(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 January 2011 (06.01.2011)

PCT

(10) International Publication Number
WO 2011/000516 A2

(51) International Patent Classification:
A61K 35/18 (2006.01)

(21) International Application Number:
PCT/EP2010/003854

(22) International Filing Date:
29 June 2010 (29.06.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0911215.2 29 June 2009 (29.06.2009) GB

(71) Applicant (*for all designated States except US*): **ERASMUS UNIVERSITY MEDICAL CENTER ROTTERDAM** [NL/NL]; P.O. Box 2040, NL-3000 CA Rotterdam (NL).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **GROSVELD, Franklin, Gerardus** [NL/NL]; Erasmus University Medical Center Rotterdam, Dr. Molewaterplein 50, NL-30015 GE Rotterdam (NL). **POURFARZAD, Farzin** [IR/NL]; Fortunastraat 8, NL-3054 PA Rotterdam (NL).

(74) Agents: **MASCHIO, Antonio** et al.; Edwards Angell Palmer & Dodge Innovations LLP, 10 Carlton Crescent, Southampton SO15 2EZ (GB).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

WO 2011/000516 A2

(54) **Title**: GAMMA GLOBIN THERAPY

(57) **Abstract**: The invention relates to a method for identifying inhibitors of γ-globin gene expression suppression, to suppressors of γ-globin gene expression, and to the inhibitors as such.

Gamma Globin Therapy

The invention relates to therapeutic targets for increasing γ-globin gene expression in patients suffering from a blood disorder, such as β-thalassemia or sickle cell anemia.
5 Moreover, the invention provides a method for identifying such targets and identifying inhibitors of such targets which are useful therapeutically, as well as to such inhibitors *per se*.

The human α and β-globin gene loci produce haemoglobin tetramers which together with haem are responsible for the gas transport in the blood. The genes are developmentally
10 regulated with one switch in the embryonic period for the α locus (from ζ to α) and β locus (from ε to γ) and a second switch in the β locus (from γ to β) around the time of birth. Conditions affecting the function of β-globin, known as β-thalassemias and sickle cell disease, are among the most frequent inherited single gene disorders in the human population. The most common treatment for these diseases is blood transfusion and iron
15 chelation therapy or treatment with small molecules such as hydroxyurea or short chain fatty acids, which affect either red cell volume and or lead to an increase of foetal γ globin gene expression. However these treatments are not very satisfactory and do not lead to a normal quality of life nor do they prevent a relatively early death. At present, bone marrow transplantation is the only effective cure, a risky procedure that is not available to the
20 majority of patients. The severity of β-thalassemia and sickle cell disease are greatly ameliorated by expression of the γ-globin genes, which is normally expressed during the foetal stage of development. Since the large majority of patients have normal γ-globin genes, reactivation of γ-globin expression in adults would provide a very elegant and attractive treatment. In particular the treatment with short chain fatty acids aims to
25 achieve such γ globin gene activation, but is only partially successful.

Despite intense research efforts by many laboratories, it is not known how the human γ-globin genes are normally suppressed around the time of birth when expression switches to the adult β-globin gene. Association studies on persons with relatively high γ globin gene expression during adult life have indicated that a number of non globin loci are
30 involved in the suppression of the γ genes (Menzel et al., 2007; Lettre et al 2008; Uda et al 2008; Jiang *et al.*, 2006; and Menzel *et al.*, 2009). The most promising of these, Bcl11A, was recently shown to lead to elevated γ gene expression, when its activity is suppressed (Sankaran et al 2008). Interestingly the promoter region of the γ globin genes

has previously been identified as the region responsible for its suppression (Yu, M., 2006); however, the Bcl11A gene binds to the region downstream of the γ globin genes. It interacts with the Gata1/Fog1/NuRD repressor complex, but whether this interaction is involved in γ globin gene suppression is as yet not clear (Sankaran et al 2008).

5 The importance of the promoter sequences has been well documented by genetic studies in mice (Yu, M., 2006) and in a number of persons with point mutations in the γ-globin promoters leading to a condition known as hereditary persistence of foetal haemoglobin (HPFH), resulting in sustained elevated γ-globin levels through adult life. The strongest HPFH known is caused by a G to A mutation at position –117, resulting in γ-globin
10 expression at therapeutic levels (Stamatoyannopoulos, 2001; Hardison *et al.*, 2002).

A number of treatments have been proposed which seek to restore foetal γ-haemoglobin production in adult patients. For example, Bianchi *et al.* have proposed the use of Angelicin, an isopsoralen derivative, as well as structural derivatives thereof (US 2006/0111433); and Susan Perrine has proposed the use of Butyrate to increase γ-globin
15 gene expression (for instance, see US 2009/0082444, US 6,403,647 and [1993] NEJM 328:81-86). The use of agents which potentiate the hypoxia-inducible factor HIFα has been proposed (WO 2005/011696), whilst gene therapy approaches have included the use of siRNA to contemporaneously downregulate deficient β-globin genes and express γ-globin transgenes (Samakoglu *et al.* [2006] Nature Biotechnology 24:89-94). However,
20 the mechanism of suppression of γ-globin genes after birth, and the nature of candidate suppressors, remains obscure.

In order to gain more insight into the mechanism of suppression, the present inventors have devised a completely novel approach to directly identify the proteins bound to the promoter of the γ globin genes by purifying the suppressed γ globin promoter *in vivo* using
25 a protein tag that binds the promoter inducibly. Since this entails the use of a substantial amount of genetically modified cells that continue to regulate the γ globin genes properly, transgenic mice are used as the model system. The γ to β switch and the HPFH mutations can be mimicked in the mouse, suggesting that it is an appropriate model system.

30

Summary of the Invention

In a first aspect, there is provided a method for identifying an inhibitor of a suppressor of γ-globin gene expression in a mammalian cell, comprising the steps of providing one or more candidate modulators of a suppressor selected from the group consisting of
5 BAF53A, CTNNBL1, ZNF148, Fancl, CHD4, NAP1L1 and NCOA2, administering said one or more candidate modulators of the suppressor to an erythroid cell in which globin gene expression is switched from γ to β; and measuring γ-globin gene expression in said cell.

In a preferred embodiment, the suppressor is ZNF148 or Fancl.

10 The method according to the invention permits the identification, and optionally isolation, of agents which increase production of γ-globin. γ-globin gene expression is downregulated when globin gene expression switches from γ to β (at around birth in humans), as a result of the expression of suppressors of γ-globin gene expression. The method of the invention identifies inhibitors of such suppressors, which are capable of
15 therapeutically increasing γ-globin gene expression. It is believed that suppression of γ-globin gene expression occurs at the transcription stage, and therefore "suppressors" as referred to herein are understood to be suppressors of gene transcription. However, where suppression is occurring at the post-trancriptional stage, the suppressors may be suppressors of mRNA translation, or RNA silencers.

20 In a preferred embodiment, the candidate inhibitors cause an increase in γ-globin gene expression by at least 2g/dl in the blood of the host when administered to a test subject.

Erythroid cells which can be used in the method of the invention may be any erythroid cell in which the γ to β switch has taken place, and γ-globin gene expression is downregulated. For example, the erythroid cell is a human erythroid progenitor (HEP)
25 cell which expresses β-globin at greater levels than γ-globin, preferably at least 1, 2 or 3 orders of magnitude greater. Advantageously, the method may be performed in transgenic non-human animals, for example mammals, such as rodents, typically mice. The animals used may be transgenic.

Exemplary inhibitors include small molecules, polypeptide and nucleic acid inhibiotors.
30 For example, the inhibitor is selected from the group consisting of an RNAi molecule (such as a siRNA of shRNA molecule), a peptide, such as a peptide aptamer, and an antibody or antibody fragment. The inhibitors according to the invention are preferably

suitable for interfering with the formation of protein complexes, which are responsible for the suppression of γ-globin gene expression.

In a second aspect, the invention provides a method for identifying a suppressor or activator of γ-globin gene expression, comprising the steps of:

5 (a) providing a gene construct comprising at least partial sequences of a human γ-globin gene, proximately linked to a binding site for a DNA binding protein;

(b) introducing the gene construct into an erythroid cell;

(c) introducing into said erythroid cell a reagent comprising a DNA binding protein capable of binding to the binding site in (a), and one or more selectable tags;

10 (d) crosslinking the reagent bound to the binding site with cellular factors with which it is in contact; and

(e) isolating the cellular factors crosslinked to the reagent by means of the one or more selectable tags.

The method is suitable for identifying suppressors or activators of γ-globin gene
15 expression. In the case of activators, they may have direct therapeutic potential, or may be useful to develop potentiators of γ-globin gene expression according to the methods described herein.

A suitable binding site is a binding site for the tetracycline repressor TetR; this is coupled with the DNA binding protein TetR as part of the reagent.

20 The selectable tags may be any tags which can be used for identification and/or isolation of the reagent, together with its cross-linked factors, in the cell. Advantageously, the tags comprise a biotinylatable tag and/or an HA tag.

In one embodiment, the reagent is encoded by a second gene construct and expressed in situ in the erythroid cell, such that the first gene construct and the reagent are present
25 contemporaneously in the same cell. In an advantageous embodiment, this may be achieved by producing transgenic lines, for example of transgenic mice, individually expressing the first and second gene constructs. The lines can then be crossed to achieve the desired contemporaneous expression.

In one embodiment, the factor and the reagent are de-crosslinked, to free the factor from the reagent polypeptide. This will assist in the isolation of the factor, and its identification. The factor is a putative suppressor of γ-globin gene expression. This can be confirmed by reducing or eliminating the expression of the factor in an erythroid cell, for example by administering a shRNA or siRNA molecule. The factor may then be used in a method according to the first aspect of the invention, to identify inhibitors of γ-globin gene suppression. Such inhibitors are useful for treating blood disorders, for example disorders which involve insufficient or aberrant production of β-globin. These include sickle cell anaemia and β-thalassemia.

Brief Description of the Figures

Figure 1 is a scheme of the identification of protein binding to the suppressed γ globin promoter by the method of the invention.

Figure 2 shows an S1 nuclease protection analysis of the 7x tetO-γ::p53 transgenic mouse.

Transgenic mice carrying two copies of the modified γ globin minilocus (line 05-23736-05) are bred and RNA prepared from fetal livers and adult blood on the indicated days. The right paneld shows the controls; 13.5dpc-ve, 13.5d liver of non-transgenic mouse RNA; human adult blood RNA; MEL cell RNA; 3x MEL cell RNA; Pac812 is a mouse carrying a complete human wt β globin locus; free probe; Puc, marker fragments. The samples were run on a 6% polyacrylamide urea gel.

Figure 3 is a schematic representation and expression of the 3TetR-G1HRD construct.

The top line shows the TetR fusion protein with CFP, 3xNLS, HA tag, Bio tag and poly A sites. The bottom left panel shows the expression of this construct in transgenic mice (13.5 fetal liver). Bottom middle and right panels Western blots of brain and fetal liver extracts of littermates of the same line, left loading control, right streptavidin detection of the biotinylated protein; wt wild type mouse; TetR mouse not containing BirA; mouse not containing TetR. The molecular weights are indicated next to the MWM lane on the right.

Figure 4 shows the affinity purification of the γ globin promoter from *3TetR-G1HRD::R26-BirA::Tet-γ::p53* cells.

Figure 4A shows a streptavidin bead purification; top left quantitation of the ChIP. The different fragment of the γ globin 7TetO-γ locus are shown in the right bottom panel. Amylase, calreticulin, -3.9GATA and necdin are fragments from the respective genes that serve as negative controls. The ChIP was carried out on cells with or without Doxycyclin. Protein G is the negative control bead. Top right shows expression and purification of TetR protein which are detected by HA antibody or Streptavidin in the bottom panels.

Figure 4B shows an HA ChIP purification using anti HA antibody coupled beads. The different bars represent the quantitaion of the Chlp bound material, the HA peptide eluted material, the material left on the beads after elution and a negative anti V5 antibody bead chip as a negative control. The panel on the bottom right shows a Western blot of the washing and elution of the beads with HA peptide, the top right panel shows the colloidal Blue staining of the gel.

Figure 5 shows the binding of factors to the γ globin promoter in HEP cells.

The different bars show the binding of KIAA1794 (FANCI), ZBP-89 (ZNF148), cdc5L and EuHMTase1 to the γ globin promoter and exon 3 of the gene after a ChIP with antibodies for the respective proteins and normalized against an IgG control ChIP antibody.

Figure 6 shows expression of the γ globin genes in HEP cells after ZNF148/ZBP-89 and FANCI knock down.

Figure 6A left panel shows the knockdown of ZNF148 with five different shRNA viruses. Top shows levels of ZNF148, bottom shows level of APEX protein control. Rights hand panels show γ globin gene expression in two different HEP cell lines (different donors) by S1 nuclease protection. The levels were determined before differentiation for HEP2 and after differentiation for both. NTsh, non targeting sh RNA. Lane HEP3shNT was reloaded later. The identity of the protected RNAs are shown on the left, the marker bands on the right.

Figure 6B left panel shows the knockdown of FANCI over time with a pool of siRNA. Top shows levels of FANCI protein, bottom shows level of APEX protein control. Top panels show γ globin gene expression in HEP2 cell line by S1 nuclease protection. The levels were determined before differentiation for HEP2. NT-days represent the time course control of a nontargeting siRNA. The lanes on the right are MEP2 cells before and after differentiation, human fetal liver, human fetal liver in culture before and after

differentiation, free probe and the marker. Bottom right is the quantitative PCR of the sample above normalized against GAPDH.

Figure 7 shows the TetO modified γ globin minilocus

The drawing on the left shows the γ globin minilocus (Dillon et al., 1991) and the position
5 of the insertion of 7 copies of the tetO operator sequences in the γ globin gene promoter. The right hand panel shows the EcoRi restriction digest of the normal γ globin minilocus (Wt LCR-Aγ-3'HS1) and the modified γ globin locus (LCR-LoxP-7xtetO-Aγ-3'HS1) cosmids. The fragment containg the TetO sequences is indicated The lanes on the left show the Southern blot of several transgenic mice containing two copies of the locus and
10 hybridized with the γ globin minilocus as a probe. The identity of the bands is indicated on the left. The fragment sizes of the marker bands are indicated on the right.

Figure 8 shows the expression of the human γ globin gene in *3TetR-G1HRD::R26-BirA::Tet-γ::p53* cells.

The left shows the expression in the mouse (as in Figure 2), The middle lanes (3 lanes
15 each) show the RNA from *3TetR-G1HRD::R26-BirA::Tet-γ::p53* cells, a cell line derived from mice without BirA and TetR expression and a wt mouse. The lanes on the right are adult mouse blood RNA, 12.5d fetal liver RNA from a mouse expressing the complete wt human locus, same RNA loaded 3x more, RNA from K562 cells, free probe and a marker lane.

20 Figure 9 shows a gel of the boiled beads after two step preparative affinity tag purifications.

The two samples were purified from cells grown with or without doxycyclin and loaded to a 4-12% SDS acrylamide gel. The gel was cut into 15 slices for mass spectrometry. The TetR indicates the 3xTetR fusion protein. It is thought that the second band is probably
25 non biotinylated 3xTetR protein. The MW are shown on the right.

Figure 10 shows chromatin precipitation of ZBP-89 on the BCL11A promoter in HEP cells.

Figure 11 shows knock down of ZBP-89 in human erythroid precursor cells.

