LOCUS CONTROL REGION ACTIVITY BY 5’HS3 REQUIRES A FUNCTIONAL INTERACTION WITH β-GLOBIN GENE REGULATORY ELEMENTS:
IDENTIFICATION OF EFFECTIVE βγ-GLOBIN MINIGENES FOR GENE THERAPY OF THE β-CHAIN HEMOGLOBINOPATHIES

by

Joel Edward Rubin

A thesis submitted in conformity with the requirements for the degree of Master of Science
Graduate Department of Molecular and Medical Genetics
University of Toronto

© Copyright by Joel Edward Rubin 2000
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
This Thesis is dedicated to my

Mother, Father, Brother, and Grandmother
Locus control region activity by 5’HS3 requires a functional interaction with β-globin gene regulatory elements:
Identification of effective β/γ-globin minigenes for gene therapy of the β-chain hemoglobinopathies

Joel Edward Rubin, 2000
Degree of Master of Science
Graduate Department of Molecular and Medical Genetics
University of Toronto

ABSTRACT

Our lab previously used single-copy transgenic mice to map the chromatin opening activity of the human β-globin locus control region to the 5’HS3 element. The study presented here illustrates that an 850 bp 5’HS3 fragment reproducibly activates β-globin transgenes in the fetal livers of transgenic mice but fails to activate Aγ-globin transgenes. Analysis of several novel β/γ-globin hybrid transgenes, demonstrates that 5’HS3 must functionally interact with β-globin intron 2 and β-globin 3’ enhancer gene sequences to reproducibly activate single-copy transgenes. In addition, this analysis demonstrates that 5’HS3, in the presence of these β-globin 3’ elements, can activate both the β- or γ-globin promoters but activation is dependent on an AT-rich region within β-globin intron 2. This study has identified ideal minigenes for adult expression of anti-sickling γ-globin coding sequences in gene therapy of the β-chain hemoglobinopathies. Retrovirus-mediated transfer of these minigenes into primitive hematopoietic stem cells is discussed.
ACKNOWLEDGEMENTS

First, I would like to thank my family and friends. A special thanks goes to Felipe Cisternas, Tino Piscione, and Cam Osborne for keeping me busy with hockey pools, coffee breaks, and nights at O’Grady’s. I would also like to thank all of the Ellis lab members for technical assistance and helpful discussions. I would especially like to thank Peter Pasceri for his extremely hard work and dedication to the lab and Franca Posteraro for her secretarial assistance. I would also like to take this opportunity to thank my student committee members, Sabine Cordes and C.C. Hui, for their guidance and valuable suggestions. Finally, my utmost respect and gratitude goes to my supervisor and friend, James Ellis, who has given me more than enough support for the past few years and has taught me a great deal.
# TABLE OF CONTENTS

**ABSTRACT** .................................................................................................................................................. III

**ACKNOWLEDGEMENTS** .................................................................................................................................. IV

**TABLE OF CONTENTS** ................................................................................................................................... V

**LIST OF TABLES** ............................................................................................................................................. VII

**LIST OF FIGURES** ........................................................................................................................................... VIII

**I. INTRODUCTION** ........................................................................................................................................ 1

1. The hemoglobinopathies ................................................................................................................................. 1

2. Targeting the hematopoietic stem cell in blood gene therapy ................................................................. 2

3. Major problems associated with gene therapy ............................................................................................ 5
   3.1. Transgene delivery .......................................................................................................................................... 5
   3.2. Transgene expression ..................................................................................................................................... 14

4. β-globin LCR ..................................................................................................................................................... 17

5. Designing βLCR/β-globin minigenes for use in gene therapy ....................................................................... 22
   5.1. Mapping the minimal βLCR elements required for reproducible transgene expression ..................... 27
   5.2. Candidate β-globin gene regulatory elements that may confer LCR responsiveness ............................. 28
   5.3. Defining the minimal β-globin regulatory elements required for LCR activity by 5'HS3 .................... 30

**II. MATERIALS AND METHODS** ................................................................................................................ 32

1. Transgene plasmid construction ..................................................................................................................... 32

2. Generation of transgenic mice ...................................................................................................................... 34

3. DNA analysis .................................................................................................................................................. 35

4. RNA analysis .................................................................................................................................................. 36

5. Recombinant retrovirus plasmid construction ............................................................................................. 37

6. Generation of recombinant retrovirus ........................................................................................................... 37

7. Determination of retrovirus titer ................................................................................................................... 38
LIST OF TABLES

Table 1: Comparison of potential vectors for transgene delivery into hematopoietic stem cells......12
Table 2: Human βLCR fragments assayed for activity in transgenic mice ..................................23
Table 3: Description of each β/γ-globin hybrid transgene and a summary of single-copy and total
mice expressing these constructs ...............................................................................................60
LIST OF FIGURES

Figure 1: Retroviral vector life-cycle ........................................................................................................8
Figure 2: The human β-globin locus and Holocomplex model of LCR activation .................................18
Figure 3: Globin constructs used to identify important β-globin gene regulatory elements required
for LCR activity by 5'HS3 ..........................................................................................................................40
Figure 4: Analysis of transgenes and determination of percent transgenicity of mice by Southern
blot analysis ..............................................................................................................................................42
Figure 5: Expression of BGT9 ...............................................................................................................45
Figure 6: Expression of BGT26 and BGT34 ......................................................................................46
Figure 7: Expression of BGT35 .............................................................................................................49
Figure 8: Expression of BGT54 ............................................................................................................50
Figure 9: Expression of BGT50 ............................................................................................................51
Figure 10: Expression of BGT76 ..........................................................................................................53
Figure 11: Expression of BGT64 ..........................................................................................................54
1. INTRODUCTION

1. The hemoglobinopathies

The hemoglobinopathies are a group of inherited diseases resulting from a deficiency in wild type hemoglobin or the production of abnormal hemoglobin. Hemoglobin (Hb) is necessary for oxygen transport by the red blood cells in the body. The Hb protein is a tetramer made up of two α-globin chains and two β-globin chains. A deficiency in producing either chain results in thalassemia referred to as α or β thalassemia depending on the gene affected. The mutations that cause these thalassemias are highly variable and the disease phenotype ranges from intermediate to severe in homozygotes. The mutations can reside outside of the structural gene and result in regulatory control defects or within the gene and produce non-functional protein products. In the most serious cases, where functional protein levels are lowest, the patient has a severe anemia resulting from the insufficient hemoglobin production and is afflicted with all of the complications associated with a deficiency in oxygen transport (Stamatoyannopoulos et al., 1994).

Structural defects in the β-globin gene also exist. The most common of these is the sickle mutation, resulting in sickle cell anemia. The single nucleotide mutation changes the sixth amino acid residue (Glu-6→Val) in the β-globin chain producing β' protein. Hemoglobin molecules polymerize when one or two β' proteins are incorporated into the Hb tetramer (α2 β'2 or α2 ββ'). Only one valine substitution per tetramer is required for intertetramer contact. However, the most severe phenotype is seen in homozygotes that produce only β'. The polymerization of HbS into
numerous fibers causes the red blood cells to look ‘sickled’ in shape when in the tense or deoxygenated state. The rapid destruction of sickled cells results in anemia (Stamatoyannopoulos et al., 1994). Furthermore, the aggregation of the affected cells interferes with blood circulation causing tissue damage. These cells tend to accumulate in and, ultimately, enlarge the spleen. The hemoglobinopathies can affect up to 1 in 600 individuals in some populations and over 100,000 affected children are born each year (Stamatoyannopoulos et al., 1994).

Current treatment of the hemoglobinopathies is by blood transfusion. However, some drugs are available such as hydroxyurea and butyrate derivatives that can partially derepress γ-globin expression (Charache et al., 1995). γ-globin can completely compensate for β-globin protein function and is normally expressed in the fetal stage of development but silent in the adult. The only available possibility of cure is by allogeneic bone marrow transplant (Lucarelli et al., 1993; Walters et al., 1996). However, most patients do not have an HLA-matched donor and those that do face a significant risk of fatal complications by the transplantation procedure. Efforts to design an alternate cure for the hemoglobinopathies via gene therapy have been underway for the last decade (Anderson, 1998; Emery and Stamatoyannopoulos, 1999; Pawliuk et al., 1998; Rivella and Sadelain, 1998; Verma and Somia, 1997).

2. Targeting the hematopoietic stem cell in blood gene therapy

Gene therapy is an attractive means of treating genetic disease. The idea is that a genetic defect can be repaired by adding therapeutic genetic material into somatic cells thereby curing or ameliorating the disease phenotype in humans. Since gene correction by homologous
recombination in human cells is not feasible as of yet, gene therapy studies for monogenic diseases have concentrated on adding a therapeutic gene to the genome. However, gene therapy is not only applicable to inherited disorders but can be applied to acquired diseases such as cancer, disease resulting from viral infection, and neurodegenerative disease. In some of these acquired diseases, the therapeutic gene can be used to kill the cells affected (‘suicide gene’) instead of replacing a defect.

One of the most attractive targets for gene therapy is a stem cell because of the ability of this cell to repopulate the tissue. In particular, transferring genetic material to primitive hematopoietic stem cells (HSC), which repopulate all blood cells, could potentially protect a patient for life. The targeting of more differentiated blood cells to express a therapeutic gene may ameliorate the symptoms of the disease but treatment would only be temporary and may not represent a cure. Tissue-specific regulatory elements would ensure that the therapeutic gene is only expressed in the correct cells (ie. T-cells, B-cells, erythroblasts). Aside from the hemoglobinopathies, other monogenic blood diseases that are potentially curable by gene therapy include severe combined immunodeficiency disease (SCID) caused by a defect in the adenosine deaminase (ADA) gene (Weinberg and Kohn, 1998), hemophilia A and B caused by a deficiency in Factor VIII and Factor IX (Kay and High, 1999), respectively, and erythropoietic protoporphyria caused by a defect in the ferrochelatase gene (Pawliuk et al., 1999). Therefore, gene therapists have been targeting the blood for years and much research is being conducted to purify populations of HSCs based on cell surface markers (Dick, 1999; Krause et al., 1996) and efficiently transfer DNA to these cells (Anderson, 1998; Verma and Somia, 1997).
Initially, *in vivo* blood gene therapy studies were limited to transferring DNA to murine HSCs and repopulating the blood system of a mouse with genetically modified cells (Dzierzak *et al.*, 1988). *In vitro* surrogate human progenitor assays have been developed to identify colony-forming cells (CFCs) (Dick *et al.*, 1991; Hock and Miller, 1986) and long-term culture-initiating cells (LTC-ICs) (Eaves *et al.*, 1992; Hughes *et al.*, 1989; Moritz *et al.*, 1994). However, these cells have failed to represent the repopulating cell fractions of the human hematopoietic system and transfer of DNA to human HSCs, the ultimate target cell in blood gene therapy, has been difficult to assess because of a lack of an *in vivo* surrogate progenitor assay. Fortunately, assay systems have now been developed for this purpose.

One such assay, based on the engraftment of non-obese diabetic (NOD)/SCID mice with human hematopoietic cells, has identified a novel population of HSCs defined as SCID-repopulating cells (SRCs) (Larochelle *et al.*, 1996). These cells are highly enriched in isolated CD34⁺CD38⁻ blood cell fractions and, more recently, have also been found to be in CD34⁺CD38⁻ blood cell fractions (Dick, 1999). These SRCs have been shown to be capable of extensive proliferation and multilineage (lymphoid and myeloid) differentiation *in vivo* (Vormoor *et al.*, 1994) and represent a more primitive cell than CFCs and LTC-ICs (Bhatia *et al.*, 1997; Cashman *et al.*, 1997; Hogan *et al.*, 1997; Larochelle *et al.*, 1996). The NOD/SCID mouse model provides gene therapists with an assay system for the development and optimization of human HSC gene transfer protocols and transgene expression analysis. In fact, SRCs from patients with sickle cell disease have been used to reconstitute the blood system of these mice giving rise to a mouse model for sickle cell anemia (Larochelle *et al.*, 1995). However, other mouse models exist where the animals
are engineered through multiple inter-transgenic line crosses to produce only human \(\alpha\) chains and \(\beta^\prime\) chains (Paszty et al., 1997; Ryan et al., 1997).

3. Major problems associated with gene therapy

In principle, the idea of gene therapy is sound. In practice, many obstacles have emerged and must be overcome before safe and effective gene therapy protocols can be routinely developed. Problems of significant importance, not only for gene therapy of blood diseases but gene therapy in general, involve lack of efficient delivery systems and lack of sustained expression of the therapeutic gene (Anderson, 1998; Emery and Stamatoyannopoulos, 1999; Miller and Whelan, 1997; Verma and Somia, 1997; Weinberg and Kohn, 1998; Pawliuk et al., 1998; Rivella and Sadelain, 1998).

3.1. Transgene delivery

For effective gene therapy, transgenes must be efficiently transferred to somatic tissues. Transfer can be accomplished in a number of ways: 1. \textit{ex vivo}, where the target cells are isolated, modified, and transplanted back into the patient; 2. \textit{in situ}, where the transgene is delivered locally to the tissue of interest; 3. \textit{in vivo}, where the transgene is delivered systemically by means of the blood. Unfortunately, there is not yet an ideal vehicle for gene delivery to human cells. However, all of the properties necessary for an ideal vehicle can be found in individual systems (Verma and Somia, 1997). There are two broad categories of gene delivery vectors: non-viral and viral. Examples of non-viral vectors include direct injection of DNA into cells or mixing DNA with polylysine or cationic lipids to facilitate passage through the cell membrane. Although there are no
packaging size constraints or immunological concerns associated with this type of delivery, these methods suffer from poor efficiency and result in transient expression (Felgner, 1997). Furthermore, the range of disease effectively treatable by gene therapy is limited using non-viral vectors.

