



# **UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

Instituto de Fisiología Celular  
Departamento de Genética Molecular

PROGRAMA DE DOCTORADO  
EN CIENCIAS BIOMÉDICAS

**Estudio de la estructura de la cromatina en dos  
elementos  
de regulación localizados en el extremo 3' del grupo de  
genes  $\alpha$ -globina de pollo**

**T E S I S**

PARA OBTENER EL GRADO DE:

**DOCTOR EN CIENCIAS**

PRESENTA:

**Q. Martín Escamilla del Arenal**

**Tutor: DR. FÉLIX RECILLAS TARGA**

**Dedicada**

**A Vanesa, la mujer que le trajo felicidad y belleza a mi vida**

**A mi niña hermosa Emilia, que llego a mi vida para darle sentido**

## **Agradecimientos Institucionales**

Este trabajo fue realizado en el Instituto de Fisiología Celular, Departamento de Genética Molecular, laboratorio 122-Norte de la Universidad Nacional Autónoma de México, bajo la tutoría del Dr. Félix Recillas Targa.

Este trabajo fue apoyado por la Dirección General de Asuntos del Personal Académico-UNAM (IN203200, IX230104, IN209403, and IN214407), Consejo Nacional de Ciencia y Tecnología, CONACyT (33863-N, 42653-Q, and 58767) y por the Third World Academy of Sciences (TWAS, 01-055 RG/BIO/LA).

Mis estudios doctorales fueron apoyados por una Beca otorgada por el CONACyT (166354) y el complemento de beca de la Dirección General de Estudios de Posgrado DGEP-UNAM.

El comité tutorial que asesoró el desarrollo del presente trabajo, estuvo conformado por:

Dr. Félix Recillas Targa  
Celular

Instituto de Fisiología

Dra. Alicia González Manjarrez  
Celular

Instituto de Fisiología

Dr. David Romero Camarena  
Genómicas

Centro de Ciencias

Durante el desarrollo experimental de esta tesis, se recibió una excelente asesoría técnica por parte de la Bióloga Georgina Guerrero Avendaño.

El jurado de examen estuvo constituido por:

Dra. Carmen Gómez Eichelman (presidente)

Dr. Félix Recillas Targa (secretario)

Dr. Miguel Lara Flores

Dr. Mario Zurita Ortega

Dr. Luís Servin González

Dra. Alicia González Manjarrez

Dr. Luís Vaca Domínguez

En la preparación de este trabajo colaboraron con valiosos comentarios la Dra. Vanesa Olivares Illana, la Dra. Patricia del Arenal y el Dr. Edgardo Escamilla. Así como los miembros del jurado.

### **Agradecimientos**

En primer lugar, agradezco a la UNAM y todo lo que representa.

Al Instituto de Fisiología Celular y toda la gente que en el trabaja, por colaborar en mi preparación día con día.

Al Dr. Félix Recillas, por ayudarme a construir una oportunidad para mi futuro.

A los Doctores Armando Gómez y la Dra. Marieta Tuene, gracias por ser el tipo de súper héroes que los aprendices de ciencia admiramos. Gracias por su sencillez y sus ganas por siempre enseñar algo nuevo.

A los integrantes del laboratorio 122-N, por que cada uno de ellos me enseñó algo y me brindo su amistad y cariño. A la Geito sin la cual el laboratorio no sería lo que es, gracias por darme consuelo cuando más lo necesitaba. A la hermana mayor, la Vivís, que nos enseñó que hacer y que no hacer en el laboratorio. A mis hermanos: Marco con quien inicie este viaje y aprendí mucho de lo que se. Inti con quien compartí muchos de mis mejores momentos en este laboratorio. Héctor y su invaluable experiencia. El Netito, compañero de rutas. Al TZK y ahora le toca al amigo tezca. A mis hermanitas Mayra y Eria, que trajeron nuevos bríos al laboratorio y toda la alegría y conocimiento que nos daban a diario. A Gianelli, si la cual el laboratorio 122N no sería lo productivo que es, gracias por inspirarme a ser una persona trabajadora y responsable. A mis tres estudiantes, a las cuales trate de enseñarles todo lo que sabía y gracias a las cuales comprendí muchas de las cosas que estaba haciendo. Alejandra y sus ganas por hacer todo bien, a Adriana y su pasión por la biología y a Itzita y sus interminables preguntas, gracias por aguantarme y enseñarme lo que creía saber. Al mismísimo Abraham a quien tuve la suerte de conocer bien al final de mi Doctorado. A la nueva generación y todas las ganas que traen por hacer las cosas bien, al Christian mejor conocido como el GTX,

al tocapito, a Pedro, a Edgar, a Meche. Gracias a todos por escucharme y ayudarme con sus consejo todos los días.

A mis padres sin los cuales esta tesis no sería posible, gracias por el cariño y el tiempo que me han dedicado, gracias por siempre estar ahí. Gracias por muchas cosas más. Los quiero mucho.

A mi hermana y su cariño incondicional. Gracias por enseñarme lo valioso que es un hermano, siempre y en tu caso la mejor.

A Silvita y su gran sentido del humor. Gracias por tantas carcajadas y quererme como a un hijo. Te quiero.

A mis tíos Eduardo y Adriana que los quiero como si fueran mis padres y me han mostrado día a día lo que significa ser familia y que decir de sus insoportables hijos, mis adorables primos, compañeros de la infancia y parte de mis recuerdos mas valiosos.

A mi prima Michele y a mi prima la Chica, por que las quiero mucho y son parte importante en mi vida. A mis tíos Beto e Iris, gracias por su cariño. A todos mis demás primos Daniel, Toño, Luís, Estela, Oli, Ernesto, Mónica, Cris, John, Lalo y Maru. Gracias por formar parte de una gran familia.

A mi abuela Mastelita, gracias por todo tu cariño y todos los consejos que siempre me has dado. A mi tío Gilberto, gracias por los momentos de alegría que me has regalado, al Fer y Mariana, por que los quiero. A mis tíos Víctor y Edith y a la prima Sofía, por que son el tipo de familia que uno sabe que están ahí, con su cariño y mejores deseos.

A la familia que conocí gracias al amor de mi vida. Gracias a Pilar, Mariana, Paola, Carmen, Mario, Mario y Mario, por ser una gran familia que desde siempre me quiso y dio su cariño. También gracias a los más de 100 integrantes que constituyen esta gran familia, llena de tradiciones y buenos deseos.

A mi hermana adoptiva, mi mejor amiga y confidente para toda la vida. Maria Ana gracias por ser como eres y quererme por siempre. Gracias tambien, por siempre estar.

A los amigos que ya son familia, a los amigos que tuve la suerte de conocer hace mas de 15 años y un día ya formaban parte de lo que mas quiero. Gracias por ser parte de una amistad que me hace una persona muy afortunada, gracias a Santiago, Marcos, Daniel, El Richi, Mónica, Dieguito, Dany, Itzam, Samanta, Yuridia, Valentina, Lucia, Julia, Nayar y Nareni. Los quiero.

A la banda de la facultad de Química, mis amigos de interminables pachangas y parte importante en mi formación académica, gracias a la Clau, Paco, Héctor, Carlos, Camilo, Emiliano, Marco, Blanca, Maligno, Miguel y muchos mas. Gracias por su cariño y amistad.

A Eréndida, a Lina Riego alias la Jefa, gracias por participar en la formación de los siemientos que ahora me permiten presentar este trabajo, gracias a ti y a la güera Estrada por tanto cariño y al Fernando.

A la Doctora Alicia González, gracias por ser ese tipo de jefes que motivan a uno. Gracias por muchas sabias palabra que me han enriquecido como persona y como investigador.

A mis compadres de bohemia Alexander y Mauricio. A la mismísima Cristina Parrandas y sus interminables interrogatorios. A Lourdes Valenzuela y a Amaranta. A todos los integrantes del 301-0te, gracias por todo lo que me enseñaron.

Esta tesis también esta dedicada a la profesionalidad y libertad de expresión con que solía trabajar Carmen Aristegui.

Resumen	9
Abstract	10
Abreviaturas	11
Introducción general	13
La Regulación Transcripcional	13
Regulación a nivel de secuencia	15
Elementos de regulación en el ADN	16
Promotores	16
Las islas CpG y los promotores	17
Elemento proximal	18
La RNA polimerasa II y los factores generales de transcripción	18
Especificidad de los enhancers por sus promotores	21
Enhancers	22
Modelos de acción	22
Enhancer y cromatina	22
Modelo de prendido o apagado versus el modelo progresivo	23
Modelo del enhanceosoma contra “billboard” (pizarra)	23
Modelo en tres dimensiones	25
Los elementos de Control del Locus ó LCR	27
Modelos de acción de los LCRs	28
Modelo de formación de asa o “Looping”	28
Modelo de encarrilamiento o “Tracking”	29
Modelo de relevo o “Linking”	29
Transcritos intergénicos	30
Delimitadores o “insulators”	31
Modelos de acción	33
Modelo de bloqueo de enhancer	33
Modelo de protección contra el efecto de posición	34
Regulación a nivel de la cromatina	35
Definición y componentes	36
Regulación epigenética	38
La metilación del ADN	39
Los complejos de remodelaje ATP-dependientes	41
Modificaciones post-traduccionales de las histonas	42
Variantes de histonas	45
Hipótesis del código de barras (“barcode”)	46
Regulación epigenética a través de los complejos Polycomb y Trithorax	47
.	47
Generación de regiones libres de nucleosomas	48
El núcleo	49
Dinámica nuclear	50
Dominios cromosómicos	52
Factores transcripcionales	54
El factor transcripcional GATA-1	54
GATA-1, cromatina y regiones libres de nucleosomas	56
La relación entre GATA-1, la RNA polimerasa II y los enhancers	58
FOG (“Friend of GATA”)	60
El factor de la familia Maf eritroide específico (NF-E2)	61
El miembro fundador de la familia “krüppel-like” (EKLF)	62

El factor Ying Yang 1 (YY1)	63
El factor general de transcripción Sp1	65
El modelo de estudio y antecedentes particulares	65
Dominios Globina	65
El dominio $\alpha$ -globina	67
El dominio $\beta$ -globina	68
El elemento silenciador/enhancer $\alpha$ -globina de pollo	69
Planeamiento del problema	71
Hipótesis	72
Objetivos generales	73
Objetivos particulares	73
Resultados	74
El sistema biológico	74
Generación de los vectores madre utilizados en este trabajo	75
Estudio in vitro del elemento enhancer	77
Función de los múltiples sitios de interacción al factor GATA-1 en el enhancer	79
Caracterización de la secuencia conservada de 120 pb	84
Estudio funcional de la secuencia conservada de 120 pb	87
Estudio in vitro del elemento silenciador	90
La conexión con la cromatina	94
Caracterización de la estructura de la cromatina en los elementos de regulación	95
Determinación del posicionamiento nucleosomal	97
Evaluación semicuantitativa de la remodelación de la cromatina	102
Mapeo de los sitios de hipersensibilidad	106
Determinación in vivo de la participación de los sitios de unión al factor GATA-1	110
Modificaciones post-traduccionales de las histonas	116
Regiones libres de nucleosomas	118
Metilación del ADN en el fragmento de 1036 pb	121
Discusión	126
Conclusiones	135
Perspectivas	136
Material y métodos	144
Apéndices	154
1. GATA-1 Modulates the Chromatin Structure and Activity of the Chicken $\alpha$ -Globin 3' Enhancer	154
2. YY1 and GATA-1 Interaction Modulate the Chicken 3'-Side $\alpha$ -Globin Enhancer Activity	166
3. Globin genes transcriptional switching, chromatin structure and linked lessons to epigenetics in cancer: A comparative overview	181
4. Participación de la Estructura de la Cromatina en la Regulación de la Expresión Génica	192
5. Transcriptional regulation: regulatory elements and their way of action	211
6. Regulación de la estructura de la cromatina y la expresión del grupo de genes $\alpha$ - y $\beta$ -globina a lo largo de la diferenciación eritroide	242



## Resumen

La estructura de la cromatina genera un impedimento estérico a los factores de regulación que interaccionan con el ADN. También juegan un papel activo en la regulación transcripcional. La participación que tiene en la función de elementos tipo enhancer, en particular en la actividad del enhancer  $\alpha$ -globina, y los eventos iniciales necesarios para acceder a sus secuencias, son temas en los cuales existen muchas incógnitas. GATA-1 es un factor transcripcional de unión al ADN, el cual se ha propuesto que participa en la abertura de la cromatina en elementos de regulación promotoras y enhancers. La mutación puntual de su secuencia “core” en el enhancer  $\alpha$ -globina de pollo, genera la pérdida en el efecto que muestra este elemento en la actividad de un gen reportero y el patrón nucleosomal gradualmente adquirido a través de la diferenciación y conservada en el desarrollo, se muestra alterado. También genera la pérdida de un grupo de sitios de hipersensibilidad (SH) a la dexosiribonucleasa I (DNasa I) formados entre las secuencias de unión al factor GATA y en la región que delimita al silenciador del enhancer. El patrón nucleosomal observado, por lo anterior, resulta indispensable para la funcionalidad del enhancer  $\alpha$ -globina. Este trabajo de tesis apoya un modelo en el cual GATA-1 estaría realizando un papel clave en la apertura de la cromatina y enucleación de factores transcripcionales y posiblemente de co-factores y actividades remodeladoras de la cromatina al enhancer  $\alpha$ -globina de pollo. GATA-1 también se encuentra modulando la actividad del enhancer al interaccionar con una región conservada de 120 pb, localizada 3' del enhancer. Por lo anterior y en base a los resultados obtenidos por nuestro grupo de trabajo y a la literatura, proponemos un modelo en el cual el factor transcripcional GATA-1 juega un papel de regulador maestro de la hematopoyesis.

## **Abstract**

The structure of the chromatin generates an esteric impediment to the regulatory factors that interact with the ADN. Also the chromatin plays an active role in the transcriptional regulation. Little is known concerning the chromatin structure contribution to enhancer's way of action. GATA-1 is a key erythroid transcriptional factor that also participates in the chromatin remodeling at erythroid promoters and enhancers. Point mutations of its recognition sequence in the chicken 3'  $\alpha$ -globin enhancer, causes the loss of its regulatory effect on a gene reporter and nucleosomal repositioning that is acquired gradually through differentiation and chicken development. We have also observed the loss of a group of DNase I hypersensitive sites located between the GATA-1 binding sequences and in the region that delimits the silencer to the enhancer elements. These results show us that specific nucleosomal positioning is needed for the proper and regulated  $\alpha$ -globin enhancer activity. This work supports a model, in which GATA-1 is playing a key role in the local chromatin remodelling and nucleation of additional transcriptional factors and co-factors required for chicken  $\alpha$ -globin enhancer regulated function. In addition, we have found that GATA-1 is capable, in collaboration with the transcription factor YY1, to modulate negatively the activity of the enhancer in early erythroid differentiation stages. Our results allow us to propose a model in which the transcriptional factor GATA-1 plays a central role, like a master regulator during haematopoiesis.

## **Abreviaturas**

acH3: Acetilación global de la histona H3

acH4: Acetilación global de la histona H4

ADN: Ácido desoxiribonucleico

ARN: Ácido ribonucleico

ARNm: Ácido ribonucleico mensajero

ADNc: ADN complementario

ChIP: Inmunoprecipitación de cromatina

CpG: Dinucleótidos C-G

CPT: Complejo de pre-iniciación de la transcripción

CTD: Dominio carboxilo-terminal de la RNA polimerasa II

ADNsa I: Desoxiribonucleasa I

FISH: Hibridación *in situ* por fluorescencia

HAT: Acetiltransferasa de histona

HDAC: Desacetilasa de histona

HDMT: Desmetilasas de histona

HMT: Metiltransferasas de histona

HP1: Proteína de heterocromatina 1

H3K9me3: Trimetilación de la lisina 9 de la histona H3

H3K27me: Metilación de la lisina 27 de la histona H3

H3K36me: Metilación de la lisina 36 de la histona H3

H4K20me3: Trimetilación de la lisina 20 de la histona H4

H3K4me2: Dimetilación de la lisina 4 de la histona H3

H3K79me2: Dimetilación de la lisina 79 de la histona H3

LCR: Región de Control del Locus

MNasa: endonucleasa micrococcal

PcG: Proteínas del complejo represor Polycomb

RNAi: RNA de interferente

SH: Sitio de Hipersensibilidad a la ADNsa I

RNA pol II: RNA polimerasa II

TAFs: Del inglés: “TBP-associated factors”

TBP: Del inglés: “TATA-binding protein”

TrxG: Proteínas del complejo activador Trithorax

3C: Captura conformacional de cromosomas

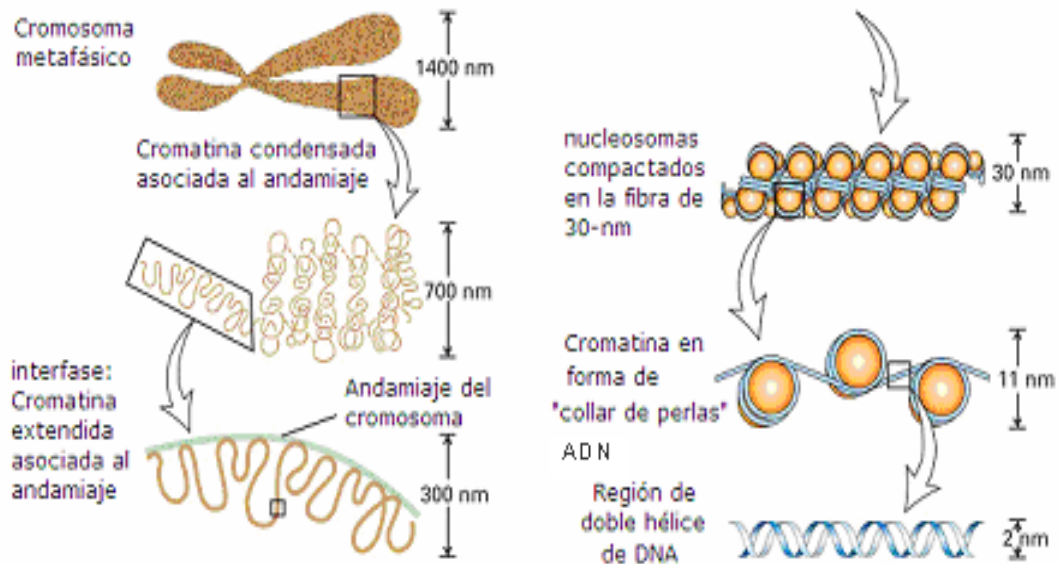
## **Introducción general**

### **La Regulación Transcripcional**

Las características que definen a una célula, vista como la unidad mínima en que están formados todos los seres vivos, están dadas por las proteínas que la constituyen, las cuales son el resultado de patrones específicos de la expresión de sus genes (4). A su vez, los genes están formados por un segmento determinado de la cadena de ADN portadora de la información genética, la cual codifica para un ARN mensajero, involucrado en la síntesis de una cadena polipeptídica, es decir, de una proteína. Un punto crucial en la determinación de los patrones de expresión, está dado por los factores transcripcionales que reconocen secuencias nucleotídicas en el ADN, principalmente en las regiones promotoras de los genes (4). La interacción de estos factores con sus secuencias reguladoras, desencadenarán eventos finamente regulados, que tienen como resultado la activación o represión de un gen ó grupos de genes. Tres tipos de elementos de regulación en *cis* se encuentran regulando la actividad: 1) Los promotores, correspondiente a la secuencia que se encuentra cercana al inicio de transcripción. 2) Los silenciadores, potenciadores (mejor conocido por su nombre en inglés, enhancer) y elementos reguladores del locus (LCR, del inglés “Locus Control Regions”), correspondiente a elementos de regulación que se pueden encontrar desde cientos de pares de bases a cientos de kilo-bases. 3) Los elementos frontera o aisladores (de su nombre en inglés “boundary elements” o “insulator” respectivamente), los cuales se ha propuesto que podrían encontrarse delimitando genes o grupos de genes y organizando al genoma en dominios transcripcionales independientes (14, 138).

Organismos superiores como la levadura y el humano están constituidos por células de origen eucariota; lo anterior significa que tienen su ADN compartimentalizado al interior de un núcleo. Este no se encuentra disperso, se estructura a través de interacciones proteína-ADN formando el complejo que conocemos como cromatina, la cual permite contener y organizar al genoma al interior del núcleo (Fig. 1). La unidad fundamental de la cromatina es el nucleosoma, el cual está constituido por el octámero de histonas (H3,H4, H2A y H2B, dos de cada una de las proteínas) alrededor del cual interaccionan 147 pares de bases (pb) de ADN. Esta estructuración en cromatina, trae como consecuencia un impedimento al libre acceso para las proteínas reguladoras que reconocen secuencias blanco en el ADN (4). Existen múltiples niveles de compactación del ADN siendo el cromosoma en metafase el nivel

más elevado (Fig. 1; 4, 5). Cabe mencionar que la gran mayoría de los eventos de remodelación de la estructura de la cromatina, se propone que ocurren a nivel de la fibra de 30 nm, también conocida como solenoide.



**Figura 1.** Esquema representando los distintos niveles de compactación de la cromatina, desde los cromosomas metafásicos, constituyendo el nivel más alto de compactación, hasta la doble hélice de ADN (Figura modificada de la cita número 6).

Evolutivamente la organización del genoma eucariota en cromatina permitió que éste pudiera ser contenido al interior del núcleo (7). Dado el elevado nivel de compactación del genoma, la estructura de la cromatina se vuelve un obstáculo para la actividad transcripcional, es decir, para la expresión de los genes. Por lo anterior, es necesario remodelar dicha estructura represora para permitir la expresión regulada de los genes. Consecuentemente, el entender como se expresa un gen depende en gran medida de la capacidad para remodelar de manera regulada la estructura de la cromatina tanto al nivel de sus regiones control como de sus secuencias codificantes y no-codificantes.

La complejidad adquirida por el genoma eucariota a través de la evolución ha requerido, a la par, el desarrollo de múltiples e inter-dependientes niveles de regulación, para poder llevar a cabo la expresión genética de manera ordenada y permitir el desarrollo de un organismo. En términos generales, la regulación transcripcional se ha dividido en 3 niveles (7). El primero se sitúa al nivel de secuencias nucleotídicas que se

encuentran codificada en la molécula de ADN. El segundo nivel involucra la estructura de la cromatina, donde la regulación de la expresión de un gen no solo depende de la información genética codificada en la secuencia del ADN, si no también depende de su organización en cromatina y la regulación epigenética<sup>1</sup> asociada a ésta (4). El tercer nivel involucra la organización espacial del genoma al interior del núcleo. A lo largo de esta introducción ampliaré los temas concernientes a los niveles en los cuales la regulación transcripcional es controlada, resultando de particular interés para este trabajo la parte relacionada a patrones de remodelación de la cromatina que llegan a formar regiones libres en nucleosomas y el papel que podría estar desempeñando el factor transcripcional eritroide específico GATA-1 via este tipo de regulación en la regulación del enhancer 3'  $\alpha$ -globina de pollo.

