similar discrepancy between theoretical and observed molecular weight in SDS-PAGE was detected for proliferation cell nuclear antigen (PCNA) (Liang et al. 1992). Both *in vitro* translation and bacterial expressed PCNA migrated at the same rate as endogenous PCNA suggesting that this discrepancy in size may be due to the inherent residue content of the PCNA (Liang et al. 1992). Since, the 30 KDa Hemogen species is the most abundant species, the faster migrating Hemogen species at approximately 19 KDa and 17 KDa may be degradation products of the 30 KDa species.

We performed Hemogen western blot with differentiated MEL nuclear extracts using rabbit anti-Hemogen serum and pre-serum as a negative control. We detected a 100

KDa species with the anti-Hemogen sera but not in the pre-immuned sera (figure 14). Since the theoretical size of Hemogen was 55KDa, we assessed whether the 100 KDa band detected by our anti-Hemogen sera corresponded to Hemogen. We expressed Flag-Hemogen in MEL cells and performed a Flag western blot in parallel with a Hemogen western blot using MEL nuclear extracts. A 100 KDa was observed in both western blots (figure 17). In addition, a commercial Hemogen antibody also recognized the same 100 KDa species (figure 14). The commercial Hemogen antibody was generated using a Hemogen primary sequence distinct from the sequence used to generate our anti-Hemogen sera. These observations suggested that the 100 KDa species detected by the anti-Hemogen serum corresponded to Hemogen. The specificity of the anti-Hemogen serum was further confirmed when we observed Hemogen as the predominant protein within Hemogen IP elution from our mass spectrometry analysis (table 3). The size discrepancy between theoretical size and observed size in a SDS-PAGE gel may be due to post-translational modifications as well as an over-representation of glutamic acid and proline residues which have low affinity to SDS. We did not detect sumoylation within Hemogen IP elution by sumo-1 and sumo2/3 western blot (data not shown). Moreover, we did not observe ubitquitin by mass spectrometry (data not shown). It is possible that sequence coverage was not comprehensive to detect ubiquitin. We are unsure whether other post-translational modifications may result in the observed size difference.

From our list of putative Hemogen associating proteins, many were involved in transcriptional regulation (table 2). Interestingly, members of the NuRD complex were identified as putative interactors of Hemogen. The NuRD complex is a multi-subunit

protein complex generally associated with gene repression. Identified members of the NuRD complex included: CHD4, MTA2, RBBP4, RBBP7, p66α and p66β. Moreover, we also identified FOG-1. The NuRD complex subunits we identified were similar to the subunits previously identified to associate with FOG-1 in MEL cells (Hong et al. 2005). Members of the erythroid NuRD complex includes: MTA-1, MTA-2, p66α, HDAC1, HDAC2, RBBP4, RBBP7, CHD4, MBD3, and MBD2 (Hong et al. 2005). Many members of this NuRD complex such as MTA-1, MTA-2, RBBP4 and RBBP7 directly interact with the N-terminus of FOG-1 (Hong et al. 2005).

Along with FOG-1 and members of the NuRD complex, we identified GATA-1. FOG-1 associates with GATA-1 (Rodriguez et al. 2005). Hemogen may be part of the FOG-1/NuRD/GATA-1 complex (Hong et al. 2005). Previous reports revealed similar FOG-1 and GATA-1 knockout phenotypes (Tsang et al. 1998 and Pevny et al. 1991) and suggested that FOG-1 mediates the expression of a large number of GATA-1 target genes. GATA-1 is considered as a master regulator of erythroid differentiation. Microarray gene expression analysis of GATA-1 induced erythroid differentiation revealed that GATA-1 is involved in the activation and repression of target genes implicated in all aspects of erythroid cell maturation. A wave of early gene repression including mitogenic genes and genes of early acting transcription factors occurs in GATA-1 mediated erythroid differentiation (Welch et al. 2004). GATA-1 proteomic screen performed in MEL cells during erythroid differentiation and ChIP data suggested that the FOG-1/NuRD/GATA-1 complex mediated the repression of mitogenic genes and genes of early acting transcription factors such as *kit* and *GATA-2* (Hong et al. 2005 and