Most of the current focus has been put on the second category, viral vectors. Viruses are attractive delivery vectors because most viruses have evolved a specific machinery to deliver their genome to cells. Furthermore, some viruses, such as Adeno-Associated Viruses (AAV) and retroviruses, can stably integrate into the host genome making these viruses prime candidates for developing ex vivo gene therapy vectors to cure monogenic disorders such as the hemoglobinopathies. In cases such as slowing the progression of tumors, in situ infection and overproduction of a ‘suicide’ gene using an Adenoviral vector (which does not integrate) may be suitable. The nature of the disease will determine which viral vector and method of infection is ideal.

One of the most efficient delivery techniques for gene transfer into HSCs is retrovirus-mediated delivery (Hughes et al., 1992; Szilvassy et al., 1989). Most studies involving transduction of HSCs have employed derivatives of Moloney Murine Leukemia Virus (MMLV), a prototypic murine retrovirus. These derivatives are replication incompetent as the coding region for the essential viral proteins, gag, pol, and env, is deleted from the viral genome. Therefore, virion generation from these vectors requires packaging in a cell line (termed packaging cell lines) that is genetically engineered to produce viral proteins in trans (Fig. 1). The virus produced can be pseudotyped with a particular envelope protein by expressing this protein in the packaging cell. The envelope protein determines the host range of the virus. An ecotropic host range refers to a viral
envelop protein that recognizes a cell surface receptor protein unique to one species such as rodents or primates. With an amphotropic host range, the virus can infect across species as it recognizes a receptor protein common to both rodents and primate cells. Most prototypic retroviruses are pseudotyped with an amphotropic envelope protein able to recognize a wide variety of host cells. With the removal of viral coding sequence, inserts up to 8 kb can be cloned into these vectors without affecting viral packaging and infectivity (Miller and Rosman, 1989).

Murine Stem Cell Virus (MSCV) vector, and derivatives thereof, have been specifically designed for the infection of HSCs (Hawley et al., 1994; Osborne et al., 1999; Robbins et al., 1998). MMLV-based vectors packaged in an amphotropic cell line can efficiently transduce murine HSCs ex vivo (Ding et al., 1996; Hawley et al., 1992; Pawliuk et al., 1999; Pawliuk et al., 1997). However, these amphotropic viruses are not as effective at transducing human HSCs (Bordignon et al., 1995; Dunbar et al., 1995; Hoogerbrugge et al., 1996; Kohn et al., 1998; Kohn et al., 1995). This is partially due to a lower number of the amphotropic receptors recognized by the virus envelope protein on human HSCs (Crooks and Kohn, 1993; Orlic et al., 1996). MMLV-derived vectors are also limited to the infection of dividing cells as they require the breakdown of the nuclear membrane that only occurs during division (Miller et al., 1990). Most HSCs are in a quiescent state (Jones et al., 1990). Therefore, for efficient infection, the HSCs that are not actively dividing must be stimulated to cycle in vitro (Miller et al., 1990). A further complication is that the HSCs must retain their pluripotency, which can be lost when the cells are stimulated to divide ex vivo (Bodine et al., 1991; Gan et al., 1997; Gothot et al., 1998; Larochelle et al., 1996; Peters et al., 1996; Tisdale et al., 1998). Culture conditions and infection protocols are currently being developed.
Retroviral vector life-cycle. Retroviral vectors are designed to be replication incompetent by deleting the coding regions for essential viral proteins such as gag, pol, env. Therefore, virion assembly takes place in a packaging cell, which transiently or stably expresses these viral proteins. The vector DNA is transfected into the cell (a.) and vector RNA is transcribed (b.). Vector RNA is selected for virion packaging by the recognition of the packaging signal (*Ψ*). Virion assembly takes place in the cytoplasm (c.), which contains the viral proteins, and the virion buds from the cell (d.). The virions can then be harvested and used to infect a target cell population. The surface (SU) moiety of the envelope protein recognizes its receptor on the target cell and allows the transmembrane (TM) moiety to initiate membrane fusion (e.). Once in the cell, the virion core breaks down and the vector RNA is reverse transcribed by the enzyme reverse transcriptase, which is packaged within the virion (f.). The double-stranded DNA genome then integrates randomly within the host genome (g.). The coding sequence of the transgene (TGN) and selectable marker gene (SM) is expressed by the cell machinery in the target cell (h.). As a provirus, the vector genome is stable and no further replication occurs. LTR, Long Terminal Repeat.

Figure 1
to increase the efficiency of transducing human HSCs and maintaining these cells in an undifferentiated state (Cheng et al., 1998; Conneally et al., 1998; Dao et al., 1998; Dunbar et al., 1996; Kiem et al., 1998; Kiem et al., 1997; Kiem et al., 1999; Schilz et al., 1998; Tisdale et al., 1998; van Hennik et al., 1998).

More recently, researchers have begun to investigate the use of lentiviral vectors, which are retroviruses that can infect both dividing and non-dividing cells (Lewis et al., 1992). This aspect of lentivirus biology has provided an attractive means of drastically increasing the efficiency of human HSC transduction. One lentivirus currently under investigation for vector design is Human Immunodeficiency Virus-1 (HIV-1) (Naldini et al., 1996a; Poeschla et al., 1996; Reiser et al., 1996). When pseudotyped with the vesicular stomatitis virus G protein (VSV-G), these vectors have been shown to infect non-dividing cells and stably integrate into the host cell genome. Long-term expression of transgenes in brain, liver, muscle, and retina have been obtained using lentiviral vectors (Blomer et al., 1997; Kafri et al., 1997; Miyoshi et al., 1997; Naldini et al., 1996a; Naldini et al., 1996b). Recently, HIV-derived vectors have also proven to stably and efficiently transduce quiescent CD34+CD38- human HSCs ex vivo (Case et al., 1999; Miyoshi et al., 1999). However, since HIV-1 is the etiological agent of Acquired Immunodeficiency Syndrome (AIDS) in humans, the use of this vector raises safety concerns. Although, like all retroviral vectors, the HIV-derived vectors are engineered to be replication incompetent, the concern is that wild type HIV-1 or replication-competent recombinants (RCRs) may be generated by DNA recombination in the packaging cell. However, pseudotyping these viral vectors with non-HIV envelop proteins or the use of non-human lentiviruses, such as Feline Immunodeficiency Virus (FIV) (Poeschla et al., 1998),
can reduce the risk of disease from RCRs. Such a risk is further minimized with the use of multiple plasmids coding for the necessary viral genes in the packaging of these vectors (Kafri et al., 1999). Human foamy virus (HFV) is non-pathogenic to humans and this retrovirus is able to transduce a wide variety of non-dividing host cells. Therefore, HFV-based vectors are also being considered. (Russell and Miller, 1996).

Another safety concern for the use of lentiviral vectors, as well as prototypic retroviruses, is the possibility of insertional activation of cellular oncogenes or inactivation of tumor suppressor genes by random integration of the vector provirus into the host genome. To partially overcome this problem, self-inactivating (SIN) MMLV- and HIV-based vectors have been described (Armentano et al., 1987; Miyoshi et al., 1998; Osborne et al., 1999; Zufferey et al., 1998). SIN vectors eliminate the activity of the viral promoter and enhancer in the 5' Long Terminal Repeat (LTR) of the provirus through the normal infection cycle of the virus. The mutation is made within the 3' LTR of the virus construct which is copied to the 5' LTR in the provirus as the 3' LTR serves as a template for the generation of the 5' LTR. Therefore, the viral promoter is active in the packaging cell but, following reverse transcription and integration into the target cell genome, the promoter is no longer active. In addition, the loss of promoter activity in the target cell further reduces the chance of generating RCRs and eliminates any cis-acting effects that the LTR may have on transgene expression such as transcriptional activation or interference. These modifications to lentiviral vectors have greatly increased their attractiveness for use in gene therapy.

The commonly used amphotropic retroviruses vectors recognize the Ram-1 protein, a ubiquitous phosphate symport, which does not provide much specificity for the target cell.
(Anderson, 1998). Such viruses are not effective for *in vivo* infection as the virus will infect most dividing cells and the titer needed to infect a large proportion of the target cells is unattainable. The non-specific infection by these viral vectors further limits their use to *ex vivo* protocols. However, these retroviruses and lentiviral vectors can be pseudotyped with other viral envelope proteins such as VSV-G protein or the gibbon ape leukemia virus (GALV) protein, using specific packaging cells lines or packaging plasmids which express these proteins. The host range can be varied with the use of these pseudotyped vectors but still remains broad and specificity is low (Anderson, 1998).

Ideally, the virus should recognize a receptor that is specific to the target cell which would dramatically increase the efficiency of infection and efforts have been devoted to modifying the virus' natural envelope protein to recognize specific targets (Kasahara *et al.*, 1994; Salmons and Gunzburg, 1993). The retroviral envelope protein complex is a homotrimer where the monomer consists of a surface (SU) moiety and a transmembrane (TM) moiety (Fass *et al.*, 1996). The SU domain is responsible for receptor recognition while the TM domain is involved in fusion of the cellular and viral membranes which is a prerequisite for virus core entry. Most of the modifications to date have altered the SU domain of the envelope protein using ligands or single-chain antibodies which have been successful for specific binding of the virus to the target cell (Kasahara *et al.*, 1994; Salmons and Gunzburg, 1993). However, these modifications have interfered with the TM domain and completely abolished the entry (fusion) function of the viral envelope protein. Clearly more structure-function information for these envelope proteins is needed before such engineering is feasible. Other experimental approaches to increase the efficiency of virus infection have focused on tethering the virus to the target cells (Cosset and Russell, 1996; Hall *et al.*, 1997; Hanenberg *et
al., 1997; Hanenberg et al., 1996). Specific targeting of the viral vector could potentially increase the efficiency of transduction of human HSCs \textit{ex vivo} or allow for the engineering of lentiviruses or AAV vectors that can specifically target these cells \textit{in vivo}.

Table 1: Comparison of potential vectors for transgene delivery into hematopoietic stem cells

<table>
<thead>
<tr>
<th>Method</th>
<th>Integration</th>
<th>Packaging capacity</th>
<th>Expression</th>
<th>Delivery</th>
<th>Safety Concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA</td>
<td>Poor</td>
<td>unlimited</td>
<td>Transient</td>
<td>\textit{ex vivo/in situ}</td>
<td>None</td>
</tr>
<tr>
<td>Liposome</td>
<td>Poor</td>
<td>unlimited</td>
<td>Transient</td>
<td>\textit{ex vivo/in situ}</td>
<td>None</td>
</tr>
<tr>
<td>Adenovirus</td>
<td>No</td>
<td>30 kb</td>
<td>Transient</td>
<td>\textit{ex vivo/in vivo/in situ}</td>
<td>Immune Response</td>
</tr>
<tr>
<td>AAV</td>
<td>Yes/No</td>
<td>4.0 kb</td>
<td>Stable/Transient</td>
<td>\textit{ex vivo/in vivo/in situ}</td>
<td>Immune Response</td>
</tr>
<tr>
<td>Retrovirus</td>
<td>Yes</td>
<td>8.0 kb</td>
<td>Stable</td>
<td>\textit{ex vivo/in situ}</td>
<td>Insertional Mutagenesis</td>
</tr>
<tr>
<td>Lentivirus</td>
<td>Yes</td>
<td>8.0 kb</td>
<td>Stable</td>
<td>\textit{ex vivo/in vivo/in situ}</td>
<td>Insertional Mutagenesis. RCRs</td>
</tr>
</tbody>
</table>

AAV, Adeno-Associated Virus; RCRs, Replication Competent Retroviruses

Wild type AAV commonly integrates at a defined location on chromosome 19 (19q13-qter) in a human host cell (Kotin et al., 1991; Kotin et al., 1990; Samulski et al., 1991) and can infect non-dividing cells (Bartlett et al., 1998; Russell et al., 1994). These viruses represent potential vectors for gene therapy and are being investigated for their ability to transduce various cell types \textit{ex vivo} and \textit{in vivo} (Clark et al., 1997; Fisher et al., 1997; Flannery et al., 1997; Halbert et al., 1997; Snyder et al., 1997). However, these vectors suffer from small packaging capacity (=4.0 kb), an inability to generate efficient packaging systems, difficulty in preparing high-quality stocks, and safety concerns involving immune responses (Verma and Somia, 1997). Furthermore, as viral
sequence is removed to increase the packaging capacity of the vector, the less efficient the virus is at integrating within the host genome (Kearns et al., 1996; Ponnazhagan et al., 1997; Russell et al., 1994; Rutledge and Russell, 1997; Walsh et al., 1992; Yang et al., 1997), which is a key aspect of an effective vector for gene therapy and required for long term expression of the therapeutic gene.