Figure 12 shows binding partners of ZBP-89. The list on the left shows a number of
30 complexes that bind to ZBP-89. The panels on the right show standard co-

immunoprecipitation experiments of some of the factors to confirm that these factors are bound to ZBP-89.

Figure 13 shows Lentiviral RNA interference-mediated gene silencing of ZBP-89.

A) Top panel immunoblot for ZBP-89 shRNA knock down in HEP cells. First lane no virus control; NT, Non-target shRNA lentiviral transduced HEP cells. Bottom panel, APEX1 staining of the same blot as loading control. **B)** S1 nuclease protection analysis of globin expression in ZBP-89 knock down HEP cells of two different donors 5 and 6 (donor#5 and 6). First lane no virus control, NT; Non-target shRNA, mZBP-89 sh3; Mouse ZBP-89 short hairpin RNA 3 also depletes human ZBP-89 and induces γ-globin gene expression. Extra controls in the right, 3XHEP 5 and 6; 3xRNA from no virus treated HEP 5 and HEP6 cells, 3xFL; Human fetal liver RNA, PUC18 MspI size marker. The position of human α, γ and β are indicated at the left of the panel. **C)** HPLC hemoglobin chain analysis of the ZBP-89 shRNA knock down in HEP6 (donor#6) cells. **D)** HbF Immunohistochemistry of the ZBP-89 shRNA knock down HEP6 cells.

Figure 14 shows Lentiviral RNA interference-mediated gene silencing of CDC5L.

A) Top panel immunoblot for CDC5L short hairpin shRNA knock down in HEP cells. First lane no virus control; NT, Non-target shRNA lentiviral transduced cells. Bottom panel, APEX1 staining of the same blot as loading control. **B)** S1 nuclease protection analysis of globin expression in CDC5L sh2 knock down HEP cells of two different donors 5 and 6. First lane no virus control, NT; Non-target shRNA. Extra controls in the right, 3XHEP 5 and 6; 3xRNA from no virus treated HEP 5 and HEP6 cells, 3xFL; Human fetal liver RNA, PUC18 MspI size marker. The position of human α, γ and β are indicated at the left of the panel. **C)** HPLC hemoglobin chain analysis of the CDC5L shRNA knock down in HEP5(donor#5) cells. **D)** HbF Immunohistochemistry of the CDC5l shRNA knock down HEP5 cells.

Figure 15 shows Lentiviral RNA interference-mediated gene silencing of APEX1.

A) Top panel immunoblot for APEX1 shRNA knock down in HEP cells. First lane no virus control; NT, Non-target shRNA lentiviral transduced cells. Bottom panel, CDC5L staining of the same blot as loading control. **B)** S1 nuclease protection analysis of globin expression in APEX1 knock down HEP cells. NT; Non-target shRNA. Extra controls in the left, 3XHEP 5 and 6; 3xRNA from no virus treated HEP 5 and HEP6 cells, 3xFL; Human fetal liver RNA, PUC18 MspI size marker. The position of human α, γ and β are indicated

at the left of the panel. **C)** HPLC hemoglobin chain analysis of the APEX1 shRNA knock down in HEP6 (donor#6) cells. **D)** HbF Immunohistochemistry of the APEX1 shRNA knock down HEP6 cells.

Detailed Description of the Invention

Unless otherwise stated, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention. Methods, devices, and materials suitable for such uses are now described. All publications cited herein are incorporated herein by reference in their entirety for the purpose of describing and disclosing the methodologies, reagents, and tools reported in the publications that might be used in connection with the invention.

The practice of the present invention employs, unless otherwise indicated, conventional methods of chemistry, biochemistry, molecular biology, cell biology, genetics, immunology and pharmacology, known to those of skill of the art. Such techniques are explained fully in the literature. See, e. g. , Gennaro, A. R., ed. (1990) Remington's Pharmaceutical Sciences, 18th ed., Mack Publishing Co.; Hardman, J. G., Limbird, L. E., and Gilman, A. G., eds. (2001) The Pharmacological Basis of Therapeutics, 10th ed., McGraw-Hill Co.; Colowick, S. et al., eds., Methods In Enzymology, Academic Press, Inc.; Weir, D. M. , and Blackwell, C. C., eds. (1986) Handbook of Experimental Immunology, Vols. I-IV, Blackwell Scientific Publications; Maniatis, T. et al., eds. (1989) Molecular Cloning: A Laboratory Manual, 2nd edition, Vols. I-III, Cold Spring Harbor Laboratory Press; Ausubel, F. M. et al., eds. (1999) Short Protocols in Molecular Biology, 4th edition, John Wiley & Sons; Ream et al., eds. (1998) Molecular Biology Techniques: An Intensive Laboratory Course, Academic Press; Newton, C. R., and Graham, A., eds. (1997) PCR (Introduction to Biotechniques Series), 2nd ed., Springer Verlag.

As used herein, an "inhibitor" is a compound or combination of compounds which decreases the effect or quantity of that which is inhibited. Specifically, and inhibitor of a γ-globin gene suppressor increases γ-globin gene expression by reducing the available amount of the suppressor or the effect of the suppressor on the γ-globin gene.

A "suppressor" of γ-globin gene expression is a cellular factor, or combination of cellular factors, which decreases the levels of γ-globin in the blood of a mammal. Typically, such suppressors are responsible for the switch from γ-globin to β-globin expression in humans at or around birth.

5 "Modulators" may be inhibitors or potentiators. Typically, in the context of the present invention, a modulator is an inhibitor.

In the context of the present invention, administration is performed by standard techniques of cell culture, depending on the reagent, compound or gene construct to be administered. For instance, administration may take place by addition to a cell culture 10 medium, introduction into cells by precipitation with calcium phosphate, by electroporation, or by other means. If the method of the invention employs a non-human mammal as the test system, the mammal may be transgenic and express the necessary reagents in its endogenous cells.

The term "operably linked" refers to a juxtaposition wherein the components described 15 are in a relationship permitting them to function in their intended manner. A control sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences.

A “candidate”, as in “candidate modulator”, is a compound which is believed to have a 20 certain function or activity and which is subjected or can be subjected to an assay, for example as described herein, to identify the presence or absence of that function or activity.

An “erythroid” cell is a cell of the erythroid lineage which expresses globin genes and, preferably, makes haemoglobin. Preferably, it is a human cell. It is not an enucleated 25 red blood cell, which does not produce haemoglobin. In assays according to the invention, it is advantageous to use cells in which the expression of γ-globin has been reduced, for example as a result of a switch to the production of β-globin as occurs in humans at around birth.

“Crosslinking” is a technique well known in the art for attaching proteins to each other in 30 vivo, so that they may be isolated together. Common crosslinking agents include the imidoester crosslinker dimethyl suberimidate, the NHS-ester crosslinker BS3 and formaldehyde. Each of these crosslinkers induces nucleophilic attack of the amino group

of lysine and subsequent covalent bonding via the crosslinker. The zero-length carbodiimide crosslinker EDC functions by converting carboxyls into amine-reactive isourea intermediates that bind to lysine residues or other available primary amines. SMCC or its water soluble analog, Sulfo-SMCC, are commonly used to prepare antibody-hapten conjugates for antibody development. Formaldehyde is the preferred crosslinking reagent to crosslink proteins to DNA.

Suppressors of γ-globin gene expression

The invention relates to suppressors of γ-globin gene expression, as well as to methods for identifying such suppressors. Suppressors may be identified in a two step procedure; the first is to identify compounds which bind to the γ-globin locus in a cell in which the switch from γ-globin to β-globin expression has taken place. The second is to ascertain the activity of the suppressor on expression of γ-globin genes in cells or animals.

Candidate suppressors identified herein include BAF53A, CTNNBL1, ZNF148, Fancl, CHD4, NAP1L1 and NCOA2. The structures of these molecules are known in the art. For example, BAF53A is defined in the pubmed database as BAF53A [Homo sapiens] gi|49065484|emb|CAG38560.1|[49065484]. The structure of the molecule can be found under the relevant GI number. Other candidate suppressors include CTNNBL1 (catenin, beta like 1) [Homo sapiens] gi|123235142 |emb|CAM27645.1|[123235142]; ZNF148 protein [Homo sapiens] gi|27370628|gb|AAH35591.1|[27370628]; FANCI protein [Homo sapiens] gi|187954505|gb|AAI40770.1|[187954505]; CHD4 protein [Homo sapiens] gi|116283563|gb|AAH25962.1|[116283563]; NAP1L1 RecName: Full=Nucleosome assembly protein 1-like 1; AltName: Full=NAP-1-related protein; Short=hNRP gi|1709337|sp|P55209.1|NP1L1_HUMAN[1709337]; and NCOA2 protein [Homo sapiens] gi|62201602|gb|AAH92442.1|[62201602].

These and other suppressors are identifiable by the methods described herein, and are candidate targets for drug development, in the generation of inhibitors of γ-globin gene suppression.

Transgenic animals

The methods of the invention are advantageously carried out in a model system which properly displays or mimics the γ-globin-β-globin switch which occurs in humans at around birth. Transgenic non-human mammals may be produced using techniques

known in the art, and are good model systems for use in the present invention. Embryo micromanipulation techniques permit introduction of heterologous DNA into, for example, fertilised mammalian ova. For instance, pluripotent stem cells can be transformed by microinjection, calcium phosphate mediated precipitation, liposome fusion, retroviral
5 infection or other means, the transformed cells are then introduced into the embryo, and the embryo then develops into a transgenic animal. In one method, developing embryos are infected with a retrovirus containing the desired DNA, and transgenic animals produced from the infected embryo. In a preferred method, however, the appropriate DNAs are coinjected into the pronucleus or cytoplasm of embryos, preferably at the
10 single cell stage, and the embryos allowed to develop into mature transgenic animals after transfer into pseudopregnant recipients. Those techniques are well known (see reviews of standard laboratory procedures for microinjection of heterologous DNAs into mammalian fertilised ova, including Hogan et al., Manipulating the Mouse Embryo, (Cold Spring Harbor Press 1986); Krimpenfort et al., Bio/Technology 9:844 (1991); Palmiter et
15 al., Cell, 41:343 (1985); Kraemer et al., Genetic manipulation of the Mammalian Embryo, (Cold Spring Harbor Laboratory Press 1985); Hammer et al., Nature, 315:680 (1985); Wagner et al., U.S. Pat. No. 5,175,385; Krimpenfort et al., U.S. Pat. No. 5,175,384, the respective contents of which are incorporated herein by reference).

Transgenic animals may also be produced by embryonic stem (ES) cell technology as
20 described for example in US 5,464,764. Using such methods the transgenic construct is stably introduced into the genome of ES cells, which are introduced into tetraploid embryos or blastocysts to generate transgenic offspring. Transgenic animals may also be produced by nuclear transfer technology as described in Schnieke, A. E. et al., 1997, Science, 278: 2130 and Cibelli, L B. et al., 1998, Science, 280: 1256. Using this method,
25 fibroblasts from donor animals are stably transfected with a plasmid incorporating the coding sequences for a polypeptide of interest under the control of regulatory sequences. Stable transfectants are then fused to enucleated oocytes, cultured and transferred into female recipients.

Analysis of animals which may contain transgenic sequences would typically be
30 performed by either PCR or Southern blot analysis following standard methods.

By way of a specific example for the construction of transgenic non-human mammals, nucleotide constructs comprising a sequence encoding a DNA binding molecule are microinjected using, for example, the technique described in U.S. Pat. No. 4,873,191, into oocytes which are obtained from ovaries freshly removed from the mammal. The oocytes

are aspirated from the follicles and allowed to settle before fertilisation with thawed frozen sperm capacitated with heparin and prefractionated by Percoll gradient to isolate the motile fraction.

The fertilised oocytes are centrifuged, for example, for eight minutes at 15,000 g to
5 visualise the pronuclei for injection and then cultured from the zygote to morula or blastocyst stage in oviduct tissue-conditioned medium. This medium is prepared by using luminal tissues scraped from oviducts and diluted in culture medium. The zygotes must be placed in the culture medium within two hours following microinjection.
Oestrous is then synchronized in the intended recipient mammals, such as cattle, by
10 administering coprostanol. Oestrous is produced within two days and the embryos are transferred to the recipients 5-7 days after oestrous. Successful transfer can be evaluated in the offspring by Southern blot.

Alternatively, the desired constructs can be introduced into embryonic stem cells (ES cells) and the cells cultured to ensure modification by the transgene. The modified cells
15 are then injected into the blastula embryonic stage and the blastulas replaced into pseudopregnant hosts. The resulting offspring are chimeric with respect to the ES and host cells, and nonchimeric strains which exclusively comprise the ES progeny can be obtained using conventional cross-breeding. This technique is described, for example, in WO91/10741.

20 Alternative methods include the use of viral vectors, such as retroviral vectors, adenoviral vectors and herpesviral vectors. Such techniques have moreover been described in the art, for example by Zhang et al. (Nucl. Ac. Res., 1998, 26:3687-3693).
Suitable viral vectors may be retroviral vectors, and may be derived from or may be derivable from any suitable retrovirus. A large number of different retroviruses have been
25 identified. Examples include: murine leukaemia virus (MLV), human immunodeficiency virus (HIV), simian immunodeficiency virus, human T-cell leukaemia virus (HTLV), Equine infectious anaemia virus (EIAV), mouse mammary tumour virus (MMTV), Rous sarcoma virus (RSV), Fujinami sarcoma virus (FuSV), Moloney murine leukaemia virus (Mo-MLV), FBR murine osteosarcoma virus (FBR MSV), Moloney murine sarcoma virus (Mo-MSV),
30 Abelson murine leukaemia virus (A-MLV), Avian myelocytomatosis virus-29 (MC29), and Avian erythroblastosis virus (AEV). A detailed list of retroviruses may be found in Coffin et al., 1997, "retroviruses", Cold Spring Harbour Laboratory Press Eds: J M Coffin, S M Hughes, H E Varmus pp 758-763.

Accordingly, transgenic animals useful in the invention can be produced by methods including but not limited to 1) microinjection of fertilized eggs; 2) transfection or infection of ES cells followed by injection of ES cells into blastocysts, resulting in chimeric offspring, or fusion of ES cells with tetraploid embryos resulting in totally ES derived offspring; and 3) cloning by nuclear transfer. Homologous recombination may occur when using ES cells or when using the method of nuclear transfer.

Gene constructs

Gene constructs are well known in the art, and methods for constructing them are known from standard references such as those cited above, and others well known to those skilled in the art. The construct according to the invention comprises, in essence, a DNA binding protein binding site, one or more isolatable tags and sequences from the γ-globin gene. Advantageously, the γ-globin sequences are the regulatory control sequences located 5' of the transcriptional start site, i.e. the promoter and upstream regions.

The DNA binding domain is preferably not transcriptionally activating, that is does not cause any change in the level of transcription of the γ-globin gene. Such DNA binding domains are well known and include the DNA binding portions of the proteins ACE1 (CUP1), lambda cl, lac repressor, jun, fos, GCN4, or the Tet repressor TetR. The GAL4 DNA binding domain may also be used.

Isolatable tags include, for example HA tags, V5 tags and biotinylatable tags. Biotinylatable tags are known in the art (see, for instance, Schatz PJ. Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in Escherichia coli. *Bio/technology (Nature Publishing Company).* 1993;**11**:1138–1143; and Beckett D, Kovaleva E, Schatz PJ. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci.* 1999;**8**:921–929). HA and V5 tags are also commonly available.