### **Regulación a nivel de secuencia**

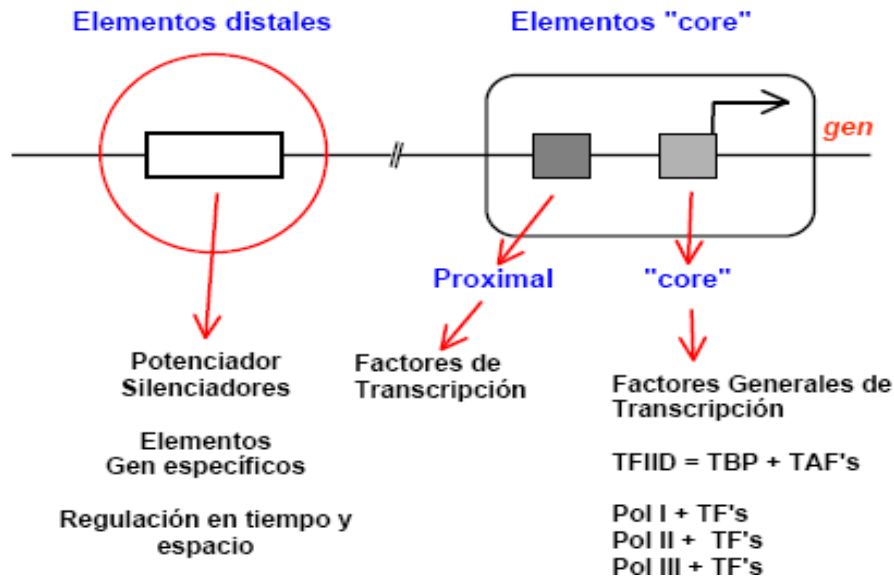
El desarrollo, crecimiento y sobre-vivencia de un organismo, requiere de la regulación y expresión coordinada de cientos de genes. La expresión de éstos es controlada a múltiples niveles y por una gran variedad de mecanismos, donde la cascada de eventos que preceden a la transcripción tiene como blanco final a los promotores y el sitio de inicio de la transcripción asociado a estos (Fig. 2) (8, 9, 10). La regulación al nivel de secuencia nucleotídica es la mejor estudiada, involucra elementos en *cis* tales como promotores, enhancers y silenciadores (entre otros). La función de estos elementos es dependiente de la atracción de factores transcripcionales que se unen a secuencias específicas en el ADN a través de interacciones proteína-ADN y proteína-proteína (13), entre co-factores y de un conjunto de moléculas cuya finalidad es remodelar la estructura de la cromatina (Fig. 2) (4, 7).

Las secuencias de regulación en *cis* generan complejidad genética, ocupando cerca de un tercio del genoma humano, controlando procesos como la expresión génica, la replicación, el apareamiento y la segregación cromosómica (14). En organismos tan simples como el erizo de mar, *Ciona intestinalis*, se estima que su genoma posee entre 10,000 y 20,000 enhancers tejido-específicos (14). Un locus típico en *Drosophila*, contiene varios enhancers (entre otras secuencias reguladoras como los delimitadores)

---

<sup>1</sup> Considerando a la regulación epigenética como los patrones heredables de expresión de un gen, que no pueden ser explicados por cambios en la secuencia del ADN (4).

distribuidos en regiones que alcanzan las decenas e incluso centenas de kilo bases (kb), asociados a regiones codificantes de tan solo 2 o 3 kb (15).



**Figura 2.** Esquema de los dos tipos de elementos de regulación que intervienen en el control de la expresión génica. El promotor está compuesto por el elemento “core” que se localiza cercano al sitio de inicio de la transcripción y donde encontramos a la caja TATA, secuencia a la cual se asocia TBP, TFIID, la RNA pol II y el resto del complejo de preiniciación e iniciación de la transcripción. En un intervalo de 50 a 200 pb se encuentra el elemento proximal donde se incorporan los factores transcripcionales tanto específicos como generales con la capacidad de formar puentes moleculares tanto con los TAFs del complejo TFIID o con los componentes del mediador asociado a la CTD de la RNA pol II. Los elementos distales pueden localizarse a grandes distancias del promotor, ya sea río arriba o abajo (10).

## Elementos de regulación en el ADN

### Promotores

De manera clásica, un promotor se define como una secuencia continua de ADN, suficiente para dirigir de manera regulada la correcta iniciación de la transcripción, mediada por la RNA pol II y su maquinaria transcripcional (8-10, 16-21). Estos elementos suelen encontrarse cercanos al sitio de inicio de la transcripción dentro de un rango que va de +20 a -30 en relación con el sitio de inicio de la transcripción (también conocido como el +1). Al elemento mínimo de un promotor se le denomina el elemento “core” (Fig. 2).

Por una herramienta experimental reciente nombrada CAGE-base (del inglés “cap analysis gene expresión”; 23), que permite el análisis de los perfiles de inicio de la



transcripción, se encontró en ratón y humano que los promotores suelen carecer del inicio clásico de la transcripción, localizado en una posición precisa. Este mismo grupo reporto (23), que la estructura típica de un promotor debería de ser vista como un arreglo de distintos inicios de transcripción repartidos en una región que va de las 50 a las 100 pb (22, 24). En estos estudios se determinó que la longitud en pb aproximada de los promotores se debería encontrar en un rango que va de las 71 a las 150 pb (24). Por lo anterior definiremos al “core” del promotor como la región que se encuentra enmarcando el cluster de inicios de la transcripción, suficiente para reclutar el complejo de iniciación, de manera regulada y dependiente de señales extra- e intra-celulares (22).

En las secuencias promotoras correspondientes al elemento “core”, existen diferentes motivos consenso de unión al ADN (17, 18), entre los más comunes encontramos a la caja TATA, elementos iniciadores (Inr), elementos de reconocimiento para el factor TFIIB (BRE) y elementos promotores río abajo (DPE). Estos motivos de unión al ADN, suelen encontrarse en secuencias promotoras, ya sea de manera individual ó en diferentes combinaciones ó acompañadas de otras secuencias (8, 17, 18). Este tipo de combinaciones constituyen lo que se conoce como la "arquitectura" del promotor la cual permite establecer una combinatoria única y selectiva para el reconocimiento de moléculas reguladoras y por lo tanto, para lograr una expresión génica regulada.

### ***Las islas CpG y los promotores***

Los genes pueden ser clasificados por sus patrones de expresión, ya sea por su tejido-especificidad o por su expresión ubicua (25). Las cajas TATA suelen encontrarse en promotores que presentan un solo inicio de la transcripción, estos promotores suelen presentar patrones de expresión tejido-específico mientras que las islas CpG<sup>2</sup> suelen encontrarse relacionadas a promotores con varios sitios de inicio de la transcripción, estos suelen presentar patrones de expresión más ubicuos (22) y codifican para proteínas extracelulares (25). Los promotores que contienen islas CpG, han sido asociados con genes de rescate (del inglés, “housekeeping genes”) (25). Desde el punto de vista de la

---

<sup>2</sup> Las islas CpG son ricas en Guaninas y Citocinas (GC) y en ellas el dinucleótido CpG se encuentra sobre representado. Estas islas se encuentran generalmente desmetiladas (21). Sus tamaños suelen variar de los 0.5 a las 2 kb. Las islas CpG suelen carecer de cajas TATA o DPE, pero contienen múltiples motivos ricos en GC a los cuales se puede unir Sp1 (21).

arquitectura de los promotores, las islas CpG contienen múltiples promotores débiles, en lugar de uno fuerte (21). El análisis de 1031 genes en humano reveló que aproximadamente la mitad de ellos, debería estar relacionado con Islas CpG (21, 26).

### ***Elemento proximal***

Inmediatamente arriba del “core” del promotor de -50 a -200 pb en relación al inicio de la transcripción, se han localizado secuencias específicas de unión a factores de transcripción. El elemento proximal se ha propuesto como un grupo de secuencias que le otorga tiempo y tejido especificidad a la actividad del promotor (20). El contacto molecular que se establece entre los elementos proximales y el “core”, esta dado por factores transcripcionales específicos y generales, que se unen a las secuencias proximales y forman puentes vía el mediador<sup>3</sup> (27). En animales superiores, la expresión tejido-específica suele ser más compleja y requerir de elementos tipo enhancers y/o LCR, que son secuencias que actúan a grandes distancias e incluso en *trans*, regulan positivamente a sus promotores (28, 17).

### **La RNA polimerasa II y los factores generales de transcripción**

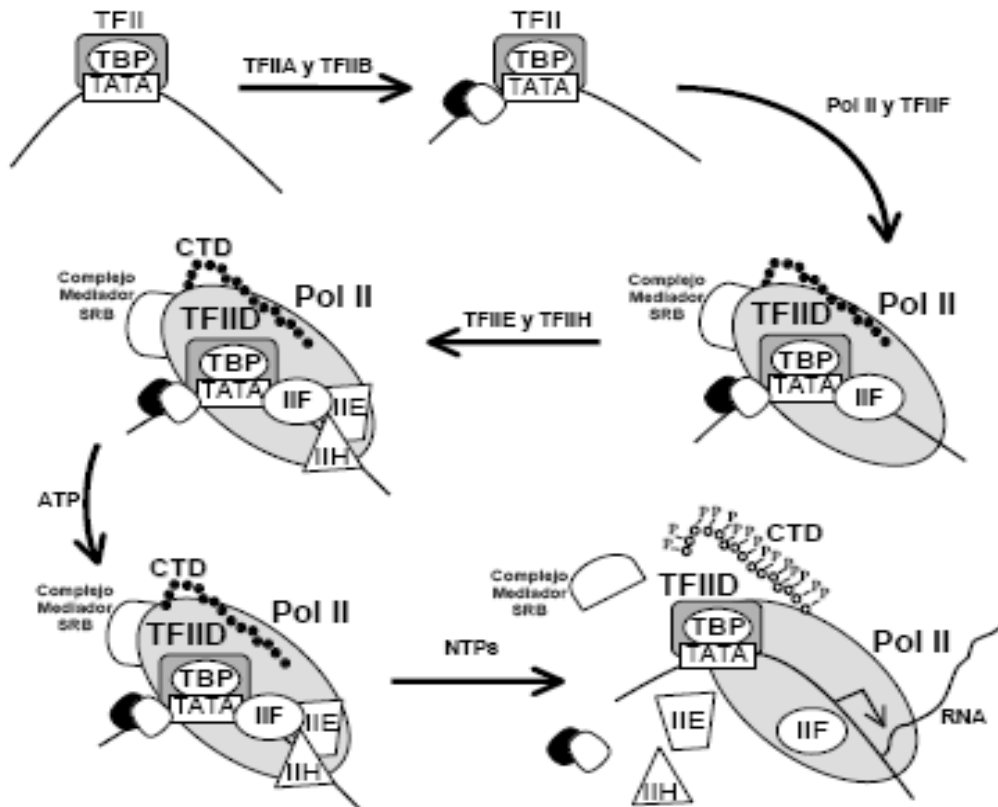
Una adecuada y eficiente transcripción, requiere de la RNA polimerasa II y un conjunto de factores peptídicos auxiliares, mejor conocidos como factores “generales” o “basales” de transcripción (TFs), entre los cuales se encuentran los complejos: TFIIA, TFIID, TFIIE, TFIIIF, TFIIF y TFIIB (9, 16, 43). Estos se unen a las secuencias promotoras, de manera secuencial para formar lo que conocemos como el complejo de pre-iniciación de la transcripción (CPT; Fig. 3).

Uno de los principales paradigmas en regulación transcripcional, se encuentra en comprender la forma en que un gen puede expresado de manera específica en tejidos y etapa del desarrollo. El descubrimiento de diferentes secuencias promotoras y tres tipo de RNA polimerasa (I, II y III), localizadas en distintos compartimentos de la célula, ayudó a comprender que la especificidad necesaria para el control fino de la regulación transcripcional, está dada por combinaciones particulares en la composición de la

---

<sup>3</sup> El elemento mediador, se define como un complejo protéico que reconoce al enhancer y las secuencias río arriba del promotor por ser activado (38).

maquinaria basal de transcripción y factores transcripcionales tanto específicos como generales (8, 16, 43). Cabe mencionar que Orphanides y Reinberg han propuesto la teoría unificadora para el control de la expresión génica donde se propone que los procesos que llevan a la activación de un gen son secuenciales e inter-dependientes y que ocurren a múltiples niveles (44).



**Figura 3.** Modelo secuencial del ensamblaje del complejo de pre-inicio de la transcripción (CPT) por parte de la RNA pol II. A través de la unión de TBP a la caja TATA se forma el complejo TFIID con la incorporación de un grupo de proteínas conocidas como TAFs, algunas de las cuales pueden ser incluso tejido específicas. Una vez asociado el complejo TFIID se incorporan TFIIA y TFIIB cuya función es la de determinar la distancia y orientación a la cual la RNA pol II se posicionará en el “core” del promotor. Posteriormente la RNA pol II escoltada por el complejo TFIIF para favorecer su incorporación al complejo de pre-iniciación de la transcripción. El siguiente paso es la incorporación de los complejos TFIIE y TFIIH, este último de suma importancia, dado que en presencia de ATP y nucleótidos, TFIIH fosforila a la región CTD de la RNA pol II permitiendo el proceso de alargue y por lo tanto la síntesis del RNA mensajero (10, 16, 43, 44).

La existencia de tres tipos de RNA polimerasas, con funciones distintas, generó una primera división en grupos de genes, que dependen del tipo de RNA polimerasa por la que son transcritos (10, 16): clase I, genes que codifican para las sub-unidades largas

de los RNA ribosomales (RNAr). Estos son transcritos por la RNA polimerasa I, la cuál se encuentra localizada en el nucleolo. Clase II, genes que codifican para algunos RNAs nucleares pequeños, y principalmente para los RNAs mensajeros (RNAm) que codifican para proteínas estructurales. Estos son transcritos por la RNA polimerasa II. Clase III, genes que codifican para RNA estructurales pequeños y componentes de la maquinaria de procesamiento o "splicing" de los RNAs pre-mensajeros. Estos son transcritos principalmente por la RNA polimerasa III.

Como se mencionó anteriormente, las polimerasas se asocian con distintos factores generales de transcripción los cuales forman el complejo CPT compuesto por más de 100 proteínas (8, 16, 43). En particular, para el grupo de genes transcritos por la RNA pol II, al menos se requiere de 6 factores generales de transcripción, que en conjunto formarán lo que conocemos como el CPT (Fig. 3). La asociación estable de la RNA pol II con la secuencia promotora, requiere de la unión previa de TFIID (en particular TBP, del inglés: "TATA-binding protein") y sus factores asociados conocidos como TAFs (TFIIB y TFIIF). Sin embargo para que la transcripción se pueda llevar a cabo se necesita de los factores transcripcionales TFIIE y TFIIH (43). El inicio de la transcripción requiere del previo reconocimiento de la caja TATA, por la proteína TBP y su unión genera una distorsión en la molécula de ADN (Fig. 3). Este complejo formado por TBP y el ADN constituye el blanco que reconocerá TFIIB. Se ha propuesto que TFIIB no reconoce una secuencia en específico, sino la curvatura generada por la interacción de TBP con la caja TATA. Se propone que la unión de TFIIB podría estar indicando la direccionalidad del promotor (16, 43). La asociación de la RNA pol II al complejo ADN-TBP-TFIIB, requiere de la interacción de TFIIF con la polimerasa, su función será la de asistir el reconocimiento del promotor. Lo anterior es el resultado de la capacidad de TFIIF para interactuar con TFIIB, la RNA pol II y el ADN. Cabe señalar que TFIIF se posiciona río debajo de la caja TATA (Fig. 3; 16, 43). Una vez que se ha establecido la unión de la RNA pol II a la región promotora, se necesita de la unión de los factores TFIIE y TFIIH para lograr el proceso conocido como "escape" del promotor (Fig. 3). Este evento permitirá liberar a la RNA pol II de la secuencia promotora e inducir el inicio de la transcripción, en un proceso mejor conocido como elongación (Fig. 3). El complejo multiprotéico TFIIH resulta ser clave, dado que es el único factor general de transcripción que presenta una actividad enzimática necesaria para el inicio de la transcripción, que consiste en la fosforilación de la región carboxilo

terminal de la RNA pol II, conocida como dominio CTD (16, 43, 45). El complejo TFIIF además de incluir a la cinasa responsable de fosforilar a la región CTD de la RNA pol II, tiene la capacidad de reparar el ADN mediante el proceso conocido como reparación por escisión de nucleótidos (45). Finalmente debemos recordar que en su contexto endógeno, la transcripción ocurre en presencia de un templado cromatinizado y que frecuentemente requiere de elementos de regulación a distancia.

### ***Especificidad de los enhancers por sus promotores***

Los promotores no solo están involucrados en la correcta iniciación de la transcripción, éstos también deberán estar participando en generar especificidad por los enhancers (21). ¿Pero donde estará la base para generar esta especificidad enhancer-promotor? Se ha demostrado que elementos de reconocimiento en los promotores (ej. cajas TATA), pueden ser substituidos por otros (ej. DPE) y seguir siendo activados por sus enhancers (21, 27). Se ha encontrado que solo un reducido número de enhancers pueden discriminar entre un promotor con caja TATA y uno con caja DPE (21, 27). En *Drosophila* los enhancers AE1 e IAB5 pueden funcionar con promotores sin caja TATA, pero muestran preferencia por promotores que contienen la caja TATA (21, 30). Una de las explicaciones que se ha dado para revelar esta especificidad, introduce el término de estructura del promotor<sup>4</sup> como factor que esté dictando especificidad (21, 31-34).

La especificidad también puede ser dada por la competencia entre múltiples promotores por un enhancer, donde el que sea él seleccionado excluirá de su interacción a los otros (33, 35, 36). Esta propuesta es claramente ilustrada en el dominio  $\beta$ -globina de pollo con los genes  $\epsilon$  y  $\beta^A$ . Durante el desarrollo embrionario, en estadios tempranos el gen  $\epsilon$  interaccionará con el enhancer  $\beta^{A/\epsilon}$  siendo activado, pero en estadios adultos será el gen  $\beta^A$  quien interaccionará con el enhancer, gracias a la unión de un factor transcripcional en la región cercana al promotor, resultando en la expresión de este gen y apagando la expresión del gen  $\epsilon$  (36, 37). Este ejemplo muestra de una manera muy

---

<sup>4</sup> Referido a los sitios de unión a factores transcripcionales en un promotor, la distancia que hay entre ellos y el contexto en que se encuentran (39).

clara como la especificidad de un enhancer por su promotor, puede ser determinada por la expresión de un factor estadio- y tejido-específico.

Resulta sencillo de imaginar pero difícil de comprobar, como un código próximo al promotor donde una combinación específica de factores se ensamblan para que un tipo específico de enhancers sea reclutado, mientras que otra combinación reclutará un enhancer distinto (40). En algunos casos las secuencias reguladoras incluso se encontrarán cercanas al promotor equivocado (40). En conclusión, existen tres mecanismos generales que explican la selectividad que presentan los enhancers por sus promotores (17).

1. La especificidad será dictada por las proteínas que se unan a éste; co-factores y factores transcripcionales (TAFs, del inglés: “TBP-associated factors” y componentes del mediador), las cuales también interaccionan con el promotor.
2. Los elementos frontera (o “insulators”) se encuentran interfiriendo con todas aquellas interacciones que no son deseadas.
3. Por competencia entre distintos promotores por un mismo enhancer.

## **Enhancers**

Históricamente los potenciadores (del inglés: enhancers) fueron definidos como secuencias virales y posteriormente celulares, que incrementan la expresión de un gen al que se encuentran asociados, de una forma independiente a su orientación y a grandes distancias en relación con el inicio de transcripción, incluso mega bases (27, 17, 31, 82). La identificación de éstos fue lograda vía aproximaciones genéticas, basadas en mutaciones que alteraban la expresión de los genes homeóticos del complejo bitorax (27, 42). Los enhancers son elementos discretos de ADN que contienen motivos específicos de unión a factores transcripcionales que interaccionan y transmiten las señales necesarias para lograr la expresión génica (17). Presentan una estructura modular, en términos de sus motivos de unión a factores transcripcionales (27), que suele tener entre 50 pb y 1.5 kb (17). Los enhancers suelen activar a los promotores que tienen más cercanos, sin embargo existen reportes, en los cuales los enhancers ignoran ciertos promotores, e interaccionan con otros más distantes (47).

## **Modelos de acción**

Con la aparición de elementos de regulación, surgió un gran interés por comprender sus mecanismos de acción. Como es de esperarse, para un elemento del cual aún se sabe poco, se han propuesto múltiples niveles de acción y mecanismos por los cuales podría estar actuando, algunos de los cuales, se discutirán a continuación.

### ***Enhancer y cromatina***

Una pregunta central en la biología del desarrollo, es ¿cómo un enhancer activa transcripcionalmente a un gen de forma tejido- y tiempo-específico, en el contexto de la cromatina? Se han sugerido que los enhancers podrían estar contrarestando, el efecto represor de la cromatina y sus proteínas asociadas (17) y de esta forma facilitar la interacción de la maquinaria basal transcripcional (27, 31). Como ejemplo de lo anterior, la estructuración en cromatina activa, accesible al corte por endonucleasas del promotor  $\epsilon$ -globina, depende de la presencia del enhancer HS2 de humano (164, 133). Alternativamente, los enhancers podrían estar promoviendo el ensamblaje de los complejos transcripcionales a un ambiente heterocromatinizado, generando ensamblajes estables y productivos, pero no participando directamente, en la remodelación de la estructura de la cromatina (17). Ambos mecanismos parecen ser relevantes *in vivo* (27).

### ***Modelo de prendido o apagado versus el modelo progresivo***

También se ha propuesto que los enhancers actúan incrementando la probabilidad para que un gen sea activado en el tiempo y tejido específicos (17). Lo anterior significa que el enhancer estaría logrando estadísticamente, que un mayor número de células activen transcripcionalmente a un gen, modelo conocido como de "prendido y apagado" (Fig. 4; 9, 17). De manera alternativa, el modelo progresivo sugiere que los enhancers aumentan los niveles transcripcionales de manera homogénea en todas las células (Fig. 4; 9, 17). Evidencias experimentales apoyan ambos modelos (9, 17, 19, 54).

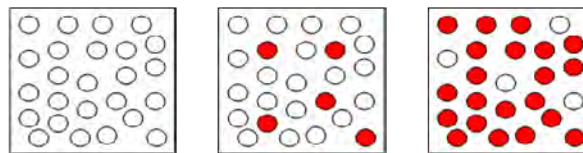
### ***Modelo del enhanceosoma contra "billboard" (pizarra)***

Uno de los problemas centrales en la regulación génica, se encuentra en poder explicar como un gen o grupo de genes durante el desarrollo y diferenciación, o en respuesta a

señales ambientales, pueden ser seleccionados para ser activados (49). El modelo del “enhanceosoma” propone que la secuencia del enhancer servirá como un blanco, en el cual se podrá ensamblar un complejo de proteínas nucleares. Una modificación en un sitio de unión a proteínas del enhancer, traería como consecuencia un efecto drástico para la actividad del enhancer (49-52). Este modelo toma en cuenta el papel cooperativo que debe de existir entre las proteínas de unión al enhancer y los co-factores asociados. En un segundo modelo mejor conocido como “billboard”, la posición de los sitios de unión a factores transcripcionales dentro del enhancer, es más flexible, la unión de los factores proteicos será un evento secuencial (52). En este modelo, el enhancer en diferentes configuraciones, podrá interactuar con la maquinaria basal de la transcripción (52). La interacción separada de cada uno de los factores, cooperará con la arquitectura funcional del enhancer. Por lo tanto, el enhancer no estará trabajando como una unidad, sino de manera cooperativa a partir de un conjunto de actividades individuales (52).

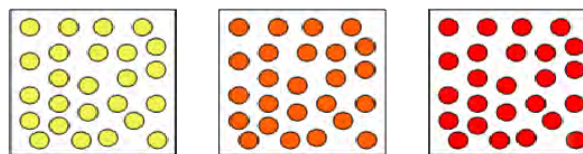
#### **A Modelo de prendido y apagado (on/off)**

El “enhancer” aumenta la probabilidad de que un gene sea transcrito, pero no afecta los niveles de transcripción en una célula



#### **B Modelo progresivo**

El “enhancer” uniformemente aumenta los niveles de transcripción en una célula de manera continua



**Figura 4.** Modelo de prendido ó apagado (on/off) contra modelo progresivo para la activación transcripcional mediada por los enhancers. **A**, los genes se encuentra en cualquiera de los estados, prendido o apagado. Los enhancers aumentaran la probabilidad de sus genes designados, de ser transcritos, pero no afectan los niveles de la transcripción. La fracción de las células en las cuales se activa el gene puede reflejar la fuerza del enhancer. **B**, en el modelo progresivo, los genes son activados uniformemente por los enhancers. Los niveles transcripcionales son proporcionales a la fuerza del enhancer (17).



En apoyo a este modelo, el enhancer  $\alpha$ -globina de pollo presenta una región conservada entre pato y pollo de 120 pb (232), la cual por sí sola no presenta ninguna actividad, y en el contexto del enhancer, parece estar modulando la actividad de éste, dependiendo del estadio eritroide de la célula (50). También en un estudio de interacciones proteína-ADN (“footprint”) del enhancer TRC $\alpha$  se encontró que en diferentes estadios incluyendo en aquellos en que este enhancer no estaba activo, se observaba la interacción de un complejo (53). Una pequeña alteración en la estructura del complejo de pre-ensamblaje del enhancer, generaba una respuesta regulatoria rápida (53). Por lo tanto, la conformación de un enhancer según el modelo de “billboard”, estaría facilitando respuestas rápidas y específicas para la expresión de ciertos genes.