It is established that AAV vectors very efficiently transduce transformed blood cell lines (Einerhand et al., 1995; Lebkowski et al., 1988; Miller et al., 1993a; Walsh et al., 1992; Zhou et al., 1996) and these vectors were initially thought to provide an alternative to using prototypic retroviruses for infection of human HSCs ex vivo. However, the ability of this vector to transduce primary hematopoietic cells is controversial. The bulk of evidence over the last few years indicate that these vectors poorly transduce HSCs ex vivo and can lead to transient transgene expression indicating that AAV vectors will not replace retroviral vectors for gene transfer into human HSCs (Hargrove et al., 1997; Malik et al., 1997; Russell and Kay, 1999). These vectors are perhaps better used for in vivo applications (Clark et al., 1997; Herzog et al., 1997; Mandel et al., 1997; Snyder et al., 1997). In fact, AAV vectors have been used to successfully treat Hemophilia B in dogs by in situ intramuscular delivery (Herzog et al., 1997) surpassing the need to infect HSCs. This is a potential treatment for hemophiliacs, however, and does not represent a cure. Viruses that do not integrate but are maintained as stable episomes, such as Herpes Simplex Virus (HSV) and Semliki Forest Virus (SFV) are also being assessed for their ability to infect HSCs and feasibility for gene therapy protocols (Berglund et al., 1993; Fink et al., 1996). To date, however, retroviruses either prototypic or lentiviral are the leading potential delivery vector for transgenes into HSCs. A comparison of possible vectors for transfer of DNA into HSCs are listed in Table 1.
The use of retroviral vectors to transfer the β-globin or γ-globin gene to HSCs has been further hampered by the generation of low titer recombinant retrovirus with multiple rearrangements detected in the transmitted proviral structure (Chang et al., 1992; Gelinas et al., 1992; Mulligan, 1993; Novak et al., 1990; Plavec et al., 1993). Initial retrovirus transfer studies for β-globin or γ-globin have shown that the intron structures of these genes are required for maximal expression and all transgenes designed should contain these sequences. Therefore, these transgenes are inserted into retroviruses in the anti-sense (AS) orientation compared to proviral transcription to preserve introns and avoid the recognition of the β-globin polyadenylation site in the packaging procedure. However, the AS sequence of the insert can be deleterious to the retroviral life-cycle. Deleterious sequences can arise from the occurrence of RNA processing recognition signals such as splice sites or polyadenylation sites within the AS insert by chance (McIvor, 1990) and the AS sequence could interfere with essential viral processes such as reverse transcription thought to be triggered by repetitive sequences (Pulsinelli and Temin, 1991). DNA viruses such as AAV have also been shown to be recombinogenic upon transfer of the β-globin gene to blood cells (Inoue et al., 1999). Low titers and instability of recombinant retroviral vectors have contributed to low HSC transduction efficiency and have limited gene therapy studies.

3.2. Transgene expression

Aside from gene delivery, another major obstacle for gene therapy protocols is in obtaining high-level, sustained expression of the therapeutic gene. Interestingly, retroviral vectors such as MMLV are completely silenced in primitive cells such as HSCs and viral sequences in the provirus are thought to silence transduced genes over time through an unknown mechanism (McCune and
Townes, 1994; Osborne et al., 1999). Since retroviruses are an attractive way to transfer genes to HSCs, this is a major concern. Our lab has shown that the MSCV (Osborne et al., 1999) and, more recently, the HIV (unpublished) LTR can silence potent globin transgenes in transgenic mice. We have since developed a derivative of MMLV that ameliorates retrovirus silencing called HSC-1 (Osborne et al., 1999). We are currently investigating the mechanism of this silencing in hopes of generating retroviral vectors that are completely devoid of silencing. In addition, we are investigating the use of insulator elements in retroviral vectors to protect the transgene from the silencing event. These insulator elements are characterized by their ability to block enhancer function when placed between an enhancer and promoter as well as protect a transgene from position effects that silence transgenes when they integrate into condensed chromatin environments (Bell and Felsenfeld, 1999). Silencing of transgenes by the AAV Inverted Terminal Repeat (ITR) has also been proposed (Chen et al., 1997).

Another way to ensure that the transgene is effectively expressed is through the correct use of transcriptional regulatory elements. Much care should go into designing therapeutic minigenes and defining the minimal elements necessary for reproducible, high-level, tissue-specific expression of a therapeutic gene. Since there is a risk of foreign promoters, such as viral promoters, to be recognized as foreign and silenced in the target cell, minigenes are best designed using the natural cis-regulatory elements of the therapeutic gene of interest (Anderson, 1998). Furthermore, in some cases such as β-globin production, overexpression of the therapeutic gene by a heterologous promoter can be toxic and strict regulation may be necessary. Most attempts to design minigenes for use in gene therapy of the hemoglobinopathies have concentrated on transgenes expressing the
actual gene defective in the disease such as β or α-globin or a compensatory gene such as γ-globin. However, alternative therapeutic genetic material can be expressed in these minigenes. One example is a transgene that expresses erythropoietin, a glycoprotein hormone, that can stimulate erythropoiesis in certain patients with hemoglobinopathies (Naffakh and Danos, 1996). Another example is a transgene that expresses a ribozyme able to repair the gene defect, such as the sickle mutation, at the level of mRNA (Weatherall, 1998).

One of the most difficult aspects of an effective minigene to achieve is reproducible expression at single-copy. Retroviruses and AAV integrate at single-copy, therefore, their transduced genes should be regulated by tissue-specific elements that function at single-copy. Locus Control Regions (LCR) are well suited for this task as they direct tissue-specific, reproducible expression from all integration sites and transgene copy numbers. For example, the human β-globin LCR (βLCR) directs high-level β-globin transgene expression in erythroid cells of transgenic mice regardless of the integration site (Grosveld et al., 1987). LCRs or LCR-like elements have also been identified in the human α-globin locus (Higgs et al., 1990) and in other tissue-specific loci such as the chicken lysozyme locus (Bonifer et al., 1990) and immunoglobulin mu locus (Forrester et al., 1994).

Before the discovery of the βLCR, several research groups attempted to transfer the β-globin gene alone with its cis-regulatory elements by a retrovirus into murine blood cells (Bender et al., 1989; Cone et al., 1987; Dzierzak et al., 1988; Karlsson et al., 1988; Karlsson et al., 1987). In a pioneering experiment, Cone et al. demonstrated the feasibility of using retroviral vectors to transfer and express an intact β-globin gene in MEL cells (Cone et al., 1987). This same group
demonstrated that these recombinant retroviral vectors could also transduce murine HSC and result in long-term expression of the β-globin gene in mouse red blood cells in vivo (Dzierzak et al., 1988). However, the use of this insert was shown to be insufficient for gene therapy applications since expression levels were too low to be therapeutic ranging from 0.1-5% compared to endogenous mouse βmaj levels on a per copy basis and expression was subject to position effects (Bender et al., 1989; Cone et al., 1987; Dzierzak et al., 1988; Karlsson et al., 1988; Karlsson et al., 1987). The discovery of the βLCR convinced researchers that β-globin levels could be boosted to full levels and protected from position effects with the addition of βLCR elements to β-globin minigenes.

4. β-globin LCR

LCR activity is operationally defined as the ability to direct high-level, position-independent transgene expression in mice. In the endogenous human β-globin locus, the βLCR is located 6 to 22 kb upstream of the five developmentally-regulated structural genes (5' e-Gγ-Aγ-δ-β 3') and comprises five DNaseI hypersensitive sites (5'H5-5) (Fig. 2), four of which (5'H5-4) are developmentally stable and erythroid-specific (Forrester et al., 1986; Grosveld et al., 1987; Tuan et al., 1985). The effects of natural occurring mutations within the βLCR (Driscoll et al., 1989) and its use to express β-globin reporter genes in mice strongly suggest that it is absolutely required for the initiation and maintenance of high-level expression of the β-globin gene locus in erythroid cells throughout development. As such, the βLCR is thought to be a 'master control' element with two functions: chromatin opening activity that directs transgene expression at all integration sites and
The human β-globin locus and Holocomplex model for βLCR activation. A. The β-globin LCR is composed of at least four tissue-specific, developmentally-stable DNaseI hypersensitive sites (5'HS1-4; indicated by arrows) located 6-18 kb upstream of the ε-globin gene. 5'HS5 (also indicated by an arrow) is considered part of the βLCR but is not restricted to erythroid tissues and its function is unknown. The region deleted in the Hispanic Deletion compared to the deletion analyzed by Reik et al., 1998 is illustrated. There are five β-like globin genes in the locus (ε, Gγ, Aγ, δ, and β) and these genes are differentially expressed during development. B. In the Holocomplex model of βLCR activation, the individual HSs are thought to fold into an active complex, termed Holocomplex, that can activate individual, competing globin promoters via DNA looping.
transcriptional enhancement activity that boosts levels of transgene expression (Bulger and Groudine, 1999; Grosveld, 1999; Li et al., 1999b). The βLCR is also thought to be a major determinant for replication timing and usage of the DNA replication origin within the β-globin gene (Aladjem et al., 1995; Aladjem et al., 1998).

Each HS of the human βLCR has been mapped by functional tests to individual 200-300 bp core elements (Philipsen et al., 1990; Pruzina et al., 1991; Talbot et al., 1990) which are further supported by conservation of core sequences illustrated in phylogenic homology comparisons between species (Hardison et al., 1997). The core elements have all been footprinted and multiple binding sites for erythroid-specific and ubiquitous transcription factors have been identified (Hardison et al., 1997; Orkin, 1995). However, the sequence flanking the cores is much less conserved but is important for the function of these HSs (Jackson et al., 1996a; Jackson et al., 1996b). Most of the initial studies on βLCR function concentrated on assigning a function to each HS. 5’HS2 was found to be a potent enhancer in transient and stable transfection assays conducted in MEL cells indicating the enhancer activity in the presence and absence of chromatin (Collis et al., 1990; Talbot et al., 1990; Tuan et al., 1989). In contrast, 5’HS3 and 4 only had enhancer activity in stable transfection of MEL cells indicating a requirement for chromatin in the activity of these elements (Collis et al., 1990; Jackson et al., 1996b; Moon and Ley, 1991). In transgenic mice, 5’HS2, 3, and 4 were all able to confer high-level expression of multicopy cosmid constructs bearing the human γ- and β-globin genes (Fraser et al., 1993). No enhancer activity in tissue-culture or transgenic mice is associated with the 5’HS1 element. However, our lab has demonstrated that full LCR activity on a β-globin transgene in mice is only achieved when the
5'HS1 element is present in the LCR fragment (Pasceri et al., 1998). Very little is known about the function of the ubiquitous 5'HS5 site. It has become clear through the years that the quantitative effects of the individual HSs on expression in transgenic mice are markedly less than the sum of the HSs.

Much of what we know about LCR activation comes from using βLCR linked β-globin locus cosmid or yeast artificial chromosome (YAC) transgenes in mice. These studies have further illustrated that the LCR is not simply a classical enhancer as it is orientation-dependent (Tanimoto et al., 1999a) and distance-dependent (Dillon et al., 1997; Hanscombe et al., 1991; Peterson and Stamatoyannopoulos, 1993). Large deletions (1-2 kb) of individual HSs in these constructs have lead to a mild reduction in gene expression, the genes are correctly expressed throughout development, and the remaining HSs are detected (Peterson et al., 1996). A similar phenotype is seen in mice where large regions encompassing individual HSs have been deleted in the endogenous locus (Fiering et al., 1995; Hug et al., 1996). Other studies have shown that individual HS can partially or fully functionally replace other HS within the βLCR. For example, in YAC transgenics where 5'HS2 is replaced with 5'HS3, the chromatin opening activity of the βLCR remains intact but expression levels of all β-like genes are drastically reduced (Bungert et al., 1999). However, other regulatory elements that are DNaseI hypersensitive in erythroid cells cannot replace the βLCR HSs (Tanimoto et al., 1999b). These experiments indicate that each HS contributes to LCR activity but that there is a functional redundancy of HSs. Furthermore, these studies support a model where HS synergize and fold together to form a Holocomplex (Bungert et al., 1995; Bungert et al., 1999; Grosveld et al., 1993). The individual globin promoters are thought
to compete for the LCR Holocomplex which can interact with individual globin promoters, via DNA looping, and activate transcription (Fig. 2). A mutually exclusive model predicts that the βLCR interacts indirectly, by providing an open, accessible chromatin conformation and that the gene-proximal regulatory elements mediate the developmental switches as well as the high levels of gene transcription (Martin et al., 1996).

In contrast to large HS deletions, small core deletions (200-400 bp) in YAC constructs lead to a severe reduction in expression of all β-like genes and the disappearance of the remaining HSs (Bungert et al., 1995; Bungert et al., 1999; Navas et al., 1998). The reasons for these discrepant results are unclear. Bungert et al. suggest that a small core deletion completely disturbs Holocomplex formation whereas larger deletions allow an alternative Holocomplex to form with a decreased ability to function (Bungert et al., 1999). The most striking studies involving βLCR function has just recently been reported. These gene targeting experiments where the whole or majority of the βLCR was deleted in the endogenous mouse or human locus, respectively, suggest that the βLCR is not required for chromatin opening of the endogenous murine or human β-globin loci but is absolutely required for expression of the locus (Epner et al., 1998; Reik et al., 1998). In the case of the human deletion (Fig. 2), the phenotype is less severe than the naturally occurring Hispanic deletion (Fig. 2). The Hispanic deletion removes 5'HS5-2 plus 30 kb upstream of 5'HS5 and results in the complete shutdown of the locus with respect to expression, chromatin structure, and DNA replication timing (Driscoll et al., 1989). These findings have prompted some researchers to look for additional regulatory elements that govern chromatin structure upstream of the known murine and human βLCRs (Bulger et al., 1999) and to propose an alternative Linking model for
βLCR function (Bulger and Groudine, 1999). This model contradicts the Holocomplex/Looping model of βLCR function and is the first to dismiss the βLCR element in chromatin opening. The model proposes that the βLCR mediates high-level expression of the individual β-like globin genes indirectly by binding specific proteins. Binding of these proteins in the βLCR region is thought to extend ("linking") to the individual globin promoters creating an active transcription complex and activating transcription during development (Bulger and Groudine, 1999).