Drug Development Targets

According to the present invention, suppresors of γ-globin gene expression used as a target to identify compounds, for example lead compounds for pharmaceuticals, which are capable of upregulating γ-globin gene expression by reducing or removing the suppression caused by the suppressors. Accordingly, the invention relates to an assay and provides a method for identifying a compound or compounds capable, directly or

indirectly, of modulating the activity a suppressor of γ-globin gene expression, comprising the steps of:

(a) incubating a suppressor of γ-globin gene expression with the compound or compounds to be assessed; and

5 (b) identifying those compounds which influence the activity of the suppressor.

Suppressor Binding Compounds

According to a first embodiment of this aspect invention, the assay can be configured to detect polypeptides which bind directly to the suppressor molecule.

The invention therefore provides a method for identifying a modulator γ-globin gene
10 expression, comprising the steps of:

(a) incubating a suppressor with the compound or compounds to be assessed; and
(b) identifying those compounds which bind to the suppressor.

For example, a reported construct may be used, in which a detectable reporter gene is driven by γ-globin promoter sequences. Exposure of this reporter construct to
15 suppressors(s) of γ-globin gene expression, and potential inhibitors of such suppressors, will allow the effects of these agents to be monitored via the reporter gene.

Preferably, the method further comprises the step of:

(c) assessing the compounds which bind to the suppressor for the activity to modulate γ-globin gene expression in a cell-based assay.

20 Binding to the suppressor may be assessed by any technique known to those skilled in the art. Examples of suitable assays include the two hybrid assay system, which measures interactions in vivo, affinity chromatography assays, for example involving binding to polypeptides immobilised on a column, fluorescence assays in which binding of the compound(s) and the suppressor is associated with a change in fluorescence of
25 one or both partners in a binding pair, and the like. Preferred are assays performed in vivo in cells, such as the two-hybrid assay.

In a preferred aspect of this embodiment, the invention provides a method for identifying a lead compound for a pharmaceutical useful in the treatment of disease involving aberrant globin gene expression, comprising incubating a compound or compounds to be

tested with an a suppressor of γ-globin gene expression, under conditions in which, but for the presence of the compound or compounds to be tested, the suppressor associates with a γ-globin gene or gene construct with a reference affinity;

determining the binding affinity of the suppressor in the presence of the compound or
5 compounds to be tested; and

selecting those compounds which modulate the binding affinity of the suppressor with respect to the reference binding affinity.

Preferably, therefore, the assay according to the invention is calibrated in absence of the
10 compound or compounds to be tested, or in the presence of a reference compound whose activity in modulating the binding of the suppressor to a globin gene locus is known or is otherwise desirable as a reference value. For example, a reference value may be obtained in the absence of any compound. Addition of a compound or compounds which increase the binding affinity of the suppressor for the γ-globin locus
15 increases the readout from the assay above the reference level, whilst addition of a compound or compounds which decrease this affinity results in a decrease of the assay readout below the reference level.

Modulators of Suppressors

In a still further aspect, the invention relates to a compound or compounds identifiable by
20 an assay method as defined in the previous aspect of the invention.

Compounds which influence the activity of γ-globin suppressors may be of almost any general description, including low molecular weight compounds, including organic compounds which may be linear, cyclic, polycyclic or a combination thereof, peptides, polypeptides including antibodies, or proteins. In general, as used herein, "peptides",
25 "polypeptides" and "proteins" are considered equivalent.

Antibodies

Antibodies, as used herein, refers to complete antibodies or antibody fragments capable of binding to a selected target, and including Fv, ScFv, Fab' and $F(ab')_2$, monoclonal and polyclonal antibodies, engineered antibodies including chimeric, CDR-grafted and
30 humanised antibodies, and artificially selected antibodies produced using phage display or alternative techniques. Small fragments, such Fv, VH domains and ScFv, possess

advantageous properties for diagnostic and therapeutic applications on account of their small size and consequent superior tissue distribution. Methods for making antibodies and antibody fragments, as well as humanised and human versions thereof, are known in the art. Since the compounds of the invention are required to act intracellularly, intracellular antibodies are preferred. Such antibodies are described, for example, in WO0054057, WO03077945, US2005272107, US2005276800, and WO03014960.

RNAi

RNA interference can be used to downregulate expression of γ-globin suppressors. RNA interference refers to the process of sequence-specific post-transcriptional gene silencing in animals mediated by short interfering RNAs (siRNAs) (Fire et al., 1998, Nature, 391, 806, Hamilton et al., 1999, Science, 286, 950-951). The process of post-transcriptional gene silencing is thought to be an evolutionarily-conserved cellular defence mechanism used to prevent the expression of foreign genes and is commonly shared by diverse flora and phyla (Fire et al., 1999, Trends Genet., 15, 358). The presence of long dsRNAs in cells stimulates the activity of a ribonuclease III enzyme referred to as dicer. Dicer is involved in the processing of the dsRNA into short pieces of dsRNA known as short interfering RNAs (siRNAs) (Hamilton et al., supra; Berstein et al., 2001, Nature, 409, 363). Short interfering RNAs derived from dicer activity are typically about 21 to about 23 nucleotides in length and comprise about 19 base pair duplexes (Hamilton et al., supra; Elbashir et al., 2001, Genes Dev., 15, 188).

The RNAi response also features an endonuclease complex, commonly referred to as an RNA-induced silencing complex (RISC), which mediates cleavage of single-stranded RNA having sequence complementary to the antisense strand of the siRNA duplex. Cleavage of the target RNA takes place in the middle of the region complementary to the antisense strand of the siRNA duplex (Elbashir et al., 2001, Genes Dev., 15, 188).

Zarbl, 1999, Molecular and Cellular Biology, 19, 274-283 and Wianny and Goetz, 1999, Nature Cell Biol., 2, 70, first described RNAi mediated by dsRNA in mammalian systems. Others have reported on various RNAi and gene-silencing systems. For instance, McCaffrey et al., 2002, Nature, 418, 38-39, describe the use of certain siRNA constructs targeting a chimeric HCV NS5B protein/luciferase transcript in mice. Randall et al., 2003, PNAS USA, 100, 235-240, describe certain siRNA constructs targeting HCV RNA in Huh7 hepatoma cell lines.

Procedures for designing siRNA molecules specific for any desired sequence are known in the art, for instance from US2009137500.

Peptides

Peptides, such as peptide aptamers, can be selected from peptide libraries by screening
5 procedures. In practice, any vector system suitable for expressing short nucleic acid sequences in a eukaryotic cell can be used to express libraries of peptides. In a preferred embodiment, high-titer retroviral packaging systems can be used to produce peptide aptamer libraries. Various vectors, as well as amphotropic and ecotropic packaging cell lines, exist that can be used for production of high titers of retroviruses
10 that infect mouse or human cells. These delivery and expression systems can be readily adapted for efficient infection of any mammalian cell type, and can be used to infect tens of millions of cells in one experiment. Aptamer libraries comprising nucleic acid sequences encoding random combinations of a small number of amino acid residues, e.g., 5, 6, 7, 8, 9, 10 or more, but preferably less than 100, more preferably less than 50,
15 and most preferably less than 20, can be expressed in retrovirally infected cells as free entities, or depending on the target of a given screen, as fusions to a heterologous protein, such as a protein that can act as a specific protein scaffold (for promoting, e.g., expressibility, intracellular or intracellular localization, stability, secretability, isolatablitiy, or detectability of the peptide aptamer. Libraries of random peptide aptamers when
20 composed of, for example 7 amino acids, encode molecules large enough to represent significant and specific structural information, and with 10^7 or more possible combinations is within a range of cell numbers that can be tested.

Preferably, the aptamers are generated using sequence information from the target suppressor.

25 In identifying an aptamer, for example, a population of cells is infected with a gene construct expressing members of an aptamer library, and the effect of the aptamers on a globin gene reporter construct are assessed. Coding sequences of aptamers selected in the first round of screening can be amplified by PCR, re-cloned, and re-introduced into naïve cells. Selection using the same or a different reporter system can then be repeated
30 in order to validate individual aptamers within the original pool. Aptamer coding sequences within cells identified in subsequent rounds of selection can be iteratively amplified and subcloned and the sequences of active aptamers can then be determined by DNA sequencing using standard techniques.

Examples

Materials and methods

DNA constructs

5 A human γ globin gene ClaI/KpnI fragment from the γ globin minilocus (Dillon and Grosveld, 1991) was modified as a subclone by insertion of a legated heptamer of the tetO oligos with the sequence TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAG into the StuI site of the γ globin gene promoter 380 basepairs upstream of the transcription initiation site (Figure 1). The modified ClaI/KpnI fragment was cloned back into the
10 minilocus by standard λ phage packaging (Stratagene) and transduction into E.Coli DH10B.. DNA was isolated and the integrity of the modified minilocus established by cleavage with EcoRI and gel electrophoresis (Figure 7).

The 642 bp tetR binding domain (Gossen and Bujard, 1992) was cloned in frame onto
15 CFP cDNA (pECFP-N1, Clontech). The insert was recloned into pEYFP-Nuc replacing eYFP (Clontech) thereby gaining the 3x NLS and SV40 polyA sequence. A biotinylation tag (de Boer et al., 2003) and HA tag (YPYDVPDYA) were PCR cloned in frame by overlapping oligonucleotides and restriction sites in between the NLS and poly A resulting in 3TetR cDNA. The 3TetR fragment was excised by Eco RI/NotI and inserted into the
20 NotI site of the G1HRD expression vector (Ohneda et al., 2002). The resulting 3TetR:G1HRD vector was tested for integrity (not shown) and standard transfection and expression in MEL cells which demonstated its localization to the nucleus by CFP fluorescence (not shown).

Transgenesis and cell line derivation

25 The minilocus DNA was subsequently cut with SalI, the insert fragment isolated by gel electrophoresis and injected into fertilized oocytes (Dillon and Grosveld, 1991) of p53-/- mice (Donehower et al., 1992).A number of γ minilocus transgenic mice was obtained, one of which contained only two copies of the minilocus (line 05-19255-00). Its integrity was was established by EcoRi digestion and Southern blots of genomic DNA (Figure 7).
30 Similarly The Asp718/PvuI 3TetR:G1HRD restriction fragment was isolated from the vector by gel electrophoresis and injected into fertilized eggs obtained from BirA

expressing mice (Driegen et al., 2005). The resulting mice, *G1HRD::R26-BirA*, expressed the 3TetR protein in the erythroid lineage (Figure 3). Finally the two mouse strains were bred together to yield the *3TetR-G1HRD::R26-BirA::Tet-γ:p53* mice. These mice were bred together and 13.5d fetal liver cells were cultured (Dolznig et al., 2001) to obtain the
5 *3TetR-G1HRD::R26-BirA::Tet-γ:p53* cell line. The transgenic mouse breeding also yielded control mice and subsequently cell lines *3TetR::Tet-γ:p53* or *G1HRD::R26-BirA::Tet-γ:p53.*

Culturing of human proerythroblasts from blood (HEPs)

Culturing of undifferentiated cells

10 The culturing was essentially as published (Leberbauer et al., 2005). 40ml blood was sampled and buffy coat isolated by centrifugation. White cell were removed after Ficoll gradient purification from the inter phase and washed. For initial expansion, $5x10^6$ cells/ml were cultivated in serum-free medium (StemSpan; Stem Cell Technologies, Vancouver, BC, Canada) supplemented with Epo (2 U/mL Erypo Janssen- Cilag, Baar,
15 Switzerland), the synthetic glucocorticoid Dex (1 μM; Sigma, St Louis, MO) SCF (100 ng/mL; R&D Systems, Minneapolis, MN), and lipids (40 μg/mL cholesterol-rich lipid mix; Sigma) for about 5 days, the cells were fed every day by partial medium changes with the same medium. After 5 days the big proerythroblasts became visible and by day 7 they start to overgrow the culture. In order to purify proerythroblasts from lymphocytes and
20 macrophages a percoll purification was carried out and the cells isolated from the interphase (followed by a wash). Homogeneous cultures of erythroid progenitors were kept in the same medium at $1.5\text{-}2x10^6$ cells/ml by daily partial medium changes. Proliferation kinetics and size distribution of the cells were monitored daily using an electronic cell counter (CASY-1; Schärfe-System, Reutlingen, Germany).

25 *Induction of differentiation*

To induce terminal differentiation, proliferating erythroblasts were washed in ice-cold PBS and reseeded at $1\text{-}1.5x10^6$ cells/ml in StemSpan supplemented with Epo (0.5mg/ml human transferring Holo (1 mg/mL; SCIPAC Ltd, UK). Differentiating erythroblasts were maintained at $2\text{–}3x10^6$ cells/ml by daily cell counts and partial medium changes.

γ globin promoter purification using chromatin pulldown

Part A. Cross linking and Chromatin preparation.

The *3TetR-G1HRD::R26-BirA::Tet-γ:p53 I/11*cells are grown in standard medium (Dolznig et al., 2001) and harvested at the density of 2-3x10^6 cells/ml. The chromatin is
5 crosslinked by adding formaldehyde directly to culture medium to a final concentration of 1%. It is incubated for 10 minutes at room temperature. The reaction is stopped by the addition of 1M glycin to final concentration 0.125M. The cells are spun at 1500 RPM for 3min at 4°C (Eppendorf centrifuge 5810R), the supernatant is decanted and the cells are washed in icecold PBS. The spin and wash are repeated twice. The cells are
10 resuspended at a density of 2x10^{-8} cells/ml sonication buffer (10 mM Tris-HCl, 1 mM EDTA , 0.5 EGTA, PH 8.0) and 1x complete protease inhibitor (Roche cat# 11873 580 001). Incubate for 10min on ice. The cells are disrupted and the DNA fragmented by sonication at 0^0C (Sanyo, SoniPrep150 large probe, 14 times, amplitude 7; 15" on; 45"off). The sample is centrifuged for 10 minutes at 13,000 rpm at 4°C, and the
15 supernatant aliquoted in 50ul per microcentrifuge tube (i.e 10x10^6 cells per tube). The pellet are discarded. The samples are snap frozen in liquid nitrogen and stored at -80°C.

(***Note:*** The sonication conditions may have to be optimized for different machines). The efficiency of the sonication is analysed by adding 8 µL 5 M NaCl and the crosslinks reversed at 65°C for 4 hours. The DNA is recovered by phenol/chloroform extraction and
20 run on an agarose gel to visualize the shearing efficiency).

Part B. Pulldown.

Note: For a negative/background control, prepare a sample to use as a no-antibody immunoprecipitation control in step 5 below. Additionally unbound DNA samples should be prepared as controls for the PCR analysisdescribed below.

25 Note: When proceeding to the PCR a portion of the diluted cell supernatant 2% (~20 µL) is kept to quantify the amount of DNA present in different samples at the PCR protocol. This sample is considered to be the input/starting material and needs to have the cross-links reversed.

75ul of the immunoprecipitating monoclonal Anti-HA-Agarose beads (Sigma Cat#
30 A2095) are washed per sample with 1 ml PBS. The agarose is pelleted by brief centrifugation for 1min at 2000 RPM and the supernatant is discarded. This wash step is

repeated once more. The beads are equilibrated with 1ml ChIP dilution buffer (0.01% SDS, 1.1% Triton X100, 1.2mM EDTA, 16.7mM tris-HCl, pH 8.1, 150mM NaCl) and the agarose pelleted by brief centrifugation for 1min at 2000 RPM. The supernatant is discarded. The equilibration step is repeated two more times.