En suma, ambos modelos parecen estar describiendo enhancers con diferentes funciones. El modelo del “enhanceosoma” se estaría ajustando a los enhancers de respuesta rápida mediante un complejo pre-ensamblado, como aquellos que participan en procesos de inflamación o infecciones (52). El modelo “billboard” parece estar describiendo a un enhancer que regula la actividad de un gen a lo largo del desarrollo, donde sus sitios de unión y el intercambio entre las proteínas que lo reconocen a través del desarrollo, se comportan de una forma dinámica (52).

### ***Modelo en tres dimensiones***

Para explicar el funcionamiento de los enhancers la mayoría de los modelos proponen una comunicación entre este y el promotor, ya sea por interacciones proteína-proteína (looping model), por el escaneo de la maquinaria transcripcional del enhancer hasta encontrarse con el promotor (scanning model), generando un dominio vía modificaciones de la cromatina y proteínas facilitadoras entre el enhancer y el promotor (linking model) (59), o finalmente por la generación de transcritos intergénicos no-codificantes que se originan a partir del enhancer hasta el promotor, facilitando la apertura de la cromatina (55). Estos modelos son abordados en detalle cuando se define a los LCR (ver mas adelante).

Los modelos de interacción a distancia, pueden ser clasificados en dos grandes categorías. Las que describen un contacto físico entre el promotor y el enhancer y los modelos que no (82). El modelo de contacto propone la formación de interacciones

estables ADN-proteína entre el enhancer y el complejo peptídico asociado a la vecindad ó sobre el promotor (56). Esta asociación podría generar una especie de memoria que facilite las ondas consecutivas de activación (56). Complementariamente, se ha propuesto que el super-enrollamiento del ADN debería estar ayudando a generar estas interacciones, acercando al promotor y al enhancer (56). Studitsky y colaboradores demostraron que este súper-enrollamiento ayudaba a acercar al enhancer a su promotor (teniendo un efecto sobre la transcripción) (56).

El modelo clásico de contacto propone que las interacciones están dadas por la afinidad proteína-proteína y de los complejos proteína-ADN formados en cada elemento, generando un asa entre estos dos elementos (59, 57). Este contacto se podría dar por colisiones al azar, donde interacciones específicas y reproducibles proteína-proteína estabilicen esta asa (59, 58). El problema que presenta este modelo, es que no estaría de acuerdo con la función que presentan los “insulators” de bloqueo de enhancer<sup>5</sup>. Un modelo alternativo asume que proteínas reguladoras (factores transcripcionales, co-activadores y la RNA pol II) serán primero reclutadas al enhancer, y posteriormente se deslizarán a lo largo del ADN hasta encontrar a las secuencias blanco. Este modelo si se ajusta al modelo de actividad de bloqueo (17, 59).

La formación de esta asa, traería como consecuencia la formación de un complejo transcripcional, también conocido como holocomplejo, que se ha propuesto, contribuye facilitando el proceso de iniciación y elongación en la que participa la RNA pol II (55, 17). Esta interacción podría generar modificaciones al nivel de la estructura de la cromatina y de una manera regulada contrarrestar el efecto represivo por parte de ésta. Otra alternativa es la de facilitar ó atraer diferentes actividades enzimáticas a la región promotora, entre ellas, modificar post-traduccionalmente las entidades a las que son reclutados (61) como, por ejemplo, la acetilación de histonas para relajar la estructura de la cromatina y favorecer la formación funcional del complejo de pre-iniciación de la cromatina (4, 17, 18, 55).

Entre los modelos clásicos que proponen que una interacción promotor-enhancer no se llevara a cabo, Reik y colaboradores (2004) proponen un modelo de formación de asas que estaría localizando a las secuencias por ser activadas a regiones diferentes del

---

<sup>5</sup> Propiedad de bloquear la actividad de un enhancer, cuando se localizan entre el enhancer y su promotor blanco (60).

núcleo<sup>6</sup> (62). Esto lo propuso para el caso del locus *Igf2/H19*, donde se sabe que uno de estos genes se impronta<sup>7</sup>, la interacción y formación de asas mandaría al gen por ser improntado a regiones represivas del núcleo y al gen por ser activado, cerca de activadores transcripcionales (53, 63). Alteraciones en la posición de un gen, en relación a su cercanía con regiones de hetero o eucromatina, pueden afectar su actividad (64). Se ha demostrado que el re-localizar a un gen de una región de eucromatina hacia un compartimiento sub-nuclear represivo cercano a regiones pericentroméricas de heterocromatina, genera un efecto silenciador (64). Por ejemplo, en levadura el silenciamiento del locus del mating-type, está asociado a su re-localización a regiones de la periferia nuclear donde proteínas de la familia Sir se encuentran concentradas (64). Más ejemplos de este último modelo, serán retomados en la sección relacionada con la dinámica nuclear.

### **Los elementos de Control del Locus ó LCR**

Al igual que los enhancers, las regiones de control del locus (LCR del inglés: "Locus Control Region") contienen múltiples sitios de unión a activadores transcripcionales. Estos elementos tienen una actividad tejido-específica, y sus blancos son genes que necesitan ser expresados en una determinada etapa del desarrollo o diferenciación (65). Un enhancer está típicamente caracterizado por la presencia un SH, mientras que un LCR está formado generalmente por grupos de SH en zonas del genoma donde la estructura de la cromatina es más laxa y por lo tanto la unión de factores transcripcionales se ve favorecida. La organización de cada uno de estos sitios de hipersensibilidad es modular (55, 66). Los LCRs están constituidos por un conjunto de SH con dos funciones principales: 1) poseen una actividad tipo potenciador y 2) fomentan la apertura de la estructura de la cromatina al nivel de un dominio (17, 55, 66, 67). Estudios, donde se ha encontrado la pérdida de la región cromosómica que comprenden al LCR y las secuencias río arriba de éste, reportan un efecto dramático en la expresión de los genes que regula y se observan una estructuración en cromatina que resulta resistente al corte por la ADNsa I (177). Por lo mismo, lo anterior se ha

---

<sup>6</sup> Es decir, que la arquitectura nuclear estaría jugando un papel fundamental en la regulación transcripcional (ver mas adelante). La participación que la estructura nuclear podría tener un efecto directo sobre la regulación transcripcional (punto que se encuentra actualmente en debate).

<sup>7</sup> La impronta genómica es un fenómeno genético y epigenético por el que ciertos genes son expresados de manera monoalélica.

interpretado como que el LCR es el responsable de generar la des-compactación de la cromatina del dominio  $\beta$ -globina de humano y ratón, y de favorecer la activación transcripcional de sus genes en diferentes etapas del desarrollo (133).

Funcionalmente, los LCR tienen la habilidad de mantener la expresión de un transgene de manera independiente al sitio de integración y dependiente del número de copias del transgene (68). En base a esta última característica, se menciona que los LCRs tienen una actividad dominante de apertura de la cromatina (178). Sin embargo diferentes grupos han reportado que el LCR no es completamente dominante sobre el efecto de posición tanto en cultivos primarios como en ratones. Incluso construcciones que contienen las 150 kb que incluyen al dominio  $\beta$ -globina, se ven sujetas al efecto de posición en ratones transgénicos (179, 227, 251). Aunque aun falta mucho por entender, un aspecto que apoya la relevancia de este tipo de secuencias, es el número creciente de dominios que presentan un LCR entre sus elementos de regulación (66, 67).

### **Modelos de acción de los LCRs**

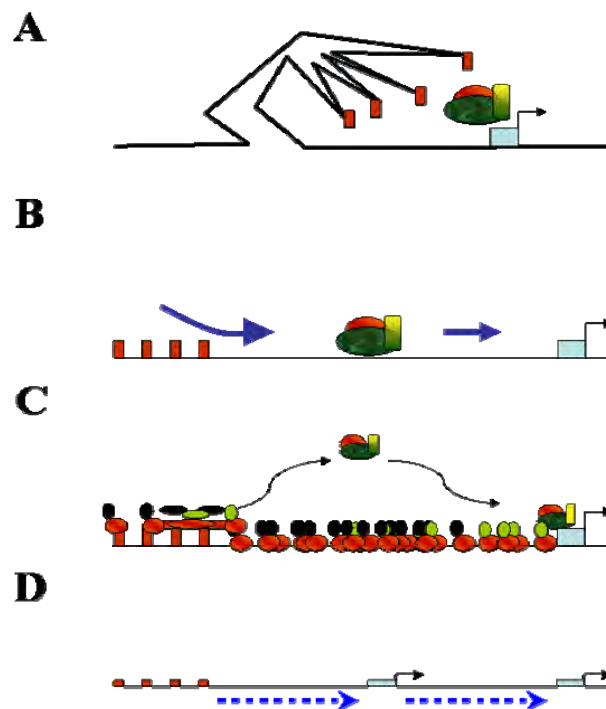
Actualmente y derivado de varios años de trabajo por un número considerable de grupos de investigación se ha propuesto cuatro modelos que tratan de explicar el funcionamiento de los LCRs:

#### ***Modelo de formación de asa o “Looping”***

Este modelo favorece la idea de interacciones físicas entre el LCR y los promotores individuales de los genes de manera regulada a lo largo del desarrollo (Fig. 5A). El LCR induce la formación de un holocomplejo para estimular la transcripción de genes individuales dentro del *locus* y a lo largo del desarrollo acercando al promotor del gen para activar formando un asa (o múltiples asas) que contribuye aumentando la concentración de la maquinaria transcripcional de manera específica (55).

En este modelo la organización del genoma en asas cromatínicas, podría permitir que elementos que se encuentran a grandes distancias se aproximen e incluso se asocien. Lo anterior se ha propuesto y confirmado en base a un estudio denominado captura

conformacional del cromosoma (3C), que permite experimentalmente demostrar que dos secuencias genómicas (linealmente distantes) se encuentran en contacto o al menos físicamente cercanas (69, 70). Mediante una estrategia experimental alternativa, llamada RNA-TRAP, se obtuvo la primera evidencia *in vivo* de la interacción entre el LCR y el locus  $\beta$ -globina (252). Estos experimentos muestran que el LCR (HS1-HS4) se encuentra físicamente cercano a los genes activos del dominio  $\beta$ -globina en células de ratón, siendo HS2 la región que se encuentra más cercano. Los genes que no se encuentran activos, por el contrario están distantes (252).



**Figura 5.** Modelos de interacción de elementos distales con promotores. **A** Un elemento a distancia (cubos rojos) a través de la formación de un asa, interacciona con su promotor blanco. **B** El enhancer y la maquinaria asociada a el rastrea sobre la cromatina hasta localizar el promotor blanco. **C** Proteínas facilitadoras ayudarían al enhancer a encontrar a su promotor blanco. **D** Los elementos distales, participa en la apertura de la cromatina, vía la generación de transcritos intergénicos (flechas azules).

### ***Modelo de encarrilamiento o “Tracking”***

En este modelo, se propone que diversos factores transcripcionales específicos y generales, así como co-factores, reconocen las secuencias presentes en el enhancer o LCR formando un complejo de activación (Fig. 5B). Dicho complejo migraría o se “encarrilaría” a lo largo de cromatina hasta topar con la maquinaria basal de la

transcripción, que estaría situada en el promotor en turno, cercano al sitio de inicio de la transcripción (58).

### ***Modelo de relevo o "Linking"***

Este modelo (Fig. 5C), sugiere una cadena o relevo de interacciones moleculares, que fomentan la formación de un "puente" molecular acercando al LCR al promotor. Esto se logra a través de interacciones entre factores expresados en un estado de diferenciación específico, que se localizan a lo largo de la fibra de la cromatina, facilitando la formación de la estructura de la cromatina de manera específica y regulada para cada gen del dominio (55).

### ***Transcritos intergénicos***

En el dominio  $\beta$ -globina humano, se han identificado transcritos intergénicos, es decir, en regiones no-codificantes (que tomaron como molde el ADN situado entre los genes y regiones 5' y 3' no codificantes del locus). Aunque se desconoce su función, se ha demostrado una correlación entre transcritos no-codificantes y la formación de la cromatina abierta. Se cree que estos podrían ser los responsables de contrarrestar el efecto represor de la cromatina compacta (Fig. 5D, 55). El dominio  $\beta$ -globina de pollo, fue uno de los primeros donde se describió la existencia de transcritos intergénicos de regiones no-codificantes (55).

Diferentes estudios han mostrado la existencia de un mecanismo de transcripción que inician en el HS2 del LCR  $\beta$ -globina de pollo, se extiende a lo largo de la región comprendida entre el elemento de regulación y el promotor a ser activado (253). En un estudio en ratón, con el fragmento HS2 integrado de manera aleatoria, encontraron que este sitio es capaz de iniciar la transcripción de manera autónoma y generar los SH, es decir, sin la necesidad de un promotor río arriba o abajo. Este resultado es consistente con la capacidad propuesta para los enhancers de generar SH a la ADNsaI de manera autónoma (255). En un trabajo realizado por el grupo de Tuan (2005), encuentran que el HS2 ensambla al complejo de transcripción de la RNA pol II, el cual puede mediar la actividad a distancia del LCR, sintetizando un transcrito intergénico no-codificante a través del ADN que separa al promotor de su enhancer. El complejo transcripcional que

se une a este enhancer HS2 parece ser distinto al que se une a los promotores (253). El HS2 presenta diferentes inicios de transcripción, incluso en regiones río abajo, a diferencia de los RNA generados en la secuencias codificantes que muestran en sus extremos 5' una caperuza (del inglés "cap"), el transcrito generado en el HS2 suele contener de 1 a 5 bases extras (es decir que no están codificadas en el templado de ADN) las cuales no contienen un cap, sugiriendo un mecanismo distinto en el procesamiento del extremo 5' por el complejo de la RNA pol II (253).

Finalmente, se encontró que el HS2 del LCR  $\beta$ -globina es transcrito en células eritroides y no en células no eritroides (256). En el contexto endógeno y con el uso de transgenes, se ha encontrado que el HS2 inicia la transcripción principalmente en la dirección de los genes localizados río abajo. Incluso la colocación de un terminador insertado en la secuencia de ADN localizada entre el LCR y los genes, reduce drásticamente la actividad del LCR (256). La acetilación de histonas y la dimetilación de H3K4 correlacionan con la presencia de los transcritos (257). La inhibición de la iniciación de la transcripción por la pol II vía mutación de la caja TATA en el promotor  $\epsilon$ -globina trae como resultado la pérdida en los patrones de modificación de las histonas. Sin embargo la transcripción intergénica y la acetilación de las regiones intergénicas no se ven afectada, indicando que éstas no son consecuencia de la transcripción mediada por el promotor (257).

La expresión de un gen constituye un evento biológico muy preciso y sofisticado donde ha sido necesario crear sistemas cada vez más complejos de control para lograr en tiempo y en espacio la transcripción de un gen. Sorprendentemente, los elementos de regulación a distancia no son los únicos responsables, el ensamblaje de complejos de iniciación de la transcripción en las zonas cercanas al inicio de ésta resulta fundamental y añade otro nivel de sofisticación.

## **Silenciadores**

Los silenciadores se han definido como elementos a los que se unen factores nucleares que tienen un efecto negativo sobre la transcripción de manera independiente de su posición y orientación en relación con el inicio de la transcripción (72, 73). Estos poseen una organización modular al igual que los enhancers. Además de los silenciadores, se han definido secuencias denominadas elementos reguladores negativos,

cuya actividad es dependiente de la posición en relación al promotor. Los mecanismos de acción de ambos elementos serían diferentes pues los silenciadores clásicos se propone que estarían interfiriendo con el ensamblaje de la maquinaria transcripcional (73). Por otra parte los elementos reguladores negativos estarían reclutando proteínas que interferirían con la unión de otros factores transcripcionales a su sitio de unión al ADN. Esto ocurre en elementos de regulación del costado 5' afectando procesos de la regulación post-transcripcional como la edición de intrones, poliadenilación, sitios de inicio de la traducción, la elongación y estructura de la cromatina (74).

### **Delimitadores o “insulators”**

El descubrimiento de elementos enhancer que pueden actuar a grandes distancias, trajo consigo múltiples cuestionamientos en cuanto a la forma en que su actividad era restringida a un gen o dominio (ver más adelante) en específico (77). En años recientes se ha propuesto que los genes transcripcionalmente activos se encuentran organizados en regiones de la cromatina “abiertas”. A estas regiones se les conoce como dominios cromosómicos transcripcionalmente activos (66). Se postula que contienen de uno a varios genes, estos deberán tener límites bien definidos, e incluir a los elementos de regulación necesarios para llevar a cabo la regulación transcripcional de sus genes (Fig. 6; 76). Pero funcionalmente, ¿cómo es que los elementos de regulación de dominios transcripcionalmente activos no alteran de manera inespecífica la expresión de dominios adyacentes? y estructuralmente ¿cómo es que dominios con estructuras de la cromatina antagónicas (heterocromatina *versus* eucromatina) adyacentes no se afectan una a la otra?

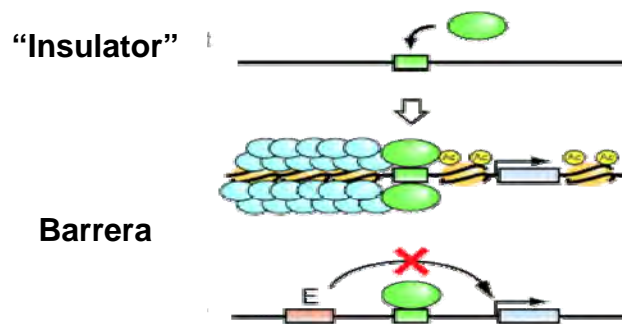
El descubrimiento de secuencias de ADN que permitían la expresión constante de un transgen de manera independiente al sitio de integración del mismo (77), ayudó a comprender los mecanismos por los cuales estos dominios se estarían aislando de su entorno. Los límites estarían constituidos por lo que conocemos como elementos aisladores o delimitadores llamados en inglés: "insulators" o "boundary elements". Los elementos frontera abarcan secuencias nucleotídicas que van de 0.5 a 3 kb, a las cuales se unen proteínas nucleares (78), pueden ser definidos como secuencias que separan dominios funcional y estructuralmente distintos, evitando la promiscuidad entre elementos de regulación, así como la modificación de la estructura de la cromatina por dominios vecinos (79). Los "insulators" han sido propuestos como secuencias



responsables de la formación y mantenimiento de un dominio o región en las cuales un gen ó grupo de genes se encuentran activos transcripcionalmente (66,28).

Funcionalmente, éstos se definen como secuencias neutras, es decir, que no afectan la transcripción y en general se encuentran enmarcando dominios génicos (81).

Los “insulators” se definen funcionalmente por dos propiedades: 1) Tienen la capacidad de bloquear la acción positiva de un potenciador sobre su promotor siempre y cuando las secuencias tipo "insulator" se localicen entre el potenciador y el promotor (Fig. 6) y 2) los “insulators” tienen la propiedad de proteger a un transgen de los efectos de posición causados por los distintos entornos cromatínicos, asociados al sitio de integración en el genoma (Fig. 6). Las actividades de protección contra los efectos de posición y bloqueo de enhancer no son siempre incluyentes, esto es, que no necesariamente todos los “insulators” poseen ambas propiedades, asimismo para ciertos “insulators” ambas funciones son separables (82). Durante los últimos años, varios grupos han identificado elementos tipo “insulators” en diversos dominios génicos, como en los dominios  $\alpha$ - o  $\beta$ -globina de pollo o el locus improntado *Igf2/H19*, entre otros (83) o incluso enmarcando genes que escapan a la inactivación del cromosoma X (258).



**Figura 6.** La propiedad de barrera, propone que la presencia de proteínas que se unen a los “insulators”, delimitan regiones de cromatina evitando la propagación de heterocromatina en regiones donde los genes tienen que estar activos. La función de bloqueo de enhancer evita que señales externas actúen sobre la expresión de los genes (modificada de 259).

Uno de los modelos mas claros para el estudio de los “insulators” y que muestra la interrelación que hay entre el estudio de estos elementos y de los enhancers, se encuentra en la realización de una serie de construcciones hechas en mini cromosomas por el grupo de Dean (90, 166). Ellos mostraron que la introducción del “insulator”

cHS4 entre el enhancer HS2 y su promotor, trae consigo la reducción en la hiperacetilación del ADN que se encuentra entre el promotor y el enhancer, reduce el reclutamiento de CBP y p300 al enhancer. Genera la acumulación de la pol II hacia la región del enhancer y del “insulator”, acompañado de una disminución en la región promotora (90). En el bloqueo del enhancer HS2 por el “insulator” cHS4, se observa la pérdida en la remodelación de la estructura de la cromatina del promotor mediada por el enhancer, la cual es requerida para su actividad transcripcional, sin embargo no aparece afectar al enhancer (90). Estos resultados muestran la importancia que tiene el entendimiento de los modelos de acción de los “insulators”, para inferir los mecanismos mediante los cuales los enhancer actúan.

## **Modelos de acción**

### ***Modelo de bloqueo de enhancer***

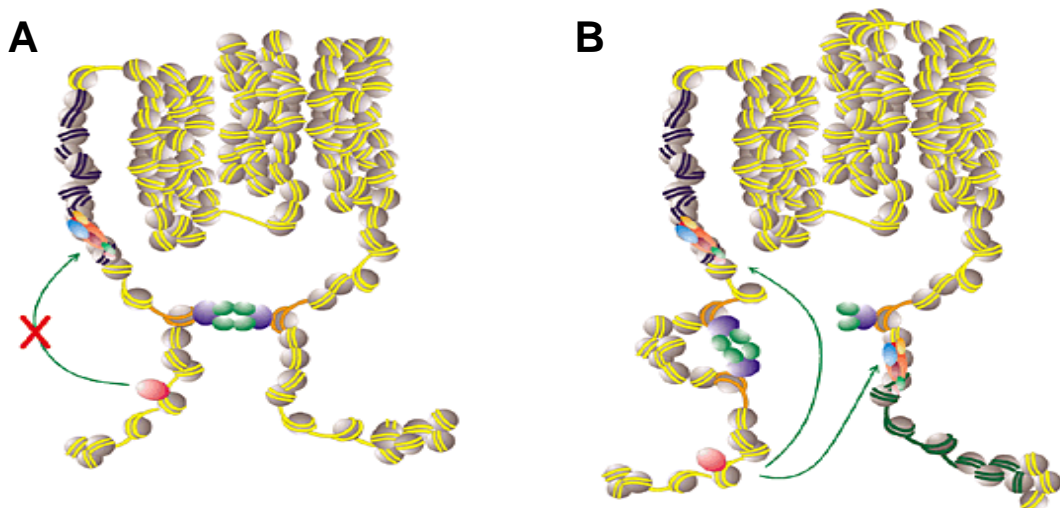
Este modelo propone que los “insulator” afectan la actividad de un enhancer sobre su promotor, siempre y cuando el “insulator” se encuentre entre estos dos elementos. Su efecto aparentemente, no afecta la habilidad del enhancer para comunicarse con elementos de los cuales no está aislado (260, 261). En particular bloqueando interacciones no deseadas entre promotores y enhancers y así participando en la selectividad de estos últimos. Este tipo de regulación se puede dar de manera regulada (ejemplo: dependiente de hormonas, 84). Los mecanismos mediante los cuales se lleva a cabo la actividad de bloqueo han generado dos modelos alternativos. El modelo de atrapado del promotor (del inglés “promoter decoy”), el cual sugiere que los “insulators” neutralizan la actividad de los enhancers vía la captura de ellos (85). El segundo modelo, el de barrera (del inglés: “chromatin boundary”), sugiere que los “insulators” deberán estar organizando la estructura de la cromatina generando barreras (28). En relación al modelo de barrera, Studitsky y colaboradores (2003) en un sistema que asemeja a los “insulators”, son capases de aislar al enhancer de su promotor mediante la formación de asas, proponiendo un modelo topológico de alineamiento (85). Esta propuesta fue después probada *in vivo* (87), vía la generación inducible de asas que atrapaban al enhancer, aislándolo de su blanco y resultando en una regulación dependiente de la estructura de la cromatina (87). Estos resultados son consistentes con el modelo de asa descrito para el locus *Igf2/H19* (62). En *Drosophila*, se ha propuesto

que las proteínas de unión al “insulator” *scs/scs*’, *Zw5* y *BEAF*, interaccionan entre ellas *in vivo* generando estructuras de aislamiento topológico (62).