Rodriguez et al. 2005). *In vivo* disruption of the FOG-1/NuRD interaction in mice also revealed abberant expression of alternative hematopoietic lineage genes within erythroid cells (Gao et al. 2009). If Hemogen associates with the GATA-1/FOG-1/NuRD complex, then Hemogen may be recruited to GATA-1 repressed target genes such as *GATA-2*, *kit* or genes of alternative hematopoietic lineages. It should be noted that GATA-1 was identified in both specific and mock IP. There is a possibility that GATA-1 was more abundantly found within our specific IP in comparison to our mock IP. Since our mass spectrometry analysis was not quantitative we cannot dispel this possibility.

Recent ChIP experiments revealed that members of the NuRD complex along with FOG-1 and GATA-1 were found not only at repressed genes but also at transcriptionally activated genes in erythroid cells. Genes actively expressed prior to erythroid differentiation were targets of the NuRD complex (Miccio et al. 2010). Moreover, active genes in megakaryocytes were also targeted by members of the NuRD complex (Miccio et al. 2010). Hemogen's possible association with the NuRD/FOG-1/GATA-1 complex does not necessarily implicate Hemogen's role as a repressor. The NuRD complex has histone deacetylation mediated by HDAC1/2 and chromatin remodeling ATPase activities mediated by CHD4 (Wolffe et al. 2000). RBBP4 and RBBP7 are structural subunits that act as an interface for protein-protein interaction (Wolffe et al. 2000). RBBP4 interacts with histone-fold domain of histone H4 (Wolffe et al. 2000). It is hypothesized that CHD4 disrupts the nucleosome to permit RBBP4 access to the histone-fold domain of histone H4 which normally lies sequestered inside the coils of nucleosomal DNA (Wolffe et al. 2000). RBBP4 can then directly interact with

HDAC1/2 and tether the deacetylase to the target site for deacetylation at the N-terminal tail of histone H4 (Wolffe et al. 2000). Perhaps deacetylation depends on efficient engagement of the NuRD complex with the chromatin (Miccio et al. 2010).

Interestingly, reporter assays revealed that NuRD may act as a co-activator (Miccio et al., 2010). Significantly reduced levels of the reporter gene expression were observed upon knockdown of members of the NuRD complex. It would be insightful to investigate whether Hemogen is enriched to transcriptionally active or to repressed genes previously observed to be targets of NuRD.

It is also possible that Hemogen may associate with NuRD in alternative complexes in the absence of GATA-1 and FOG-1. Primarily, we would need to perform co-IP experiments on Hemogen and members of the NuRD complex to validate our mass spectrometry identifications. Subsequently, to determine whether Hemogen associates with NuRD independent of FOG-1 and GATA-1, we can perform sequential co-IP. GATA-1 and NuRD do not exist in a complex without FOG-1 within MEL cells (Rodriguez et al. 2005). Immuno-depleted FOG-1 nuclear extracts may be used for IP of a member of the NuRD complex followed by Hemogen western blot. This will determine the existence of the NuRD/Hemogen interaction in the absence of FOG-1 and GATA-1. In MEL cells, NuRD preferentially associated with MafK before erythroid differentiation (Brand et al. 2004). Hemogen may be associated to NuRD through interaction with MafK. Since we used MEL nuclear extracts after erythroid differentiation, perhaps the majority of the MafK complexes do not contain members of the NuRD complex.

We observed TAL-1 as a putative associating partner of Hemogen. TAL-1 is part of a multi-protein pentameric complex consisting of LMO2, E47, Ldb1 and GATA-1 (Wadman et al. 1997). The pentameric complex recognizes TAL-1 and GATA-1 composite motifs in regulatory elements of many erythroid specific genes (Anderson et al. 1998; Valverde-Garduno et al., 2004 and Fujiwara, T. et al. 2009). A recent report suggested that TAL-1/GATA-1 interaction is associated with active eythroid genes (Tripic et al. 2009). If Hemogen associates with the TAL-1 containing pentameric complex, then Hemogen may be involved in gene expression.