5 The beads are blocked for 1 hour with 1% Fish skin gelatin (Sigma G7765), and 0.2mg/ml Chicken egg albumin and 40ul sonicated salmon sperm DNA (10mg/ml). The agarose is pelleted by brief centrifugation (1 min) at 2000 RPM the supernatant discarded.

Concurrent with preparing the beads a number of frozen tubes are thawed on ice. Each sample is taken up in 1 ml ChIP Dilution Buffer adding 1X complete protease inhibitors
10 as above. The chromatin is precleared with 75 µL of Protein A Agarose/Salmon Sperm DNA (50% Slurry), (Upstate # 16-157C) for 30 minutes at 4°C with rotation, to reduce the nonspecific background. The agarose is pelleted by brief centrifugation and the supernatant collected. Add the precleared supernatant to the blocked beads and incubate at 4 °C overnight with rotation.

15 For a negative control, a no-antibody immunoprecipitation was performed by incubating the supernatant fraction with same amount of Protein A Agarose/Salmon Sperm DNA washed, equilibrated and blocked in the same way. A non-expressing TetR-HA-biotag fusion protein cell lysate or Doxycycline treated cell lysate was also used as a negative control treated the same as the experimental sample.

20 The agarose was pelleted by centrifugation (1min at 2000 RPM) and the supernatant carefully removed (it contains the unbound, non-specific DNA). The protein A agarose/antibody/ *TetR-HA-biotag* /DNA complex is washed for 3-5 minutes on a rotating platform with 1 mL of each of the buffers listed in the following order, all at 4^{0}C; a Low Salt Immune Complex Wash Buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM
25 Tris-HCl, pH 8.1, 150 mM NaCl), one wash;a High Salt Immune Complex Wash Buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8.1,

500 mM NaCl), one wash; a LiCl Immune Complex Wash Buffer (0.25 M LiCl, 1% NP-40, 1% deoxycholic acid (sodium salt), 1 mM EDTA, 10 mM Tris, pH 8.1), one wash; TE Buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0), two washes.

30 The sample is at this stage a protein A/antibody/histone/DNA complex ready for either an Immunoprecipitation/Immunoblot assay (Section I) or Polymerase Chain Reaction (PCR) assay (Section II) or further purification.

Immunoprecipitation/Immunoblot protocol to detect histone.

Following washing of the beads the immunoprecipitated complex can be analyzed by immunoblot analysis. 25 µL of 1X Laemmli buffer is added per sample and boiled for 10 minutes. 20 µL is loaded per lane and the immunoblot procedure carried out as described per appropriate antibody.

PCR protocol to amplify DNA bound to the beads

500µl of freshly prepared elution buffer (1%SDS, 0.1M NaHCO3) is added per sample of the pelleted protein A agarose/antibody/ *TetR-HA-biotag/DNA* complexed beads. The sample is vortexed briefly to mix and then incubated at room temperature for 15 minutes with rotation. 20 µL 5 M NaCl is added to the elution (500 µL) and the protein-DNA crosslinks are reversed by heating at 65°C for more than 4 hours or over night. The sample can be stored overnight at 20°C.

Note: Include the input/starting material (*20* µL of the -80°C frozen sample plus 480 µl of elution buffer and 20 µl 5 M NaCl heating to 65 °C for 4 hours)

10 µL of 0.5 M EDTA 20 µL 1 M Tris-HCl, pH 6.5 is added plus 2 µl of 10 mg/mL Proteinase K to the eluates and incubated for one hour at 45°C. The DNA is cleaned by two phenol/chloroform and one chloroform extraction. 1ul of 20 µg glycogen is added followed by 6 volumes of isopropanol. After 30min at -20°C the sample is centrifuged and the pellets washed with 70% ethanol and air dried. They are resuspended into water, the input into 400 µl and pull down sample into 200µl H2O for PCR. 4µl is used per PCR reaction.

HA elution and further purification by streptavidin pulldown

After binding to the HA beads and washing the sample was eluted in low salt wash buffer with 0.5mg/ml of HA peptide (SciLight Biotechnology, LLC). The eluate was diluted 10x with ChIP dilution buffer plus protease inhibitor and bound to streptavidin paramagnetic beads (Dynabeads, M280, In Vitrogen) overnight at 4°C on a rotating wheel. Beads are separated by magnet, washed with 4 buffers (low salt washing buffer, high salt washing buffer, LiCl washing buffer and 2x TE buffer).and boiled in 1X Laemmli sample buffer. The sample is run on a NuPAGE 4-12% gradient SDS acrylamide gel (InVitrogen) as shown in Figure 9.

Mass spectrometry

The mass spectrometry was carried out after gel electrophoresis (Fig. 9) Each gel was cut into 15 slices and analysed by LC-MS/MS in an orbiTrap. The peptides were identified using the Mascot search engine.

5 **RNAi and S1 nuclease protection**

shRNA vectors were obtained from the TRC Mission human and mouse library from Sigma as bacterial glycerol stocks. Packaging and transfections in HEPs were as described using puromycin selection , as set forth on the biomics page on the Erasmus MC website. For FANCI siRNA was purchased (Darmacon, ON-TARGET plus SMART
10 pools) and transfected as described (POLYplus, INTERFERin™. S1 nuclease was as described (Hanscombe et al., 1991).

Example 1

Experimental design

15 We opted for an unbiased approach by introducing Tet operator (TetO) sequences in the upstream γ-globin promoter. Upon adition of doxycyclin, these sequences are bound by a tagged TetR protein, that is used as a hook to fish out the γ-globin promoter region by affinity purification. Three basic parameters are important in such an approach: 1. The hook and its binding site in the promoter should not disturb the normal expression pattern
20 of the γ-globin genes. 2. Since the γ-globin promoter fragment of interest constitutes less than 1 millionth of the genome, a large number of cells is needed to enable the purification of sufficient amounts of material. 3. The control should be comparable to the experimental sample to allow the elimination of contaminants.

We used mouse transgenesis to generate a model system meeting these requirements.
25 The basic design of the approach is schematically shown in Fig. 1. Binding sites for the bacterial tetracycline repressor protein (TetR) were inserted into a human β-globin minilocus (Grosveld et al., 1987) that contained the locus control region (LCR), the γA gene and the 3' hypersensitive site 1 (3'HS1). Previous experiments have shown that such a γ-globin minilocus is regulated properly in the mouse (Dillon and Grosveld, 1991).
30 The γ-globin gene is expressed in the embryo and early mouse fetal liver, and is silenced around day 14-16 in the fetal liver and in the adult (Peterson et al., 1998)(Strouboulis et

al., 1992). Furthermore, the γ-globin gene remains active at the adult stage, when the –117 G to A HPFH mutation is introduced in the promoter (Berry et al., 1992). The modified minilocus (*Tet-γ*) was introduced into transgenic mice that were bred in a p53 null background to facilitate the derivation of immortalized cell lines (Donehower et al.,
5 1992) (Dolznig et al., 2001). We introduced a triple tag in the TetR protein (3TetR), that enables sequential purification on HA and streptavidin affinity beads. The 3TetR cDNA was cloned in the hematopoietic expression vector pIE3.9IntpolyAA (Ohneda et al., 2002) and this construct (3TetR-G1HRD) was introduced in transgenic mice. The *3TetR-G1HRD* transgenewas bred into mice expressing the *E. coli* BirA biotin ligase from the
10 *Rosa26* locus (Driegen et al., 2005). BirA will efficiently biotinylate theBio tag of the 3TetR protein.. Crossing of the *3TetR-G1HRD::R26-BirA* mice with the *Tet-γ::p53* mice resulted in quadruple *3TetR-G1HRD::R26-BirA::Tet-γ::p53* mice that were used for the generation of fetal-liver derived erythroid progenitor cell lines.

Transgenic mice

15 First the Tet binding sites were tested in electrophoretic mobility shift assays with *3TetR-G1HRD::R26-BirA* foetal liver cell extracts. This showed strong binding of 3TetR to the Tet operator sequences, and some weak but specific Gata1 binding (not shown). We found no evidence for binding of other proteins. Next, seven copies of the sequence were introduced into the StuI site 380 bp upstream of the transcription initiation site of the γ-
20 globin gene in the γ-globin minilocus (Fig. 7). This *Tet-γ* locus was used for transgenesis, and mice *Tet-γ* transgene were selected for further analysis (Fig. 7). The developmental expression pattern of γ-globin from the *Tet-γ* locus was analysed by quantitative S1 nuclease protection. The results obtained with line 05-19255-00, which contains two copies of the transgene, are shown in Fig. 2. From these data we conclude that the *Tet-γ*
25 locus is expressed properly. It is active during the embryonic period, suppressed during the late fetal liver stage and remains inactive in the adult.

In parallel, the 3TetR cDNA was cloned in the pIE3.9IntpolyAA vector (Ohneda et al., 2002) resulting in the 3TetR-G1HRD construct. This vector recapitulates the hematopoietic expression pattern of the mouse *Gata1* gene, driving expression at all
30 developmental stages of erythroid cells, i.e also when the γ-globin gene is suppressed. The 3TetR protein is composed of the TetR DNA binding domain coupled to CFP, three copies of a nuclear localization signal, an HA tag and a biotinylation tag (Fig. 3). The 3TetR-G1HRD construct was introduced into transgenic mice expressing the BirA biotin ligase (Driegen et al., 2005). Several mice were obtained expressing 3TetR protein in the

nucleus of erythroid cells (Fig. 3). The protein was specifically biotinylated by BirA, as shown by probing Western blots of crude nuclear extracts with HRP-streptavidin (Fig. 3).

Interbreeding of the mouse lines resulted in quadruple *3TetR-G1HRD::R26-BirA::Tet-γ::p53* mice that were used for the generation of fetal-liver derived erythroid progenitor cell lines.

Purification of the γ globin promoter

Erythroid progenitor cell lines (Dolznig et al., 2001) were derived from 13.5d fetal livers, cloned and expanded in vitro. A stable cell line was selected that has repressed the human γ-globin gene, but expresses high levels of the mouse α- and β- globin genes (Fig. 8). This cell line formed the basis for the subsequent biochemical optimization to purify the γ-globin gene promoter. First we confirmed that the γ-globin gene promoter was bound by the 3TetR protein, and that it was released upon the addition of doxycyclin to the medium (Fig. 4A). We found that 3TetR is efficiently biotinylated by BirA and efficiently pulled down by streptavidin beads (Fig. 4A). Subsequently all parameters for the purification of the promoter were optimized. This included fixation conditions, crosslinking, type of pull down beads, buffers, blocking agents to prevent non-specific binding, large scale culture and importantly the elution conditions (Table 1).

The optimization resulted in conditions that gave over 50-fold purification for both the HA- and the biotin pulldown (Figure 4A, B). Precise estimates were very difficult because the yield of the IgG control pulldown was very low. Both pulldowns show an enrichment of the sequences immediately upstream and downstream of the 3TetR sites including all of the promoter sequences. Sequences further upstream (HS5) and downstream (exon 3) show no enrichment when compared to the controls. Thus, the approach appears to work, however a single pulldown before mass spectrometry analysis would be insufficient to identify proteins bound to the γ-globin promoter. We therefore tested whether the chromatin captured with the HA beads could be eluted efficiently by the addition of HA peptide, allowing a subsequent pulldown step on the eluted chromatin. The result (Fig. 4B) shows that all of the bound material could be eluted efficiently by HA peptide. The yield of the purification steps was measured both as 3TetR protein or γ globin promoter DNA recovered after the subsequent HA and biotin pulldowns, which showed a yield of ~2% of the starting material.

Because the sensitivity of mass spectrometry is in the order of >100 attomolar, we estimated to need approximately $3x10^9$ cells. Thus 1.2 liter of *3TetR-G1HRD::R26-*

BirA::Tet-γ:p53 cells was grown to a density of $3x10^6$ cells/ml with or without doxycyclin. The cells were crosslinked with 1% formaldehyde, chromatin was prepared, sheared and subjected to an HA pulldown/elution followed by streptavidin/biotin pulldown. The material was decrosslinked and the proteins separated by gel electrophoresis (Fig. 9). The gels were sliced and each slice was subjected to mass spectrometry analysis. We performed three independent experiments. In the slices of the samples without doxycyclin present, we identified 534 proteins with a Mascot score >60. Of these proteins, 495 were also found in the slices of the doxycyclin-treated samples. Thus collectively, we found 39 possible candidate proteins bound to the γ-globin promoter. A number of these were very unlikely candidates to be involved in the suppression of the γ genes, such as DNA directed RNA polymerase, glutathione S transferase or ribosomal RNAs. These were relegated to the bottom of the list, leaving approximately 20 promising candidates. This included Cdc5L, Actl6a and Chd4 that have been reported previously to bind to the γ-globin promoter or Locus Control Region in vitro (Bernstein and Coughlin, 1997) (Mahajan et al., 2005) (Harju-Baker et al., 2008).

• Titration of cross-linking agent	→ 1% formaldehyde
• Pre-clearing of the Chromatin	→ NO
• Blocking the beads	→YES
• Sonication buffer Vs. SDS lysis buffer	→ Sonication buffer
• Number of cells / 100ul of HA beads	→ 210^{-7} cells/100ul beads
• Optimal streptavidin beads	→ M280
• HA peptide elution titration	→ 0.5mg/ml
• Cells grow in fermentor	→Good
• DNA shearing	→Sonication

Table 1

Functional identification of suppressor proteins

A number of the top suppressor protein candidates were subsequently tested for binding to the γ-globin gene promoter in cultured human erythroid progenitor (HEP) cells (Leberbauer et al., 2005). ChIP assays were carried out for 10 factors. FANCI, ZNF148, CDC5L (Fig.5), BAF53A, CTNNBL1, CHD4, NAP1L1, NCOA2 scored positive for binding to the promoter while Eu HMTase1 (Fig. 5) and SPT5 were negative. Two of the proteins that were positive in the ChIP assays were functionally tested through shRNA- or siRNA-mediated knockdown in HEP cells. Five shRNA constructs for ZNF148 and a pool of 4 siRNAs for FANCI were tested for their ability to lower the respective protein levels. This

showed that a reduction of the levels of FANCI, ZBP-89/ZNF148 in HEP cells increased the level of γ globin gene expression (Fig. 6).

Example 2

Analysis of suppressor proteins

5 Using standard chromatin precipitation analysis followed by massive parallel sequencing (ChIP/seq, Soler et al Genes Dev. 2010 Feb 1;24(3):277-89.) of the precipitated sequences, we analysed the interaction of ZNF148 with the BCL11A gene. We showed that ZBP-89/ZNF148 binds to the promoter region of the BCL11A gene, which is known to repress the expression of the γ-globin genes before and after differentiation of erythroid 10 cells. These results were confirmed in human erythroid cells by a chromatin immunoprecipitation using a standard protocol (Soler et al Genes Dev. 2010 Feb 1;24(3):277-89). The results (Figure 10) show that the ZBP-89 protein binds to a BCL11A promoter fragment using a known binding site (GATA-1, Yamamoto et al J Biol Chem. 2009 Oct 30;284(44):30187-99.) as a positive control and the human erythroid β globin 15 LCR hypersensitive site 5 as the negative control.