5. Designing βLCR/β-globin minigenes for use in gene therapy

The βLCR including all four tissue-specific HSs is encompassed on a 20 kb fragment which is much too large to include in gene therapy minigenes. However, a 6.5 kb LCR fragment, termed μlocus, that contains 5'HS1-4 has been reported to contain enough sequence information for full LCR activity in transgenic mice (Ellis et al., 1996) (Table 2). Furthermore, smaller recombinant forms of the βLCR that retain partial activity, such that they can direct reproducible expression of β-globin transgenes at single-copy but expression levels are not maximal, have also been identified (Ellis et al., 1997) (Table 2). The levels of β-globin gene expression are therapeutic making these βLCR derivatives suitable for β-globin gene expression in gene therapy vectors. Unfortunately, the size of the μlocus and these derivatives are still too large to include in retrovirus gene therapy vectors that contain a drug resistance gene for ex vivo or in vivo expansion of transduced HSCs (Allay et al., 1998; Allay et al., 1997; Sorrentino et al., 1999; Sorrentino et al., 1992). Researchers have concentrated on smaller βLCR derivatives that incorporate core sequences of the individual HSs so that these derivatives can be incorporated into gene therapy viral vectors.
(Einerhand et al., 1995; Emery et al., 1998; Hargrove et al., 1997; Leboulch et al., 1994; Sadelain et al., 1995; Takekoshi et al., 1995).

Table 2: Human βLCR fragments assayed for activity in transgenic mice

<table>
<thead>
<tr>
<th>5’HS</th>
<th>Size (kb)</th>
<th>Multicopy</th>
<th>Single-copy (% expr.)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-4</td>
<td>6.5</td>
<td>Yes</td>
<td>Yes (100%)</td>
<td>(Ellis et al., 1996)</td>
</tr>
<tr>
<td>2-4</td>
<td>3.0</td>
<td>Yes</td>
<td>Yes (45%)</td>
<td>(Ellis et al., 1997)</td>
</tr>
<tr>
<td>2</td>
<td>1.5</td>
<td>Yes</td>
<td>No</td>
<td>(Ellis et al., 1996)</td>
</tr>
<tr>
<td>3</td>
<td>1.9</td>
<td>Yes</td>
<td>Yes (26%)</td>
<td>(Ellis et al., 1996)</td>
</tr>
<tr>
<td>4</td>
<td>2.1</td>
<td>Yes</td>
<td>No</td>
<td>(Ellis et al., 1996)</td>
</tr>
<tr>
<td>2 core</td>
<td>0.215</td>
<td>Yes</td>
<td>No</td>
<td>(Ellis et al., 1997; Ellis et al., 1993)</td>
</tr>
<tr>
<td>3 core</td>
<td>0.120</td>
<td>Yes</td>
<td>No</td>
<td>(Ellis et al., 1997; Philipsen et al., 1993)</td>
</tr>
<tr>
<td>4 core</td>
<td>0.280</td>
<td>Yes</td>
<td>ND</td>
<td>(Pruzina et al., 1991)</td>
</tr>
</tbody>
</table>

Multicopy refers to expression in multicopy transgenic mice reflecting copy number-dependent expression. Single copy refers to expression in single-copy transgenic mice reflecting position-independent expression. Single copy expression represents the average from all integration sites tested represented as the percentage of murine β<sub>ma</sub> on a per copy basis. ND, not determined.

As mentioned earlier, β-globin inserts in the AS orientation are deleterious to the retrovirus life-cycle leading to low titers of the recombinant retrovirus and instability problems. With the inclusion of these mini βLCR derivatives, many groups have reported that the instability of recombinant retroviruses is exaggerated (Chang et al., 1992; Gelinas et al., 1992; Mulligan, 1993; Novak et al., 1990; Plavec et al., 1993). Other groups have investigated the use of γ-globin transgenes for adult expression (Emery et al., 1998; Li et al., 1999a; Ren et al., 1996; Rixon et al., 1990) and, in some cases, have looked at the possibility of using other LCR elements such as the
LCR-like element in the α-globin locus, HS-40, in retrovirus inserts (Li et al., 1999a; Ren et al., 1996). Inserts based on γ-globin expression have also been shown to be deleterious to the retrovirus life-cycle and result in genetic instability of the retrovirus (Emery et al., 1998; Emery et al., 1999; Li et al., 1999a; Rixon et al., 1990).

These deleterious effects of βLCR/β-globin minigenes on the retrovirus life-cycle have limited minigene generation. Only after multiple arrangements and orientations of βLCR HS cores were studied and extensive mutagenesis of cryptic RNA processing recognition sites in the AS β-globin gene sequence have these deleterious effects been partially avoided and potential βLCR/β-globin minigenes for retrovirus-mediated gene therapy been designed (Leboulch et al., 1994; Sadelain et al., 1995). In these studies, β-globin intron 2 was highlighted to be the major contributor from the β-globin gene to the low titer and instability problems seen in recombinant retroviral vectors containing these βLCR/β-globin inserts. These effects were partially eliminated by deleting an AT-rich region within β-globin intron 2 and stable retrovirus was generated with certain arrangements of the βLCR core elements (Leboulch et al., 1994; Sadelain et al., 1995). These minigenes were expressed to high levels in MEL cells indicating that the mutations did not perturb β-globin expression in vitro. However, the expression of the transgenes was variable from cell clone to clone and subject to position effects indicating that the mini βLCR derivatives used were not able to confer reproducible expression on β-globin transgenes at single-copy. These minigenes were further shown to be inadequate for gene therapy purposes since they were not sufficient for sustained expression in vivo and most often transcriptionally inactive several months after bone marrow transplantation in mice (Rivella and Sadelain, 1998).
The poor \textit{in vivo} expression of these minigenes reflects the fact that these mini $\beta$LCR elements were designed from expression data obtained in multicopy transgenic mice and in tissue culture. Single-copy transgenes are more susceptible to position effects than are multicopy transgenes arranged in a concatamer. Therefore, reproducible expression from single-copy transgenes can be used as an assay for a chromatin opening domain within the transgene and help to identify potential globin minigenes for gene therapy applications. It has now become clear that these mini $\beta$LCR elements and the use of core HS elements are not sufficient for reproducible single-copy expression of the $\beta$-globin transgene and potential minigene should be tested in single-copy transgenic mice. Furthermore, assaying expression of transgenes in tissue culture requires selection for transfected cells. Since selection requires expression from a drug resistance gene, the selection procedure indirectly selects for the integration of transgenes in transcriptionally-permissive chromatin environments. In addition to the fact that transfection of cells results in high copy numbers, the selection procedure further limits the usefulness of tissue culture assays to identify potential minigenes that are expressed reproducibly. Also, although the HS-40 element of the $\alpha$-globin locus seems to be less recombinogenic (Ren \textit{et al.}, 1996), inserts based on this element are not likely to be effective for long-term expression \textit{in vivo} as this element does not confer reproducible expression on the $\alpha$-globin gene in transgenic mice (Higgs \textit{et al.}, 1990). To date, most potential minigenes have been identified based on their ability to be transmitted faithfully by a retrovirus and were then tested for expression in relevant cells \textit{in vitro} or \textit{in vivo}.

Perhaps a better approach to $\beta$LCR/$\beta$-globin minigene design is to define the minimal regulatory elements necessary for reproducible expression at single-copy in mice, which is required
for effective retrovirus-mediated gene therapy, and then test potential minigenes for their effects on the retrovirus life-cycle. Furthermore, most groups who have designed βLCR/β-globin minigene have concentrated solely on the βLCR. However, simple addition of minimal LCR elements to a transgene does not necessarily result in reproducible expression at all sites and copy numbers. In fact, recent findings show that a complete βLCR requires linked β-globin gene sequences for activity, and suggest that the activity of the βLCR is not dominant but limited to expression of the β-globin gene. For example, the βLCR cannot confer reproducible transgene expression in mice on other gene sequences such as the LacZ marker gene (Guy et al., 1996; Tewari et al., 1996), γ-globin genes (Li and Stamatoyannopoulos, 1994; Roberts et al., 1997; Stamatoyannopoulos et al., 1997), or even β-globin genes that lack a 3' element (Liu et al., 1997; Pasceri et al., 1998).

To successfully use LCR elements in gene therapy vectors that integrate at single copy, it will be crucial not only to map the minimal LCR elements necessary for reproducible transgene expression in mice but also to characterize the gene proximal cis-acting sequences that functionally interact with the LCR to make any given expression cassette LCR-responsive. Defining the β-globin regulatory elements required by the βLCR not only highlights the importance of gene proximal regulatory elements in LCR activation but allows for the engineering of βLCR-linked hybrid genes coding for other therapeutic proteins regulated by β-globin regulatory elements. Furthermore, the smaller the minigene, the less chance that the AS sequence of the insert with have deleterious effects on the retrovirus life-cycle.
5.1. Mapping the minimal $\beta$LCR elements required for reproducible transgene expression

Although the requirement for the human $\beta$LCR in chromatin opening at the endogenous locus is not clear, the ability of small $\beta$LCR elements to open chromatin at ectopic sites and direct high-level $\beta$-globin transgene expression has great promise for the design of gene therapy minigenes to treat the hemoglobinopathies. Since large $\beta$LCR derivatives are effective for transgene expression but are too large to incorporate into viral vectors and small $\beta$LCR derivatives are ineffective for transgene expression in vivo, our lab has dissected the $\beta$LCR to determine which HS(s) is sufficient for chromatin opening and reproducible single-copy transgene expression in mice (Table 2). It was originally thought that 5\textsuperscript{'}/HS2, because of its enhancer properties in tissue culture, was solely responsible for copy-number dependent expression of $\beta$-globin transgenes in mice mediated by the $\beta$LCR. In fact, a number of minigenes designed for use in retrovirus or AAV vectors contain this element alone linked to a $\beta$-globin transgene (Chang et al., 1992; Miller et al., 1993b; Novak et al., 1990; Takekoshi et al., 1995; Walsh et al., 1992). However, we have found that a 5\textsuperscript{'}/HS2-linked $\beta$-globin transgene displays copy-number dependent expression characteristics only in multicopy concatamers (Ellis et al., 1993). In other words, 5\textsuperscript{'}/HS2 did not protect transgene expression from negative position effects resulting from the neighboring chromatin structure in single-copy transgenes. This data indicates that this element alone is not ideal for inclusion in gene therapy cassettes.

We have found that chromatin opening by the $\beta$LCR does not require all four tissue-specific HSs by using single-copy transgenic mouse lines to map chromatin opening activity of the $\beta$LCR to a 1.9 kb 5\textsuperscript{'}/HS3 element when it is linked to the $\beta$-globin gene (Ellis et al., 1996).
Expression levels are reduced to about 26% compared to the endogenous murine $\beta_{maj}$ on a per copy basis. DNaseI digestion experiments demonstrated that 5'HS3, but not 5'HS2, could open chromatin at all integration sites tested (Ellis et al., 1996). This study revealed the LCR activity associated with 5'HS3 and the importance of including this element in gene therapy vectors. Although gene expression was not to full levels, all single-copy transgenic animals tested expressed $\beta$-globin at levels of therapeutic value. The study mentioned above did not indicate whether 5'HS3 was sufficient for reproducible single-copy transgene expression or whether this element requires a functional interaction with $\beta$-globin gene sequences for LCR activity.

5.2. Candidate $\beta$-globin gene regulatory elements that may confer LCR responsiveness

There are several gene proximal tissue-specific regulatory elements spread throughout the human $\beta$-globin locus. YAC constructs containing the complete $\beta$-globin locus but excluding the LCR are correctly regulated in mice but the expression levels of the individual genes are drastically reduced and subject to position effects. This result indicates that the gene proximal regulatory elements are responsible for the tissue-specific and developmentally-regulated expression of the locus. There are several candidate regulatory elements within the $\beta$-globin gene that might conceivably functionally interact with 5'HS3 to confer LCR responsiveness. These include the $\beta$-globin promoter element (Antoniou et al., 1988), the intragenic enhancer (Antoniou et al., 1988; Behringer et al., 1987) and matrix attachment region (MAR) (Jarman and Higgs, 1988) in intron 2, and the 3' enhancer (Antoniou et al., 1988; Behringer et al., 1987; Trudel and Costantini, 1987; Wall et al., 1988). The $\beta$-globin promoter and enhancers have been extensively analyzed, but their role in LCR activation has not been systematically evaluated in single-copy transgenic mice. MARs
are genomic DNA fragments that have the ability to bind to isolated nuclear matrices in vitro. The overall function of MARs is unclear. However, these elements have been proposed to represent structural boundaries of individual chromatin domains (Dillon and Grosveld, 1994; Namciu et al., 1998), regulate chromosome shape and maintenance (Hart and Laemmli, 1998; Strick and Laemmli, 1995), and function in the regulation of gene expression (Bode et al., 1996; Laemmli et al., 1992).

The minimal β-globin promoter maps to a 103 bp fragment that is inducible by the βLCR in stable transfection studies and which has been fully footprinted for transcription factor binding sites (Antoniou and Grosveld, 1990). The full 1500 bp β-globin promoter is required by the μlocus βLCR fragment for full level expression at all integration sites and copy numbers (Pasceri et al., 1998). However, the 815 bp β-globin promoter is sufficient for reproducible expression in single-copy transgenic mice regulated by a 3.0 kb (5′HS2-4) βLCR cassette (Ellis et al., 1997) but β-globin expression is not to full levels. We noted that the 265 bp β-globin promoter, commonly used in gene therapy minigenes, does not direct single-copy expression linked to the same βLCR cassette. Similarly, βLCR activation of multicopy γ-globin transgenes is dependent on the length of the γ-globin promoter (Stamatoyannopoulos et al., 1993). These findings suggest that cis-acting sequences in proximal globin promoters may have a role in reproducible expression directed by βLCR elements and should be included in gene therapy cassettes.