Hemogen also putatively associated with BRG-1 the catalytic subunit of the SWI/SNF-like Baf complex (Wu et al. 2009). BRG-1 occupies the enhancers and promoters of *β-globin* and *α-globin* loci (Kim et al. 2007 and Kim, S-I. et al. 2009). BRG-1 is recruited to the major-regulatory element (MRE) enhancer and to the *α-globin* promoter within the *α-globin* locus. BRG-1 can also be found in the *β-globin* locus. BRG-1 occupies HS4, HS3, HS2 and the *β-major* promoter (Kim et al. 2007). BRG-1 is required for chromatin looping that facilitates an interaction between MRE and *α-globin* promoter (Kim, S-I. et al. 2009). The putative interaction between Hemogen and BRG-1 may suggest Hemogen's recruitment to the two globin loci.

Interestingly, we also identified smcHD1 as a putative interactor of Hemogen. Six unique peptides were identified in our Hemogen specific IP. The identification probability according to Proteinprophet was 100%. SmcHD1 is an uncharacterized protein with an N-terminal ATPase domain and a C-terminal SMC hinge domain (Blewitt

et al. 2008) normally found in SMC proteins (Hirano et al. 2006). Evidence suggested that SMC members of the cohesin complex play an active role in regulating gene expression by facilitating interactions between far upstream enhancers and promoters (Rollins et al. 2004). It is unknown whether smcHD1 is recruited to the *β-globin* locus; however its unique protein structure implicates a possible role in facilitating chromatin looping along with BRG-1 within the globin loci.

The transcriptional co-factor, DNA methyltransferase-1 (DNMT-1) was observed as a possible interactor of Hemogen. DNMT-1 is involved in the maintenance of DNA methylation patterns set during embryogenesis (Jones and Liang 2009). DNMT-1 interacts with PCNA during DNA replication to methylate nascent DNA at the replication fork (Jones and Liang 2009). However recent evidence showed that the function of DNMT-1 is not always linked with DNA replication (Spada et al. 2007). The disruption of PCNA and DNMT-1 interaction did not abrogate DNMT-1's catalytic activity suggesting that the maintenance of DNA methylation by DNMT-1 is not strictly linked to its interaction with replication machinery (Spada et al. 2007). We would confirm the DNMT-1 and Hemogen the interaction by co-IP. Interestingly, SmcHD1 mouse mutants are associated with DNA hypomethylation of x-inactivated genes (Blewitt et al. 2008). The assessment of whether SmcHD1 and DNMT-1 are associating proteins might provide insight into the DNA methylation defect in smcHD1 mutants.

To determine the size and number of Hemogen complexes present within differentiated MEL nuclear extracts, we fractionated the nuclear extract by size using size

exclusion chromatography (SEC). Using Hemogen antibody to detect Hemogen within each fraction, we observed a single, broad, Hemogen elution peak (figure 25). We performed SEC using molecular weight standards of known size to approximate the molecular weight of the fractionated Hemogen complexes. The majority of Hemogen eluted between 382 KDa to 117.46 KDa. The size of Hemogen complexes was relatively small if we consider Hemogen's putative association partners, NuRD/FOG-1 complex to have an approximate size of 1.3 MDa (Rodriguez et al. 2005). It is conceivable that the association between Hemogen and NuRD is weak and may be disrupted during the process of SEC.