In order to reduce the levels of ZBP-89/ZNF148 in cells, we tested four shRNAs. Human erythroid progenitor cells were transduced with four different shRNAs, of which sh4 and sh5 showed a decrease in ZBP-89 protein levels (Bottom panel Fig 11). A control gene actin is not affected in the same cells (middle panel Fig 11). The top panel of Fig 11 20 shows that in the cells that show a decrease of ZBP the level of BCL11A has also been diminished, which shows that the ZBP protein binding to the BCL11A locus (Figure 10) acts as an activator of the BCL 11A gene.

The sequences of the shRNA molecules are as follows:

sh1:
25 CCGGGCAATGCGTAATAACAAGTTACTCGAGTAACTTGTTATTACGCATTGCTTTTT

sh2:
CCGGCCCACCTAAGTTAGTTCTCAACTCGAGTTGAGAACTAACTTAGGTGGGTTTTT

sh3:
CCGGGCTTTCGATCAGGAATGAATTCTCGAGAATTCATTCCTGATCGAAAGCTTTTT

sh4:
CCGGGCATAGACGAAATGCAGTCTTCTCGAGAAGACTGCATTTCGTCTATGCTTTTT

sh5:
CCGGGCAAAGTTTCAAAGTATGCTTCTCGAGAAGCATACTTTGAAACTTTGCTTTTT

5 The target sequences in ZNF148 bound by sh1-sh5 are as follows

Sh1 GCAATGCGTAATAACAAGTTA

Sh2 CCCACCTAAGTTAGTTCTCAA

Sh3 GCTTTCGATCAGGAATGAATT

Sh4 GCATAGACGAAATGCAGTCTT

Sh5 GCAAAGTTTCAAAGTATGCTT

We next analysed the proteins that interact with ZBP-89/ZNF148 in erythroid cells by a protein pulldown assay and analysis of the interacting proteins as described by Soler et al Genes Dev. 2010 Feb 1;24(3):277-89 again using a biotinylation tag. This shows a 10 number of well known proteins to bind to ZBP-89 as shown in Fig 12, including the complete NuRD1 complex, a number of other transcription factors and the SWI/SNF complex. The interactions were confirmed by co-immunoprecipitation assays (Fig 12). The complete set of binding proteins is shown in Table 2.

Using Lentiviral delivery of shRNA, we then knocked down ZBP89ZNF148 in HEP cells, 15 and analysed the levels of γ-globin gene expression. As shown in Fig 13, knock-down of ZBP89 leads to an increase in γ-globin expression.

We also tested a number of the other genes identified above in knock down experiments (Figure 14, CDC5L; Figure 15, APEX1). They show a clear knock down of the respective protein after treatment with shRNAs (panels A) resulting in an upregulation of gamma 20 globin RNA (panel B) and protein (panels C peak F) and staining of the big erythroid cells with an antibody specific for hemoglobin F (γ globin). The larger, stained cells are the cells that are differentiating while the small cells are differentiated and hence never exposed to the shRNA while differentiating.

hit number	Mascot score	symbol	accession	description (Entrez)
4	1647	Zfp148	IPI00123531	zinc finger protein 148
7	1357	Chd4	IPI00396802	chromodomain helicase DNA binding protein 4
10	962	Thrap3	IPI00556768	thyroid hormone receptor associated protein 3
13	781	Gatad2b	IPI00128615	GATA zinc finger domain containing 2B
17	720	Mta1	IPI00776055	metastasis associated 1
18	670	Mta1	IPI00624969	metastasis associated 1
19	669	Gatad2a	IPI00229784	GATA zinc finger domain containing 2A
22	591	Serbp1	IPI00471476	serpine1 mRNA binding protein 1
23	574	Zfp219	IPI00469594	zinc finger protein 219
27	556	Mta2	IPI00128230	metastasis-associated gene family, member 2
29	549	Rbbp4	IPI00122696	retinoblastoma binding protein 4
30	523	Rbbp7	IPI00122698	retinoblastoma binding protein 7
34	488	Top2a	IPI00122223	topoisomerase (DNA) II alpha
39	443	Ilf3	IPI00130591	interleukin enhancer binding factor 3
40	443	Pdcd11	IPI00551454	programmed cell death 11
43	418	Hdac2	IPI00137668	histone deacetylase 2
44	415	Hdac1	IPI00114232	histone deacetylase 1
48	396	Chd3	IPI00551435	chromodomain helicase DNA binding protein 3
52	379	Rcc2	IPI00222509	regulator of chromosome condensation 2
54	364	Ilf2	IPI00318550	interleukin enhancer binding factor 2
55	364	Sfrs1	IPI00420807	splicing factor, arginine/serine-rich 1 (ASF/SF2)
62	346	Mta3	IPI00221805	metastasis associated 3
77	289	Lig3	IPI00124272	ligase III, DNA, ATP-dependent
92	246	Dnmt1	IPI00469323	DNA methyltransferase (cytosine-5) 1
96	241	Dnmt1	IPI00474974	DNA methyltransferase (cytosine-5) 1
102	227	Smarca5	IPI00396739	SWI/SNF related, matrix associated
115	199	Bclaf1	IPI00169477	BCL2-associated transcription factor 1
116	198	Nop14	IPI00353010	NOP14 nucleolar protein homolog (yeast)
131	162	Chd8	IPI00858099	chromodomain helicase DNA binding protein 8
132	161	Ebna1bp2	IPI00111829	EBNA1 binding protein 2
136	158	Elavl1	IPI00108271	ELAV (embryonic lethal,Drosophila)-like 1
137	157	Numa1	IPI00263048	nuclear mitotic apparatus protein 1
138	157	Caprin1	IPI00121515	cell cycle associated protein 1
140	156	Zc3h18	IPI00648513	zinc finger CCCH-type containing 18
147	146	Rfc2	IPI00124744	replication factor C (activator 1) 2
148	145	Gtpbp4	IPI00117642	GTP binding protein 4
152	139	Mbd3	IPI00131067	methyl-CpG binding domain protein 3
154	136	Luc7l	IPI00410804	Luc7 homolog (S. cerevisiae)-like
155	136	Plk1	IPI00120767	polo-like kinase 1 (Drosophila)
158	134	Runx1	IPI00750028	runt related transcription factor 1
166	121	Sltm	IPI00229571	SAFB-like, transcription modulator
179	106	Fxr1	IPI00122521	fragile X mental retardation gene 1
181	104	Aplf	IPI00110815	aprataxin and PNKP like factor
183	103	Thoc4	IPI00114407	THO complex 4
184	102	Mov10	IPI00130328	Moloney leukemia virus 10
185	102	Sfrs5	IPI00314709	splicing factor, arginine/serine-rich 5
186	101	Cbx3	IPI00677454	chromobox homolog 3 (Drosophila HP1 gamma)
187	101	Hp1bp3	IPI00342766	heterochromatin protein 1, binding protein 3
188	100	Ssrp1	IPI00407571	structure specific recognition protein 1
193	96	Pelp1	IPI00321597	proline, glutamic acid and leucine rich protein 1
195	95	Mela	IPI00224370	melanoma antigen
196	95	Ybx1	IPI00120886	Y box protein 1
204	90	Gm6816	IPI00116049	predicted gene 6816
207	87	Cdk1	IPI00114491	cyclin-dependent kinase 1
208	87	Kif23	IPI00330356	kinesin family member 23
211	87	Pnn	IPI00317891	pinin
212	85	Supt16h	IPI00120344	suppressor of Ty 16 homolog (S. cerevisiae)
225	74	Fip1l1	IPI00110716	FIP1 like 1 (S. cerevisiae)
234	68	Rbbp4	IPI00881032	retinoblastoma binding protein 4
235	68	Acin1	IPI00121136	apoptotic chromatin condensation inducer 1
236	67	Cdk2ap2	IPI00132178	CDK2-associated protein 2
239	66	Ccar1	IPI00135207	cell division cycle and apoptosis regulator 1
247	64	Baz1a	IPI00461396	bromodomain adjacent to zinc finger domain 1A
252	61	Bag2	IPI00130304	BCL2-associated athanogene 2
266	54	Mbd2	IPI00131088	methyl-CpG binding domain protein 2
285	45	Cyfip1	IPI00330476	cytoplasmic FMR1 interacting protein 1

Table 2

References

Bank, A. (2006). Regulation of human fetal hemoglobin: new players, new complexities Blood *107*, 435-443.

Bernstein, H.S., and Coughlin, S.R. (1997). Pombe Cdc5-related protein. A putative
5 human transcription factor implicated in mitogen-activated signaling. J Biol Chem *272*, 5833-5837.

Berry, M., Grosveld, F., and Dillon, N. (1992). A single point mutation is the cause of the Greek form of hereditary persistence of fetal haemoglobin. Nature *358*, 499-502.

de Boer, E., Rodriguez, P., Bonte, E., Krijgsveld, J., Katsantoni, E., Heck, A., Grosveld,
10 F., and Strouboulis, J. (2003). Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. Proc Natl Acad Sci U S A *100*, 7480-7485.

Dillon, N., and Grosveld, F. (1991). Human gamma-globin genes silenced independently of other genes in the beta-globin locus. Nature *350*, 252-254.

15 Dolznig, H., Boulme, F., Stangl, K., Deiner, E.M., Mikulits, W., Beug, H., and Mullner, E.W. (2001). Establishment of normal, terminally differentiating mouse erythroid progenitors: molecular characterization by cDNA arrays. Faseb J *15*, 1442-1444.

Donehower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery, C.A., Jr., Butel, J.S., and Bradley, A. (1992). Mice deficient for p53 are developmentally normal but
20 susceptible to spontaneous tumours. Nature *356*, 215-221.

Driegen, S., Ferreira, R., van Zon, A., Strouboulis, J., Jaegle, M., Grosveld, F., Philipsen, S., and Meijer, D. (2005). A generic tool for biotinylation of tagged proteins in transgenic mice. Transgenic Res *14*, 477-482.

G Stamatoyannopoulos, F.G. (2001). Hemoglobin switching. In: G Stamatoyannopoulos,
25 P Majerus, R Perlmutter and H Varmus, Editors, Molecular Basis of Blood Diseases (3rd ed), W Saunders Publishing Company, Philadelphia, 135–182.

Gossen, M., and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc Natl Acad Sci U S A *89*, 5547-5551.

Grosveld, F., van Assendelft, G.B., Greaves, D.R., and Kollias, G. (1987). Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell *51*, 975-985.

Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N., and
5 Grosveld, F. (1991). Importance of globin gene order for correct developmental expression. Genes Dev *5*, 1387-1394.

Hardison, R.C., Chui, D.H., Giardine, B., Riemer, C., Patrinos, G.P., Anagnou, N., Miller, W., and Wajcman, H. (2002). HbVar: A relational database of human hemoglobin variants and thalassemia mutations at the globin gene server. Hum Mutat *19*, 225-233.

10 Harju-Baker, S., Costa, F.C., Fedosyuk, H., Neades, R., and Peterson, K.R. (2008). Silencing of Agamma-globin gene expression during adult definitive erythropoiesis mediated by GATA-1-FOG-1-Mi2 complex binding at the -566 GATA site. Mol Cell Biol *28*, 3101-3113.

Inusa, E.C.a.B. (2007). Sickle Cell Anemia: Targeting the Role of Fetal Hemoglobin in
15 Therapy

Clinical Pediatrics *46*, 386 - 391.

Jiang, J., Best, S., Menzel, S., Silver, N., Lai, M.I., Surdulescu, G.L., Spector, T.D., and Thein, S.L. (2006). cMYB is involved in the regulation of fetal hemoglobin production in adults. Blood *108*, 1077-1083.

20 Leberbauer, C., Boulme, F., Unfried, G., Huber, J., Beug, H., and Mullner, E.W. (2005). Different steroids co-regulate long-term expansion versus terminal differentiation in primary human erythroid progenitors. Blood *105*, 85-94.

Lettre, G., Sankaran, V.G., Bezerra, M.A., Araujo, A.S., Uda, M., Sanna, S., Cao, A., Schlessinger, D., Costa, F.F., Hirschhorn, J.N., *et al.* (2008). DNA polymorphisms at the
25 BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. Proc Natl Acad Sci U S A *105*, 11869-11874.

Li, Q., Duan, Z.J., and Stamatoyannopoulos, G. (2001). Analysis of the mechanism of action of non-deletion hereditary persistence of fetal hemoglobin mutants in transgenic mice. Embo J *20*, 157-164.

Mahajan, M.C., Narlikar, G.J., Boyapaty, G., Kingston, R.E., and Weissman, S.M. (2005). Heterogeneous nuclear ribonucleoprotein C1/C2, MeCP1, and SWI/SNF form a chromatin remodeling complex at the beta-globin locus control region. Proc Natl Acad Sci U S A *102*, 15012-15017.

5 Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., Foglio, M., Zelenika, D., Boland, A., Rooks, H., *et al.* (2007). A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. Nat Genet *39*, 1197-1199.

Ohneda, K., Shimizu, R., Nishimura, S., Muraosa, Y., Takahashi, S., Engel, J.D., and
10 Yamamoto, M. (2002). A minigene containing four discrete cis elements recapitulates GATA-1 gene expression in vivo. Genes Cells *7*, 1243-1254.

Peterson, K.R., Navas, P.A., and Stamatoyannopoulos, G. (1998). beta-YAC transgenic mice for studying LCR function. Ann N Y Acad Sci *850*, 28-37.

R. Mankidy, D.V.F., R. Mabaera, C. H. Lowrey, M. S. Boosalis, G. L. White, S. A.
15 Castaneda, and S. P. Perrine (2006). Short-chain fatty acids induce {gamma}-globin gene expression by displacement of a HDAC3-NCoR repressor complex. Blood *108*, 3179 - 3186.

Sankaran, V.G., Menne, T.F., Xu, J., Akie, T.E., Lettre, G., Van Handel, B., Mikkola, H.K., Hirschhorn, J.N., Cantor, A.B., and Orkin, S.H. (2008). Human fetal hemoglobin
20 expression is regulated by the developmental stage-specific repressor BCL11A. Science *322*, 1839-1842.

Stamatoyannopoulos, G. (2005). Prospects for developing a molecular cure for thalassemia. Hematology *10*, 255-257.

Starck, J., Sarkar, R., Romana, M., Bhargava, A., Scarpa, A.L., Tanaka, M., Chamberlain,
25 J.W., Weissman, S.M., and Forget, B.G. (1994). Developmental regulation of human gamma- and beta-globin genes in the absence of the locus control region. Blood *84*, 1656-1665.

Strouboulis, J., Dillon, N., and Grosveld, F. (1992). Developmental regulation of a complete 70-kb human beta-globin locus in transgenic mice. Genes Dev *6*, 1857-1864.

30 Thein, S.L., and Menzel, S. (2009). Discovering the genetics underlying foetal haemoglobin production in adults. Br J Haematol *145*, 455-467.

Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V.G., Chen, W., Usala, G., Busonero, F., Maschio, A., Albai, G., *et al.* (2008). Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. Proc Natl Acad Sci U S A *105*, 1620-1625.

5 Yu M, H.H., Xiang P, Li Q, Stamatoyannopoulos G. (2006). Autonomous silencing as well as competition controls gamma-globin gene expression during development. Mol Cell Biol *26*, 4775-4781

10

CLAIMS

1. A method for identifying an inhibitor of a suppressor of γ-globin gene expression in a mammalian cell, comprising the steps of providing one or more candidate modulators of a suppressor selected from the group consisting of BAF53A, CTNNBL1, ZNF148, Fancl,
5 CHD4, NAP1L1 and NCOA2, administering said one or more candidate modulators of the suppressor to an erythroid cell in which globin gene expression is switched from γ to β; and measuring γ-globin gene expression in said cell.