En vertebrados, uno de los principales responsables de llevar a cabo una función similar es la proteína de unión a la secuencia *CCCTC*, conocida como *CTCF*, la cual se encuentra en la mayoría de los “insulators” (88, 89). Este factor posee un dominio de unión al ADN constituido por 11 dedos de zinc, los cuales usa de manera diferencial para interaccionar con el ADN. La incorporación de una secuencia de unión a *CTCF*, entre un enhancer y su promotor blanco en cromosomas artificiales, provoca un bloqueo en el traslado de la maquinaria transcripcional reclutada inicialmente al potenciador (90). Por otra parte, este factor tiene la capacidad de heterodimerizarse, lo cual ha permitido sugerir que esta proteína unida en los extremos de un dominio interacciona entre sí para favorecer la formación de asa que permitan definir estructural y topológicamente a un dominio (Fig. 7). Este modelo ha sido recientemente confirmado experimentalmente mediante el ensayo de captura conformacional de cromosomas (91-93).

### ***Modelo de protección contra el efecto de posición***

Los efectos de posición son la serie de eventos que el ambiente cromatínico impone a un transgen de acuerdo al sitio en el genoma al cual se ha integrado. Los “insulator” son elementos capaces de proteger la expresión de un transgen independientemente del sitio de integración, tanto de efectos positivos como negativos (94). Por ejemplo, estos elementos tienen la capacidad de evitar la propagación de heterocromatina, siempre y cuando éste elemento se encuentre posicionado entre la secuencia que promueve la heterocromatina y el transgene sensible a ésta (83).



**Figura 7.** Modelo de acción de delimitadores cromatínicos. **A** Los “insulators” (círculo verdes) forman barreras estructurales que impiden que los enhancers actúen sobre un promotor o evitan que la cromatina de regiones vecinas se extienda sobre el dominio que se regula. **B** Si un “insulator” no protege a un dominio, los elementos de regulación tienen libre acceso a los promotores, lo que conlleva a la activación transcripcional.

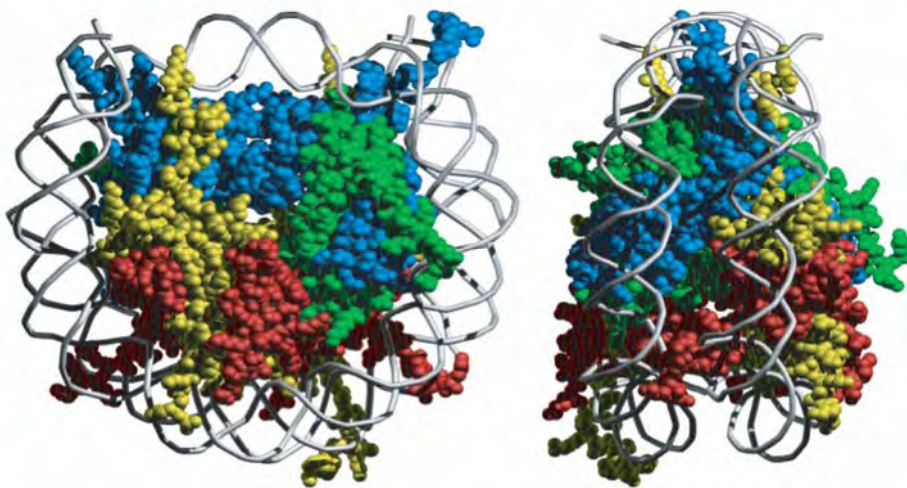
En resumen los elementos barrera o "insulator" parecen contribuir a la formación de dominios transcripcionalmente activos, por lo que los consideramos como uno de los primeros niveles de regulación. Lo anterior basado en el hecho de que una vez definido el dominio, se llevarán a cabo al interior de éste los eventos finos de regulación. Cabe señalar que no todos los genes o grupos de genes se encuentran de manera natural enmarcados por secuencias tipo "insulator" y que posiblemente será la organización del genoma la que dicte su presencia en regiones donde el ambiente cromatínico pueda influir en la regulación génica (28).

### **Regulación a nivel de la cromatina**

Como se mencionó en la introducción, el genoma eucariota se encuentra organizado en cromatina donde la unidad mínima de estructuración y por lo tanto de compactación está dada por el nucleosoma (Fig. 8; 4). En la actualidad la regulación de la expresión génica no puede ser entendida sin tomar en consideración a la estructura de la cromatina.

## Definición y componentes

Una célula de vertebrados posee alrededor de 2-4 metros de ADN en forma lineal contenida en una estructura de apenas 10  $\mu\text{m}$  de diámetro -el núcleo-. Para alcanzar dicha compactación, las células organizan el ADN en cromatina, lo cual consiste en la asociación del ADN con un conjunto de proteínas, denominadas histonas, con la finalidad de compactar el genoma y lograr contenerlo al interior del núcleo (98). La unidad fundamental de la cromatina es el nucleosoma, está compuesto por 147 pb de ADN que dan 1.65 vueltas al octámero de histonas (conformado por los dímeros de H3, H4, H2A y H2B); donde se describen 14 puntos de contacto entre el ADN y las histonas (11, 262). Esta serie de interacciones hacen del nucleosoma uno de los complejos con el ADN que muestran mayor estabilidad en condiciones fisiológicas; por lo anterior resulta indicado para empaquetar el ADN, pero no hay que olvidar que no se trata de una unidad estable, sino de una unidad dinámica que es finamente regulada por complejos proteicos (11).

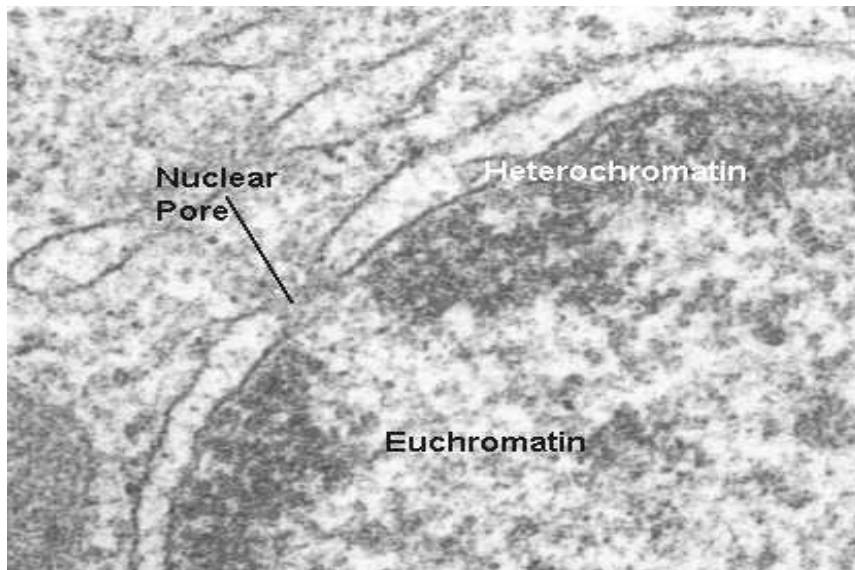


**Figura 8.** Representación de la estructura de un nucleosoma basada en la estructura cristalográfica. Diagrama visto de frente (izquierda) o visto del lado derecho (derecha). El ADN se muestra en gris, H2A en amarillo, H2B en rojo, H3 en azul y H4 en verde (6).

El nucleosoma representa el nivel primario de organización del genoma, compacta alrededor de 6 veces al ADN. Cada nucleosoma está asociado al nucleosoma vecino por un fragmento de ADN que va de 40-60 pb, región denominada internucleosomal. A nivel microscópico, dichos nucleosomas dispuestos sobre el ADN forman una fibra cuyo grosor es de alrededor de 11 nm (Fig. 1) (98). El siguiente nivel de organización está constituido por un conjunto de 6 nucleosomas cuya estructura posee un diámetro de

30 nm, conocida como solenoide. Dicho conjunto de nucleosomas provoca una compactación de más de 50 veces. Esta estructura es estabilizada por una quinta histona que no forma parte del nucleosoma, conocida como la histona H1 (ver más adelante). Evidencias experimentales, respaldadas por cálculos matemáticos, proponen que dicha fibra es en donde se llevan a cabo los principales procesos de regulación génica. La fibra de 30 nm se organiza y/o estructura favoreciendo la formación de asas cromatínicas abarcando áreas genómicas de incluso cientos de kb, las cuales pudieran constituir dominios génicos. Posteriormente, múltiples niveles de compactación más complejos contribuyen a la formación de una fibra de 240 nm, la cual puede ser identificada como la base para la formación de los cromosomas metafásicos. Esta organización permite compactar el ADN alrededor de 10,000 veces (98). El compactar el ADN en cromatina causa *per se* un impedimento o control sobre diversos procesos, lo anterior como consecuencia de una reducida accesibilidad a ciertas secuencias o regiones del genoma. Aunque esto parece una desventaja, el proceso de compactación es necesario para ocultar secuencias regulatorias de genes que no son requeridos en un momento dado, lo cual sugiere que este proceso está finamente regulado.

Desde el punto de vista funcional la cromatina ha sido dividida en eucromatina y heterocromatina (Fig. 9). La eucromatina corresponde a regiones que poseen genes transcripcionalmente activos (o potencialmente) con una estructura de la cromatina más relajada en el núcleo en interfase. Las secuencias reguladoras en estas regiones son accesibles al corte por nucleasas, las islas CpG suelen no estar metiladas y las histonas H3 y H4 están hiperacetiladas, finalmente la eucromatina suele replicarse al principio de la fase S (264). Por el otro lado, la heterocromatina se refiere a regiones transcripcionalmente inactivas, con una estructura altamente condensada, metilación del dinucleótido CpG e hipoacetilación de las histonas. A su vez la heterocromatina puede dividirse en heterocromatina constitutiva y facultativa (264, 265). La primera presenta un bajo contenido de genes y un alto número de secuencias repetidas (tipo satelitales, pericentromérica y centroméricas). Esta se replica tarde en la fase S, las histonas H3 y H4 suelen estar metiladas en los residuos de aminoácido K9 y K20, respectivamente. En células de humano y ratón suele encontrarse a la HP1 unida a la histona H3-K9me3 (266, 267).



**Figura 9.** Foto de microscopía electrónica mostrando una región de heterocromatina (heterochromatin), una de euromatina (euchromatin) y el poro nuclear de la célula (nuclear pore) (modificada de <http://cellbio.utmb.edu/cellbio/nucleus2.htm>).

Cambios en la estructura de la cromatina son necesarios para lograr una correcta regulación de los procesos celulares, entre ellos la activación génica y el silenciamiento, reparación del ADN, segregación durante la mitosis y la meiosis, la inactivación del cromosoma X y la compactación de la cromatina durante la apoptosis. A continuación hablaremos de este tipo de alteraciones de la cromatina.

### **Regulación epigenética**

La regulación epigenética gira en torno a la estructuración del genoma en cromatina y su efecto regulador sobre la expresión de los genes. Se consideran procesos de regulación epigenéticos, todos aquellos que influyen en la expresión heredable de los genes sin que ocurran cambios en la secuencia del ADN (99). Los elementos que participan en la regulación epigenética responsables de permitir que la expresión génica tenga un vínculo muy estrecho con la estructura de la cromatina son básicamente cinco<sup>8</sup>: 1) La metilación del ADN, 2) Los complejos de remodelaje ATP-dependientes, 3) Las modificaciones post-traduccionales de las histonas 4) Las variantes de histonas y, 5) Los complejos represor y activador, Polycomb y Trithorax (PcG y trxG).

---

<sup>8</sup> Estos elementos pueden estar participando de manera conjunta, como sería el caso para la proteína *brm* del grupo *trx*, esta parece estar cooperando con el complejo remodelador dependiente de ATP SWI/SNF en la regulación homeótica (319).

### ***La metilación del ADN***

La metilación del ADN es sinónimo de represión transcripcional mediada por la formación de una cromatina compacta. En términos generales la metilación del ADN contribuye a mantener la estabilidad del genoma (ver mas adelante) (106). La metilación del ADN juega un papel central en la transmisión de la información epigenética a través de la incorporación de un grupo metilo a la posición 5 de la citosina siempre y cuando se encuentre en el contexto del dinucleótido CpG (107). Lo anterior es hecho por un grupo de proteínas especializadas denominadas metiltransferasas de ADN (Dnmts). La Dnmt1 se encarga de la manutención de la metilación del ADN durante la replicación del ADN, esta enzima modifica la citosina cuando el ADN se encuentra hemimetilado<sup>9</sup> (99). Las Dnmt3a y b son fundamentales para la metilación *de novo*, éstas pueden incorporar un grupo metilo a los dinucleótidos CG del ADN no metilado en etapas tempranas del desarrollo (99). Ambas enzimas usan la 5-adenosil-metionina como sustrato, con el fin de transferir un grupo metilo a la citosina (108, 109).

La metilacion del ADN se encuentra localizada en regiones repetidas del genoma y transposones en eucariotas. Los promotores suelen no estar metilados ya que contienen islas con el motivo CpG, éstas son refractarias a la metilacion (263). Esta modificación se presenta en organismos más complejos, por ejemplo las levaduras carecen de este proceso, en moscas y hongos existe una reducida presencia de esta modificación, con funciones poco claras hasta el momento. En cambio, en mamíferos una reducción en los niveles de citosinas metiladas desencadena una mayor inestabilidad genómica y re-expresión de genes silenciados, aspecto intimamente relacionado con un número considerable de procesos asociados al origen y desarrollo del cáncer (110).

Los mecanismos propuestos mediante los cuales la metilación de ADN deberá de estar generando silenciamiento son:

---

<sup>9</sup> Se refiere a que solo una de las dos cadenas de ADN, se encuentra metilada.



1) La metilación del ADN puede impedir estéricamente la unión de un activador transcripcional a sus secuencias de reconocimiento, evitando que ejerzan su efecto regulatorio sobre el gen blanco (110).

2) Existe una familia de proteínas con la capacidad de unirse al ADN metilado conocidas como MBDs ó MeCPs (del inglés: "Methyl-Binding Proteins or Methyl-CpG-binding Proteins", respectivamente) (113). Al interactuar con el ADN metilado y con la ayuda de co-represores reclutan actividades de remodelaje de la cromatina que inducen la formación de una cromatina compacta, en particular, a través de la acción de desacetilasas de histonas (113).

Las MPT de las histonas se encuentran epigenéticamente ligadas a la metilación del ADN en algunos organismos. En la planta *Arabidopsis thaliana* y en el hongo *Neurospora crassa* se ha demostrado que la metilación de la lisina 9 de la histona H3 constituye una señal molecular para atraer a metil-transferasas de ADN y metilar localmente el ADN (114). Incluso en una publicación reciente se demostró en *Neurospora crassa* que el homólogo de la proteína de heterocromatina HP1 es atraída por la metilación de la lisina 9 de la histona H3 evento necesario para la metilación del ADN (115). Lo anterior demuestra no sólo la interdependencia entre las modificaciones covalentes de las histonas sino también con otros procesos epigenéticos como el caso de la metilación del ADN. En la actualidad no existen datos contundentes que sugieran que este mismo proceso ocurra en mamíferos.

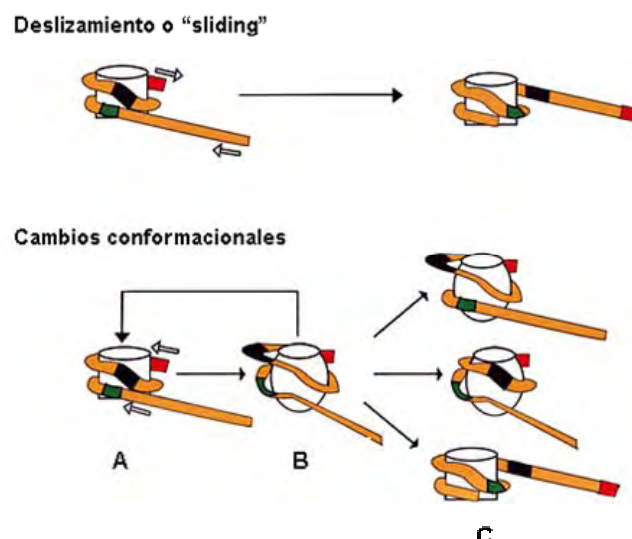
El dogma central nos dice que la metilación del ADN es un proceso irreversible. Sin embargo en años recientes ha surgido la idea de que la metilación del ADN es reversible aunque se desconoce su mecanismo y por lo anterior se necesita hacer una clara distinción entre metilación *de novo* y mantenimiento (111, 112). Ahora existen claras evidencias de un proceso activo, replicación independiente de desmetilación de ADN durante etapas tempranas del desarrollo. Se han reportado genes que se encuentran altamente metilados en el genoma parental y son rápidamente desmetilados en el cigoto, tan solo unas horas después de haber sido fertilizado, antes de que comience la primera ronda de replicación (290). Otros ejemplos donde se han encontrado procesos de desmetilación activa son los siguientes: en el locus Ig durante la maduración de células B (291), en el gen músculo-específico  $\alpha$ -actin (292), en el gen de miosina en mioblastos en diferenciación (293), en el gene Il2 durante la activación de células T (294) y en el gen de interferón  $\gamma$  (295). Sin embargo no queda excluido por

completo el que la desmetilación pasiva esté jugando un papel importante en todos estos procesos. En adición a lo anterior, se reportó que la proteína MBD2 posee actividad de desmetilasa tanto *in vitro* (296) e *in vivo* (297, 298).

En resumen, la metilación es una marca epigenética necesaria para programar la expresión génica. El reconocimiento de co-represores, así como su maquinaria asociada, permiten la regulación negativa de genes durante periodos en los cuales éstos no son requeridos.

### ***Los complejos de remodelaje ATP-dependientes***

Para activar o reprimir la expresión de un gen, es necesario que factores transcripcionales tanto generales cómo específicos tengan acceso, reconozcan y puedan unirse a sus secuencias blanco, en la molécula de ADN. Se han descrito un grupo de modificadores de la cromatina que utilizan la hidrólisis del ATP para alterar las interacciones ADN-histona (Fig. 10). Vía la ruptura transitoria de las uniones presentes en los extremos del complejo ADN-histona, formando un asa, o moviendo al nucleosoma a diferentes posiciones translacionales (resbalado, del inglés “sliding”; Fig. 10), finalmente modificando la accesibilidad al ADN por los factores generales de la transcripción (11).



**Figura 10.** Mecanismos de acción de los Complejos de Remodelaje dependientes de ATP (CRDA). Dichos complejos utilizan ATP para relajar la fibra de 30 nm o para inducir su formación. Se ha propuesto que los CRDA podrían participar deslizando (“sliding”) a los nucleosomas sobre el ADN, favoreciendo o impidiendo la exposición de secuencias específicas. También podrían fomentar cambios conformacionales, en el cual el ADN asociado a un nucleosoma (A) es alterado

transitoriamente (B), desestabilizando al nucleosoma. Esto permite reposicionar al octámero de histonas en una nueva región del ADN, permitiendo o impidiendo la exposición de secuencias (C), incluso esta actividad se ha visto asociada al remplazo e incorporación de variantes de histonas de manera regulada (286). Figura tomada de (116).

Estos desplazamientos ocurren a distancias cortas, se cree que se llevan a cabo sin que los nucleosomas se disocien por completo de sus interacciones con el ADN (117). El complejo SWI/SNF originalmente estudiado en levadura y *Drosophila* constituye el complejo pionero seguido por NURF, CHRAC y ACF entre otros (118). La sub-unidad conservada entre todos los complejos de remodelaje es la sub-unidad con actividad de ATPasa; dicha sub-unidad se encuentra embebida en complejos multiproteicos que incluyen componentes necesarios para su reclutamiento a la cromatina y/o su regulación. Cabe mencionar que estos no pueden reconocer y unirse al ADN, por lo que deben de ser reclutados por factores transcripcionales o co-factores (120). El complejo Mi-2/NuRD está compuesto por desacetilasas de histonas, proteínas de unión al ADN metilado entre otros, por lo tanto posee tres actividades remodeladoras en un solo complejo (119).

Finalmente se han encontrado procesos cooperativos entre complejos de remodelaje y acetilasas de histonas, sin que exista un orden predominante en la acción de estas proteínas sobre la región de control (118, 122). Lo que parece volverse una constante es que estas dos actividades de remodelaje de la cromatina son cooperativas (118, 122). Diversas evidencias apoyan de manera substancial la participación de los complejos remodeladores ATP-dependientes en regulación transcripcional así como en replicación, reparación y recombinación del ADN, cohesión de cromátidas hermanas durante la replicación e integración de transposones dentro del genoma (121). En resumen los complejos de remodelaje ATP-dependientes representan un eslabón más dentro de la regulación epigenética. Aún falta mucho por comprender en relación con estos complejos ya que su complejidad los hace un tema difícil pero a la vez fascinante de investigación.

### ***Modificaciones post-traduccionales de las histonas***

Existen diferentes tipos de modificaciones post-traduccionales (MPT) de las histonas (Tabla 1), entre ellas la metilación de residuos de arginina (R), la metilación,

acetilación, ubiquitinación, ADP-ribosilación y sumorilación de residuos de lisina (K), y la fosforilación de residuos de serina y treonina (11). Estas modificaciones decoran al tetrámero de histonas (H3, H4, H2B y H2A) así como a las variantes (H3.1, H3.2 y H2A.Z-entre otras) (268) en sus extremos amino y carboxilo terminales, así como en el dominio globular (11, 269). Con la excepción de la metilación, todas las modificaciones de las histonas resultan en un cambio en la carga neta de los nucleosomas, lo que puede generar la pérdida de interacciones inter- o intra-nucleosomales ADN-histona, esto ha sido encontrado en estudios *in vitro* (11, 270-272) e *in vivo* (11, 273, 274).

Existen diferentes evidencias de que la acetilación se encuentra relacionada a la activación transcripcional y la SUMOilación a la represión, éstas dos principalmente mediante un impedimento estérico. Por el otro lado la metilación y la ubiquitinación tienen diferentes efectos dependiendo del residuo modificado y el contexto en que se encuentra. Como ejemplo, la ubiquitinación del extremo terminal de las histonas H2B y H2A correlaciona con activación y represión, respectivamente (268). Sin embargo el descubrimiento de modificaciones centroméricas y el requerimiento de la di metilación de H3K9 en la elongación, nos indica que algunos patrones de modificaciones simplemente parecen carecer de interpretaciones globales (275, 276).

Tabla 1 <b>Modificaciones de la cromatina</b>		
marca	localización genómica	efecto sobre la transcripción
<b>Metilación del ADN</b>		
citocina metilada	Isla CpG	represión
<b>MPT de las histonas</b>		
lisina acetilada (acK)	H3 (9, 14, 18, 56), H4 (5, 8, 13, 16), H2A, H2B	activación
fosforilación de Serinas/Treoninas (S/Tph)	H3 (3, 10, 28), H2A, H2B	activación
metilación de Argininas (meR)	H3 (17, 23), H4 (3)	activación
metilación de Lisinas (meK)	H3 (4, 36, 74) H3 (9, 27), H4 (20)	activación represión
ubiquitinación de Lisinas (ubK)	H2B (123/120)) H2A (119)	activación represión

sumoilación de Lisinas (suK)	H2B (6/7), H2A (268)	represión
Isomerización de Prolinas (isomP)	H3 (30-38)	activación/represión
Entre paréntesis se muestra el número de aminoácido modificado. Tabla tomada y modificada de (268)		

La mayoría de las MPT son observadas durante la regulación de la transcripción. Pero este tipo de procesos también puede extenderse a procesos tan diversos como la progresión del ciclo celular y replicación del ADN. Como ejemplo la fosforilación de la serina 10 de la histona H3 es crucial para la condensación cromosómica y progresión del ciclo celular durante la mitosis y meiosis (277).