Since we observed MafK to associate with Hemogen by co-IP in differentiated MEL nuclear extracts, we assessed whether MafK co-eluted out of the SEC column with Hemogen. MafK eluted from the SEC column slightly prior to Hemogen. There is however overlap between the elution profile of MafK and Hemogen. This observation suggested that MafK may not consistently interact with Hemogen. Hemogen complexes range in size from approximately 382 KDa to 117.46 KDa. Therefore we cannot exclude the possibility that Hemogen may exist in different complexes. It is conceivable that Hemogen may exist in at least two different complexes: one that contain MafK and one that does not. Since we observed five DNMT-1 peptides with Hemogen IP in our mass spectrometry analysis, we studied the elution profile of DNMT-1 to assess whether it co-elutes with Hemogen. DNMT-1 was present at every fraction containing Hemogen. This observation is a preliminary indication that Hemogen may be associated with DNMT-1. A co-IP of Hemogen and DNMT-1 would further confirm this observation.

For future work, we would perform reporter assays in an erythroid background such as MEL cells to assess Hemogen's possible role as a transcriptional regulator. Reporter assays of exogenous Hemogen within COS-7 and 293 cells suggested Hemogen as a transcriptional activator (Li et al. 2007). The cellular backgrounds of these reporter assays do not entirely reflect the erythroid background. Erythroid proteins identified in this report to putatively associate with Hemogen such as FOG-1 and TAL-1 are not expressed in COS-7 or in 293 cells.

Since Hemogen protein levels increase with erythroid differentiation, it would be interesting to assess whether the composition of Hemogen complexes also change during differentiation. Moreover, possible Hemogen target genes may vary throughout erythroid differentiation. Expression profile of Hemogen knock-down MEL cells as compared to wild-type MEL cells before and after erythroid differentiation may provide candidate target genes. MafK is essential in *β-major* expression and is enriched in *β-globin* locus of MEL cells. Since Hemogen associated with MafK, it would be interesting to assess Hemogen's localization to the *β-globin* locus and whether *β-major* expression is affected by Hemogen knock-down.

It should be noted that aberrant Hemogen expression is associated with leukemogenesis (An et al. 2005). Primary Ter 119 positive cells purified from murine bone marrow cells do not express Hemogen transcripts (Li et al. 2001). The murine Ter 119 surface antigen is used as a specific marker of erythroid lineage and is expressed from the proerythroblast stage to late stage erythroid cells (Kina et al. 1999). However,

Hemogen transcripts are found within the red pulp of adult mouse spleen where erythropoiesis predominates (Li et al. 2001). Moreover, the rat homolog of Hemogen, RP59 is found within erythroblasts of the adult spleen expressing hemoglobin (Kruger et al. 2002). Perhaps Hemogen is expressed within erythroid precursors of the spleen. Since we used an immortalized cell line, Hemogen expression within primary erythroid cells may be different from MEL cells. Investigating Hemogen's target genes may require reference to primary erythroid cells. The identification of possible Hemogen target genes may provide insight into Hemogen association with leukemogenesis.

Our results suggest that Hemogen may function as a transcriptional regulator. Our immuno-staining results implicate that Hemogen is localized in the nucleus before and after erythroid differentiation. From our proteomic screen of Hemogen after erythroid differentiation, we also observed many transcriptional factors such as TAL-1 and E2A as well as co-factors such as the NuRD complex, DNMT-1, BRG-1 and FOG-1 as putative Hemogen associated proteins. The SMC protein may be involved in *β-globin* locus looping mediated by BRG-1. For future work, we would confirm these interactions by co-IP. Increasing evidence suggests an association between Hemogen deregulation and hematological neoplasm. The implication that Hemogen may function as a transcriptional regulator may inspire the elucidation of Hemogen target genes as a means of understanding disease states caused by Hemogen deregulation.

Reference:

An, L.-L., Li, G., et al. 2005. High expression of EDAG and its significance in AML. Leukemia **19**: 1499-1502.

Anderson, K.P., Crable, S.C., and Lingrel, J.B. 1998. Multiple proteins binding to a GATA-E box-GATA motif regulate the erythroid Kruppel-like factor (EKLF) gene. *J Biol Chem* **273**: 14347-14354.

Andrews, N.C., Erdjument-Bromage, H., et al.1993. Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. *Nature* **362**: 722-728.

Becker, P.B. and Horz, W. 2002. ATP-dependent nucleosome remodeling. *Annu Rev Biochem* **71**: 247-273.