2. A method according to claim 1, wherein γ-globin gene expression is measured by means of a reporter gene which is operably linked to γ-globin control sequences.

10 3. A method according to any preceding claim, wherein γ-globin gene expression is increased by at least 2g/dl.

4. A method according to any preceding claim, in which the erythroid cell is a human erythroid progenitor (HEP) cell which expresses β-globin at greater levels than γ-globin.

5. A method according to any preceding claim, wherein the inhibitor is selected from
15 the group consisting of an RNAi molecule (such as a siRNA of shRNA molecule), a peptide and an antibody or antibody fragment.

6. A method according to any preceding claim, further comprising isolating the candidate inhibitor.

7. A method for identifying a suppressor or activator of γ-globin gene expression,
20 comprising the steps of:

(a) providing a gene construct comprising at least partial sequences of a human γ-globin gene, proximately linked to a binding site for a DNA binding protein;

(b) introducing the gene construct into an erythroid cell;

(c) introducing into said erythroid cell a reagent comprising a DNA binding protein
25 capable of binding to the binding site in (a), and one or more selectable tags;

(d) crosslinking the reagent bound to the binding site with cellular factors with which it is in contact; and

(e) isolating the cellular factors crosslinked to the reagent by means of the one or more selectable tags.

8. A method according to claim 7, wherein the binding site is a binding site for the tetracycline repressor TetR and the DNA binding protein is TetR.

5 9. A method according to claim 7 or claim 8, wherein the selectable tags comprise a biotinylatable tag and/or an HA tag.

10. A method according to any one of claims 7 to 9, wherein the reagent is encoded by a second gene construct.

11. A method according to any one of claims 7 to 10, wherein at least one of the gene
10 constructs is introduced by transgenesis into a non-human mammal.

12. A method according to any one of claims 7 to 11, further comprising de-crosslinking the cellular factor from the reagent and identifying said cellular factor.

13. A process for identifying a compound or compounds which is an inhibitor of a suppressor of γ-globin gene expression identifiable according to any one of claims 7 to
15 12, comprising the steps of:

(a) incubating a suppressor of γ-globin gene expression with the compound or compounds to be assessed; and

(b) identifying those compounds which influence the activity of the suppressor.

14. A process according to claim 13, wherein the suppressor of γ-globin gene
20 expression acts to suppress the expression of a reporter gene operably linked to γ-globin control sequences, and the inhibitor acts to increase the expression of the reporter gene.

15. An inhibitor of a suppressor of γ-globin gene expression identified by a process according to any one of claims 1 to 6, 13 or 14, for use in the treatment of a blood disorder involving a globin gene deficiency.

25 16. An inhibitor according to claim 15, wherein the blood disorder is sickle cell anaemia or β-thalassaemia.

17. A method for upregulating expression of γ-globin in an erythroid cell in which globin gene expression is switched from γ to β, comprising administering to such a cell an

inhibitor of a suppressor selected from the group consisting of BAF53A, CTNNBL1, ZNF148, Fancl, CHD4, NAP1L1 and NCOA2.