Se han propuesto dos mecanismos mediante los cuales las modificaciones post traduccionales estarían actuando. El primero menciona que la estructuración del ADN en cromatina se ve alterada por éstas, cambiando las interacciones electrostáticas o vía interacciones internucleosomales, estabilizando o desestabilizando el complejo ADN-nucleosoma y regulando el acceso a factores transcripcionales al ADN (263). El otro propone la generación de superficies con modificaciones individuales o patrones de varias modificaciones en una misma histona, los cuales serán leídos por otras proteínas que influyen en la función y dinámica de la cromatina (11, 252, 278, 279).

Recientemente se ha demostrado que estas modificaciones favorecen la intercomunicación entre nucleosomas afectando y regulando la estructura de la cromatina localmente pero también al nivel de un dominio e incluso de un cromosoma completo (60). El efecto local se considera como la relación con regiones cercanas a los genes. Uno de los ejemplos más claros se presenta en el dominio  $\beta$ -globina de pollo el cual constituye una familia multigénica que se expresa de manera diferencial a lo largo del desarrollo (100). En su costado 5' se ha identificado un elemento tipo "insulator" seguido río arriba de 16 kb de cromatina compacta y un gen que codifica para uno de los receptores de folato con una expresión eritroide específica que se encuentra activo, en una etapa previa a la expresión de los genes globina (101). En el costado 3', se identificó otro "insulator" que delimita al dominio  $\beta$ -globina de un gen que codifica para uno de los receptores olfatorios con una expresión cerebro específico (102). Este es uno de los pocos ejemplos claros, que muestran la organización del genoma en

dominios, cuyos límites se encuentran definidos por secuencias tipo "insulator" y con un patrón bien definido en cuanto al estado de acetilación de las histonas (103, 104). Finalmente a nivel de un cromosoma completo, los cromosomas sexuales son el ejemplo clásico, con la inactivación del cromosoma X en hembras (105). En el centro de inactivación del cromosoma X (Xic), se inicia la transcripción de un RNA intergénico de 17 kb y una cadena de eventos, que traerán como consecuencia final, el establecimiento de MPT de cromatina cerrada y el silenciamiento de los genes, en gran parte, del cromosoma X inactivo (280)

En los últimos 10 años, gracias a una intensa búsqueda se ha logrado un gran avance en la comprensión de las proteínas que dirigen la modificación de las histonas. Se han identificado enzimas involucradas en la acetilación (281), metilación (282), fosforilación (269, 277), ubiquitinación (283), sumoilación (284), ADP-ribosilación (285) e isomerización de prolinas (286). La mayoría de estas modificaciones han mostrado ser dinámicas y las enzimas que remueven estas modificaciones también han sido encontradas. Todas las MPT son reversibles. Existen desacetilasas de histonas (HDACs) las cuales remueven grupos acetil; las fosfatasa de Ser/Thr, remueven el grupo fosfato. Un grupo de proteasas remueven la mono-ubiquitinación de H2B. La metilación de las argininas es alterada por desaminasas, las cuales convierten esta modificación en citrulina. Dos tipos de desmetilasas de citrulina han sido recientemente descritas: las del tipo LSD1/BHC110 (que remueven la H3K4me1 y me2) (287, 288) y las del tipo jumonji (que remueven H3K4me2 y me3, H3K9me2 y me3 y H3K36 me2 y me3) (289).

Las modificaciones post-traduccionales de las histonas generan una rama de posibilidades mediante las cuales la transcripción puede ser regulada en el contexto de la cromatina. En un inicio se propuso que estas en conjunto con la metilación del ADN, eran las encargadas de guardar la información de los estados transcripcionales a través de la diferenciación, sin embargo, esta propuesta parece no ser más la regla sino la excepción, como el caso de la fosforilación de la serina 10. Recientemente, ha surgido una idea de ver a las MPT como señales capaces de conservar la información de los estados transcripcionales en periodos cortos del ciclo celular.

### *Variantes de histonas*

Las histonas canónicas son una de las proteínas más abundantes en eucariontes. Por mucho tiempo fueron consideradas como el componente universal de todos los nucleosomas (11). Estas son depositadas en la orquilla de replicación durante la fase S del ciclo celular (299). Sin embargo estudios detallados mostraron la existencia de pequeñas diferencias entre organismos (11). Las variantes de histonas se distinguen principalmente por su expresión fuera de la fase S del ciclo celular e incorporación en la cromatina de manera replicación-independiente (11). Las diferencias entre las histonas variantes y las canónicas se pueden encontrar en sus extremos amino y carboxilo, en el dominio globular, e incluso por el cambio de unos cuantos residuos de aminoácido (11). Las variantes de histonas pueden presentar entre ellas pequeñas diferencias (entre H3.1 y H3.2 solo hay un aminoácido de diferencia) o alteraciones drásticas que cambian la naturaleza de la histona (como la incorporación de la histona macroH2A; ver más adelante) (262). Estas pueden ser depositadas en el ADN vía chaperonas de histonas, requeridas para el ensamblaje y dinámica de las histonas durante la replicación y la transcripción (300, 301) y por diferentes complejos dependientes de ATP que tienen la habilidad de quitar o intercambiar dímeros de histonas (302).

Interesante mente, la histona H4 parece no variar durante la evolución en eucariotas (303), solamente las otras histonas presentan diferentes variantes con diferentes actividades. Entre ellas y como ejemplo de las demás familias, encontramos a las variantes H2A (para mayores referencias consultar 263). Esta familia es por mucho la más diversa y larga. Hasta la fecha en mamíferos se han encontrado 5 variantes. H2A.X es una variante a la cual se le ha encontrado una función como sensor de daño a ADN y como es de imaginar, esta conservada en todos los organismos analizados a la fecha (263). Esta varía con respecto a la H2A canónica, en que solo presenta un motivo SQ en su extremo C-terminal. Ratones deficientes en esta variante son viables, sin embargo sus fibroblastos y sus células T y B presentan un nivel elevado de inestabilidad cromosómica (304). Presentan defectos cromosómicos tales como rompimiento de cromatina, cromosomas di-céntricos y translocaciones al azar (263). Las funciones que pueden estar jugando las diferentes variantes de histonas son muchas, sin embargo dentro de sus papeles más novedosos encontramos el que se discute en el siguiente capítulo.

### ***Hipótesis del código de barras (“barcode”)***

La alteración de la estructura nucleosomal, afecta la accesibilidad al ADN y puede generar dominios especializados en el genoma. Tales cambios pueden ser inducidos por modificaciones post-traduccionales de las histonas, por complejos remodeladores de la cromatina o por la incorporación de variantes de histonas (299). La incorporación de variantes de histonas no solo permite variar la composición de nucleosomas individuales, también le permite a la célula ampliar el abanico de modificaciones post-traduccionales (263). ¿Como es que el estado transcripcional de los genes es transmitido de generación en generación? Las modificaciones post-traduccionales de histonas y el código de David Allis (252) ya no pueden explicar más esta memoria epigenética a través de varias divisiones celulares.

Recientemente se propuso que los patrones de las histonas no son activamente duplicados durante la replicación, sino que son reestablecidos cada ciclo celular Henikoff (307). Diferentes modelos se han propuesto para entender como es que la memoria epigenética puede ser heredada. Se ha propuesto que los nucleosomas paternos, son simplemente dispersados (pero intactos) entre las células hijas. En este modelo la herencia sería llevada a cabo por los diferentes tiempos de replicación (306, 248). En el modelo semiconservativo se propone que el nucleosoma se partirá en dos y en lugar de que los nucleosomas se depositen como tetrámeros en las células hijas, se depositaran como dímeros de H3-H4 (310).

La hipótesis de “barcode” propone un modelo en el que genoma de mamíferos se encuentra indexado por variantes de la histona H3 (y tal vez con la contribución de otra) en una forma no aleatoria. La histona H3.3 será incorporada en regiones transcripcionalmente activas, mientras que la histona H3.2 será depositada en zonas silenciadas que pueden ser activadas, dependiendo de los procesos celulares requeridos, la histona H3.1 deberá encontrarse de manera constitutiva en regiones silenciadas del genoma y finalmente la histona CENP-A se encontró en regiones muy especializadas tipo centrómeros (305). Para complementar esta hipótesis, se proponen que las MPT y el código de histonas, servirán para regular la respuesta inmediata de los genes a señales externas, pudiendo ser este patrón de modificaciones conservado, por periodos cortos, a través del desarrollo. La hipótesis propone que los genes podrán ser prendidos o apagados durante el ciclo celular en respuesta a los patrones de MPT (305).



En conclusión, al menos tres diferentes códigos estarán dictando los procesos celulares, la información genética, las MPT y las variantes de histonas (305).

### ***Regulación epigenética a través de los complejos Polycomb y Trithorax***

Los grupos de genes Polycomb y Trithorax fueron originalmente descritos en *Drosophila* como proteínas reguladoras de la expresión de los genes homeóticos. El grupo de genes Polycomb (PcG) codifica para un conjunto de proteínas que regulan negativamente a los genes homeóticos mientras que Trithorax (TrxG) lo hace positivamente (123-125). Las proteínas del grupo Polycomb y Trithorax, son proteínas conservadas que reprimen o activan a sus blancos tanto en plantas como animales (308). Se ha postulado que PcG y TrxG son parte de los mecanismos que contribuyen a la memoria epigenética. En particular las proteínas del grupo PcG son responsables de silenciar no sólo a los genes homeóticos fomentando una estructura de la cromatina represora, evitando así la incorporación de factores de transcripción y por lo tanto la transcripción (123, 124). Para mayor información con respecto a este tema revisar las citas 123-125 y 308. Como se puede observar, varios son los mecanismos propuestos mediante los cuales la información puede ser heredada epigenéticamente. Sin embargo, ya no queda más claro, que tan heredable es cada una de las marcas antes mencionadas.

### **Generación de regiones libres de nucleosomas**

Durante mucho tiempo se pensó que el remodelamiento de la cromatina no implicaba la completa disociación de las histonas. Fue justo después del descubrimiento de las hipersensibilidades observadas a la digestión con la DNasa I (311), que se postuló que éstas correspondían a elementos de regulación tipo promotores y enhancers, los cuales se encontraban como segmentos libres de nucleosomas en el ADN. Reportes recientes apoyan la idea de que los nucleosomas responden dinámicamente a los diferentes procesos implicados en la transcripción vía mecanismos que incluyen MPT de las histonas, remodelamiento de la cromatina ATP-dependiente, chaperonas de histonas y hasta variantes de histonas (311).

Muchos han sido los esfuerzos por comprender si las regiones transcripcionalmente activas se encuentran estructuradas en cromatina o como regiones

libres de nucleosomas. Diferentes grupos han propuesto que la densidad nucleosomal en promotores es menor que en la región codificante (29, 238, 240). Aunque estos resultados también pueden ser explicados por el intercambio constante de histonas que ocurre en genes activos (312). En un estudio realizado por el grupo de Rando (205), mediante el uso de microarreglos y anticuerpos anti-histonas, mapearon 2278 nucleosomas, la mayoría de ellos posicionados (210). Interesante encontraron que la mayoría de los promotores transcritos por la RNA pol II contenían un espacio libre de nucleosomas de 200 pb sobre el promotor y al mismo tiempo se encontraban flanqueados por dos nucleosomas fijos. Estos espacios suelen coincidir con la interacción de factores transcripcionales. Los nucleosomas que se encontraban enmarcando estos espacios parece que se encuentran desacetilados y enriquecidos por la histona H2A.Z (203, 201). La histona H2A.Z se ha relacionado con genes silenciados (196). ¿Cuál será su función? ¿Tendrá que ver con que estas variantes de histonas estabilizan (fijan) al nucleosoma? ¿Cómo se generan las regiones libres de nucleosomas? Esta pregunta surge cuando se sabe que el complejo ADN histona es el más estable *in vitro* y en condiciones fisiológicas (311).

Por estudios matemáticos se ha propuesto un modelo en el cual la información necesaria para posicionar un nucleosoma, se encuentra a nivel de la secuencia del ADN (11). Segal (2006) ha demostrado la existencia en levadura de secuencias cercanas a los promotores que estarían dictando el posicionamiento de los nucleosomas, con una tendencia sugerida, donde las secuencias promotoras serían excluidas de los promotores (121). Por lo anterior y recordando las características que definen a un complejo remodelador de la cromatina ATP-dependiente la función de éstos, de acuerdo a este modelo, será la de estructurar a la cromatina en un arreglo consecutivo de nucleosomas (170). En particular en regiones donde la expresión de los genes no es requerida por causas tiempo-tejido específicas.

Pero, ¿Quién está generando la especificidad para liberar o dejar libres ciertas regiones del genoma? En particular a lo largo del desarrollo, en regiones de la cromatina que se encuentran en un arreglo represivo. Se ha mostrado que la unión de factores transcripcionales al ADN genera la desestabilización del nucleosoma y lo hacen susceptible a ser desplazado. Para lo anterior, los factores transcripcionales podrían estar haciendo uso de complejos remodeladores de la cromatina dependientes de ATP, así como de chaperonas de histonas (168).

En conclusión y en contra de lo que se pensaba anteriormente, cada vez parece mas evidente que la existencia de regiones libres de nucleosomas, podrían constituir un paso importante en la activación de un gen, ya sea facilitando el reconocimiento de las secuencias blanco de elementos de regulación a distancia y de secuencias promotoras, ó a través del proceso de transcripción de la cadena de ADN por parte de la RNA pol II. El proceso mediante el cual se generan las regiones libres de nucleosomas, podría involucrar los procesos antes mencionados y/o la interacción de factores transcripcionales con características particulares que estarían generando un relajamiento en la estructura ADN-histonas; como el caso del factor transcripcional GATA-1 del cual hablare más adelante.

### **El núcleo**

El núcleo parece estar implicado en la regulación transcripcional. Se han propuesto la formación de dominios cromosómicos, fábricas transcripcionales, eventos de relocalización etcetera (ver mas adelante), como mecanismos que podrían estar participando en la regulación de un gen o grupo de genes y que posiblemente podrían tener una consecuencia epigenética.

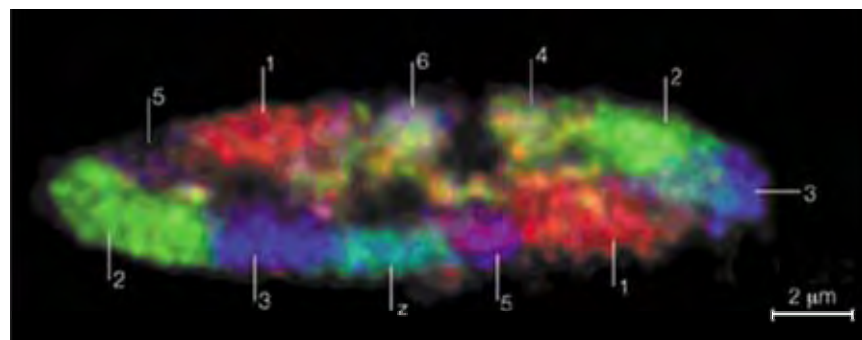
### **Dinámica nuclear**

Durante años, el núcleo celular fue considerado como un organelo sub-celular con la única función de contener en su interior a la información genética de la célula. Dos puntos de vista han surgido acerca de la relevancia que la estructura nuclear tiene. El primero propone que la estructura nuclear simplemente esta reflejando los procesos que se están llevando al interior de éste. El segundo punto de vista propone que ésta resulta en un importante factor de regulación de las funciones del genoma. Un punto de vista donde el genoma se encuentra “estructurado”; no en una red rígida sino dinámica, organizada con genes que muestran programas de expresión similares. Esta organización claramente cambia durante el desarrollo y la diferenciación y es por si sola reprogramable (por ejemplo en un proceso de clonación o transferencia nuclear) (156, 158).

Apoyados en los avances de microscopía electrónica y de fluorescencia, ha sido posible determinar que el núcleo celular es una entidad dinámica, que participa en la

regulación de la expresión génica (Fig. 11; 128, 129). La unidad mayor de organización del genoma es el cromosoma (141). Los cromosomas presentan localización nuclear preferencial en referencia a la periferia del nucleolo (137). Por lo anterior el vecindario que deberá tener un cromosoma, se encuentra lejos de ser aleatorio. El núcleo de una célula eucariota se encuentra compartimentalizado en zonas de heterocromatina y eucromatina pero también por estructuras sub-nucleares tales como los cuerpos de Cajal, los cuerpos promielocíticos (PML), los "speckles" y el nucleolo. Cada uno de estos compartimientos poseen funciones bien definidas que van desde almacenar factores de transcripción, factores ligados al procesamiento de los RNA pre-mensajeros hasta los genes ribosomales en el caso del nucleolo, entre otros. De manera general, la función que deberán estar realizando estos compartimientos es el llevar a cabo, de manera regulada, la expresión génica y el procesamiento de los RNAs mensajeros (166).

Uno de los aspectos más novedosos en la organización al interior del núcleo es la organización del genoma en territorios. Se ha propuesto que en las regiones más centrales, es decir al interior de un territorio, la estructura de la cromatina se encuentra en un alto grado de compactación. Así mismo ciertas zonas de un territorio se encuentran estructuradas en heterocromatina, mayoritariamente asociadas a la periferia nuclear y perinucleolar (Fig. 11; 128).

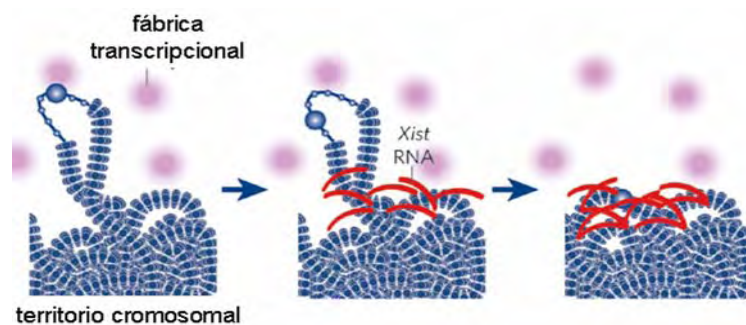


**Figura 11.** Núcleo de fibroblastos de pollo utilizando hibridación in situ de fluorescencia en donde se observa el TC de cada cromosoma así como la presencia de los cromosomas homólogos en diferente ubicación dentro del núcleo (Modificada de 70).

Por otro lado, se han encontrado genes que se localizan a grandes distancias de su dominio cromosómico. Esta relocalización ocurre en dominios que expresan de manera constitutiva sus genes (137, 130), o en algunos casos, cuando la expresión génica es inducida (125, 97). La relocalización de estos genes fuera de sus territorios se ve

reducida cuando la expresión de la RNA pol II se ve inhibida experimentalmente (137, 86), sugiriendo que estos eventos deberán ser regulados en parte por la transcripción. Los genes activos transcripcionalmente han sido localizados en puntos de concentración de la RNA pol II, nombrados “fábricas transcripcionales” (Fig. 12). El número de este tipo de fábricas vistas al interior de un núcleo parece ser menor al número de genes que se estaría expresando, sugiriendo la co-localización de múltiples genes en estas fábricas, genes pertenecientes a la misma región cromosómica o incluso a distintas regiones he incluso cromosomas (75).

En células eritroides se ha demostrado que los factores transcripcionales NF-E2p45 el cuál se dimeriza con el factor NF-E2p18 se encuentran en compartimientos distintos al interior del núcleo (133). En particular el factor NF-E2p45 se encuentra mayoritariamente asociado a zonas de heterocromatina al interior del núcleo cuando los genes globina no se encuentran activos. Al diferenciarse las células eritroides y los genes globina volverse activos transcripcionalmente NF-E2p45 se re-localiza en regiones activas, asociándose con el factor NF-E2p18 y uniéndose a sus secuencias blanco llevando a cabo sus funciones regulatorias (133).



**Figura 12.** Exclusión de los genes inactivados en el cromosoma X de las fábricas transcripcionales. Durante el proceso de silenciamiento del cromosoma X, se propone que los genes podrían estar siendo excluidos de las fábricas transcripcionales que incluyen a la RNA pol II (motes púrpuras) y siendo incluidos en el compartimento represivo de Xist RNA (figura tomada y modificada de 166).

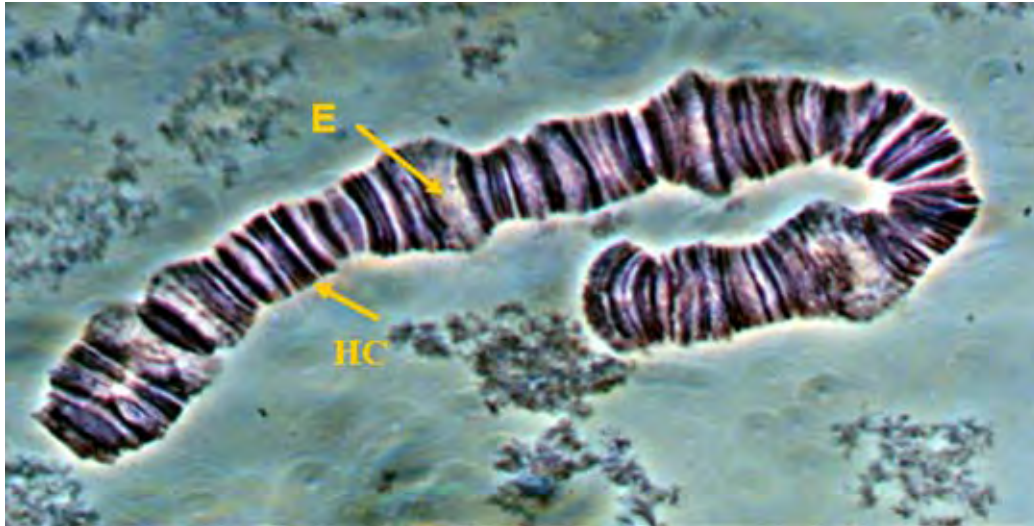
Uno de los modelos mas aceptados, para explicar la dinamica nuclear, propone que estructuralmente, el genoma eucariota estaría organizado en asas de cromatina que estrían unidos a una matriz nuclear. El significado de este arreglo es objeto de discusión (82). Sin embargo se han descrito grandes asas que sobresalen por varios micrones de

los cromosomas activos en mamíferos (80). Estas asas contienen dominios de arreglos de genes que son activamente transcritos (80). Este modelo resulta ser muy atractivo dado que cito-localiza al ADN a ser transcrito, los elementos de regulación y de procesamiento; al generarse un transcrito maduro, éste podría ser a su vez transportado por los "canales" pericromatinianos dirigiendo y facilitando su transporte a través de los poros nucleares hacia el citoplasma para que dicho mensaje sea traducido en una proteína funcional (131, 132).

### **Dominios cromosómicos**

El núcleo y los cromosomas tienen arquitecturas y dinámicas necesarias para la función de diferentes procesos (336). Los cromosomas en un núcleo en interfase son diferenciados en dos tipos por sus características citológicas y funcionales, en eucromatina y heterocromatina (336). Patrones de modificaciones post traduccionales, están asociados a estos dominios y estrechamente relacionados con funciones como regulación genética, herencia cromosómica y tiempo de replicación (336). Como ejemplo encontramos a regiones con un silenciamiento mediado vía heterocromatina, ricas en metilación de H3K9 y la carencia de acetilación de histonas y regiones de la cromatina que se encuentran altamente expresadas, que presentan regiones con histonas altamente acetiladas y la presencia de la metilación en H3K4 (336).

Durante los últimos años, diversas evidencias apoyan la idea de que el genoma de una célula eucarionte se encuentra organizado en dominios. Lo anterior en base a observaciones citológicas, a través del uso de tinciones específicas para el ADN, que permitieron delimitar regiones cromosómicas en donde el genoma de una célula en interfase se encuentra compactado y en otras, en donde se observa una cromatina más relajada (127). Esta idea se ve reforzada por la caracterización citológica de cromosomas politénicos que mostraron un patrón de bandas correspondientes a dominios transcripcionalmente activos y no activos en *Drosophila melanogaster* (Fig. 13). Un dominio es una región del genoma que posee todos los elementos estructurales y de regulación, tanto positivos como negativos, que permiten la adecuada e independiente expresión de los genes contenidos en ella. De manera funcional, se propone que existen dominios transcripcionalmente activos y no activos.