In the absence of βLCR elements, β-globin gene enhancers have been shown to contribute to the expression of the human β-globin transgenes in induced murine erythroleukemia (MEL) cells (Antoniou et al., 1988) and can confer tissue- and stage-specific expression on the Aγ-globin gene (Behringer et al., 1987) or heterologous genes (Magram et al., 1989) in transgenic mice.
However, these enhancers have no role in βLCR-mediated induction of the β-globin promoter in stable transfection studies (Antoniou and Grosveld, 1990; Collis et al., 1990; Li et al., 1999a) which led to the omission of the 3' enhancer from most β-globin gene therapy vectors (Einerhand et al., 1995; Leboulch et al., 1994; Sadelain et al., 1995). We have recently shown that reproducible expression from single-copy β-globin transgenes regulated by a 4.0 kb (5'HS1-4) βLCR cassette requires 1.65 kb of 3' sequence that includes the known 260 bp β-globin 3' enhancer (Pasceri et al., 1998). In addition, deletion of the β-globin 3' enhancer in YAC transgenic mice causes a reduction in β-globin gene expression, indicating that the β-globin 3' enhancer influences globin switching (Liu et al., 1997). These in vivo data suggest that 3' sequences may have a role in LCR activation by 5'HS3 and should also be included in gene therapy vectors. The importance of β-globin intron 2 sequences (that include the intragenic enhancer and MAR) for LCR activity by 5'HS3 has yet to be determined in transgenic mice.

5.3. Defining the minimal β-globin regulatory elements required for LCR activity by 5'HS3

The work presented here demonstrates that an 850 bp 5'HS3 element reproducibly activates linked β-globin transgenes in the fetal liver tissue of mice, which further maps the sequences necessary for reproducible transgene expression directed by this element. However, 5'HS3 is incapable of similarly activating linked Aγ-globin transgenes, demonstrating a requirement for β-globin sequences in the reproducible activation of transgenes by 5'HS3. In order to identify the critical β-globin gene sequences, Aγ-globin sequences were replaced with various combinations of candidate β-globin regulatory elements. These hybrid transgenes all retain the Aγ coding sequences and are linked to the 850 bp 5'HS3 element. The results indicate that the β-globin intron 2 and the
260 bp 3’ enhancer are involved in LCR activity mediated by the 5’HS3 element and are required for reproducible single-copy transgene expression. In addition, this study extends the utility of the β-globin LCR to include expression of the Aγ-globin gene. Such a β/γ-globin hybrid gene with single-copy expression characteristics is ideally suited for gene therapy of sickle cell anemia because γ-globin protein has better anti-sickling properties than β-globin (Takekoshi et al., 1995).
II. MATERIALS AND METHODS

1. Transgene plasmid construction

Transgene constructs were derived from the plasmids pGSE1758 (Talbot et al., 1989), pBGT14 (Ellis et al., 1997), p141 (Leboulch et al., 1994), and pAγ-globin (provided by S. Philipsen). pGSE1758 contains a polylinker 5' of the 4.2 kb HpaI-EcoRV β-globin gene fragment regulated by the 815 bp promoter. pBGT14 contains a 3.0 kb LCR cassette and the 4.2 kb HpaI-EcoRV β-globin gene fragment regulated by the 815 bp promoter. The 3.0 kb LCR contains the 1.15 kb StuI-SpeI fragment of 5'HS4, the 0.85 kb SacI-PvuII fragment of 5'HS3, and the 0.95 kb SmaI-StuI fragment of 5'HS2. p141 contains a SnaBI-PstI β-globin gene fragment which includes a 372 bp RsaI-RsaI deletion in intron 2.

BGT9 was constructed by inserting the 850bp SacI-PvuII fragment of 5'HS3 into the XhoI site of pGSE1758 using XhoI linkers. The 5.0 kb transgene including the entire β-globin gene was purified as an EcoRV fragment.

BGT26 replaces all the β-globin gene sequences in BGT9 between the SalI-XbaI sites with the 4.3 kb BspHI fragment of Aγ-globin by blunt end ligation. This includes the Aγ-globin 700 bp promoter and 3' enhancer. The 5.2 kb transgene was purified as an EcoRV fragment.

BGT34 inserts the 1.9 kb Ncol-HindIII fragment from Aγ-globin into the Ncol-EcoRV sites of BGT9 using an Nhel linker at the incompatible HindIII and EcoRV overhangs. The Aγ-globin
sequences extend from the ATG translation start site located at the Ncol site used for subcloning, to 375 bp 3' of exon 3 including the polyA site but not the Aγ-globin 3' enhancer. The end result is that BGT34 contains the 815 bp β-globin promoter controlling the entire Aγ-globin coding sequences and both introns. The 3.7 kb transgene was purified as an EcoRV-Nhel fragment.

BGT35 inserts the β-globin intron 2 sequences as a BamHI–EcoRI fragment into the compatible BamHI–EcoRI sites of BGT34. These changes also replace 4 Aγ-globin codons (101-104) with β-globin exon 2 sequences, and 16 Aγ-globin codons (105-119) with β-globin exon 3 sequences. Of these 20 β-globin codons, 17 encode the same amino-acid as Aγ-globin. The three altered codons are described in the text. The end result is that BGT35 contains the 815 bp β-globin promoter and the β-globin intron 2 sequences controlling the Aγ-globin coding sequences. The 3.7 kb transgene was purified as an EcoRV-Nhel fragment.

BGT50 contains a polylinker at the Nhel site of BGT35 that adds EcoRV, AgeI, and Clai sites 3' of the hybrid globin gene. The 260 bp PstI fragment containing the β-globin 3’ enhancer was cloned into the Nhel site using linkers. The end result is that BGT50 contains the 815 bp β-globin promoter, the β-globin intron 2, and the β-globin 3’ enhancer controlling the Aγ-globin coding sequences. The 3.9 kb transgene was purified as a Clai fragment.

BGT54 contains the 3.0 kb Clai–EcoRI fragment of BGT34 linked to the 850 bp EcoRI-ClaI fragment of BGT50. The end result is that BGT54 contains the 815 bp β-globin promoter and the β-globin 3’ enhancer controlling the entire Aγ-globin coding sequences and both introns. The 3.9 kb transgene was purified as a Clai fragment.
BGT76 contains the 2.1 kb *ClaI-BamHI* fragment of BGT26 linked to the 1.7 kb *BamHI-ClaI* fragment of BGT50. The end result is that BGT76 contains the 700 bp *Aγ*-globin promoter, the β-globin intron 2, and β-globin 3' enhancer controlling the Aγ-globin coding sequences. The 3.8 kb transgene was purified as a *ClaI* fragment.

BGT64 contains the 2.2 kb *ClaI-BamHI* fragment and the 820 bp *EcoRI-ClaI* fragment of BGT50 linked by the 540 bp *BamHI-EcoRI* β-globin intron 2 fragment of p141. The end result is that BGT64 is a derivative of BGT50 that contains a 372 bp *RsaI-RsaI* deletion in β-globin intron 2. The 3.5 kb transgene was purified as a *ClaI* fragment.

2. Generation of transgenic mice

Transgene DNA was prepared using Plasmid Maxi Kits (Qiagen). Transgene fragments were liberated from their plasmid backbones by digestion with the stated restriction enzymes. DNA fragments were recovered from 0.7% TAE agarose gel slices using GeneClean II or GeneClean Spin Column Kits (Bio101) and Elutip-d columns (Schleicher and Schuell), and resuspended in injection buffer (10 mM Tris-HCl pH 7.5, 0.2 mM EDTA). DNA concentration was determined by comparison with DNA standards run on agarose gels, and the injection fragment was diluted to 0.5-1 ng/μl in injection buffer. The diluted DNA was pre-spun for 20 minutes and aliquots removed for microinjection into fertilized FVB mouse eggs. Injected eggs were transferred into recipient CD1 female animals. 15.5 days post-injection, fetal mice were dissected and DNA extracted from head tissue while the fetal livers were saved frozen in two halves for future analysis. Head DNA was extracted by Proteinase K digestion overnight, a single phenol/chloroform extraction and
isopropanol precipitation. Transgenic fetuses were identified by slot blot hybridization with the 5'HS3 probe using standard procedures.

3. DNA analysis

Southern blot analysis was used to determine the transgene copy number and intactness and was also used to calculate the percent transgenicity of the individual founders. Southern transfer and hybridization were carried out by standard procedures and quantification of band intensity was performed using a Molecular Dynamics PhosphorImager. Single-copy animals showed a single random sized end-fragment in BamHI and EcoRI digests hybridized with the 5'HS3, β-globin intron 2 or Ay-globin 3' probes. These probes correspond to the XhoI 5'HS3 fragment, the BamHI-EcoRI β-globin intron 2 fragment, and the EcoRI Ay-globin exon 3 fragment used in cloning the constructs. With multicopy animals, the intensity of the end-fragment was defined as one transgene copy and was used to calculate the copy number of the multicopy junction-fragment in the same lane.

The intactness of the transgene in the DNA sample was verified using two sets of digests, appropriate for each construct, which together span the entire transgene. The resulting fragments were hybridized to either 5'HS3, β-globin intron 2, or Ay-globin 3' probes. Non-intact transgenes were not included in the calculation of copy number. Finally, DNA derived from one half of the frozen fetal livers was digested with AccI (BGT9) or PstI (Ay-globin transgenes) and the transgene detected by the β-globin intron 2 or 5'HS3 probes respectively to determine the proportion of transgenic cells in the fetal liver. The intensity of the transgene band per copy was compared to the
intensity of the single-copy bred line control band to calculate percent transgenicity. The quantification of the transgene band was corrected for DNA loading differences by dividing the intensity of the transgene band by an endogenous mouse gene band detected by a separate probe. Only animals containing at least one intact transgene were analyzed and highly mosaic animals (<10% transgenic cells in the fetal liver) were excluded from this study.

4. RNA analysis

Fetal liver (embryonic day 15.5) RNA was extracted using Trizol Reagent (Gibco BRL). 1 μg was hybridized to [γATP]-labeled double-stranded 5’DNA probe for human β-globin detection (Antoniou et al., 1988) or a [αdATP]-labeled double-stranded 3’ DNA probe (the EcoRI fragment containing exon 3) for detection of human Aγ-globin (Kollias et al., 1986). A [γATP]-labeled double-stranded 5’DNA probe was used for mouse β-globin major detection as a loading control (Antoniou et al., 1988). RNA/DNA hybrids were subsequently digested with 75 U S1 nuclease (Boehringer Mannheim) and run on a 6% sequencing gel as described (Antoniou et al., 1988). Probe excess was demonstrated by including a sample containing 3 μg fetal liver RNA. Specific activities of human β-globin (Hβ) or human Aγ-globin (Hγ) relative to the mouse β major (Mβ) probe is described for each S1 nuclease experiment in the corresponding figure legend. The protected 170 nt Hγ, 160 nt Hβ, and 95 nt Mβ bands were quantified on a Molecular Dynamics PhosphorImager and the % expression levels calculated according to the formula (Hβ or γ / Mβ) x 100% and corrected for specific activity differences between the probe preparations. Expression per
transgene copy was calculated as (2 \( \text{M}^{\beta} \) genes / number \( H^{\beta} \) or \( \gamma \) transgenes) x (% expression) / (% transgenicity) x 100%.

5. Recombinant retrovirus plasmid construction

All globin inserts were cloned into the MSCV vector plasmid (Clontech) that contains the neomycin resistance gene (pMSCV neo). Inserts were cloned in the reverse orientation as compared to proviral transcription (explained in text).

MSCV/BGT54, MSCV/BGT64, and MSCV/BGT50 contain the 3.9 kb BGT54 \( Clai \) fragment, the 3.5 kb BGT64 \( Clai \) fragment, and the 3.9 kb BGT50 \( Clai \) fragment, respectively, cloned into the unique \( Clai \) site in pMSCV neo.

6. Generation of recombinant retrovirus

PA317 cells were grown under standard conditions. For each transfection, 2x10^5 PA317 cells were plated 18-24 hrs prior to transfection in 60 mm plates. Transfections were carried out using 4-8 \( \mu \)g of plasmid (linearized with \( NdeI \)) and 35 \( \mu \)l of Lipofectamine reagent (Gibco BRL). The liposome-DNA complexes were added to the cells in a total volume of 3 ml OptiMem serum-free medium (Gibco BRL) and the cells were incubated for 4-6 hrs. The medium was replaced with 5 ml of regular growth medium and cells were incubated for two days. Selection of transfectants (\( \approx \) 10 days) was carried out by supplementing growth medium with 1 mg/ml Geneticin (G418; Gibco
BRL). Virus was harvested in 5 ml of regular medium from a confluent plate of G418-resistant cells after incubation overnight.

7. Determination of retrovirus titer

3T3 cells were grown under standard conditions. 1x10^3 3T3 cells were plated 24 hrs prior to infection in 100 mm plates. Viral supernatants were spun at 1500 RPM, 4°C for 3 minutes to pellet contaminating producer cells and several serial dilutions were prepared in medium supplemented with 8 μg/ml Polybrene (Sigma). 3 ml of each dilution were added to a plate of cells and incubated for 3 hrs. 7 ml of regular medium were added and cells were incubated overnight. Selection of infectants was carried out as above. The number of G418-resistant colonies on the various plates was used to calculate Colony Forming Units (CFUs)/ml of viral supernatant.
III. RESULTS

1. Transgenes designed to identify a requirement for β-globin gene sequences by 5'HS3

In order to determine whether the 850 bp 5'HS3 element can direct high-level position-independent transgene expression when linked to the β-globin gene, the BGT9 construct was created (Fig. 3). The β-globin gene sequences in BGT9 include the 815 bp β-globin promoter, the entire β-globin coding sequences including both introns and 1.65 kb of 3' sequences including the 260 bp 3' β-globin enhancer. To determine whether the 850 bp 5'HS3 element requires β-globin gene sequences for LCR activity, 5'HS3 was also linked to the Aγ-globin gene (BGT26 construct; Fig. 3). BGT26 includes the 700 bp Aγ-globin promoter, the entire Aγ-globin coding sequences including both introns, and 2.0 kb of 3' sequences including the 3' Aγ-globin enhancer.