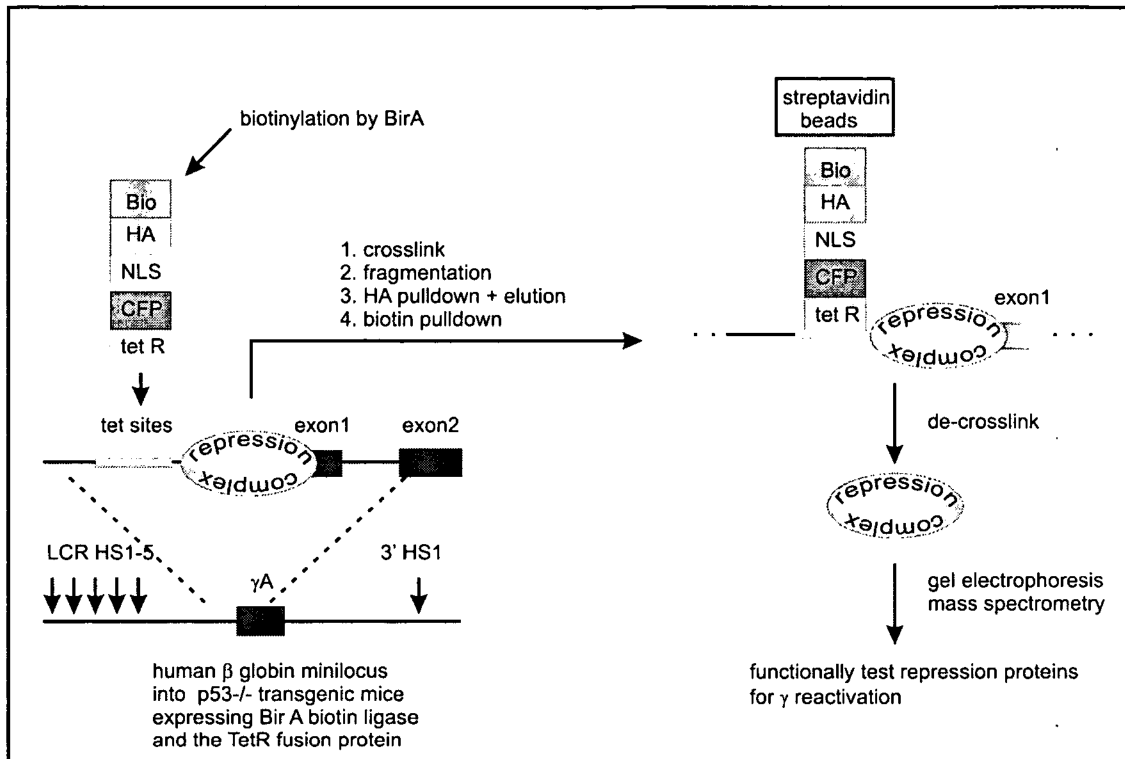

Figure 1

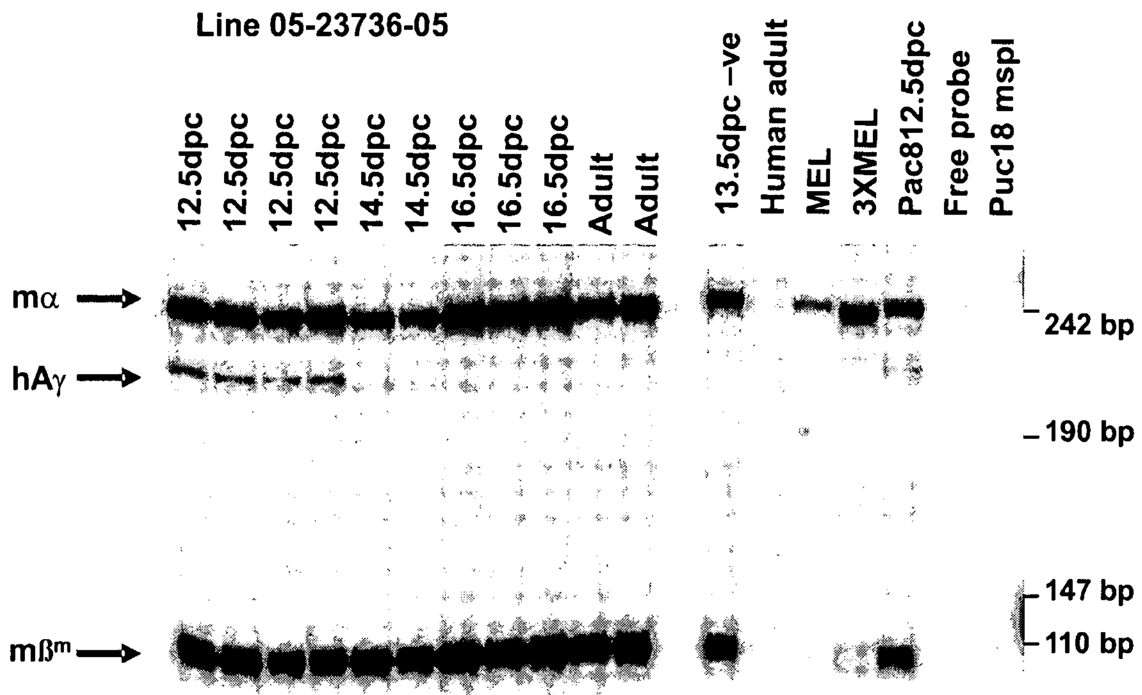

Figure 2

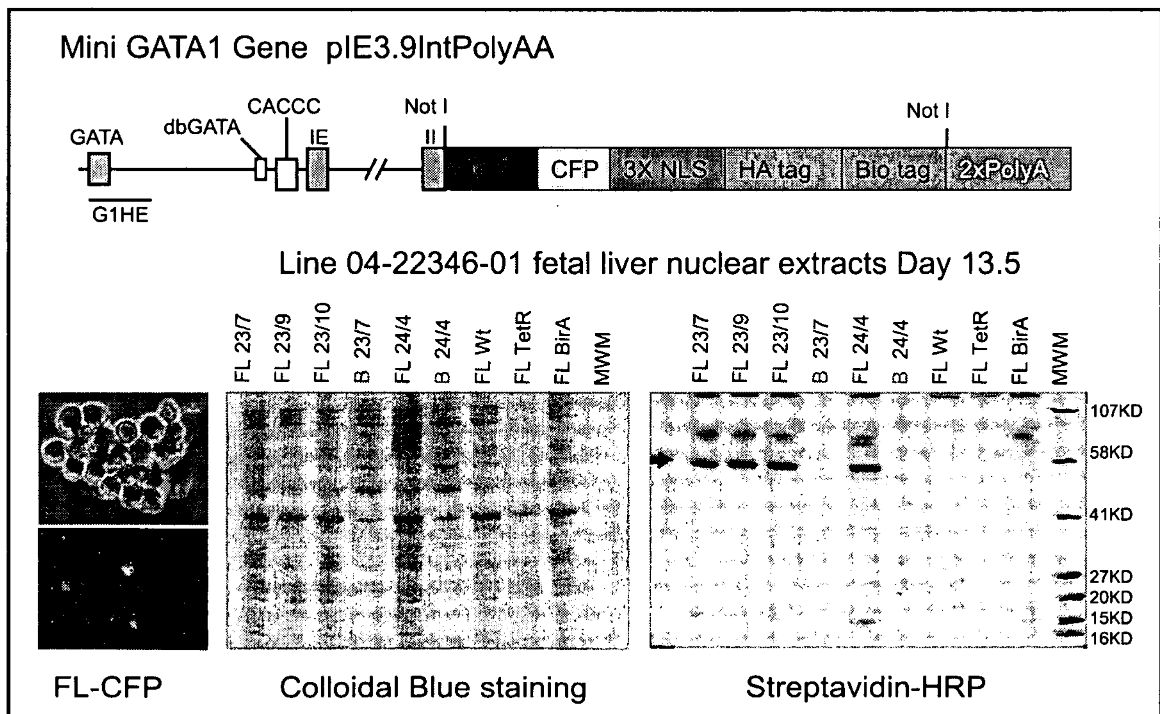

Figure 3

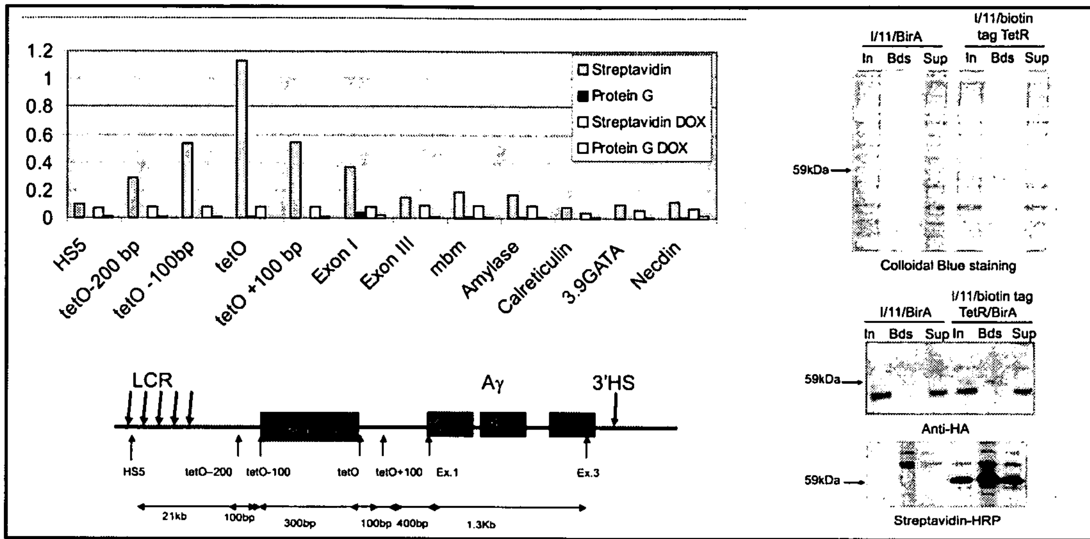

Figure 4A

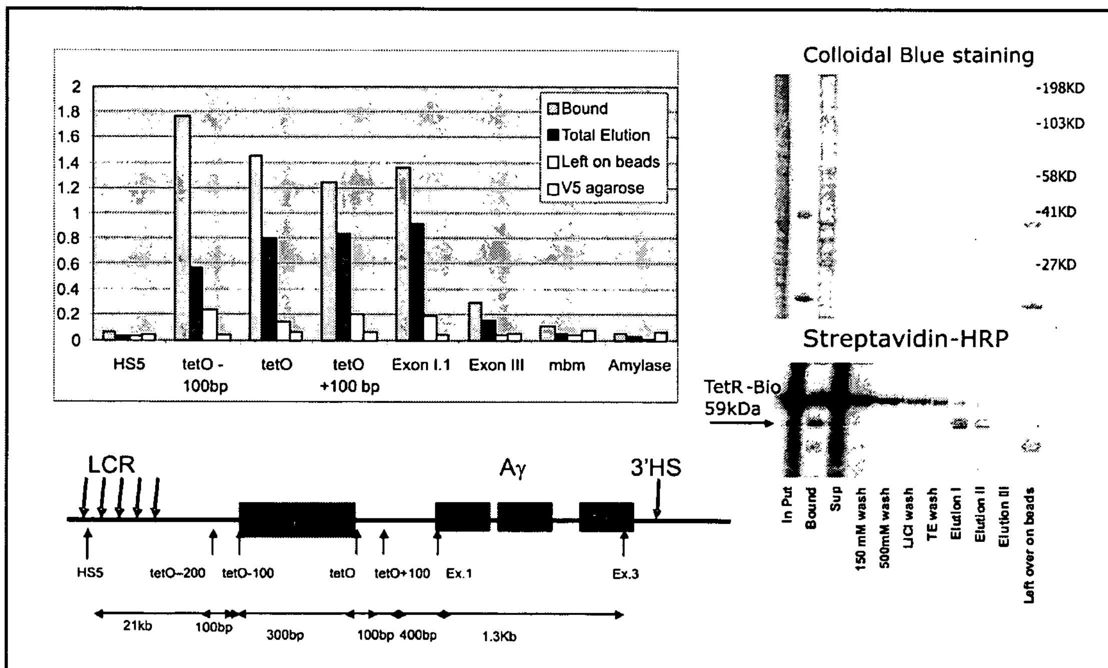

Figure 4B

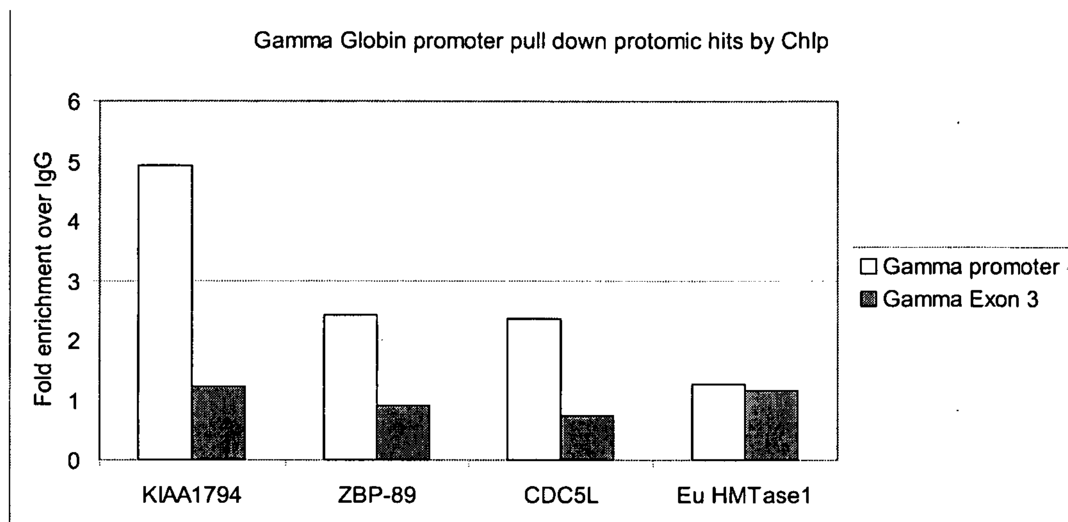

Figure 5

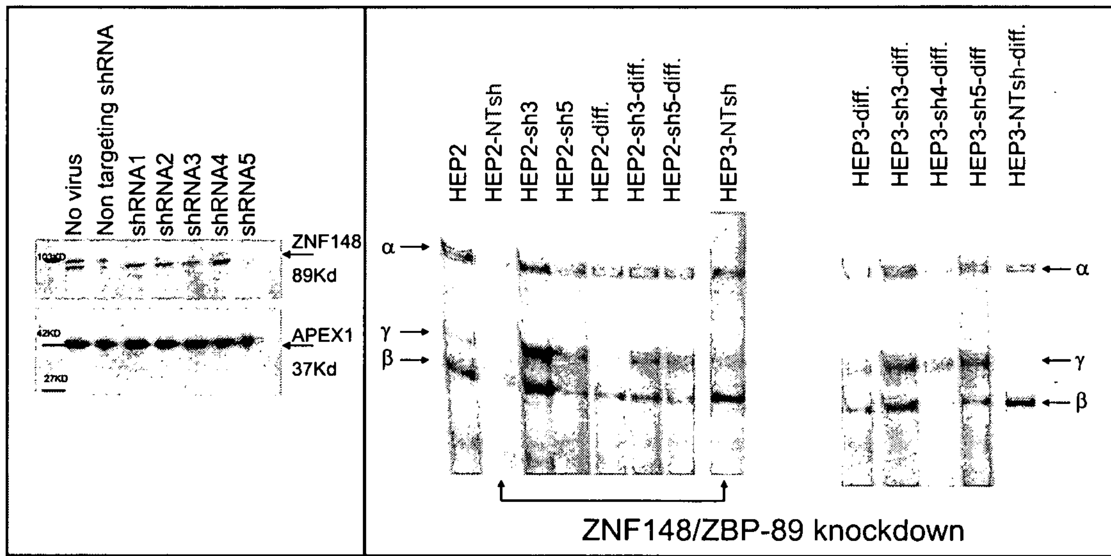

Figure 6A

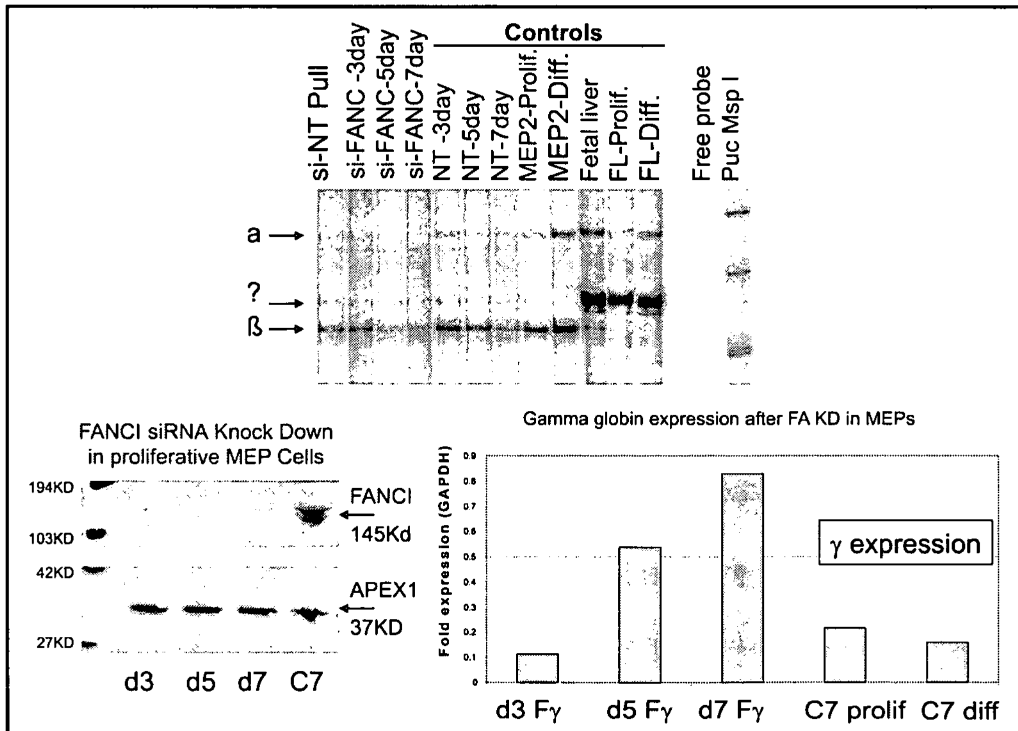

Figure 6B

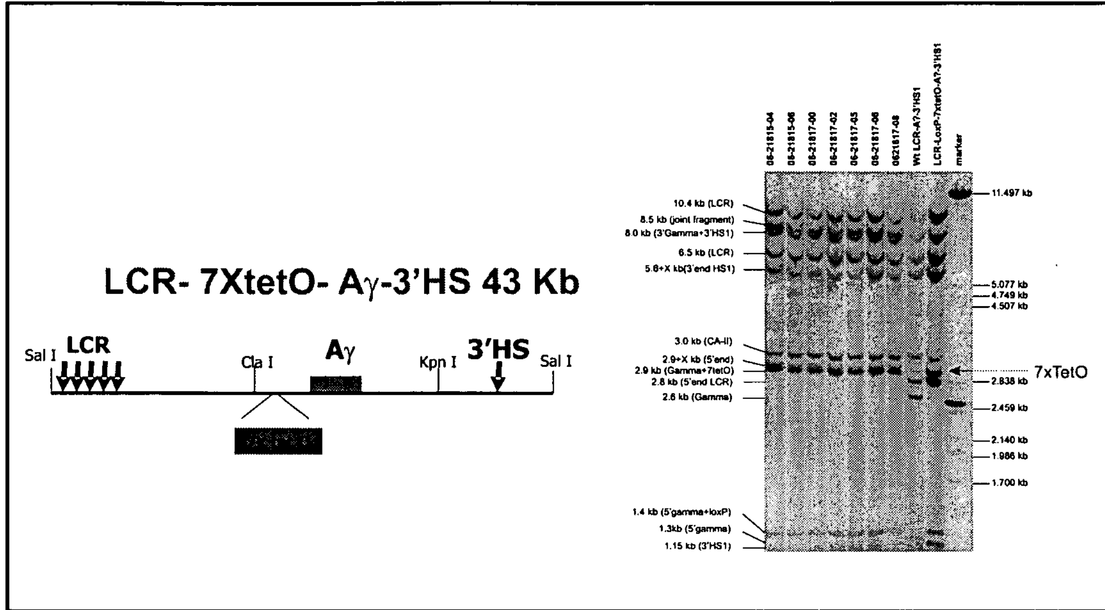