Brendel, V., Bucher, P., et. al. 1992. Methods and algorithms for statistical analysis of protein sequences. *Proc Natl Acad Sci USA* **89**: 2002-2006.

Blewitt, M., Gendrel, A.-V., et al. 2008. SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat Genetics* **40**: 663-669.

Brand, M., Ranish, A.J., et. al. 2004. Dynamic changes in transcription factor complexes during erythroid differentiation revealed by quantitative proteomics. *Nat Struc Mol Biol* **11**: 73-80.

Cairns, B., 2009. The logic of chromatin architecture and remodeling at promoters. *Nature* **461**: 193-198.

Chaturvedi, C. P., Hosey A.M., et al. 2009. Dual role for the methyltransferase G9a in the maintenance of beta-globin gene transcription in adult erythroid cells. *Proc Natl Acad Sci U S A* **106**: 18303-18308.

Choe, K.S, Padparvar, F., et al. 2003. Reversal of tumorigenicity and the block to differentiation in erythroleukemia cells by GATA-1. *Cancer Res* **63**: 6366-6369.

Delgado, M.D., Hallier, M., et al. 1994. Inhibition of Friend cells proliferation by spi-1 antisense oligodeoxynucleotides. *Oncogene* **9**: 1723-1727.

Demer, C., Chaturvedi, C.P., et al. 2007. Activator-mediated recruitment of the MLL2 methyltransferaase complex to the beta-globin locus. *Mol Cell* **27**: 573-584.

Eng, J. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of American Society for Mass Spectrometry* **5**: 976-989.

Fischle, W., Wang, Y., and Allis, C.D. 2003. Histone and chromatin crosstalk. *Curr Opin Cell Biol* **15**: 172-183.

Friend, C., Patuleia, M.C., and De Harven, E. 1966. Erythrocytic maturation in vitro of murine (Friend) virus-induced leukemic cells. *Natl Cancer Inst Monogr.* **22**:505-522.

Fujiwara, T., O'Geen,et al., 2009. Discovering hematopoietic mechanisms through genome-wide analysis of GATA factor chromatin occupancy. *Mol Cell* **36**: 667-681.

Fujiwara, Y., C. P. Browne, et al. 1996. Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc. Natl. Acad. Sci. USA* **93**: 12355-12358.

Fuks, F. 2005. DNA methylation and histone modifications: teaming up to silence genes. *Curr Opin Genet Dev* **15**: 490-495.

de Laat, W. and Grosveld, F. 2003. Spatial organization of gene expression: the active chromatin hub. *Chromosome Res* **11**: 447-459.

Johnson, K., Christensen, H., et al. 2001. Distinct Mechanisms Control RNA Polymerase II Recruitment to a Tissue-Specific Locus Control Region and a Downstream Promoter. *Mol Cell* **8**: 465-471.

Jones, P. A. and G. Liang. 2009. Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet* **10**: 805-811.

Keller, A., Nesvizhskii, A.I., et al. 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* **74**: 5383-5392.

Kiefer, C.M., Hou, C., et al. 2008. Epigenetics of β-globin gene regulation. **647**: 68-76.

Kim, J. K., S. M., et al. 2009. Epigenetic mechanisms in mammals. *Cell Mol Life Sci* **66**: 596-612.

Kim, S-I, Bresnick, E., et al. 2009. Brg1 directly regulate nucleosome structure and chromatin looping of the alpha globin locus to activate transcription. *Nuc Acid Res* **37**: 6019-6027.

Kim, S-I, Bultman, S., et al. 2007. Dissecting molecular steps in chromatin domain activation during hematopoietic differentiation. *Mol Cell Biol* **27**: 4551-4565.

Kim, S-I and Bresnick, E.H., 2007. Transcriptioanl control of erythropoiesis: emerging mechanism and principles. *Oncogene* **26**: 6777-6794.

Kina, T., Ikuta, K., et al. 2000. The monoclonal antibody Ter-119 recognizes a molecule associated with gylcophorin A and specifically marks the late stages of murine erythroid lineage. *Brit J Haematology* **109**: 280-287.