**Figura 13.** Cromosoma Politénico. Tinción de Giemsa para identificar regiones de heterocromatina (HC) y eucromatina (EC). Las zonas de eucromatina pueden presentar rearrreglos de grandes dimensiones ante la activación de los genes contenidos en dicha región, conocidos como “Puff”.

La primera indicación de que la organización cromosómica podría estar afectando la expresión de los genes, viene del descubrimiento del efecto de variegación (PEV del inglés: position effect variegation) en *Drosophila* (336). El PEV describe la inactivación epigenética o silenciamiento de un gen eucromático al ser localizado cerca de una región de heterocromatina.

En resumen resulta indudable que la visión actual del núcleo es mucho más dinámica de lo antes imaginado y con una mayor ingerencia en la regulación de la expresión de los genes. Uno de los mejores ejemplos de lo anterior lo encontramos en el proceso de silenciamiento del cromosoma-X (para mayores referencia consultar la cita 280).

### **Factores transcripcionales**

Diferentes proteínas entre ellas los factores transcripcionales, co-factores, etc, se necesitan para poder reconocer la información codificada en la hebra de ADN y poder transmitir la información que en ella se encuentra. En el dominio  $\alpha$ -globina de pollo y en especial en la región del enhancer, multiples son las proteínas que se han descrito podrían tener una función, entre ellas el factor transcripcional GATA, EKLF, NF-E2, YY1, Sp1 y sus factores asociados. Para lograr entender como se regula el enhancer  $\alpha$ -globina, es necesario saber las características que las proteínas que lo reconocen tienen.

## **El factor transcripcional GATA-1**

GATA-1 es un factor transcripcional hematopoyético que juega un papel crítico en la diferenciación y maduración de células eritroides y megacariocitos (139). Los sitios de unión a GATA-1 han sido descritos en la mayoría de los genes eritroides, incluyendo promotores, silenciadores, enhancers y en el LCR del dominio  $\beta$ -globina (140). Fue la primera proteína eritroide-específica en ser clonada (149). GATA-1 tiene 2 dedos de zinc conservados que median su unión al sitio consenso (A/T)GATA(A/G) e interacciones proteína-proteína con diversos factores transcripcionales, entre ellos están Sp1, CBP/p300, FOG-1 y el factor eritroide específico EKLF e incluso consigo mismo (155). El papel de GATA-1 en la parte terminal de diferenciación eritroide incluye la supresión en la expresión de GATA-2, su auto regulación y la sobre expresión de factores eritroides específicos, incluyendo la de EKLF, NF-E2, los genes  $\alpha$ - y  $\beta$ -globina (149, 150). Esta familia presenta dedos de zinc homólogos entre ellos, pero a nivel de secuencia de aminoácidos fuera de estos dominios la homología es muy baja. GATA-1, -2 y -3, constituyen una subfamilia, ya que los tres son expresados en células hematopoyéticas (143). GATA-1 es expresado en eritrocitos (145), megacariocitos (144), eosinófilos (143) y células de mama (144). GATA-2 es expresado en células de mama, megacariocitos y progenitores de diferentes líneas (145, 146). GATA-3 es expresado principalmente en células-T (147), mientras que GATA-4, -5 y -6 se han encontrado como reguladores de los genes que regulan el desarrollo del corazón, hígado e intestino (139, 148).

El “knock-out” del gen de *GATA-1* en ratón, genera una producción de eritrocitos incompleta, el bloqueo en la diferenciación al estadio de pro-eritroblasto y una anemia severa. (149). Mutaciones naturales en humano se han descrito en diversos síndromes poco comunes, entre los que se encuentran la anemia disertropoítica congénita (151), macrotrombocitopenia (152, 153) y leucemia megacariocítica en niños con síndrome de Down (153). Estudios recientes han mostrado que GATA-1 tiene un papel igualmente importante en activación, así como en represión durante la diferenciación eritroide (154). El gen que codifica para la proteína GATA-1 se encuentra localizado en el cromosoma X (139). Inicialmente se postuló que este gen se autorregulaba (139). Posteriormente se comprobó que la transcripción *in vivo* e *in vitro* de GATA-1, se da a través de la regulación coordinada de diferentes factores eritroides, necesitando un



incremento en la expresión de los factores transcripcionales EKLF, SCL y NF-E2, y una represión de GATA-2, EpoR y Myb (139). A pesar de lo anterior, la expresión del gen GATA-1 en pollo está regulada principalmente por su propio producto (3).

La regulación de GATA-1 por modificaciones covalentes, resulta un aspecto novedoso de la función de GATA-1. El grupo de Boyes y colaboradores (1998) reporta que GATA-1 en células humanas se acetila *in vitro* por p300 (una de las acetiltransferasas de histonas) en dos motivos altamente conservados, ricos en lisinas, las cuales se encuentran en el dominio de unión al ADN. La acetilación de estos residuos estimula la actividad de GATA-1, incrementando significativamente su unión al ADN. En pollo, GATA-1 también se acetila *in vivo*, estimulando la transcripción (157). En ratón se reportó que la acetilación *in vitro* por CBP, no afecta la interacción del complejo GATA-1-ADN. Sin embargo, *in vivo*, mutantes en los residuos blanco de acetilación de la proteína GATA-1, resultaron incapaces de inducir la diferenciación celular (158). Sin embargo, la forma en la que la acetilación de GATA-1 interviene en la actividad transcripcional no queda clara (158). GATA-1 también puede ser fosforilada en su extremo carboxilo-terminal. A esta modificación se le ha atribuido también un incremento en la afinidad por un sitio de unión al ADN (139).

Dentro de los modelos de acción propuestos para GATA-1, se favorece la interacción peptídica con diferentes factores incluyéndose a si misma. Entre los diferentes factores que podrían unirse a este dominio amino estarían las acetil transferasas de histonas (HATs, por sus siglas en inglés) las cuales generarían una estructura de la cromatina favorable para que diversos factores de la maquinaria transcripcional interaccionen, incluyendo a la RNA pol II (162). Algunos factores que también podrían interaccionar con este dominio serían EKLF, él cual se ha observado que interacciona con HATs, FOG-1, él cual podría participar en la modulación de la actividad de GATA-1 y PU.1 involucrado en el bloqueo de la eritropoyesis (como se verá más adelante).

Como se puede ver muchas son las funciones que el factor transcripcional GATA-1 podría estar realizando en la regulación de los genes eritroides. Un papel novedoso en la regulación de los elementos de control a distancia y promotores, podría ser la generación de regiones de libres de nucleosomas.

## **GATA-1, cromatina y regiones libres de nucleosomas**

Una de las preguntas fundamentales concerniente a la función de un LCR (enhancer) es el mecanismo mediante el cual son capaces de interactuar y regular a los genes globina a distancia. Una de las hipótesis predice que la estructura de la cromatina característica de estos elementos es necesaria para su interacción funcional (163, 164). Durante el desarrollo eritroide, el LCR es requerido para generar una configuración abierta de la cromatina del locus  $\beta$ -globina, lo anterior es fácilmente detectable por la accesibilidad diferencial de las células eritroides a nucleasas exógenas (165). Se ha propuesto un modelo en el cual el enhancer HS2 de manera GATA-1-dependiente (entre otros factores) recluta HATs y forma un dominio de acetilación en la región promotora vía un modelo de progresión que puede ser bloqueado por el “insulator” (165) Uno de los primeros reportes (167) que le da un papel importante al posicionamiento nucleosomal en la regulación de promotores y enhancers, fue realizado en células de cerebro, encontrando al enhancer HS2 y al promotor  $\beta$ -globina estructurados en un arreglo constante de nucleosomas. Cuando repiten el mismo estudio en cultivos primarios de eritrocitos de embriones de pollo de 15 días (eri-15) donde se sabe que los genes globina están activos, ahora reportan a estas dos secuencias como libres de nucleosomas (2). Es importante remarcar que manejan un patrón de nucleosomas estructurado, cuando se ve la escalera de nucleosomas y como nucleosomas activos o ausentes, cuando ven un patrón difuso, sin tener pruebas directas de la presencia o ausencia de nucleosomas.

¿Pero quien genera la apertura de la cromatina en estas secuencias? Para que un factor transcripcional pueda tener acceso a su secuencia blanco, necesita tener la habilidad de poder mover o quitar los nucleosomas que están cubriendo las secuencias blanco (157). Se han propuesto 2 modelos para explicar la contribución de los factores transcripcionales específicos en la apertura de la cromatina. El primero propone que la interacción de factores aditivamente va a alterar la estructura de la cromatina. El segundo propone que la presencia de los factores en el sitio de interacción, incrementa la probabilidad de mantener a un nucleosoma fuera (169), en un equilibrio entre los factores transcripcionales y el nucleosoma, que cuando la concentración de factores transcripcionales se viera incrementada, el nucleosoma sería desplazado generando la formación de una HS y por lo mismo de una región libre de nucleosomas (164).

Como ejemplo del primer modelo, en el enhancer HS2, mediante un sistema de ratones transgénicos y células MEL transfectadas establemente, se encontró que el grupo de sitios AP-1, Sp-1 y GATA-1 son necesarios para generar una alteración en la estructura de la cromatina que genere un patrón de activación completa del enhancer. La eliminación de cualquiera de estos sitios genera un decremento en la intensidad de la HS (170). Este grupo propone que el desplazamiento de un nucleosoma, es uno de los eventos primarios que ocurren como consecuencia de la unión consecutiva y formación de un complejo de diferentes factores transcripcionales (170).

Sin embargo la presencia de un activador transcripcional, capaz de interactuar con el complejo de la cromatina, podría ser suficiente para generar un sitio SH al corte de la DNasa I y de esta forma facilitar el acceso al resto de la maquinaria transcripcional.

La expresión del factor transcripcional eritroide específico GATA-1 en un ambiente no eritroide, genera la formación de los sitios a la DNasa I, SH1-4 del LCR  $\beta$ -globina, en sus posiciones endógenas. A nivel de las modificaciones de las histonas, se genera una hiperacetilación característica de las líneas eritroides. Por lo anterior se propuso que GATA-1 es suficiente para estructurar en cromatina activa del LCR y generar un patrón de expresión de los genes a similitud del endógeno (140). ¿Podría ser GATA-1, en el caso de las líneas eritroides, el encargado de iniciar la cadena de asociaciones aditivas que generaran la formación de la SH? Cuando el grupo de Layon y colaboradores (2007) expresa al factor transcripcional GATA en un ambiente no-eritroide, no queda claro si esta proteína por sí sola es suficiente para generar un SH, ellos atribuyen el efecto observado a la cooperación con factores no-eritroides (140). Pero hay que recordar que GATA-1 interactúa en la mayoría de los promotores involucrados en la eritropoyesis. Por lo cual el efecto observado podría ser efecto de la activación de algún otro factor eritroide.

La expresión de algunos miembros de la familia de proteínas GATA, se encuentra altamente restringida. Se les ha propuesto como factores con una función decisiva que estaría controlando la transcripción programada a través del desarrollo (170-173). GATA-1 ha sido implicada en la generación de estructuras de la cromatina activa en prácticamente todos los promotores eritroides (157, 163). Se le ha propuesto hasta como un regulador maestro de la eritropoyesis (140).

En un estudio *in vitro*, el grupo de Felsenfeld (1998) reporta la unión de GATA-1 a un nucleosoma generando un complejo estable en solución. Esta unión genera la ruptura entre los enlaces establecidos entre el ADN y el nucleosoma en diferentes regiones. De hecho encuentran un patrón de corte a la DNasa I muy similar al visto cuando GATA esta unida a ADN libre de nucleosomas (157). Estos datos apoyan el modelo de interacción cooperativa de factores al complejo ADN/nucleosoma. Se sugiere que parte de la energía requerida para romper las interacciones electrostáticas entre el ADN y el nucleosoma, podrían ser compensadas por un nuevo tipo de interacciones entre el ADN y los residuos básicos del factor GATA-1 (157). Lo anterior me lleva a cuestionarme si los sitios de hipersensibilidad podrían ser regiones libres en nucleosomas o este tipo de complejos. Para el caso de los sitios de hipersensibilidad generados en los genes eritroides, GATA-1 podría ser la responsable.

### **La relación entre GATA-1, la RNA polimerasa II y los enhancers**

Existe evidencia para apoyar la idea de que los enhancer reclutan a la RNA pol II y en algunos casos la RNA pol II es transferida al promotor vía la formación de un asa entre los dos elementos (174, 175). La RNA pol II también a sido encontrada en la región intergénica comprendida entre el enhancer y el promotor, apoyando el movimiento de ésta desde el enhancer al promotor (176). Estudios recientes han mostrado que los factores generales de transcripción y la pol II son reclutados a elementos tipo enhancer, siendo esta unión un evento temprano en la activación génica y funcionando como centro de atracción para el ensamblaje del complejo de pre-iniciación. De esta forma, se ha propuesto que los enhancers podrían estar regulando el “timing” de la expresión génica durante el desarrollo, diferenciación y ciclo celular (180).

La expresión de GATA-1 induce el reclutamiento de la RNA pol II a los sitios HS1, HS2 y HS3 del LCR humano y murino, mientras que el reclutamiento a HS4 resulta ser GATA-1-independiente (181). Lo que nos muestra la existencia de diferentes mecanismos para atraer a la RNA pol II a sus sitios de acción. Se proponen que la encargada de atraer a RNA pol II puede ser GATA-1 pero en una forma dependiente de NF-E2 (181). A partir de nuestras investigaciones, resulta muy interesante que el enhancer  $\alpha$ -globina presenta esta combinación de sitios de interacción a factores de transcripción y la interacción *in vivo* de GATA-1 y la RNA pol II (ver resultados).

Un mecanismo alternativo mediante el cual la RNA pol II pudiera ser reclutada a secuencias tipo enhancer, sería a través del complejo mediador<sup>10</sup>. Se ha propuesto que el mediador es crucial para conectar a activadores que se unen al enhancer con factores generales de transcripción asociados al complejo de inicio de la transcripción, entre ellos la RNA pol II (182). El reclutamiento del mediador a secuencias enhancers, *in vivo*, es activador-dependiente, en condiciones en que la maquinaria básica de la RNA pol II no está asociada aún al promotor, indicando que el reclutamiento del mediador es un paso previo a la activación (183, 184). La visión universal propone que el mediador es un puente entre las proteínas activadoras y la maquinaria basal de la RNA pol II (185-191).

Un modelo alternativo mediante el cual el reclutamiento de la RNA pol II a los enhancers ha sido estudiado, son las construcciones en minicromosomas hechas por el grupo de Ann Dean (2004, 2006). El bloqueo del enhancer cHS2 por el “insulator” HS4, causa la concentración de la RNA pol II se vea disminuida en el promotor (así como los niveles de transcrito reportados) y por el contrario, se vea incrementada en el enhancer y secuencias aledañas. Este acumulamiento de RNA pol II en el enhancer, no se lleva a cabo en la ausencia del factor transcripcional GATA-1, proponiendo un papel directo por parte de GATA-1 en el reclutamiento y estabilización de RNA pol II (166, 192).

Como se ha mostrado hasta ahora, muchas son las funciones y los papeles que podría estar realizando GATA-1 en la regulación de la eritropoyesis a través de la diferenciación y desarrollo. Por lo anterior propongo que el factor transcripcional GATA-1, podría ser ese activador inicial que al ser expresado es capaz de reconocer a su secuencia blanco en el contexto de la cromatina e iniciar la cadena de eventos necesarios para remodelar la estructura de la cromatina en los enhancers, reclutar a la RNA pol II y mediante alguno de los mecanismos mencionados activar a sus promotores blanco. Para tener un panorama más amplio de las características que posea este activador, hablaré de manera resumida acerca de los factores proteínicos con los cuales podría estar regulando la actividad del enhancer.

### **FOG (“Friend of GATA”)**

Mediante un ensayo de doble híbrido se identificó la interacción entre FOG-1 y GATA-1 (194), de ahí el nombre de FOG (por sus siglas en inglés, “Friend Of GATA”). FOG-1

---

<sup>10</sup> El mediador es un complejo altamente conservado de aproximadamente 25 proteínas, se asocia a la RNA pol II y es reclutada a enhancer transcripcionalmente activos *in vivo*.(182)

y GATA-1 se expresa en eritroblastos embrionarios y megacariocitos. FOG-1 coopera con GATA-1 en promover la diferenciación a eritrocitos y megacariocitos (195). En ratones deficientes a FOG-1, se ha observado una interrupción en la diferenciación, similar al que se ha reportado en ratones que no expresan GATA-1 (198). Por otra parte, una mutante para GATA-1, incapaz de interactuar con FOG, genera una desregulación en los genes blanco de GATA-1, entre ellos los genes  $\alpha$ - y  $\beta$ -globina, impidiendo llegar a la maduración eritroide completa (197). FOG-1 posee nueve dedos de zinc, se ha propuesto que son el producto de la duplicación del dominio de dedos de zinc de GATA-1 (139). La relación GATA:FOG en la regulación tiene un factor contexto-dependiente muy importante, GATA-1 regula la activación transcripcional de manera dependiente de FOG-1 (198), independiente de FOG (198), la represión dependiente de FOG (198), independiente de FOG (199) y la activación independiente de FOG unido a un represor dependiente de FOG (200). Se sabe que este complejo estimula la actividad transcripcional en el promotor de p45 NF-E2 en células eritroídes y megacariocíticas, y reprime la actividad en el promotor de EKLF (93). Este tipo de regulación podría ser regulado por secuencias externas, en el promotor  $\alpha$ IIb en células de megacariocíticas, se encontró que la secuencias tipo Ets son suficientes para convertir a FOG-1 de un represor de GATA-1 a un activador. Por inmunoprecipitaciones se logró determinar que a esta secuencia Ets se une el factor transcripcional Fli-1. Esta proteína es capaz de activar la expresión del promotor  $\alpha$ IIb en presencia de GATA-1 y FOG-1 (198). FOG-1 también puede reclutar proteínas represoras como al co-represor CtBP, las cuales se ha visto que ayudan a FOG-1 a modular la actividad de GATA-1 negativamente (139).

Ya que la familia de FOG no se puede unir al ADN por si sola, son reclutadas por la familia de GATA en una forma que asemeja una relación co-activador/co-represor (202, 155). También se ha propuesto que FOG podría estar participado en la interacción de GATA a su secuencia blanco de una forma contexto dependiente, con una mutante condicional de GATA-1 que no puede unir a FOG, pero que si se puede unirse al ADN *in vitro*. Encontraron que al no permitir que FOG interacte con GATA-1, se pierden los patrones de acetilación observados anteriormente en el promotor  $\beta$ -globina, esto lo atribuyen a la pérdida en la interacción de GATA-1 con este promotor, sin embargo su interacción a HS3 resulta FOG-independiente (153). FOG solo necesita 4 de sus nueve dedos de zinc para interactuar con GATA-1, dejando libres 5 que podrían

interaccionar con otro factor GATA-1 y formar un puente proteico necesario para que el LCR interaccione con sus promotores. Sin embargo existe un estudio que mostró que FOG con la presencia de un solo dedo de zinc (suficiente para que interaccione con GATA-1), puede inducir la diferenciación (202, 204). Sin embargo, estudios recientes por captura conformacional 3C, se demostró la interacción física entre el LCR y el promotor  $\beta$ -major del dominio  $\beta$ -globina, se necesita de la presencia de los factores GATA-1 y FOG-1 (202).

### **El factor de la familia Maf eritroide específico (NF-E2)**

NF-E2 es una proteína tipo “leucine-zipper” (bZIP) perteneciente a la familia de proteínas Maf. Se une a sitios tipo AP-1 en el LCR y distintos elementos de regulación de los genes globina (205). Esta compuesta por dos proteínas bZIP, p45NF-E2 que es eritroide-específica y p18NF-E2 que es una proteína ubicua. La especificidad de su unión esta dada por p18NF-E2, sin embargo no es una proteína esencial. Lo anterior se ha demostrado en megacariocitos y células eritroides mediante la generación de mutantes “null”, las cuales contrario a lo esperado no muestran ningún fenotipo característico, esto indica probablemente que p18NF-E2 puede ser substituida por otras proteínas de la familia Maf (148).

Diferentes evidencias experimentales apoyan un papel por parte de NF-E2 en la regulación de la expresión de los genes globina: La expresión de la sub-unidad p45 está restringida a linaje hematopoyético, en las que se ha encontrado en gran abundancia al factor NF-E2 unido a elementos de reconocimiento Maf (MARE) que presenta la secuencia (T/C)GCTGA(C/G)TCA(T/C) de unión al ADN (205, 206); En la línea eritroide de ratón CB3, la cual no expresa a la sub-unidad p45 de NF-E2, se observa un severo defecto en la expresión de los genes  $\alpha$ - y  $\beta$ -globina, efecto que se corrige con la sobre-expresión de NF-E2 (206); NF-E2 se une al sitio HS2 en células eritroleucémicas y de hígado de ratón en estado fetal. Su unión es dependiente de la subunidad p45 y de la secuencia MAREs (206). NF-E2 también podría estar participando en la regulación de los genes globina, a través de su interacción con la maquinaria basal transcripcional, pues se ha reportado su interacción con TAF<sub>II</sub>130, proteína asociada a TBP que contribuye a la especificidad del complejo de iniciación de la transcripción (148).

Se ha visto que NF-E2 y GATA-1 son dos activadores hematopoyéticos específicos necesarios para la expresión del dominio  $\beta$ -globina. En ratón se ha reportado que estos factores reclutan a la RNA polimerasa a los promotores de genes adultos (207). Por estudios *in vitro*, se ha mostrado que los sitios de unión a NF-E2, son importantes para la actividad de remodelaje de la cromatina y la correcta formación del sitio de hipersensibilidad en el HS2 en minicromosomas (208, 209). Mutaciones puntuales en el enhancer HS2 no pueden cortar su efecto potenciador y es solo hasta que generan una serie de mutaciones en los sitios de interacción a GATA-1 y CACCC en el enhancer y promotor se logra llevar la expresión a casi cero, por lo cual se propuso que el número de sitios de interacción está generando estabilidad al complejo enhancer-promotor (133). Se propone un modelo en el cual NF-E2 sea el centro de enucleación del enhancer HS2, GATA-1 estaría mediando la comunicación con el promotor  $\epsilon$ -globina y los sitios conservados CACCC estarían dando estabilidad al complejo enhancer-promotor (133).

Una función novedosa que presenta esta proteína fue reportada por el Groudine (2001). Localizan en una región de heterocromatina a la sub-unidad p18 y en una región de eucromatina a la sub-unidad p45. Al inducir las células a diferenciación, se observa que la sub-unidad p18 se relocaliza hacia zonas de eucromatina. También observan que el locus  $\beta$ -globina se re-localiza a un compartimiento de eucromatina, seguido de su activación (211). Estos resultados le darían un papel a NF-E2 muy importante en la regulación transcripcional vía la dinámica nuclear.

### **El miembro fundador de la familia “krüppel-like” (EKLF)**

Esta proteína pertenece al grupo de proteínas que reconoce el motivo CACCC (148). Nombrado EKLF por la gran similitud que poseen sus dedos de zinc con el gen “krüppel” (212). Las proteínas de esta familia, reciben el nombre del sitio en que se expresan mayoritariamente como es el caso de EKLF (“Erythroid Krüppel Like Factor”), que como su nombre lo indica, esta restringida a líneas eritroides (213). Se le ha dado un papel relevante en el fenómeno de “switching” (o “encendido” de los genes globina), donde sin importar que la abundancia de esta proteína es similar en todos los estadios de diferenciación en células en cultivo, esta proteína funge como un activador crítico de la expresión del gen  $\beta$ -globina de adulto, a diferencia del gen fetal, (214).