Figure 7

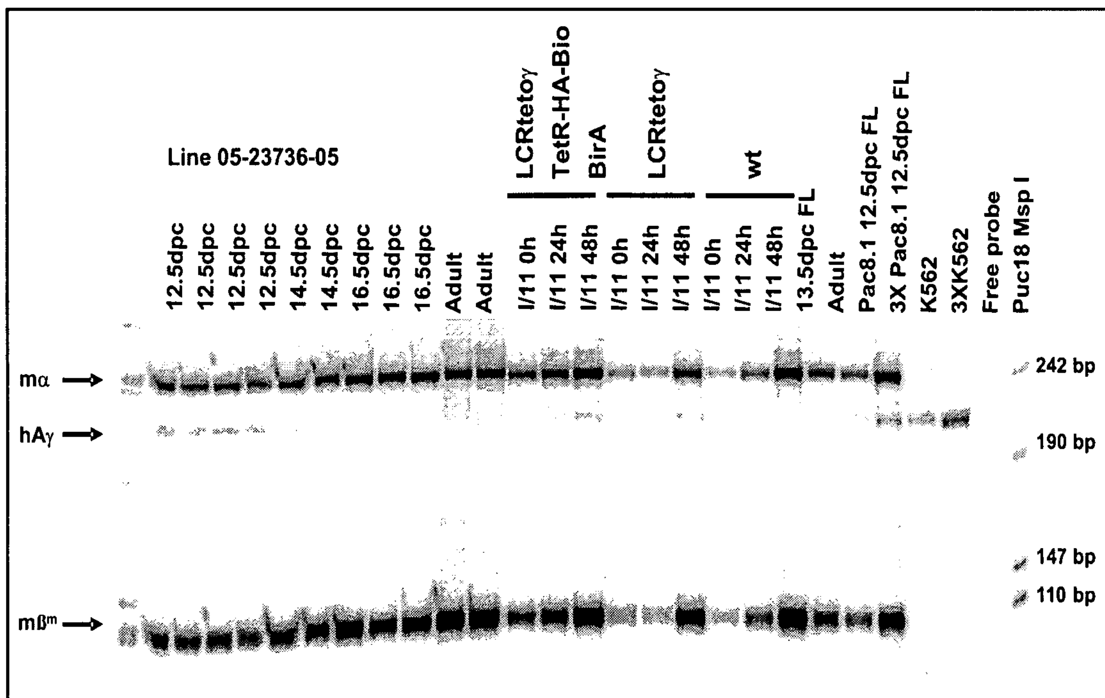

Figure 8

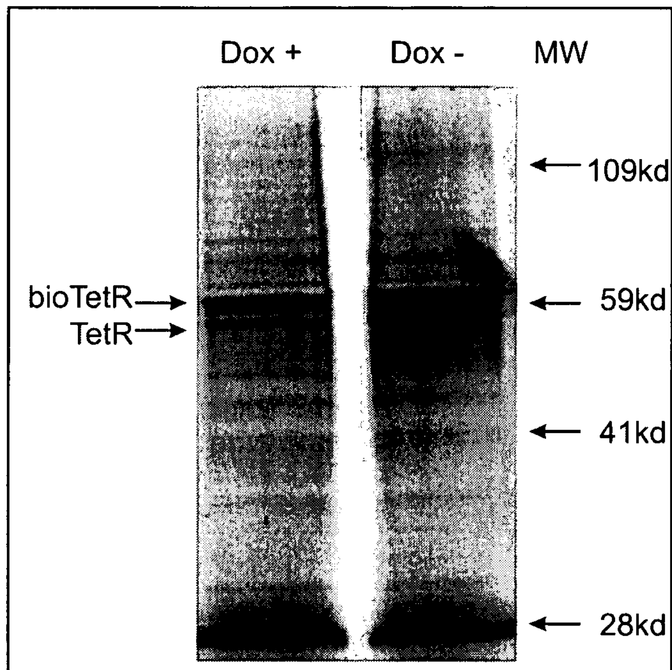

Figure 9

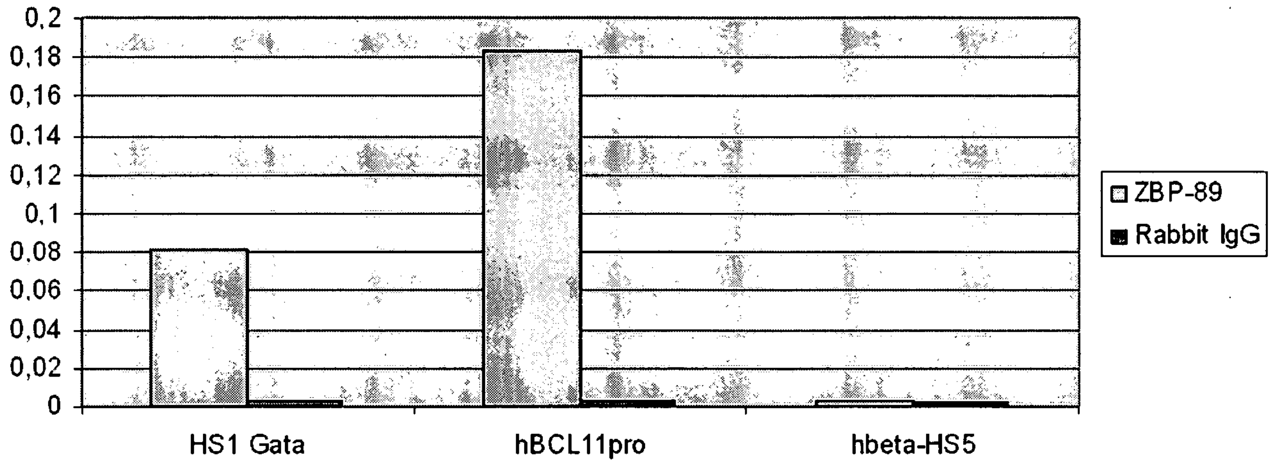

Figure 10

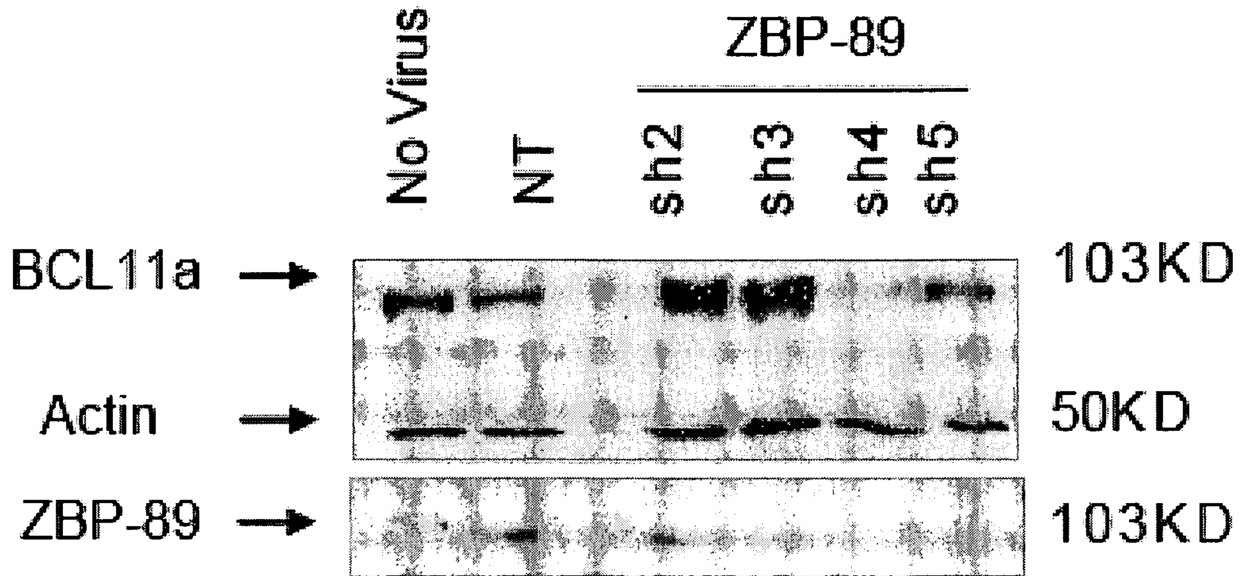

Figure 11

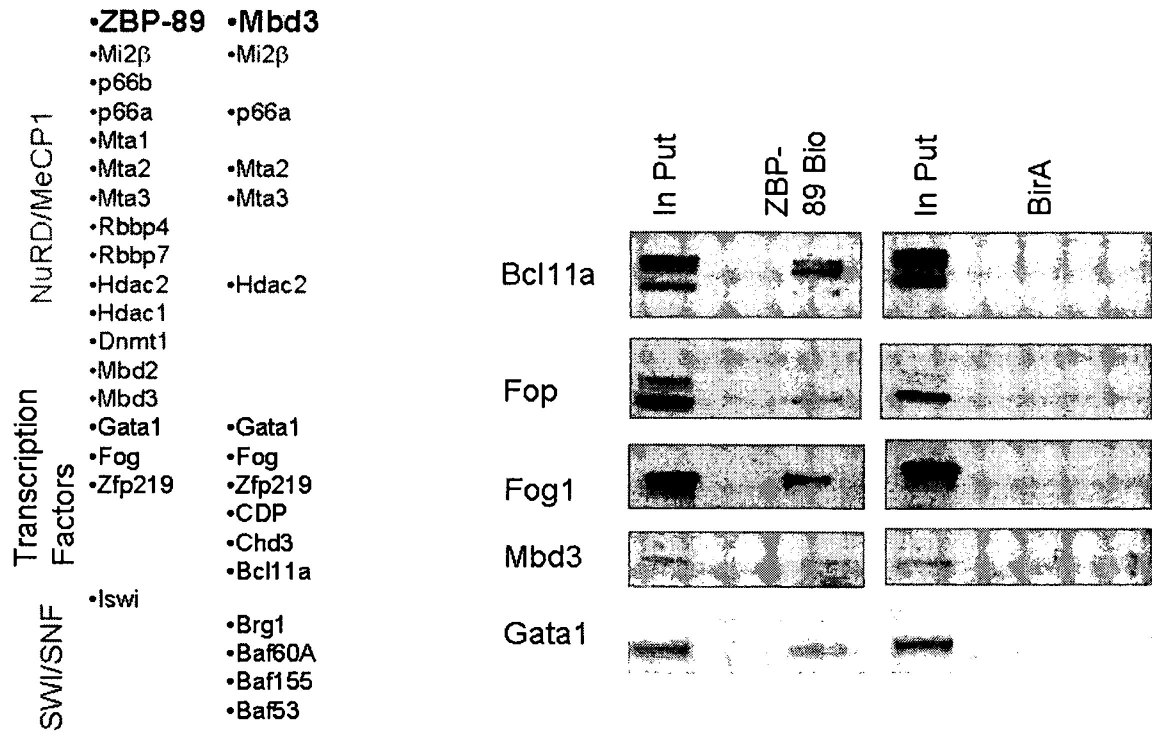

Figure 12

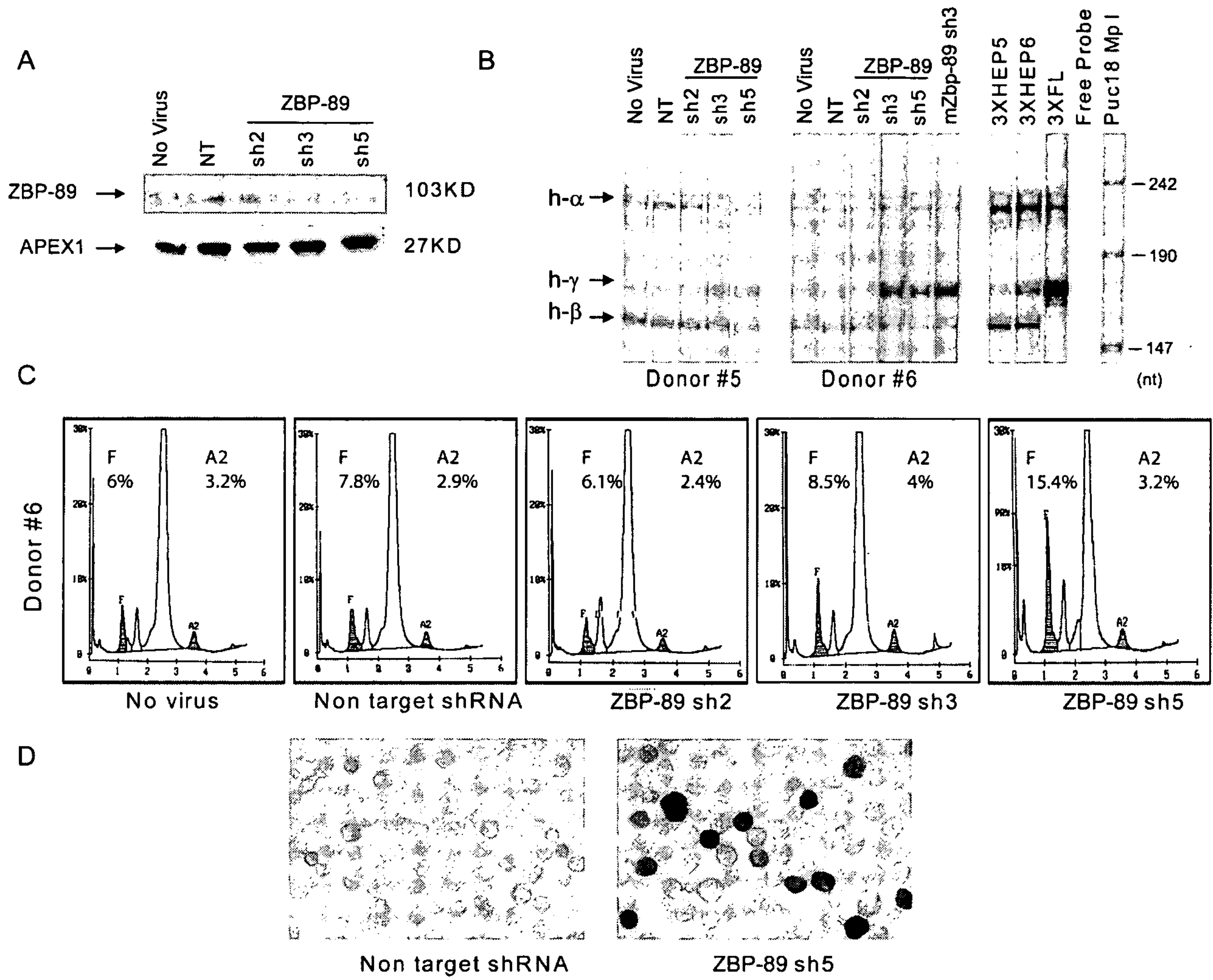

Figure 13

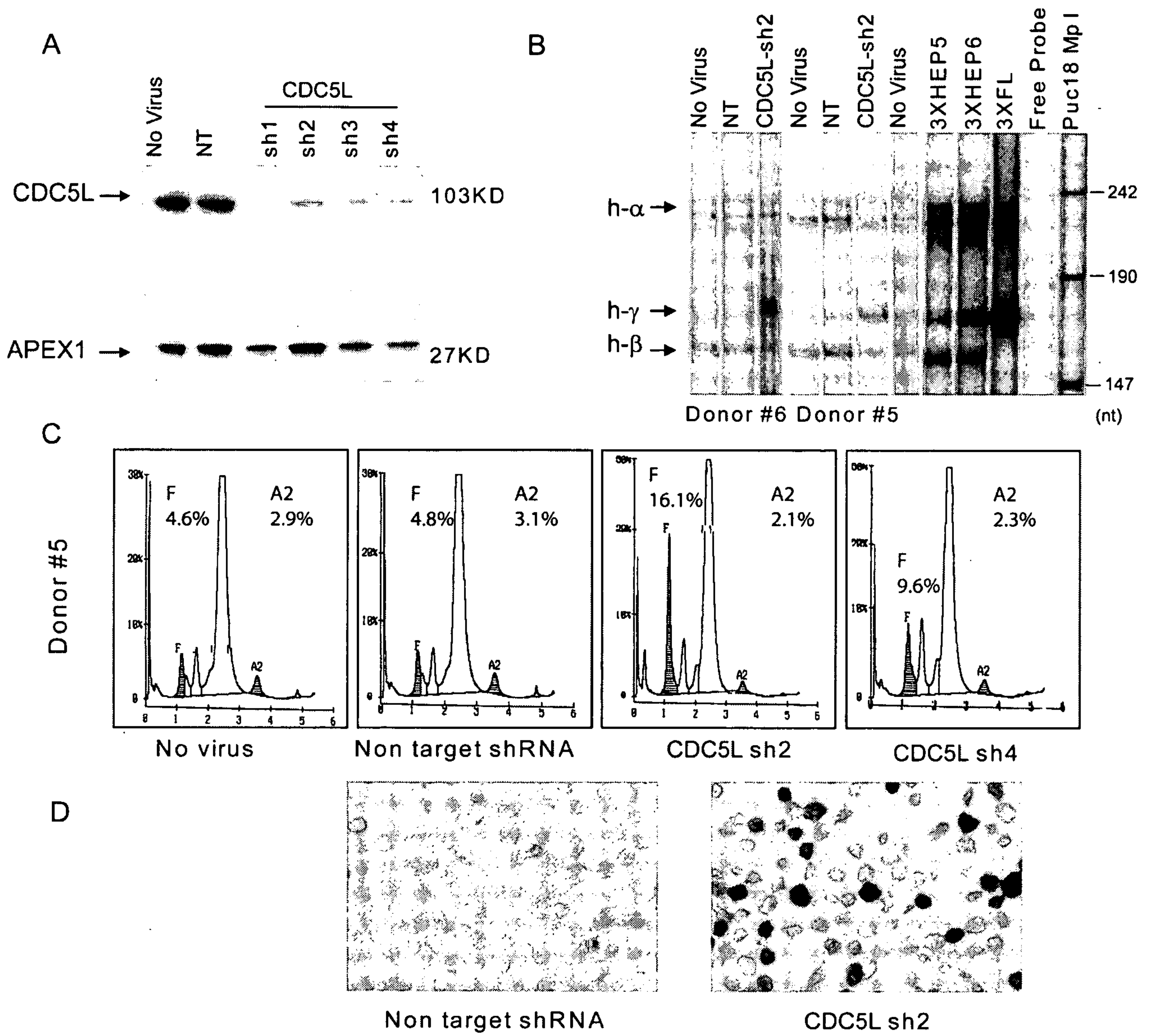

Figure 14

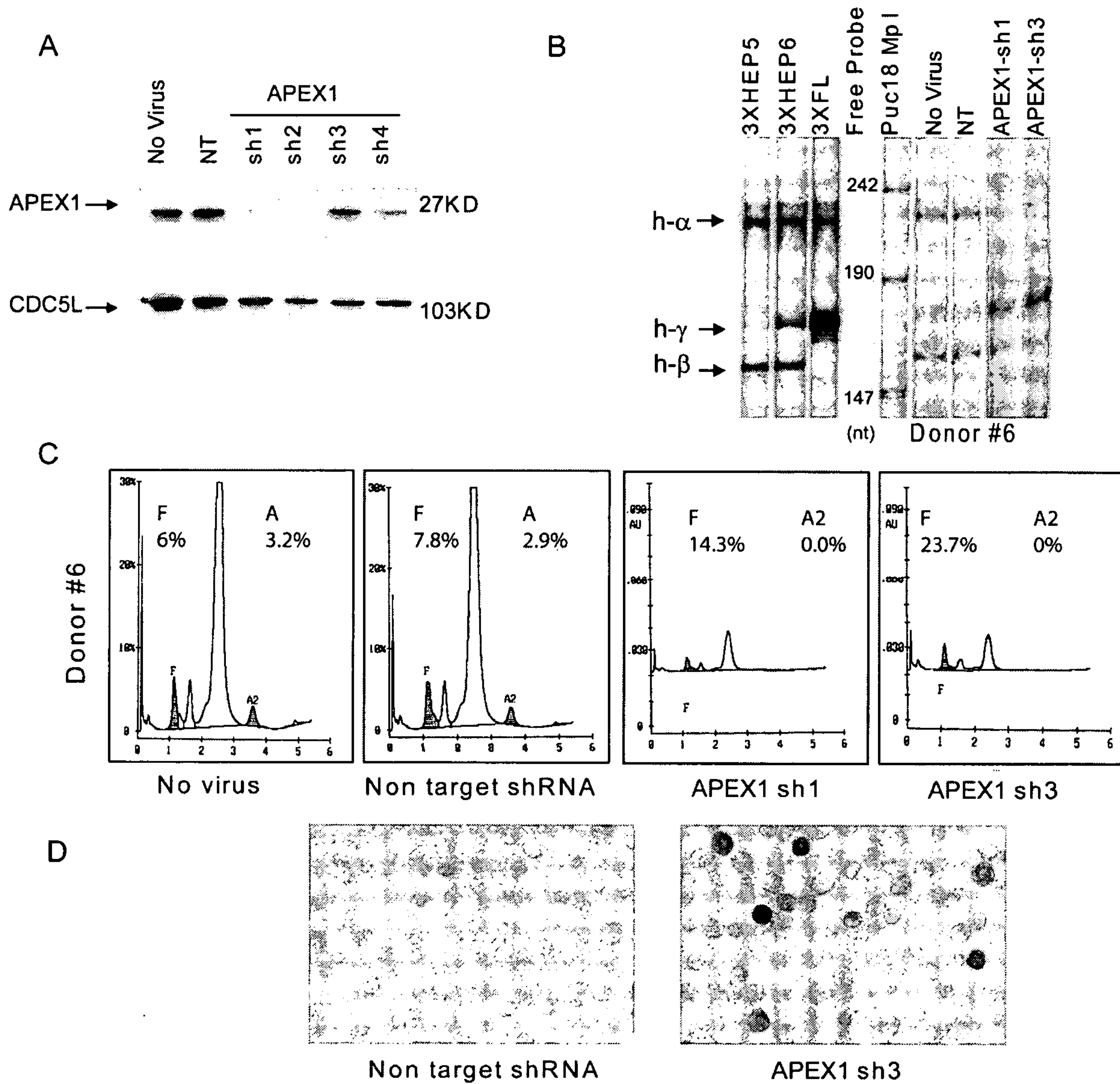

Figure 15