Kouzarides, T. 2007. Chromatin modifications and their function. *Cell* **128**: 693-705.

Kotkow, K., and Orkin, S.H. 1995. Dependence of globin gene expression in mouse erythroleukemia cells on the NF-E2 heterodimer. *Mol Cell Biol* **15**: 4640-4647.

Koury, J.M., Sawyer, S.T., and Brandt, S.J. 2002. New insight into erythropoiesis. *Current Opinion in Hematology* **9**: 93-100.

Kruger, A., C. Ellerstrom, et al. 2002. RP59, a marker for osteoblast recruitment, is also detected in primitive mesenchymal cells, erythroid cells, and megakaryocytes. *Dev. Dyn* **223**: 414-418.

Gao, Z.G., Huang, Z., et al. 2009. FOG-1-mediated recruitment of NuRD is required for cell lineage re-enforcement during haematopoiesis. *EMBO J* **29**: 457-468.

Greig, K.T., Carotta, S., and Nutt, S.L. 2008. Critical roles of c-myb in hematopoietic progenitors cells. *Seminars in Immunology* **20**: 247-258.

Grosveld, F. 1999. Activation by locus control regions? *Curr Opin Genet Dev* **9**: 152-157.

Hirano, T. 2006. At the heart of the chromosome: SMC proteins in action. *Nat. Rev. Mol. Cell. Biol.* **7**: 311-322.

Hoffman, Furie B, Benz EJ, McGlave P, Silberstein LE, and SJ, S. 2005. *Hematology – Basic Principles and Practice*. Churchill Livingstone, New York.

Hoffman, Furie B, Benz EJ, McGlave P, Silberstein LE, and SJ, S. 2008. *Hematology – Basic Principles and Practice*. Churchill Livingstone, New York.

Hong, W., Nakzawa, et al. 2005. Fog-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. *EMBO J* **24**: 2367-2378.

Hopp, T.P., and Woods K.R. 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. USA* **78**: 3824-3828.

Ingley, K., Tilbrook K.A., and Klinken S.P. 2004. New insights into the regulation of eythroid cells. *IUBMB Life* **56**: 177-184.

Iwasaki, H., Mizuno, S., et al. 2003. GATA-1 converts lymphoid and myelomonocytic progenitors into megakaryoctye/erythrocyte lineages. *Immunity* **19**: 451-462.

Li, C.-Y., Fang, F., et al. 2008. Suppression of EDAG gene expression by phorbol 12-myristate 13-acetate is mediated through down-regulation of GATA-1. *Biochimica et Biophysica Acta* **1779**: 606-615.

Li, C.-Y., Zhan, Y.-Q., et al. 2007. Overexpression of a hematopoietic transcriptional regulator EDAG induces myelopoiesis and suppresses lymphopoiesis in transgenic mice. *Leukemia* **21**: 2277-2286.

Li, C. Y., Zhan, Y. Q., et al. 2004. EDAG regulates the proliferation and differentiation of hematopoietic cells and resists cell apoptosis through the activation of nuclear factor-kappa B. *Cell Death Differ* **11**: 1299-308.

Liang, C.-P., Lee, Y.-C., and Liu, Y.-C. 1992. Deletion studies to reveal size discrepancy in proliferation cell nuclear antigen. *Electrophoresis* **13**: 346-353.

Liang, S., Moghimi, B., et al. 2008. Locus control region mediated regulation of adult beta-globin gene expression. *J Cell Biol* **105**: 9-16.

Ling, B., Zhou, Y., et al. 2007. Down-regulation of EDAG expression by retrovirus-mediated small interfering RNA inhibits the growth and IL-8 production of leukemia cells. *Oncol Rep* **18**: 659-64.

Orkin, S.H. and Zon, L.I. 2008. Hematopoiesis: An evolving paradigm for stem cell biology. *Cell* **132**: 631-644.