EKLF es un regulador estado-específico de los genes adultos  $\beta$ -globina, el cual influye a nivel de la estructura de la cromatina. A esta proteína se le ha relacionado vía su extremo C-terminal con la generación de patrones de la cromatina abierta, posiblemente vía su interacción con BRG1, un componente del complejo remodelador de la cromatina SWI/SNF (215, 216). Se ha descrito su interacción con la proteína de unión a CREB (CBP), la cual presenta actividad de acetilasa de histonas (219). Se ha demostrado que EKLF en conjunto con el complejo remodelador de la cromatina ATP-dependiente, E-RC1, puede generar sitios de hipersensibilidad en el promotor  $\beta$ -globina humano. A este complejo se le ha descrito como un complejo de remodelación tipo SWI/SNF ATP-dependiente (217, 218). Se han reportado  $\beta$ -talasemias producidas por mutaciones en esta secuencia, en las cuales se observa una reducción en los niveles de expresión atribuida a una disminución en el reconocimiento de la proteína EKLF por su secuencia blanco (214).

### **El factor Ying Yang 1 (YY1)**

YY1 es una proteína de 65 kDa cuya secuencia peptídica es homóloga a la de la familia de factores Krüppel de *Drosophila*, por lo anterior se le ha identificado como miembro de esa familia (220, 221). YY1 es un factor transcripcional ubicuo (también conocido como NF-E1, UCRBP y CF1), miembro de la familia Polycomb (222), se le ha descrito como una proteína tanto activadora como represora, dependiendo del contexto del promotor y de los co-factores que se le asocian (223). Se ha demostrado que se une a los promotores de los genes  $\epsilon$  y  $\gamma$  del dominio  $\beta$ -globina de humano en estados embrionario y fetal (148). El papel fundamental que juega YY1 en procesos biológicos durante el desarrollo, se muestra con un “Knockout” en ratón, que resulta en la muerte durante el proceso de pre-implantación (224).

La naturaleza modular de la proteína apoya su papel biológico bi-funcional. YY1 contiene 4 dedos de zinc, dos dominios de represión, uno en el cuerpo de la proteína (residuos 170-200) y otro en el extremo C-terminal, y un dominio de activación en el extremo N-terminal (222, 225, 226). El grupo de Seto (224) mostró en líneas celulares, que YY1 se acetila en su dominio central por dos acetil-transferasas de histonas de humano PCAF y p300, y en el extremo C-terminal por la acetil-transferasa PCAF. La acetilación del dominio central es necesaria para la actividad represora, haciendo mas

eficiente su unión a HDACs. La unión de HDAC's trae como consecuencia la desacetilación de las regiones que YY1 regula. Por otro lado, la acetilación del extremo C-terminal le permite a YY1 reclutar HDAC's, lo cual le confiere una actividad de desacetilasa de la histona H4. Cabe resaltar que el extremo C-terminal parece no poder ser desacetilado (224).

YY1 interacciona con diversos reguladores transcripcionales clave, entre ellos TBP, TFIIB, TAFII55, Sp1 y E1A. Se ha postulado que la relevancia de esta interacción influye en la determinación de la función que desempeñará esta proteína en un promotor específico, siendo el grupo de las desacetilasas de histonas y las actil-transferasa de histonas, las encargadas de conferirle a YY1 una actividad represora o activadora (223). Se ha determinado que YY1 interacciona con las HATs p300 y CBP y con las HDACs HDAC1, HDAC2 y HDAC3 (224).

Existen diferentes mecanismos mediante los cuales YY1 podría reprimir la actividad transcripcional, entre ellos propone un modelo en el cual YY1 estaría interfiriendo con la unión de *trans*-activadores por un impedimento estérico (220, 225). Este modelo resulta interesante en el contexto de nuestras investigaciones, ya que hemos localizado un sitio de unión a YY1, cercano al sitio de unión al factor transcripcional GATA-1 en la región del enhancer  $\alpha$ -globina (227, ver resultados). Por otro lado, se ha mostrado que la interacción de YY1 con co-represores estaría actuando a nivel de la estructura de la cromatina, desacetilando localmente las histonas y como consecuencia generando una estructura de la cromatina compacta y por lo tanto represora para la actividad transcripcional (223, 227). Sin embargo existe la interrogante de cómo actúan estos co-represores para generar una estructura de la cromatina compacta si los factores siguen posicionados en sus secuencias blanco, la respuesta a lo anterior podríamos encontrarla en las observaciones hechas por diferentes grupos, donde se ha descrito que el estado de acetilación de estos factores, altera la estabilidad de la interacción proteína-ADN (157).

Dentro de los modelos que se han propuesto para explicar su función activadora, se ha propuesto que la interacción directa de YY1 con los factores generales de transcripción estaría estimulando la transcripción. Esta proteína puede interactuar con factores generales de la maquinaria transcripcional como TBP, TFIIB y TAFII55 (223). La inhibición de una actividad represora o el desenmascaramiento de su dominio de activación por alguna proteína tipo la proteína de adenovirus E1A (223). Finalmente vía

acetil transferasas de histonas o HAT's. A partir del descubrimiento de la interacción entre YY1 con las HATs CBP y p300, se ha postulado la posibilidad de que el reclutamiento de estas enzimas, podría estar generando una estructura de la cromatina laxa, favoreciendo el ensamblaje de la maquinaria transcripcional (223).

### **El factor general de transcripción Sp1**

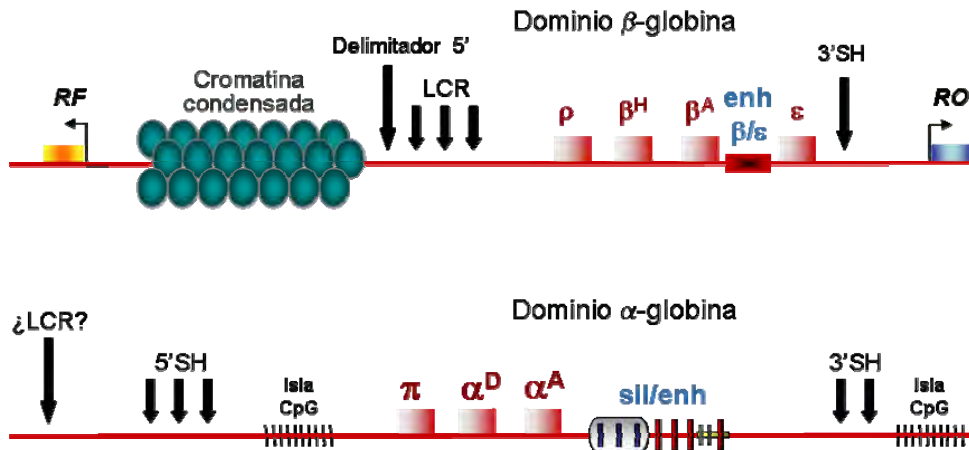
Sp1 pertenece al grupo de proteínas que reconoce el motivo CACCC (228). También reconoce el motivo CCCGCC (228). Entre las características de la familia a la que Sp1 pertenece, está el triple dominio de dedos de zinc con que cuenta (219). Sp1 es una proteína nuclear abundante en la mayoría de las células, por lo cual se le ha considerado como un factor transcripcional general. Sin embargo, se ha reportado que los niveles de expresión varían a lo largo del desarrollo y de los tipos celulares (219).

Se ha propuesto que Sp1 juega un papel en la activación de los genes globina mediada por el LCR (229). Se han encontrado sitios potenciales de unión a Sp1 en combinación con sitios GATA, recordando que GATA-1 y Sp1 (229). En ratones transgénicos, la activación de los genes globina mediada por LCR, es dependiente de la presencia de los sitios Sp1. En un intento por comprender el papel que juega Sp1 en la regulación transcripcional, se generaron ratones homocigotos para la mutación de Sp1 ( $Sp1^{-/-}$ ). A nivel embrionario los ratones muestran un retraso en el desarrollo, un gran rango de anomalías y mueren alrededor del onceavo día de gestación (229). Resulta interesante, el que los genes  $\alpha$ - y  $\beta$ -globina se sigan expresando normalmente, por lo que Sp1 no parece ser una proteína clave en la eritropoyesis, por lo menos a nivel embrionario (229). Finalmente, diferentes trabajos han sugerido un papel por parte de Sp1 en el mantenimiento de islas CpG libres de metilación (229).

## El modelo de estudio y antecedentes particulares

### Dominios Globina

El transporte de oxígeno es esencial en el desarrollo de cualquier organismo superiores, sin embargo los requerimientos varían durante cada etapa del desarrollo. Para poder trasladar el oxígeno a lo largo del organismo se requiere de la hemoglobina, la cual es una proteína sintetizada en los eritrocitos (134). La hemoglobina es una proteína tetramérica compuestas de dos tipos de cadenas polipeptídicas, denominadas  $\alpha$  y  $\beta$ . Cada cadena es producida a partir de dos dominios multigénicos localizados en diferentes cromosomas, pero co-regulados a lo largo del desarrollo. Dichos dominios están compuestos de genes embrionarios ( $\epsilon$  y  $\rho$  en el dominio  $\beta$ , y  $\pi$  en el dominio  $\alpha$ ) y genes adultos ( $\beta^A$  y  $\beta^H$  en el dominio  $\beta$ , y  $\alpha^D$  y  $\alpha^A$  en el dominio  $\alpha$ ) (134).

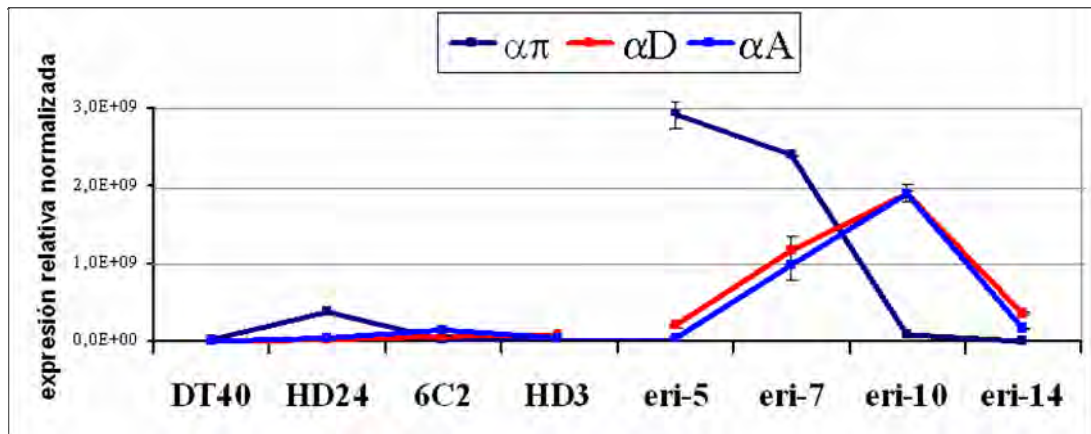


**Figura 14.** El dominio  $\beta$ -globina de pollo (arriba), abarca alrededor de 30 kb en el cromosoma 1, esta compuesto por los genes globina  $\rho$ ,  $\beta^H$ ,  $\beta^A$  y  $\epsilon$ . Dichos genes se encuentran bajo la influencia de un elemento tipo LCR y del enhancer  $\beta^{A/\epsilon}$ . Dos sitios de hipersensibilidad al corte por la DNAsa I enmarcan el dominio  $\beta$ -globina, conocidos como  $\text{CHS4}$  y  $3'HS$ , los cuales poseen funciones de delimitadores cromatínicos o “insulators”. Río arriba del dominio se encuentra una región de 16 kb de cromatina condensada, formada principalmente de secuencias repetidas CR1. Los dominios génicos que rodean al dominio  $\beta$ -globina son el del gen receptor de folato ( $RF$ ) y un gen de receptor olfatorio ( $RO$ ). El dominio  $\alpha$ -globina de pollo (abajo), abarca una región aproximada de 40 kb, aunque no se conocen los límites exactos del mismo. El dominio contiene a los genes globina  $\pi$ ,  $\alpha^D$  y  $\alpha^A$ . Los principales elementos de regulación hasta el momento descritos son, el silenciador ( $sil$ ) y el enhancer ( $enh$ ). Los genes se encuentran enmarcados por islas CpG; la ubicada en el extremo 5' corresponde al sitio de inicio de la transcripción para el gen  $ggPRX$ . Río arriba se encuentran 3 sitios de hipersensibilidad, dentro de los cuales se ha caracterizado un elemento con actividad de bloqueo de enhancer. 5 kb arriba de dichos sitios se localiza una secuencia conservada entre vertebrados, a la cual se le ha asignado una posible función de LCR. Río abajo se localizaron una serie de SH con función desconocida.

Durante el desarrollo del embrión de pollo, las células eritroides primitivas, las cuales aparecen en la circulación entre los días 4 y 5 post-fecundación, expresan los genes embrionarios (135). A partir del 5º día del desarrollo el linaje definitivo (adulto) inicia la expresión de los genes adultos ( $\beta^A$  y  $\beta^H$  en el dominio  $\beta$ , así como  $\alpha^D$  y  $\alpha^A$  en el dominio  $\alpha$ ). A este proceso se le conoce como recambio en la expresión de los genes globina (del inglés: “hemoglobin switching”) (134). Con el fin de sintetizar las distintas hemoglobinas en el momento específico del desarrollo en el que son necesarias, la expresión de los genes globina se presenta en al menos tres etapas. En la primera etapa se expresan los genes embrionarios, en la etapa fetal se co-expresan los embrionarios y adultos, y en la tercera etapa se expresan sólo los adultos. En estudios recientes en el dominio  $\beta$ -globina de humano, han determinado que la expresión sucesiva de los genes está determinada por la formación de tres sub-dominios de cromatina, y que estos sub-dominios parecen estar delineados por la presencia de transcritos intergénicos (230). Se ha sugerido que este proceso está dado por una estrecha relación entre la regulación genética y la epigenética. Los dominios globina de pollo representan un modelo interesante para el estudio de la regulación transcripcional, a nivel de dominio y a nivel local, a nivel de secuencia, estructura de la cromatina y dinámica nuclear, lo anterior con base en que dichos dominios han sido claves en la identificación y caracterización, tanto estructural como funcional, de elementos de regulación génica.

### **El dominio $\alpha$ -globina**

El dominio  $\alpha$ -globina se encuentra altamente conservado en vertebrados, su expresión se observa sólo en líneas celulares eritroides que se expresan de manera diferencial a lo largo del desarrollo y diferenciación eritroide, sin embargo los mecanismos que regulan la apertura del dominio así como la expresión secuencial de los genes ha sido muy poco estudiados.



**Figura 15.** Expresión relativa determinada por PCR duplex en extractos totales de RNA en la línea linfóide DT40, en las líneas eritroides HD3, 6C2 y HD3 y en diferentes estados del desarrollo (Guerrero G *et al*; en preparación).

El dominio  $\alpha$ -globina de pollo (Fig. 14), se encuentra contenido en una región de aproximadamente 40 kb, donde el gen embrionario se expresa en líneas primitivas y dos genes adultos  $\alpha^D$  y  $\alpha^A$ , que se expresan en células eritoides definitivas (231). La expresión de los genes  $\alpha$ -globina, fue detalladamente caracterizada, en las diferentes líneas eritroides que se utilizaron en este trabajo y a través del desarrollo (Fig. 15; Guerrero G *et al*; en preparación). En el extremo 3' se localizaron un silenciador y un enhancer con motivos de unión a factores transcripcionales específicos (232). Varios sitios de hipersensibilidad (SH) han sido caracterizados a lo largo del dominio. De 11 a 15 kb río arriba de los genes se encuentran tres sitios SH (233) que se han sugerido como el límite superior del dominio. En el extremo 3' del dominio no se ha definido un límite, sin embargo estudios recientes en el laboratorio sugieren la presencia de tres sitios HS eritroide específicos río abajo de los elementos silenciador/enhancer (Furlan-Magaril *et al*. en preparación).

A diferencia del locus  $\beta$ -globina, el locus  $\alpha$ -globina se encuentra en una región de la cromatina constitutivamente abierta (231). Es interesante que esto también ocurra en los dominios  $\alpha$ -globina en humano y ratón, es decir, que se localicen en una región cromosómica constitutivamente abierta (66).

## **El dominio $\beta$ -globina**

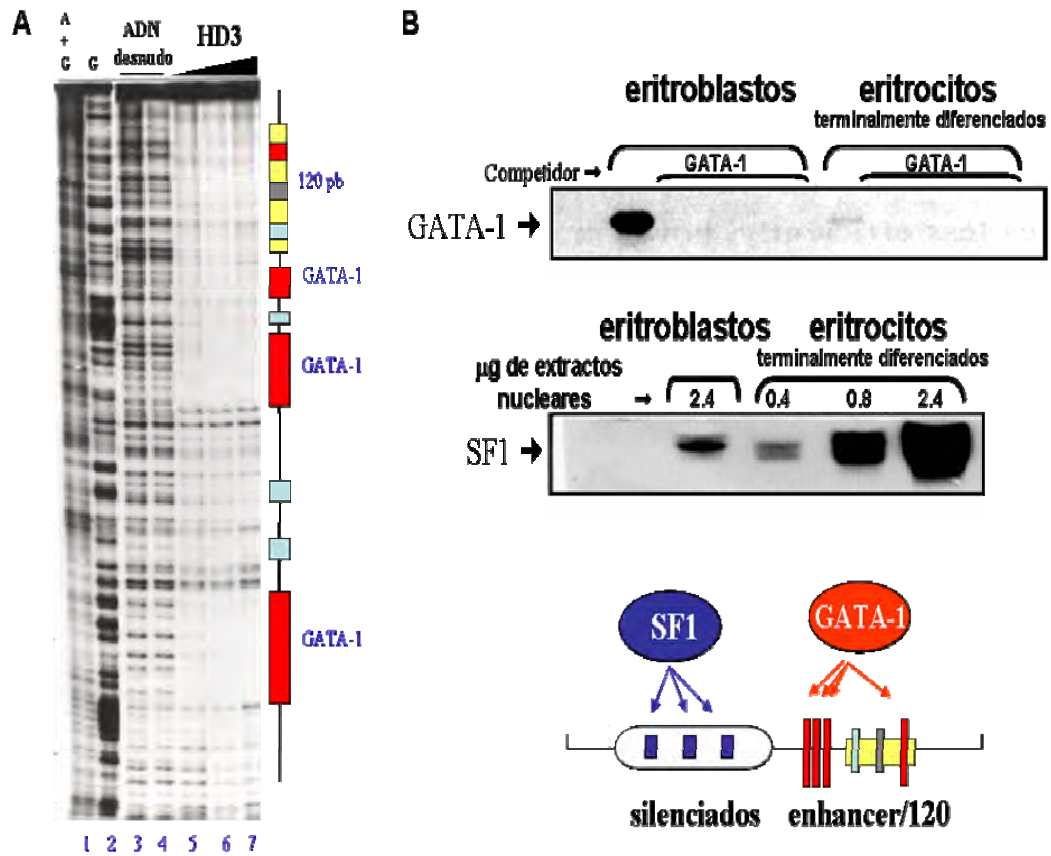
Este dominio está conformado por los genes que codifican para las cadenas  $\beta$ -globina de la hemoglobina. La regulación del dominio  $\beta$ -globina ha sido intensamente estudiado en mamíferos, principalmente en ratón y humano ya que se encuentra altamente conservado y diversas deleciones a lo largo del dominio conllevan a enfermedades como las anemias conocidas como  $\beta$ -talasemias (230, 234, 235). Tanto en el ratón como en los humanos el orden de los genes en el dominio corresponde con la etapa del desarrollo en que deben expresarse (236). El dominio  $\beta$ -globina de pollo incluye una región de 33 kb, donde se localizan un gen embrionario  $\epsilon$  y un gen fetal  $\rho$ . Estos genes se expresan en líneas celulares primitivas, mientras que los genes adultos  $\beta^A$  y  $\beta^H$ , se expresan en células eritroides definitivas (230). En el extremo 3' entre el promotor  $\beta^A$  y  $\epsilon$ , se localizó un enhancer considerado como fuerte, nombrado enhancer  $\beta^A/\epsilon$  (237). En el extremo 5' se localizó una región de cuatro SH a la DNasa I, tres de los cuales y junto con el enhancer  $\beta^A/\epsilon$ , constituyen el LCR del dominio (208).

Por medio de diversas técnicas como el RNA TRAP y la captura conformacional de cromosomas (242, 243) se ha determinado que diferentes elementos dentro de un dominio se encuentran en contacto con otros y por lo tanto permiten reconstruir un arreglo espacial de una región del genoma. Gracias a estos estudios se logró determinar que existe una estructura denominada nodo de cromatina activa (ACH por "Active Chromatin Hub") (Figura 20; 243). Esta estructura permite la interacción directa de los elementos del LCR con sus genes blanco en células adultas eritroides. La estructura del ACH depende de la unión de factores eritroides a sus sitios blancos, como es el caso de EKLF, GATA-1 y FOG (93, 245).

## **El elemento silenciador/enhancer $\alpha$ -globina de pollo**

El alineamiento de las secuencias primarias de pollo y pato mostró una región central con una alta homología, en particular: tres sitios consenso para la unión del factor eritroide específico GATA-1 fueron identificados y su unión confirmada por experimentos de interacción proteína-ADN (Fig. 16; 246, 232). Una región de homología denominada 4ª caja de homología, correspondiente a un fragmento de 120 pb (Fig. 17; 232, 247; datos no publicados). Dentro de esta región se encuentra una cuarta

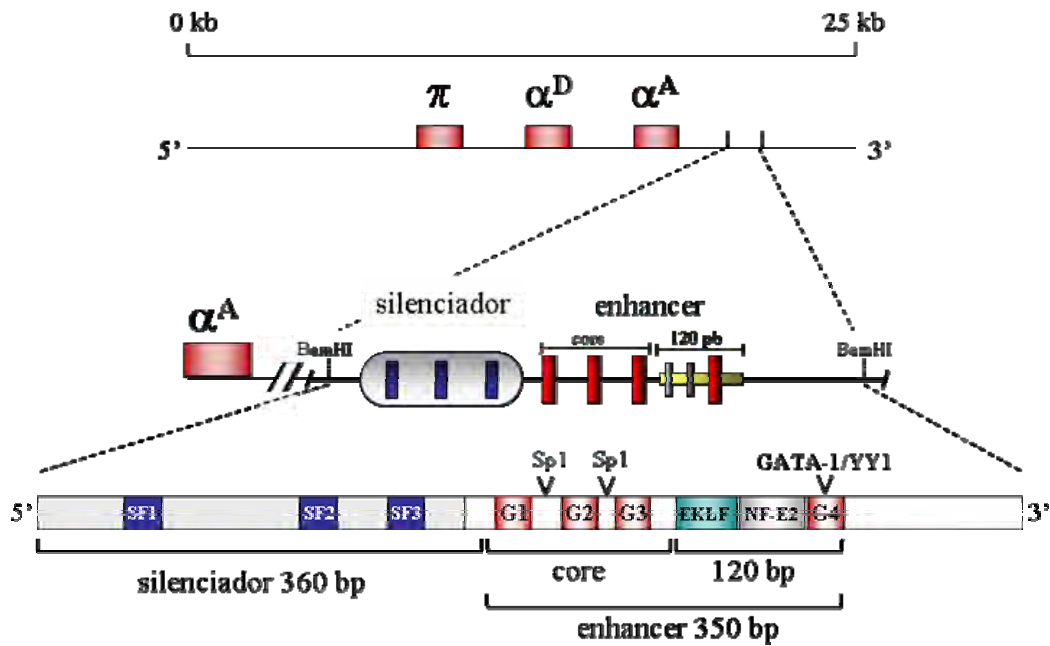
secuencia consenso de unión para el factor GATA-1, un sitio putativo AP-1/NF-E2 y secuencias repetidas GGTGG ó CACC a la cual probablemente se una el factor EKLf.



**Figura 16.** A, Ensayo de protección a la DNasa I *in vitro*, donde se muestran las ventanas de protección al corte, generadas por la interacción de diferentes proteínas en la región del “enhancer. B, Geles de retardo (arriba) mostrando la interacción proteica diferencial, que se presenta cuando se utilizan las sondas (abajo) para GATA-1 ó para SF1 (232).

Estudios funcionales y de interacción proteína-ADN identificaron la unión de factores nucleares asociados al elemento tipo silenciador, demostrando la del factor SF1 cuya secuencia codificante y producto peptídico se desconocen a la fecha (232). Por estudios de retardo se determinó la interacción de este factor al silenciador  $\alpha$ -globina, mostrando una mayor abundancia en extractos de células terminalmente diferenciadas, en los cuales coincidentemente los genes globina ya no se expresan (Fig. 16; 232).