2. Generation of transgenic mice and analysis of transgenes

These DNA constructs were purified as linear fragments and microinjected into fertilized FVB mouse eggs in order to create transgenic mice. The fetuses derived from these eggs were dissected at embryonic day 15.5 and genomic DNA extracted from head tissue, while the fetal livers were frozen in two halves for future analyses. Positive transient transgenic founder animals were identified by slot blot hybridization. The number of intact transgenes and the percent transgenicity for each transgenic founder was deduced by Southern blot analysis (data not shown).
Figure 3

Map of transgene constructs designed to determine the importance of specific human β-globin gene sequences in Locus Control Region (LCR) activity conferred by the 5'HS3 element. The sequences used in each construct are outlined in the Materials and Methods section. The 850 bp 5'HS3 fragment is indicated by a gray rectangle. The β-globin gene is indicated as a thick line with exons as black boxes. β-globin gene regulatory elements are indicated by black arrows and include the 815 bp β-globin promoter (Prom), the AT-rich region (AT) that coincides with a known Matrix Attachment Region and intragenic enhancer (Enh) located in intron 2, and the 260 bp 3' enhancer (Enh) located downstream of the gene. Aγ-globin gene sequences are represented as thin lines, and unfilled boxes (exons) or arrows (regulatory elements). Southern probe fragments are indicated as A, XhoI 5'HS3 fragment of BGT9; B, BamHI-EcoRI β-globin intron 2 fragment of BGT9; C, EcoRI fragment of BGT26.
Figure 4 shows an example of a Southern blot used to calculate transgene copy number and a blot used to determine transgene intactness and percent transgenicity of the founder mice. In the copy number blot (Fig. 4A.), head DNA from three BGT26 founders was digested with EcoRI (which cuts once within the transgene) and hybridized to the 5'HS3 probe. Founder 11 shows one band indicating that this animal contains one copy of the transgene. Founder 12 shows two bands of equal intensity, none of which reflects the size expected from a concatamer arrangement, indicating two independent integration events. Founder 50 shows two bands of different intensities. The more intense of the two is 5.2 kb which represents a head-to-tail concatamer band (the size of the transgene) and the least intense represents the random sized end-fragment. The concatamer band is three times the intensity of the end-fragment band, which indicates that the founder contains four copies of the transgene arranged in a head-to-tail concatamer. Copy numbers were confirmed using a BamHI digest (which also cuts once within the transgene) and hybridization to the 3' γ-globin probe.

In the intactness and percent transgenicity blot (Fig. 4B.), liver DNA from eight BGT50 founders was digested with PstI (which cuts twice within the transgene) and hybridized to the 5'HS3 probe. A DNA sample from a bred line control animal, known to have one intact copy of the BGT50 transgene (50-225), was also digested with PstI and included in the blot. Furthermore, a probe specific for an endogenous murine gene was used as a loading control (lowest band). The PstI digest was used to verify whether the 3.0 kb fragment of the BGT50 transgene was present within the animal and determine the number of non-intact transgenes which are subtracted from the total number of transgenes. Founders 50-207 to 48 all contain one copy of the transgene (data not
Figure 4

Determination of transgene copy number and intactness and determination of percent transgenicity of the transgenic mice by Southern Blot analysis. A. Transgene copy number blot for the BGT26 mice using an *Eco*RI (E) digest and hybridization to the 5'HS3 probe. The size of the concatamer band (5.2 kb) is indicated. B. Transgene intactness and percent transgenicity blot for the BGT50 mice using a *PstI* (P) digest and hybridization to the 5'HS3 probe. Mouse 50-225 is a representative of a single-copy bred line containing the BGT50 transgene. The size of the expected intact fragment (3.0 kb) is indicated. Non-intact copies of the transgene are subtracted from the total copy number. Tg, transgene band; mThy-1, endogenous mouse band (loading control).
shown) and show a single band of the expected size indicating that the transgene is intact for the region spanned in the PstI digest. Both founders 50-57 and 193 were determined to have two copies of the transgene in a copy number blot (data not shown) and each show two bands of equal intensity. One band is of the expected size and one is of a random size indicating that these animals contain one intact transgene and a partial transgene, which is not included in the copy number. Founder 50-197 was determined to contain eleven copies of the transgene (data not shown) and shows a band of the expected size. The intensity of the band indicates that multiple copies of the transgene are intact. However, one band, of random size, is detected indicating one non-intact transgene and reduces the copy number to ten. Intactness of the remaining region, not spanned by the PstI digest, was verified with another digest and probe (in this case BsmI and the 3' γ-globin probe). The PstI digest was also used to quantify the intact transgene band (standardized by dividing the intensity over the DNA control band and correcting for copy number) and compare the resulting intensity to that of the 50-225 DNA sample. The percent transgenicity values of the founders are listed.

3. Expression of BGT9 and BGT26 illustrates a requirement for β-globin gene sequences by 5’HS3

Expression from BGT9 and 26 was assayed in transgenic mice by S1 nuclease protection analysis in the fetal liver of 15.5 day transient transgenic mice. For the BGT9 construct, expression was analyzed using Hβ and Mβ probes (Fig. 5). As a standard sample for quantification of Hβ RNA levels, fetal liver RNA from a μD14 transgenic mouse that expresses to approximately 100% levels of murine βeq on a per copy basis (Ellis et al., 1996) was included in the assay. The BGT9
construct expresses significant levels of human β-globin mRNA in 17/17 transgenic mice. Expression from 3/3 single-copy BGT9 mice ranges from 46-109% per copy of the Mβ RNA, demonstrating that the 850bp 5’HS3 element directs reproducible single-copy transgene expression when linked to the entire β-globin gene. In addition, the BGT9 construct appears to express to a higher level at single copy than the previously described 26% levels per copy from the 1.9 kb 5’HS3 element (Ellis et al., 1996).

Similar expression analysis was performed on the BGT26 transgene using the Hy and Mβ probes (Fig. 6). As a standard RNA sample for all the Hy S1 nuclease experiments, BGT50-48 RNA (labelled C in all figures) was included in the assay. Expression by BGT50-48 is equivalent to the highest expressing single-copy BGT9 mouse, in that it contains one intact copy of the BGT50 construct (described later) and expresses γ-globin at 109% per copy of the level of Mβ (mean value from 8 experiments). S1 analysis of fetal liver RNA from BGT26 transgenic mice shows that low to moderate level expression is obtained in 2/3 animals with an undetectable level in the third animal. These data demonstrate that 5’HS3 cannot direct position independent transgene expression on a linked Δγ-globin gene and suggest that LCR activity by 5’HS3 requires a functional interaction with β-globin gene sequences.

4. Design of novel 5’HS3 β/γ-globin hybrid transgenes to define important β-globin sequences required for LCR activity by 5’HS3

To identify the β-globin gene sequences required for LCR activity by 5’HS3, several novel hybrid globin transgenes were created (Fig. 3). BGT34 contains 5’HS3 linked to the 815 bp
Figure 5

Expression of human $\beta$-globin mRNA in transgenic mice containing the BGT9 construct. S1 nuclease analysis on fetal liver RNA of 15.5 day founder BGT9 transgenic mice showing that BGT9 is expressed in all 17 animals including 3 single-copy mice. These data show that the 850 bp $5'\text{HS3}$ element can express reproducible levels of $\beta$-globin transcripts. Relative specific activity of $H\beta/M\beta$ probes is 1:1. $H\beta$, human $\beta$-globin protected probe fragment; $M\beta$, mouse $\beta$ major protected probe fragment; N, nontransgenic; $\mu$D, one copy $\mu$D14 microlocus line (discussed in text); $\%\text{TGN}$, percent transgenicity of the individual founders as determined by Southern blot; Tg, transgene.
**Figure 6**

Expression of human \(\gamma\)-globin mRNA in transgenic mice containing the BGT26 and BGT34 constructs. S1 nuclease analysis on fetal liver RNA of 15.5 day founder transgenic mice showing that BGT26 is expressed to low or undetectable levels and that BGT34 is expressed at significant levels in 4/7 transgenic mice. These data show that \(\beta\)-globin gene sequences are required for reproducible single-copy transgene activation by 5'HS3 and that the \(\beta\)-globin promoter element is not sufficient for this activity. Relative specific activities of Hy/M\(\beta\) probes is shown. Hy, human \(\gamma\)-globin protected probe fragment; M\(\beta\), mouse \(\beta\) major protected probe fragment; N, nontransgenic; C, 50-48, the highest expressing BGT50 single-copy transgenic mouse (discussed in text; see Fig. 9); 3X, probe excess control. Copy \#=1*, one intact copy plus a partial copy of the transgene; \%TGN, percent transgenicity of the individual founders as determined by Southern blot; Tg, transgene.
β-globin promoter and the Aγ-globin coding sequences terminating 375bp downstream of exon 3. This construct includes Aγ-globin intron 2 which has no known enhancer activity but excludes the Aγ-globin 3’ enhancer. The remaining constructs are modifications of BGT34 that all retain the β-globin promoter. BGT35 has a replacement of the Aγ-globin intron 2 sequences with β-globin intron 2. This adds the β-globin intron 2 enhancer and MAR, but also alters three amino acids in the Aγ-globin coding sequences to their equivalents in the β-globin gene (K104R, T112K, I116H). These changes do not alter amino-acids known to be important for anti-sickling effects (Takekoshi et al., 1995). BGT54 is essentially the same as BGT34 but with the addition of the 260 bp β-globin 3’ enhancer 375 bp downstream of the Aγ-globin coding sequences. Finally, BGT50 contains both the β-globin intron 2 sequences and the 260 bp β-globin 3’ enhancer inserted 375 bp downstream of the Aγ-globin exon 3. BGT50 also contains the three amino acid alteration in the Aγ-globin coding sequences.

5. Expression of 5’HS3 βγ-globin hybrid transgenes illustrates a requirement for β-globin 3’ elements

Expression from each of the 5’HS3 βγ-globin constructs was assayed in transgenic mice by S1 nuclease protection analysis in the fetal liver of 15.5 day transient transgenic mice as above. Human Aγ-globin mRNA was detected in 4/7 BGT34 transgenic mice (Fig. 6). Of these 7 animals, 5 animals contained a single intact copy of the BGT34 transgene and one animal (34-71) contained an intact copy and a partial transgene (indicated by the asterisk). Only 3/6 of these single-copy
BGT34 animals expressed detectable levels, with a range of 0-74% per copy. This finding demonstrates that the β-globin promoter is not sufficient to rescue LCR activity by 5'HS3.

Other candidate β-globin sequences that may functionally interact with 5'HS3 include the elements in β-globin intron 2 and 3' of the gene. Therefore, 5'HS3-linked Aγ-globin transgenes containing the 815bp β-globin promoter and either β-globin intron 2 (BGT 35) or the 260 bp β-globin 3' enhancer (BGT 54) or both (BGT 50) were analyzed. A total of 6/7 BGT35 animals expressed the transgene, including 2/3 single-copy animals that ranged from 0-39% per copy (Fig. 7). These data show that the β-globin promoter and β-globin intron 2 are not sufficient to rescue LCR activity directed by 5'HS3.

Similarly, BGT54 transgenes express significant levels of Aγ-globin mRNA in 5/6 single-copy animals (range 0-76% per copy) and in 8/9 multicopy animals (Fig. 8). Of the 6 single-copy BGT54 animals, two contained an intact transgene and a partial transgene (indicated by asterisks), and one sample represents a single-copy bred line where RNA expression was assayed in the adult blood (indicated as the "line"). However, one single-copy and a three copy animal express undetectable levels of Aγ-globin mRNA. This finding demonstrates that the β-globin promoter and β-globin 3' enhancer are not sufficient to rescue LCR activity by 5'HS3.

Finally, the BGT50 construct was tested in transgenic mice and significant expression was detected in 8/8 transgenic mice (Fig. 9). Seven single-copy animals were generated, two of which contained an intact and a partial transgene (indicated by asterisks). One animal (50-225) is a single-copy bred line sample where RNA was assayed in the fetal liver. The average expression of single-
Expression of human $\alpha$-globin mRNA in transgenic mice containing the BGT35 construct. S1 nuclease analysis on fetal liver RNA of 15.5 day founder transgenic mice showing BGT35 is expressed at significant levels in 5/7 transgenic mice. These data show that the $\beta$-globin promoter element and intron 2 sequence are not sufficient for reproducible single-copy transgene expression. Abbreviations as described in Fig. 6.
Figure 8

Expression of human $\alpha$-globin mRNA in transgenic mice containing the BGT54 construct. S1 nuclease analysis on fetal liver RNA of 15.5 day founder transgenic mice showing that BGT54 is expressed at significant levels in 13/15 transgenic mice. These data show that the $\beta$-globin promoter element and 3' enhancer sequence are not sufficient for reproducible single-copy transgene expression. Abbreviations as described in Fig. 6.
Figure 9

Expression of human \( \alpha \)-globin mRNA in transgenic mice containing the BGT50 construct. S1 nuclease analysis on fetal liver RNA of 15.5 day founder transgenic mice showing that BGT50 is expressed in all 8 animals including 7 single-copy mice. These data show that the \( \beta \)-globin intron 2 and 3' enhancer elements are sufficient for reproducible transgene expression when linked to the 5'HS3 element and the \( \beta \)-globin promoter. Abbreviations as described in Fig. 6.
copy BGT50 transgenes is 71% per copy of Mβ levels and ranges from 40%-109% per copy. Only BGT50 was reproducibly expressed in all transgenic mice including 7/7 single-copy animals indicating that LCR activity by 5'HS3 on the β-globin promoter requires a functional interaction with both β-globin intron 2 and the 260 bp β-globin 3’ enhancer.