Maley, F. and Guarino, D.U. 1977. Differential binding of sodium dodecyl sulfate to amino acids as evidenced by elution from Sephadex. *Biochem Biophys Res Commun.* **77**: 1425-1430.

Miccio, A., Wang, Y.H., et al. 2009. NuRD mediates activating and repressive functions of GATA-1 and FOG-1 during blood development. *EMBO J* **1**: 1-15.

Mikkola, H.K., Klintman, J., et al. 2003. Haematopoietic stem cells retain long-term repopulation activity and multipotency in the absence of stem-cell leukaemia SCL/tal-1 gene. *Nature* **421** : 547-551.

Mignotte, V., Navarro, S., et al. 1990. The extinction of erythroid genes after tetradecanoylphorbol acetate treatment of the erythroleukemic cells correlates with down-regulation of the tissue-specific factors NF-E1 and NF-E2. *J Biol Chem* **265**: 22090-2292.

Moi, P. and Kan, Y.W. 1990. Synergistic enhancement of globin gene expression by activator protein-1-like proteins. *Proc Natl Acad Sci USA* **87**: 9000-9004.

Motohashi, H., O'Connor, T., et al. 2002. Integration and diversity of the regulatory network composed of Maf and CNC families of transcription factors. *Gene* **294**: 1-12.

Nagai, T., Igarashi, K., et al. 1998. Regulation of NF-E2 activity in erythroleukemia cell differentiation. *J Biol Chem* **273**: 5358-5365.

Ney, P.A. 2006. Gene expression during terminal erythroid differentiation. *Curr Opin in Hematology* **13**: 203-208.

Ney, P.A., Sorrentino, B.P., et al. 1990. Tandem AP-1-binding sites within the human beta-globin dominant control region function as an inducible enhancer in erythroid cells. *Genes Dev* **4**: 993-1006.

Palis J. 2008. Ontogeny of Erythropoiesis. *Current Opinion in Hematology* **15**: 155-161.

Palis, J. (2009). Molecular Biology of Erythropoiesis. Molecular Basis of Hematopoiesis. B. K. Amittha Wickrema. New York, Springer: 73-93.

Palstra, R.J. Tolhuis, B., et. al. 2003. The beta-globin nuclear compartment in development and erythroid differentiation. *Nat Genet* **35**: 190-194.

Paul, R., Schuetze, S. et al. 1989. A common site for immortalizing proviral integrations in Friend erythroleukemia: molecular cloning and characterization. *J Virol* **63**: 4968-4961.

Perkins, D.N., Pappin, D.J., Creasy, D.M., and Cottrell, J.S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**: 3551-3567.

Perry, C., and Soreq, H. 2002. Transcriptional regulation of erythropoiesis. *Eur J biochem* **269**: 3607-3618.

Pevny, L., Simon, M.C., et al. 1991. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* **349**: 252-260.

Rao, G., Rekhtamn, N., et. al. 1997. Deregulated expression of the PU.1 transcription factor blocks murine erythroleukemia cell terminal differentiation. *Oncogene* **14**: 123-131.

Rodriguez, P., Bonte, E., et al. 2005. GATA-1 forms distinct activating and repressive complexes in erythroid cells. *Embo J* **24**: 2354-66.

Rollins, R.A., Korom, M., et al. 2004. Nipped-B protein supports sister chromatid cohesion and opposes the stromalin/Scc3 cohesion factor to facilitate long-range activation of the cut gene. *Mol Cell Biol* **24**: 3100-3111.

Ruthenburg, A.J., Li, H.T., et al. 2007. Multivalent enagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol* **8**: 983-994.

Saha, A., Wittmeyer, J., and Cairns B.R. 2006. Chromatin remodeling: the industrial revolution of DNA around histones. *Nat Rev* **7**:437-447.

Saxena, R.K. and Khandelwal. 2009. Aging and destruction of blood erythrocytes in mice. *Current Science* **97:** 500-507.

Shimizu, R., Engel, J.D., and Yamamoto, M., 2008. GATA1-related leukemias. *Nat Rev* **8**: 279-287.

Spada, F., A. Haemmer, et al. 2007. DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells. *J Cell Biol* **176**: 565-71.

Swiers, G., Patient, R., and Loose, M. 2006. Genetic regulatory networks programming hematopoietic stem cells and erythroid lineage specification. *Genomes & Dev Control* **294**: 525-540.

Tenen, D.G. 2003. Disruption of differentiation in human cancer: AML shows the way. *Nat Rev Cancer* **3**: 89-101.

Tolhouis, B., Palstra, R.J., et al. 2002. Looping and interaction between hypersensitive sites in the active β-globin locus. *Mol Cell* **10**: 1453-1465.

Tripic, T., Deng, W.L., et al. 2009. SCL and associated proteins distinguish active from repressive GATA transcription factor complexes. *Blood* **5**: 2192-2201.

Tsang, A.P., Visvader, J.E., et al. 1998. Failure of megatkaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. *Genes Dev* **12**: 1176-1188.

Tsiftsoglou, A.S., Pappas, I.S., and Viziriankis, I.S. (2003a). Mechanisms involved in the induced differentiation of leukemia cells. *Pharmacology and Therapeutics* **100**: 277-290.

Tsiftsoglou, A.S., Pappas, I.S., and Viziriankis, I.S. (2003b). The developmental Program of Murine Erythroleukemia Cells. *Oncology Research* **13**: 339-346.

Tsiftsoglou, A.S., Viziriankis, I.S., and Strouboulis, J. 2009. Erythropoiesis: Model Systems, Molecular Regulators, and Developmental Programs. *IUBMB Life* **61**: 800-830.

Tsiftsoglou, A.S., and Wong, W. 1985. Molecular and cellular mechanisms of leukemic hematopoietic cell differentiation: an analysis of the Friend system. *Anticaner Research* **5**: 81-100.

Tuan, D., Solomon, W., Li, Q., and London, I.M. 1985. The "beta-like globin" gene domain in the human erythroid cells. *Proc Natl Acad Sci USA* **82**: 6384-6388.

Schuetze, S., Paul, R., et al. 1992. Role of the PU.1 transcription factor in controlling differentiation of Friend erythroleukemia cell. *Mol Cell Biol* **12**: 2967-2975.

Stamatoyannopoulos, G. 2005. Control of globin gene expression during development and erythroid differentiation. *Experimental Hematology* **33**: 259-271.

Valverde-Garduno, V., Guyot, B., et al. 2004. Differences in the chromatin structure and cis-element organization of the human and mouse GATA-1 loci: implications for cis-element indentification. *Blood* **104**: 3106-3116.

Wadman, I.A., Osada, H., et al. 1997. The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb/NLI proteins. *EMBO J* **16**: 3145-5234.

Weaver, R. 2005. *Molecular Biology*. McGraw Hill, New York.

Welch, J., Watts, J.A., et al. 2004. Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood* **104**: 3136-3146.

Wolffe, A.P., Urnov, F.D., and Guschin, D. 2000. Co-repressor complexes and remodeling chromatin for repression. *Biochem Society Transactions* **28** part 4: 379-386.

Wu, J.I., Lessard, J., and Crabtree, J. 2009. Understanding the words of chromatin regulation. *Cell* **36**: 200-206.

Vakoc, C.R., Sachdeva, M.M., et al. 2006. A profile of histone lysine methylation across transcribed mammalian chromatin. *Mol Cell Biol* **26**: 9185-9195.

Yang, L.V., Nicholson, R.H., et. al. 2001. Hemogen is a noval nuclear factor specifically expressed in mouse hematopoietic development and its human homologue EDAG maps to chromosome 9p22, a region containing breakpoints of hematological neoplasms. *Mech Dev* **104**: 105-111.

Yang, L., J. Wan, et al. 2006. The GATA site-dependent hemogen promoter is transcriptionally regulated by GATA1 in hematopoietic and leukemia cells. *Leukemia* **20**: 417-425.