6. 5'HS3 activation of the Ay-globin promoter

In order to determine whether the β-globin promoter is required, or that β-globin intron 2 sequences and the β-globin 3’ enhancer are themselves sufficient for LCR activity by 5’HS3, the BGT76 construct was created (Fig. 3). BGT76 is essentially the same as BGT50 with the replacement of the 815 bp β-globin promoter with the 700 bp Ay-globin promoter. Expression from BGT76 was assayed in transgenic mice by S1 nuclease protection analysis in the fetal liver of 15.5 day transient transgenic mice as above. 17/17 BGT76 transgenic mice expressed detectable levels of human Ay-globin mRNA (Fig. 10). Five single-copy animals were generated, three of which also contained a partial transgene (indicated by asterisks). These single-copy transgenic mice expressed a mean 46% Mβ levels ranging from 12%-72%. This finding demonstrates that the LCR activity by 5’HS3 is not dependent on the presence of the β-globin promoter and that the Ay-globin promoter can also be activated by a functional interaction between 5’HS3, β-globin intron 2, and the 260 bp β-globin 3’ enhancer.
Figure 10

Expression of human Ay-globin mRNA in transgenic mice containing the BGT76 construct. S1 nuclease analysis on fetal liver RNA of 15.5 day founder transgenic mice showing that BGT76 is expressed in all 17 animals including 5 single-copy mice. These data show that activation of single-copy transgenes by a functional interaction between 5'HS3, the β-globin intron 2, and 3' enhancer elements is not specific for the β-globin promoter. Abbreviations as described in Fig. 6.
Figure 11

Expression of human Aβ-globin mRNA in transgenic mice containing the BGT 64 construct. S1 nuclease analysis on fetal liver RNA of 15.5 day founder transgenic mice showing that the BGT64 is expressed at significant levels in 13/17 transgenic mice. These data show that the AT-rich region in β-globin intron 2 is necessary for the single-copy expression characteristics of BGT50. Abbreviations as described in Fig. 6.
7. Requirement for an AT-rich region within \(\beta\)-globin intron 2

In order to determine whether the AT-rich region in \(\beta\)-globin intron 2 (located within a known MAR element) is required for LCR activity by 5'HS3, a derivative of BGT50 called BGT64 was created (Fig. 3). BGT64 contains a 372 bp deletion in \(\beta\)-globin intron 2. The deleted sequence is deleterious to the retrovirus life-cycle when included in \(\beta\)-globin inserts as part of recombinant retroviral vectors (Leboulch et al., 1994; Sadelain et al., 1995). Expression from BGT64 was assayed in transgenic mice by S1 nuclease protection analysis in the fetal liver of 15.5 day transient transgenic mice as above. 13/17 BGT64 transgenic mice expressed detectable levels of \(\mathrm{H}_{\gamma}\) mRNA (Fig. 11) including 2/4 single-copy animals (range 0-22% per copy). The two expressing single-copy animals contained an intact transgene and a partial transgene (indicated by asterisks). Of the four mice that did not express detectable levels of transcript, two animals contained a single intact copy of the transgene and two animals contained multiple copies of the transgene. These data show that the 372 bp AT-rich segment deleted from \(\beta\)-globin intron 2 in BGT64 is important for LCR activity by 5'HS3.

8. Generation of recombinant retrovirus containing \(\beta/\gamma\)-globin hybrid inserts

BGT54, 64, and 50 were cloned into the pMSCVneo vector in the AS orientation giving rise to MSCV/BGT54, MSCV/BGT64, and MSCV/BGT50, respectively, and recombinant retrovirus was generated in the amphototropic packaging cell line, PA317. The biological virus titer of supernatants harvested from PA317 cells was determined by infection of NIH3T3 cells with serial dilutions of viral stocks and counting of G418-resistant colonies after selection. Virus titer
generated from populations of PA317 cells transfected with MSCV/BGT54, 64, and 50 were all similar for a given experiment and ranged from \(10^2\)-\(10^3\) infective virions per ml compared to \(10^3\) for the MSCV neo virus control (a total of five experiments were carried out). Therefore, the inclusion of these globin inserts drastically reduced virus titer 100 to 1000 fold. These results indicate that each insert is deleterious to the retrovirus life-cycle and that specific sequences or sequence arrangements, common to all inserts, are responsible for these effects.
IV. DISCUSSION

The human $\beta$LCR can reproducibly activate single-copy $\beta$-globin transgenes in the fetal erythroid tissue of mice. This activity illustrates that this element is a useful tool to include in gene therapy vectors designed to treat the $\beta$-chain hemoglobinopathies in humans. However, gene therapy cassettes should contain the minimal amount of regulatory elements necessary to be effective since retroviral vectors have insert size constraints (Miller, 1992) and insert sequences can be deleterious to the retrovirus life-cycle (Leboulch et al., 1994; Sadelain et al., 1995). The 850 bp 5'HS3 fragment discussed here is much smaller than other LCR fragments that function at single-copy (Table 2) and can express therapeutic levels of $\beta$-globin transcripts from single-copy transgenes making this element a far better $\beta$LCR-derivative to be used in gene therapy cassettes than $\beta$LCR elements that include multiple core HSs.

This study illustrates that 5'HS3 cannot reproducibly activate an $\alpha\gamma$-globin transgene in the fetal erythroid tissue of mice and must functionally interact with $\beta$-globin regulatory elements to do so. Expression of $\gamma$-globin protein chains in patients with $\beta$ thalassemia has a therapeutic effect and a potent antisickling effect in patients with sickle cell anemia (Takekoshi et al., 1995). Even low level expression of $\gamma$-globin can have beneficial effects in these patients. Therefore, transgenes that express reproducibly at single-copy but code for $\gamma$-globin protein are important minigenes to engineer for treatment of the $\beta$-chain hemoglobinopathies. The definition of the minimal combination of $\beta$-globin regulatory elements that must functionally interact with 5'HS3 to reproducibly express $\alpha\gamma$-globin transgenes serves the dual purpose of: 1) examining the functional
and cooperative interactions of these elements; as well as 2) creating a γ-globin expression cassette whose expression characteristics are well suited for gene therapy purposes.

1. Regulatory elements and reproducible transgene expression

1.1. 5'HS3 LCR activity requires linked β-globin gene sequences

The BGT9 construct demonstrates that the 850 bp fragment of 5'HS3 confers LCR activity on both multicopy and single-copy β-globin transgenes. This fragment is considerably smaller than the 1.9 kb 5'HS3 element previously shown to confer chromatin opening activity on single-copy β-globin transgenes. In addition, the 850 bp 5'HS3 element directs approximately 70% levels of β-globin gene expression at single-copy. As reproducible single-copy transgene expression is not obtained from a 125 bp minimal 5'HS3 core element (Ellis et al., 1997), this data demonstrates that sequences flanking the minimal core in the 850 bp 5'HS3 element are important for LCR function. Flanking sequences extending beyond the 850 bp element may actually reduce single-copy β-globin transgene expression to approximately 26% per copy as observed with the 1.9 kb 5'HS3 fragment (Ellis et al., 1996). This idea is further supported by the finding that small core deletion of HSs in β-globin locus YAC constructs in mice that leave the flanking sequences intact show more severe effects than when the core HS and flanking sequences are removed (Bungert et al., 1999). These data can be explained by a silencer activity present within the flanking sequence (Bulger and Groudine, 1999). Results from the BGT26 transgenic mice establish that the 850 bp 5'HS3 element cannot reproducibly express the Aγ-globin gene in the fetal liver of mice. This data agrees
with published findings (Li and Stamatoyannopoulos, 1994; Navas et al., 1995) and demonstrate that 5'HS3 LCR activity requires linked β-globin gene sequences.

1.2. β-globin intron 2 and 3' enhancer synergize to confer 5'HS3 responsiveness

A series of β/γ-globin hybrid genes were constructed to define the minimal combination of β-globin gene sequences required for LCR activity directed by 5'HS3. The hybrid transgenes and their expression in transgenic mice are summarized in Table 1. From this analysis, it is clear that 5'HS3 responsiveness is not dictated by the promoter alone. For example, BGT34 expresses in only 3/6 single-copy mice. In the additional presence of either the β-globin intron 2 or the 3' enhancer, expression improves to 2/3 (BGT35) and 4/5 (BGT54) single-copy animals, respectively. Both of these constructs provide more evidence that 5'HS3 responsiveness is not an inevitable consequence of the β-globin promoter, but rather suggest that the likelihood of expression at any given integration site is increased by the presence of intragenic or downstream elements.

In contrast to the effect of the β-globin promoter, position-independent transgene expression directed by 5'HS3 was conferred by the simultaneous presence of the β-globin intron 2 and 3' enhancer in the BGT50 and BGT76 constructs. These constructs expressed in 7/7 and 5/5 single-copy animals, respectively, and at high levels ranging from 40-109% per copy for BGT50 and 12-72% per copy for BGT76. The range of single-copy transgene expression from BGT50 is equivalent to that from BGT9 (46-109%) that contains the entire β-globin gene, showing that no other β-globin gene sequences are required for full 5'HS3 responsiveness. Together, these results show that a functional interaction between 5'HS3, the β-globin intron 2, and 3' enhancer elements is sufficient for single-copy transgene activation from globin promoters in these constructs. These
findings are the first to support a role for $\beta$-globin intron 2 in LCR activity using transgenic mice and agree with reports that LCR activity requires sequences that are 3' of the $\beta$-globin gene. For example, our lab have previously demonstrated that single-copy $\beta$-globin transgene expression from a 4.0 kb (5'HS1-4) LCR requires a fragment that includes 1.65 kb of 3' $\beta$-globin sequences (Pasceri et al., 1998). In the present study, the mapping of the required 3' sequences are refined to the minimal 260 bp 3' $\beta$-globin enhancer.

Table 3: Description of LCR and $\beta$-globin gene regulatory elements present in each $\beta$/$\gamma$-globin hybrid transgene and a summary of single-copy and total mice expressing these constructs

<table>
<thead>
<tr>
<th>BGT</th>
<th>5'HS3</th>
<th>Prom</th>
<th>$\beta$ivs2</th>
<th>$\beta$3'enh</th>
<th>Mice Expressing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single-copy</td>
</tr>
<tr>
<td>34</td>
<td>+</td>
<td>$\beta$</td>
<td>-</td>
<td>-</td>
<td>3/6 (0-74%)</td>
</tr>
<tr>
<td>35</td>
<td>+</td>
<td>$\beta$</td>
<td>+</td>
<td>-</td>
<td>2/3 (0-39%)</td>
</tr>
<tr>
<td>54</td>
<td>+</td>
<td>$\beta$</td>
<td>-</td>
<td>+</td>
<td>5/6 (0-76%)</td>
</tr>
<tr>
<td>50</td>
<td>+</td>
<td>$\beta$</td>
<td>+</td>
<td>+</td>
<td>7/7 (40-109%)</td>
</tr>
<tr>
<td>76</td>
<td>+</td>
<td>$\gamma$</td>
<td>+</td>
<td>+</td>
<td>5/5 (12-72%)</td>
</tr>
<tr>
<td>64</td>
<td>+</td>
<td>$\beta$</td>
<td>$\Delta$</td>
<td>+</td>
<td>2/4 (0-22%)</td>
</tr>
</tbody>
</table>

The range of expression in single-copy animals for each construct is listed as a percentage of $\alpha$-$\gamma$-globin transcripts to endogenous murine $\beta$-major transcripts on a per copy basis. Prom, promoter (either $\beta$ or $\gamma$); $\beta$ivs2, $\beta$-globin intron 2; $\beta$3'enh, $\beta$-globin 3' enhancer; $\Delta$, $\beta$-globin intron 2 present but containing a 372 bp AT-rich region deletion.

1.3. Silencing of the $\gamma$-globin promoter in the fetal erythroid tissues of mice

In the endogenous human $\beta$-globin locus, expression from the duplicated $\gamma$-globin genes ($G\gamma$ and $A\gamma$) is active at 16 weeks of gestation when the site of hematopoiesis becomes the fetal liver (Stamatoyannopoulos et al., 1994). By birth, the major site of hematopoiesis is the bone marrow.
and these cells express only δ- and β-globin while the γ-globin genes are silenced. The human locus is also developmentally regulated in YAC transgenic mice, where γ-globin expression is limited to the primitive cells of the embryonic yolk sac and silenced in the definitive cells of the fetal liver after the single embryonic to adult globin switch in these animals (Gaensler et al., 1993; Peterson et al., 1993). Furthermore, when a γ-globin transgene is linked with or without an LCR element in transgenic mice, expression is limited to the yolk sac (Chada et al., 1986; Dillon and Grosveld, 1991; Kollias et al., 1986) and some groups have proposed that the γ-globin genes are autonomously silenced by cis-acting sequences in adult tissues (Dillon and Grosveld, 1991). In particular, the 700 bp promoter, used in this study, has been highlighted as containing the cis-acting silencer sequences (Stamatoyannopoulos et al., 1993). Taking these studies into consideration, it might be expected that the BGT26 transgene is not active in the fetal liver of mice at day 15.5.

However, a low to moderate level of γ-globin expression was detected in 2/3 mice indicating that the BGT26 transgene was not completely silenced and capable of adult expression. In fact, other groups have reported fetal expression of LCR-linked γ-globin transgenes in mice and only recovered developmental control by adding a β-globin gene to the construct (Behringer et al., 1990; Enver et al., 1989; Enver et al., 1990). These data indicate that aside from gene specific elements, gene competition for the LCR also plays a role in hemoglobin switching and perhaps the γ-globin genes are not fully silenced autonomously in the fetal erythroid tissues of mice.

The expression analysis of the BGT34 construct indicates that removing the silencing activity of the γ-globin promoter and replacing this sequence with the β-globin promoter does not drastically affect the reproducibility of transgene activation (2/3 mice for BGT26 compared to 3/6
mice for BGT34). However, the levels of transgene expression may be higher for the BGT34 construct compared to the BGT26 construct. The higher level expression in the BGT34 transgenic mice may reflect the removal of the Ay-globin promoter silencing activity or the addition of a transcriptional enhancement activity in the β-globin promoter. Similar to the comparison between BGT26 and 34, the reproducible transgene expression seen for the BGT50 construct (8/8 mice) is not affected by the addition of the Ay-globin promoter in the BGT76 construct (17/17 mice). Again, the absence of the Ay-globin promoter or addition of the β-globin promoter may account for the slightly higher levels of expression detected in the BGT50 and BGT9 mice.

1.4. 5'HS3 responsiveness requires AT-rich β-globin intron 2 sequences

In order to more finely map the sequences within the β-globin intron 2 that are required for LCR activity by 5'HS3, a 372 bp AT-rich region in the BGT64 construct was deleted (Table 1). The results show expression in only 13/17 BGT64 mice including 2/4 single-copy animals. These findings demonstrate that a combination of the intragenic enhancer and the 3' enhancer is not sufficient to direct LCR activity from 5'HS3 even in the presence of the β-globin promoter. These conclusions therefore differ from previous reports that AT-rich deletions do not disturb β-globin expression in murine erythroleukemia (MEL) cells (Leboulch et al., 1994; Sadelain et al., 1995). The AT-rich deletion lies outside of sequences that are required for splicing or polyadenylation of β-globin transcripts (Antoniou et al., 1998). As the AT-rich region is completely embedded within a 540 bp BamH1-DraI fragment that has MAR activity (Jarman and Higgs, 1988), our results may suggest that a MAR element is required for reproducible single-copy transgene expression in conjunction with the presence of the 3' β-globin enhancer. Similar functional interactions between
MARs/facilitators and enhancers have been described for the immunoglobulin mu (Jenuwein et al., 1997) and ADA LCRs (Aronow et al., 1995).

1.5. Mechanism for single-copy expression characteristics

It is clear that the 5'HS3 element is a potent regulatory element at ectopic chromosomal locations. The single-copy expression characteristics of BGT 9, 50, and 76 suggest that these transgenes contain a chromatin opening domain. However, since founder transgenic animals were used in this study, the chromatin structure at the site of transgene integration could not be analyzed as transgenic lines are required for such studies. It is a possibility that 5'HS3 is sufficient for chromatin opening at the site of transgene integration but the presence of enhancer elements are absolutely required for high-level expression at all integration sites. Although it is generally believed that an open chromatin conformation is correlated with transcription, several groups have reported that an open chromatin conformation does not necessarily result in transcription (Feng et al., 1999; Reik et al., 1998; Vyas et al., 1992).

The precise mechanism of these functional interactions between LCR elements such as 5'HS3, MARs, and enhancers has not been elucidated to date. The Holocomplex model of βLCR function does not predict a requirement for gene-specific elements other than for promoters. Our finding that β-globin intron 2 and the 3' enhancer sequences are required for LCR activity by 5'HS3 at ectopic sites either suggests that these elements are involved in chromatin opening, or participate in additional DNA looping events with the LCR or promoter. Binding by trans-acting factors to elements spread throughout the β-globin gene is compatible with Linking models of LCR function. The Holocomplex and Linking models are not necessarily mutually exclusive as a
physical interaction between LCR elements and promoters would not interfere with a functional requirement for factors bound through the rest of the gene.

The function of the AT-rich region of the β-globin intron 2 in LCR activity remains elusive. A MAR activity is located in this region and may be important in vivo for the transgene DNA to interact with the nuclear matrix. However, the Aγ-globin 3' enhancer fragment also contains a MAR (Cunningham et al., 1994) but is not sufficient to rescue LCR activity in the BGT26 construct, suggesting that the AT-rich region in β-globin intron 2 may have other important functions. One possibility to consider is that AT-rich regions play roles in both transcription and the initiation of DNA replication (Cook, 1999; DePamphilis, 1999).

1.6. Reproducible single-copy expression by globin transgenes may require an origin of DNA replication

An origin of DNA replication lies within an 8 kb initiation region (IR) of the β-globin gene locus that is early replicating in erythroid tissue and late replicating in other somatic tissues. Recently, Aladjem et al. have confirmed that this biochemically-defined IR does contain a DNA replication origin using a functional assay to identify mammalian replicators on the basis of intrachromosomal location (Aladjem et al., 1998). Using various fragments from this 8 kb region, they were able to dissect the IR into a core element and auxiliary regions for DNA replication and implicates specific DNA sequences in the initiation of mammalian DNA replication.

I have noted that these critical regions co-localize with β-globin gene sequences present within the BGT9 transgene and encompass the important regulatory elements defined in this work.
to be important for reproducible expression of globin transgenes by 5'HS3. The presence of a functional origin of DNA replication within the BGT9 construct may play a role in its single-copy expression characteristics. It follows that BGT50 and 76 may contain sufficient β-globin sequence to recreate a functional origin of DNA replication. Since both processes involve an 'unwinding' of the DNA, it is conceivable that there is a link between the two processes and perhaps a functional origin of replication aids, in conjunction with 5'HS3, in reproducibly expressing these transgenes at single-copy. This hypothesis is testable and the β/γ-globin hybrid transgenes described in this thesis may provide a useful tool in further dissecting the minimal β-globin sequences necessary for a functional origin of DNA replication using the system described by Aladjem et al. This work is currently being carried out by members of our laboratory.

2. Gene therapy of the hemoglobinopathies

2.1. Novel 5'HS3 β/γ-globin transgenes for gene therapy

The ability of our novel hybrid globin transgenes to express high levels of transcripts at all integration sites and at single copy can be applied to both erythroid-specific transgene expression cassettes in animals and for gene therapy. BGT9 would be effective for gene therapy of β thalassemia and sickle cell anemia as it expresses high levels of β-globin transcripts, the defective gene in these diseases. BGT50, on the other hand, is the first description of Aγ-globin coding sequences controlled by β-globin regulatory elements and expresses high levels of γ-globin. γ-globin protein is as effective as β-globin in the treatment of β thalassemia but much more effective than β-globin in the treatment of sickle cell anemia. Both transgenes are well suited for DNA- or
viral-mediated gene therapy. Although BGT76 is expressed reproducibly at single-copy, its use as a gene therapy minigene for treatment of the hemoglobinopathies is unclear. It is possible that the 700 bp γ-globin promoter used in this construct will be silenced in adult blood cells in humans. It is unknown if the β-globin gene regulatory elements would be dominant over this silencing effect. However, there have been mutations identified within the γ-globin promoter associated with high-level γ-globin production in the adult blood, a condition termed Hereditary Persistence of Fetal Hemoglobin (HPFH) (Stamatoyannopoulos et al., 1994). Introducing HPFH mutations in the γ-globin promoter of the BGT76 transgene may allow such a derivative to be an effective minigene.

Given that the BGT50 hybrid transgene is expressed reproducibly in the fetal livers of transgenic mice, it is likely that this construct will also direct high-level expression of γ-globin in adult murine erythrocytes after transplantation of mice with transduced HSCs and in human adult erythrocytes in the event of a gene therapy protocol. Its 3.9 kb size is small enough for insertion into any currently designed viral vector for gene therapy. However, a prerequisite for further gene therapy studies using such a construct is to generate recombinant virus.

2.2. Generation of recombinant retrovirus

Our lab believes that the leading viral vector for transduction of HSCs is retroviruses. It was therefore crucial to generate recombinant retrovirus containing the BGT50 insert for transduction of biologically relevant cells and subsequent expression analysis. I attempted to generate recombinant retrovirus containing BGT54, 64, and 50 as inserts to determine what effects these inserts had on retrovirus titer and stability. These globin inserts differ only in their intron 2 content. β-globin intron 2 was previously reported to be deleterious to the retrovirus life-cycle when included in β-
globin inserts (Leboulch et al., 1994; Sadelain et al., 1995) and is present within the BGT50 transgene. The 372 bp deletion in intron 2 present within the BGT64 transgene was reported to abolish the deleterious nature of this intron (Leboulch et al., 1994). Ay-globin transgenes have been reported to be less recombinogenic in retroviral vectors (Ren et al., 1996), however, the intron 2 of this gene has also been reported to be deleterious (Li et al., 1999a; Rixon et al., 1990). All three of the constructs tested gave rise to low titer virus when included as inserts. Interestingly, the effect on titer was not specific for MSCV/BGT50 which contains β-globin intron 2 and the deleterious effects were not eliminated with the 372 bp AT-rich deletion present in the BGT64 construct. In fact, MSCV/BGT50 did not seem to generate lower titers than the other constructs. These results indicate that other specific sequences or arrangement of sequences, common to all transgenes, are deleterious to the retrovirus life-cycle.

It is well established that AS insert sequences can be deleterious to the retrovirus life-cycle leading to low titer retrovirus with multiple rearrangements detected in the transmitted proviral structure. This aspect further necessitates the design of minigenes containing minimal regulatory elements in order to decrease the chance of including deleterious sequences. I have noted that the distal 200 bp of the 815 bp β-globin promoter (present within all transgene constructs) contains a cluster of AAUAAA polyadenylation recognition sites in the AS orientation. Recognition of any of these sites would lead to packaged viral RNA that was non-functional as important sequences necessary for LTR formation upon reverse transcription would be lost but the viral genomic RNA would still be packaged as the packaging signal would still be present. In fact, I have identified viral supernatants from individual PA317 clones, for all recombinant retroviruses tested, that contain a
large proportion of packaged RNA, assayed by viral RNA dot blot (Murdoch et al., 1997), but have very low biological titers. However, the inclusion of the 815 bp promoter alone in the AS orientation as a retrovirus insert does not affect virus titer (data not shown). There remain other potential sequences present within all transgenes tested, such as the 850 bp 5’HS3 fragment, that may be deleterious. However, I have shown that this fragment alone also does not affect virus titer (data not shown).

The deleterious nature of inserts may not simply be the result of individual sequences but may be the result of the arrangement of sequences. For example, a cryptic 5’ splice site present at one end of the transgene on its own would not be deleterious. However, in the context of a full transgene, a 3’ splice site in another region may cause an RNA splicing event in the packaging procedure leading to the deletion of important sequences within the transgene. It is therefore more reasonable to look at the effect of full length potential minigenes for their effect of retrovirus titer than to isolate individual sequences.

Since BGT54, 64, and 50 are deleterious to the retrovirus life-cycle, possible experiments include looking at the integrity of the recombinant viral RNA by RT-PCR and sequencing the products to identify potential RNA processing events. Since recombination events can occur during reverse transcription and integration, analysis of the integrity of the provirus in the target cell by similar methods or Southern blot may also lead to the identification of deleterious sequences. This work is currently being carried out by members of our lab.
3. Conclusions

I have extended the utility of the βLCR in transgenic mice to include reproducible expression of Ay-globin transgenes and have identified β-globin gene regulatory elements that are important for LCR activation at ectopic sites. The approach of testing potential minigenes in single-copy mice before recombinant retrovirus generation has proven to be useful and informative for designing effective minigenes for in vivo expression. I postulate that any gene could be expressed to high levels in erythroid cells by inserting its cDNA or genomic exon/intron sequences (from the ATG site to 3' untranslated sequences) between the NcoI and BamHI sites of BGT50. In this manner, the β-globin intron 2 would function as part of a 3' untranslated region, and the β-globin polyadenylation sites would be used for transcription termination. Such an expression cassette may be extremely useful for directing high-level erythroid expression of non-globin transgenes in animals. Furthermore, the design of cassettes for the expression of α-globin is of therapeutic interest for the cure of the α-chain hemoglobinopathies via gene therapy.

The deleterious sequences within BGT50 are likely specific to retroviruses. Until the deleterious sequences within this transgene are identified and mutated without disturbing gene expression characteristics, viral-mediated transfer of BGT50 and BGT50 derivatives for gene therapy might be best accomplished with alternative viral vectors such as SFV- or Herpes simplex virus-based vectors. In addition, recent advances in direct DNA microinjection into human HSC may allow for efficient gene transfer (Davis et al., 2000). However, creative ways of packaging MMLV-based vectors using SFV biology are currently being developed (Li and Garoff, 1998). Since SFV replicates in the cytoplasm, the generation of retroviral genomes using the SFV
replication cycle eliminates the postulated deleterious RNA processing events in the packaging procedure. In fact, a collaborating lab has generated high titer, stable recombinant retrovirus containing BGT 50 using this packaging scheme and is now looking at expression of BGT50 in vivo after transfer into primitive erythroid cells (manuscript in preparation). Furthermore, lentiviruses and other viruses contain RNA export signals within their genome that can be used to decrease RNA processing in the packaging procedure and, thus, instability problems (Mikaelian et al., 1996; Najera et al., 1999; Rizvi et al., 1997). The limiting factor to further gene therapy studies is stable transfer of these minigenes to primitive HSCs. Once gene transfer is attained, then these modified cells could be used to reconstitute the blood system of a β thalassemic or sickle cell mouse model to look for rescue of the phenotype.

Although several clinical gene therapy trials have been underway for the last ten years, there have been no tales of success. Nonetheless, the promise of gene therapy is still great and many specific gene therapy protocols are being designed for individual diseases based on the nature of the illness. In terms of gene therapy of the β-chain hemoglobinopathies, the problems and needs associated with this type of therapy have been identified and are currently being addressed. There has been progress in gene transfer to human HSCs with better understanding of the biology of these cells and the use of lentivirus-based vectors. The results shown here indicate that better minigenes for in vivo expression can be constructed. Furthermore, the use of insulators may block viral silencing in transduced HSCs and perhaps increase the reproducibility of expression of these minigenes. In view of the advances that have been made in just the last few years, routine treatment of the β-chain hemoglobinopathies via gene therapy has become a much more realistic goal.
V. REFERENCES


78


81


82