**Figura 17.** En la parte superior se muestra un esquema del dominio  $\alpha$ -globina de pollo, siendo  $\pi$  el gen embrionario, y  $\alpha^D$  y  $\alpha^A$  los genes adultos que se expresan en células eritroides definitivas. En la parte inferior se muestra el fragmento de ADN de 1036 pb que comprende al silenciador (azul) con los sitios de unión a factores SF (1, 2 y 3), el enhancer (rojo) compuesto por tres sitios de unión al factor eritroide específico GATA-1 y la región de 120 pb de homología (amarillo) que presenta un cuarto motivo de unión a GATA-1 (rojo) y sitios de unión para EKLf (verde) y AP1/NF-E2 (Gris; 232).

El enhancer  $\alpha$ -globina fue inicialmente descrito como una secuencia de 300 pb, compuesta por tres sitios de unión al factor GATA-1 (Fig. 17). Dicho elemento puede *trans*-activar a los promotores  $\alpha$ -globina, sugiriendo que sólo actúa de manera dependiente al estado del desarrollo y del tipo de hemoglobina necesaria en cada etapa (135). El enhancer contiene dos sitios de reconocimiento al factor ubicuo Sp1 intercalado entre los sitios GATA-1. Su función no se ha estudiado, pero se propone que los sitios Sp1 coadyuvan en la *trans*-activación dirigida por GATA-1, además de tener una demostrada participación en procesos de remodelaje de la estructura de la cromatina (75, 249). En nuestras condiciones experimentales, se ha demostrado que la concentración del factor transcripcional GATA-1 que se une al enhancer, es inversamente proporcional a la del factor SF1 que se une al silenciador (232). Esto sugiere que la función de estos dos factores pudiera ser antagonista y específica de las distintas etapas del desarrollo. ¿Cuáles serán los elementos de regulación, necesarios para que dos elementos con actividades antagónicas puedan ser regulados? y ¿cuál será

la contribución de la estructura de la cromatina? son dos de las preguntas que nuestro grupo de trabajo esta intentando contestar.

## **Planeamiento del problema**

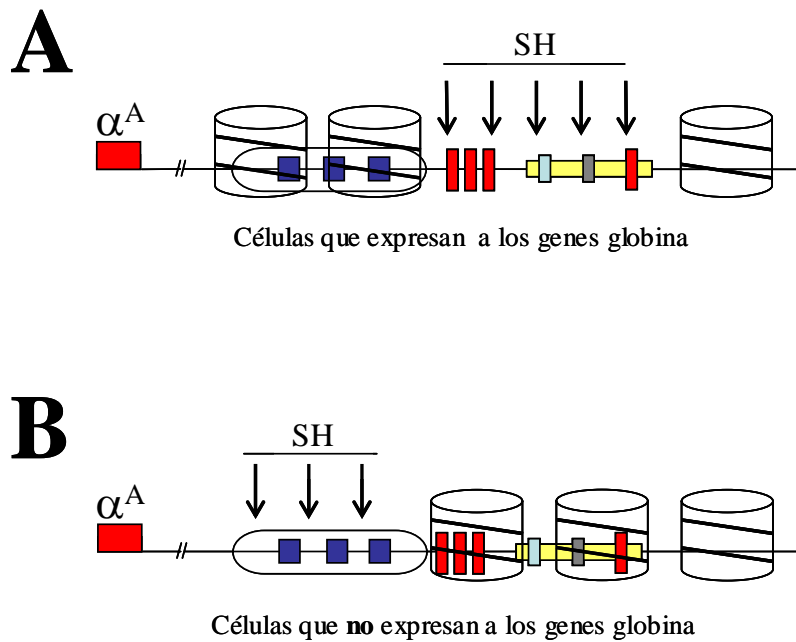
Los seres vivos están constituidos por diferentes proteínas resultado de la expresión en tiempo y tejido específica de los genes contenidos en el genoma. La expresión de cada gen es regulada a tres niveles, a nivel de secuencia, por los diferentes motivos de unión a factores transcripcionales que contienen las secuencias promotoras y los elementos a distancia que les confieren tiempo y tejido especificidad. A nivel de la estructura de la cromatina, determinante para almacenar al genoma al interior del núcleo y regular positiva y negativamente la expresión de los genes, exponiendo ó bloqueando la accesibilidad de sus secuencias y constituyendo un blanco sujeto a modificaciones post-traduccionales, con consecuencias epigenéticas. Finalmente a nivel de la estructura del núcleo (este último nivel no será abordado en este trabajo).

Entre los diferentes elementos de regulación a distancia encontramos a los enhancers. Se ha propuesto que su efecto se lleva acabo de manera tiempo y tejido específica, incrementando los niveles de expresión y/o aumentando la probabilidad de que los genes que regulan sean expresados. Los mecanismos mediante los cuales son regulados y regulan a sus genes, no son conocidos. Se ha propuesto que los diferentes sitios de interacción a factores transcripcionales y su estructuración en cromatina activa son vitales para su regulación y efecto a distancia. Una de las grandes incógnitas respecto a su funcionamiento, es la forma en la que un activador transcripcional puede reconocerlos e iniciar la cadena de eventos transcripcionales que terminarán en la síntesis de una cadena polipeptídica.

Este proyecto de investigación se apoya en un modelo único de estudio que nos permite analizar a lo largo la diferenciación y del desarrollo la actividad de un elemento silenciador y uno enhancer, desde el punto de vista de su efecto en la expresión de los genes globina y su relación con la estructura de la cromatina. En particular este trabajo de tesis pretende entender la participación de la estructura de la cromatina en la regulación de estos elementos.

## Hipótesis

Proponemos que la estructura de la cromatina y la actividad de los elementos silenciador y enhancer, localizados en el costado 3' del dominio  $\alpha$ -globina de pollo, debe ser regulada diferencialmente en distintas etapas del desarrollo eritroide. En los estadios en los cuales los genes globina se encuentren activos, la estructura de la cromatina de los elementos silenciador y enhancer, deberá remodelarse y volverse accesible al corte por nucleasas. Por el contrario, cuando los genes globina se encuentren inactivos, la estructura de la cromatina deberá observarse inaccesible al corte por nucleasas (Fig. 18). Por lo tanto, la estructura de la cromatina deberá estar contribuyendo en la actividad regulada del enhancer 3'  $\alpha$ -globina de pollo.



**Figura 18.** Esquema representativo de la hipótesis planteada en este trabajo de tesis doctoral. Cuando se expresen los genes globina, proponemos que se observará remodelada la región del enhancer (esta contiene cuatro sitios de interacción a GATA-1, cajas rojas, uno a NF-E2 caja verde y uno a EKLf caja gris) evidenciado por la generación de hipersensibilidades (SH y flechas negras). La ausencia de nucleosomas en la zona remodelada, significa que esperamos que la cromatina se encuentre más accesible y no ausente de ellos. Cuando los genes globina ya no se estén expresando, ahora el silenciador (cajas azules) será el que se encuentre en una estructura de la cromatina accesible a nucleasas y por lo tanto activo.

### **Objetivos generales**

Establecer un modelo integrativo, que proponga un mecanismo en el contexto de la cromatina para la función diferencial de los elementos silenciador y enhancer localizados en el costado 3' del grupo de genes  $\alpha$ -globina de pollo durante la diferenciación eritroide y desarrollo del pollo.

### **Objetivos particulares**

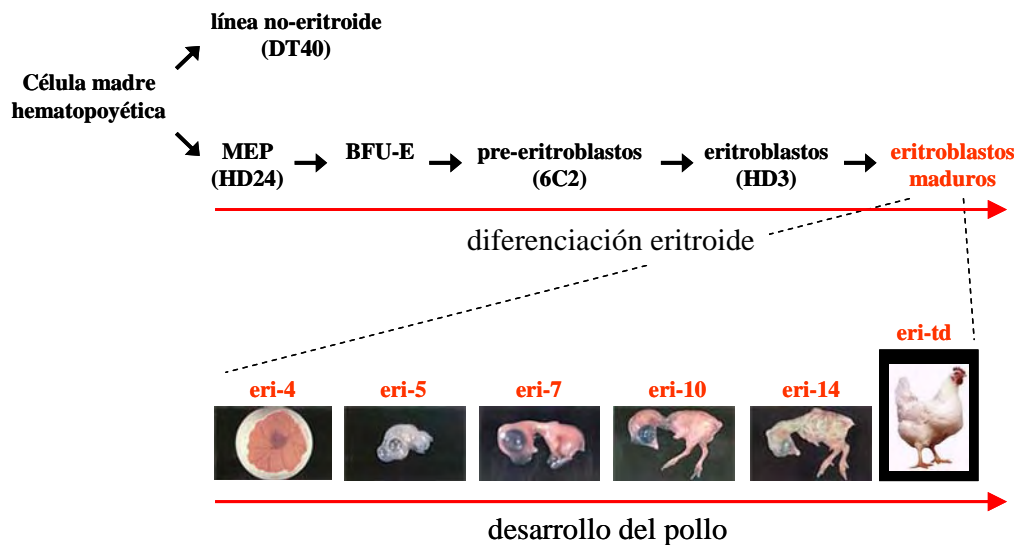
1. Caracterizar funcionalmente la región de 1036 pb que comprende a los elementos silenciador, enhancer y a la región homóloga de 120 pb.
2. Caracterizar a lo largo de distintas etapas de diferenciación eritroide, la estructura de la cromatina en la región de 1036 pb que comprende los elementos de regulación silenciador, enhancer y a la región homóloga de 120 pb.
3. Estudiar los mecanismos de remodelación diferencial de la estructura de la cromatina en el fragmento de ADN de 1036 pb que comprende los elementos de regulación silenciador, enhancer y a la región homóloga de 120 pb.
4. Determinar el estado de metilación en la región silenciador, enhancer y a la región homóloga de 120 pb.

## Resultados

La pregunta global que nuestro grupo de trabajo ha abordado en los últimos 6 años es: ¿Cual es la participación del enhancer  $\alpha$ -globina en la regulación de la expresión de sus genes? Gracias a una serie de estudios *in vitro*, los cuales se muestran en la primera parte de los resultados, hemos logrado caracterizar funcionalmente a los elementos localizados en el fragmento de ADN de 1036 pb. En paralelo se realizó la caracterización de la estructura de la cromatina de esta región. Gracias a estos resultados se logró realizar una caracterización funcional de los elementos silenciador y enhancer+120 pb, en el contexto de la cromatina, resultados que nos permitieron comprender la relación íntima que existe entre la estructura de la cromatina y la modulación de estos elementos.

## El sistema biológico

Para realizar los estudios propuestos en este trabajo, contamos con un modelo de estudio de líneas eritroides arrestadas en diferentes estadios de la diferenciación y cultivos primarios de eritrocitos (Fig. 19). Como control no eritroide se utilizó la línea linfoide DT40.



**Figura 19.** Modelo de estudio. En la parte superior se muestran las diferentes etapas de la diferenciación eritroide, así como la línea celular que corresponde a cada etapa: DT40 (linfoblastos), HD24 (precursor mielóide-eritroide), 6C2 (pre-eritroblastos), HD3 (eritroblastos). En la parte inferior se muestra el desarrollo embrionario del pollo desde eritrocitos de 4 días (eri-4) hasta eritrocitos terminalmente diferenciados (eri-td).

Una de las particularidades de este sistema, es que las células no pierden sus núcleos cuando llegan a eritrocitos, permitiéndonos técnicamente hacer extracciones nucleares para realizar diferentes estudios *in situ*<sup>1</sup>.

### **Generación de los vectores madre utilizados en este trabajo**

El enhancer  $\alpha$ -globina de pollo esta conformado por tres sitios de interacción al factor transcripcional GATA-1, los cuales se encuentran distribuidos en un fragmento de 120 pb, con una separación entre ellos de 60 pb (Fig. 20). En la región de 120 pb encontramos un cuarto sitio de interacción al factor transcripcional GATA-1; el cual de manera interesante se encuentra a 60 pb del último sitio GATA-1 del enhancer, lo anterior significa que entre cada sitio tenemos un fragmento de 60 pb. Entre los sitios de interacción al factor GATA-1 del enhancer, se encontraron sitios de unión al factor Sp1 (ver más adelante). En las 120 pb se localizaron sitios de interacción para los factores EKLf y NF-E2. Finalmente, el silenciador presenta tres sitios de interacción a proteínas (SF1-3; Fig. 20), que a la fecha se desconoce su identidad.

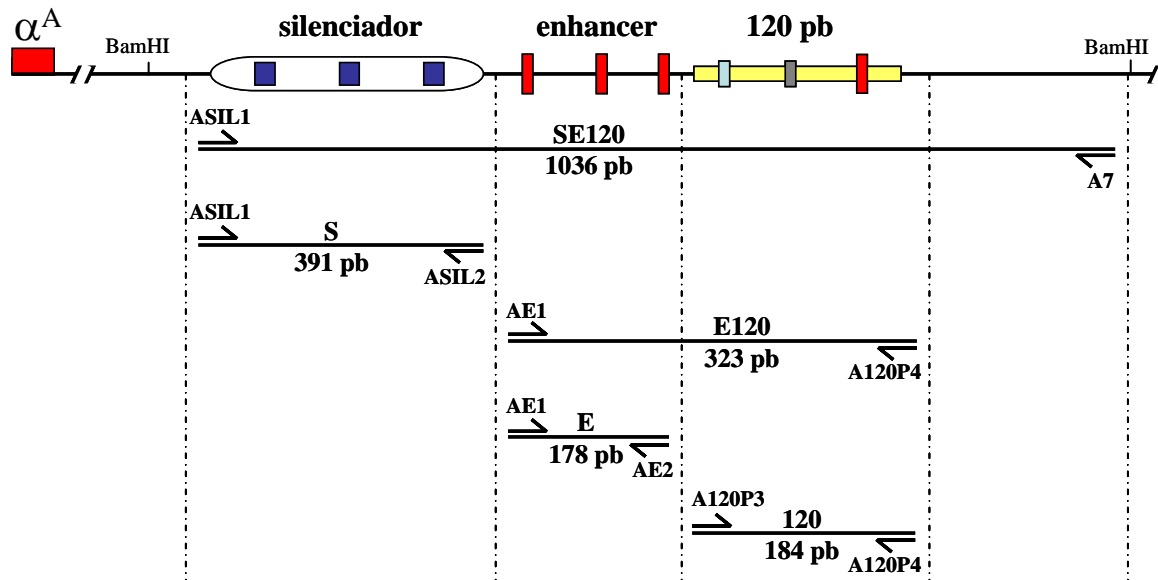
Con la intención de comprender cual es la participación de cada uno de los elementos contenidos en el fragmento de 1036 pb, decidimos evaluar la capacidad de *trans*-activación de cada uno de los fragmentos de ADN señalados (Fig. 20), sobre el promotor  $\alpha^D$ . Para lograr lo anterior, comenzamos por clonar las regiones que se indican en la figura 20 en el vector pCAT- $\alpha^D$  (donado amablemente por el Dr. Sergey Razin, del Instituto de Biología del Gen de la Academia de Ciencias de Rusia)<sup>2</sup> en el sitio *Bam*HI localizado río abajo del gen reportero (generando las construcciones, pCAT $\alpha^D$ , pC $\alpha^D$ E, pC $\alpha^D$ E120, pC $\alpha^D$ SE120 y pC $\alpha^D$ 120; anexo vectores). Adicionalmente, se generó el vector pL $\alpha^D$  por la clonación del promotor  $\alpha^D$  en el sitio *Bg*III del sitio de multiclonación del vector pGL3-basic (Promega). Este vector contiene al gen de la luciferasa como reportero, el cual resultó ser un mejor gen reportero al incrementar su sensibilidad. A este vector también se le clonaron cada uno de los fragmentos que se

---

<sup>1</sup> Término que adoptamos para referirnos a un estado que no es *in vivo*, pero tampoco en *in vitro*. Es un estado intermedio donde se pierde la membrana celular, pero la integridad del núcleo se conserva y la estructuración del genoma en cromatina no se ve afectada.

<sup>2</sup> Esta construcción proviene del vector pCAT Basic (Promega), el cual contiene al reportero de la cloramfenicol acetiltransferasa (CAT), bajo el efecto del promotor  $\alpha^D$  clonado en el sitio *Sma*I.

muestran en la figura 20 en el sitio de restricción *Bam*HI localizado río abajo del reportero de luciferasa, generando las construcciones ( $pL\alpha^D$ ,  $pL\alpha^DE$ ,  $pL\alpha^{DE4}$  y  $pL\alpha^{DSE4}$ , ver anexo vectores).



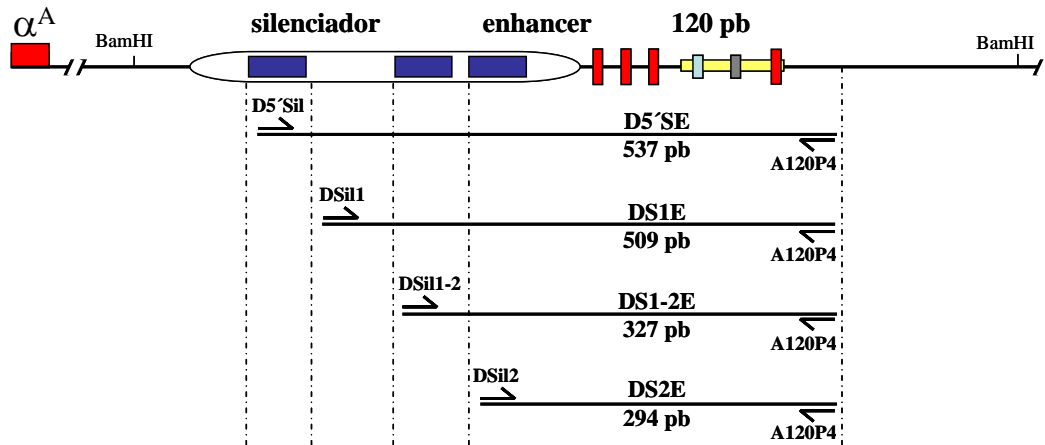
**Figura 20.** Diseño experimental para la generación de las construcciones que abarcan el fragmento de ADN de 1036 pb. Esquema no a escala, del diseño y localización de los oligonucleótidos que se utilizaron para extraer de ADN genómico de pollo las regiones señaladas. SE120 (silenciador/enhancer + 120 pb), S (silenciador), E120 (enhancer + 120 pb), E (enhancer) y 120 (región homóloga 120 pb). Cajas azules, sitio de interacción a las proteínas SF1, SF2 y SF3 (mostradas en ese orden) en la región del silenciador (ovalado). Rectángulos rojos, cajas de interacción a GATA-1. Rectángulo amarillo, secuencia conservada de 120 pb. Rectángulo verde, caja de interacción a EKLf. Rectángulo gris, caja de interacción a NF-E2. Este código de colores y formas, se conservará durante todas las figuras mostradas en esta tesis.

En el caso de las construcciones que fueron utilizadas para el estudio del silenciador, el fragmento de 391 pb (Fig. 20), fue clonado en el vector pGEM11-Z(+) (Promega), con la enzima de restricción *Bam*HI (plásmido pGSil), con el objetivo de facilitar las construcciones subsecuentes. Para obtener las construcciones con el silenciador localizado río abajo del promotor en ambas orientaciones (construcciones  $pL\alpha^{D5'Sil}$  y  $pL\alpha^{D3'Sil}$ ; Fig. 21) el silenciador se obtuvo por una restricción enzimática con *Bam*HI del vector pGSil y se clonó en el sitio *Bam*HI del vector  $pL\alpha^D$ .

Con el fin de separar al silenciador del enhancer, se extrajo del vector pGSil vía una digestión enzimática *Sac*I y *Xba*I, y se clonó en *Sac*I y *Nhe*I del vector  $pL\alpha^{DE4}$ ,



aprovechando que los sitios de restricción *Xba*I y *Nhe*I son compatibles, resultando en las construcciones con orientaciones diferentes pL5'Sil $\alpha^D$ E y pL3'Sil $\alpha^D$ E (Fig. 21).



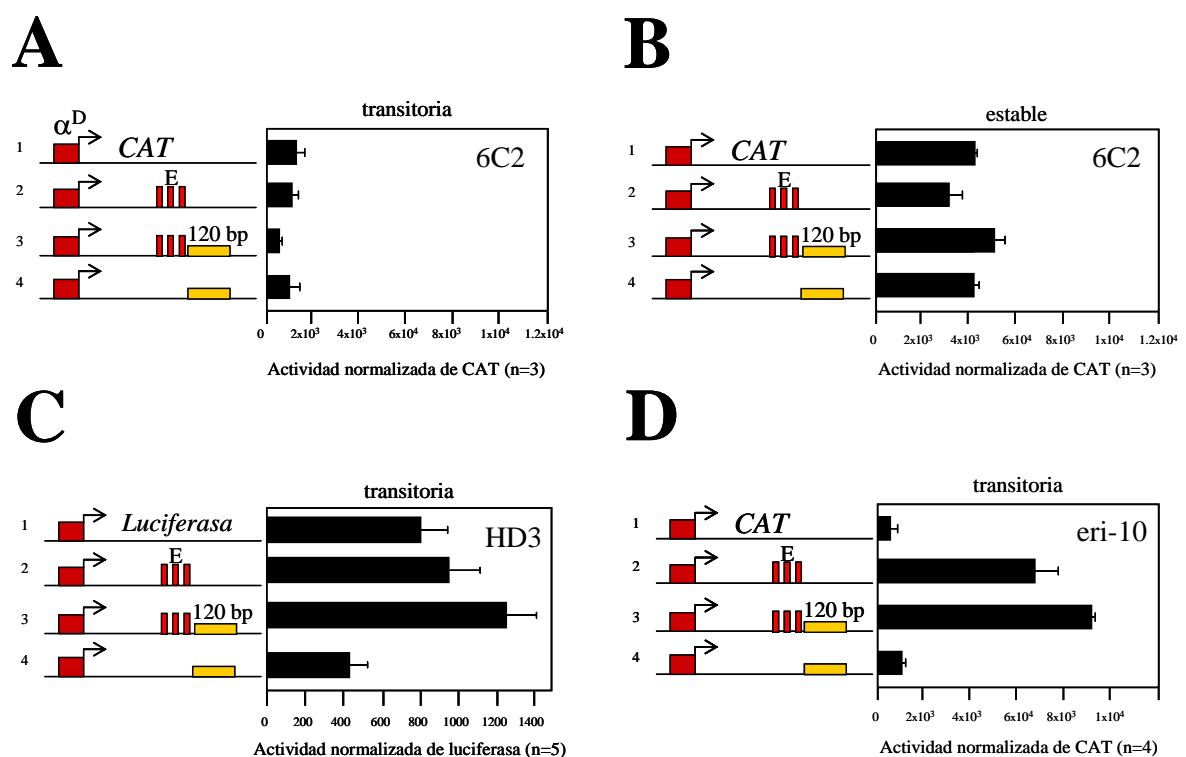
**Figura 21.** Diseño experimental del análisis por deleciones secuenciales del silenciador. Esquema no a escala, del diseño de los oligogonucleótidos que se utilizaron para aislar las secuencias de interés a partir de ADN genómico de pollo.

Para localizar las secuencias relevantes en la función del silenciador, se construyeron plásmidos con deleciones secuenciales del silenciador. El diseño del experimento se observa en la figura 21. Los productos de PCR obtenidos se clonaron en el sitio *Bam*HI del vector pL $\alpha^D$ . Los vectores obtenidos se denominaron pL $\alpha^D$ 5'(DS1E), pL $\alpha^D$ 5'(DS1-2E) y pL $\alpha^D$ 5'(DS2E) mostrando la orientación endógena. Las construcciones con la orientación invertida muestran un 3' en lugar del 5'. Todas las construcciones que se muestran en esta tesis, fueron secuenciadas por la Unidad de Biología Molecular del Instituto de Fisiología Celular, para corroborar su integridad.

### Estudio *in vitro* del elemento enhancer

Las construcciones que contienen al enhancer (Fig. 20), fueron transfectadas transitoriamente en las líneas 6C2 (pre-eritroblastos) y HD3 (eritroblastos), y en cultivos primarios de eritrocitos de pollo de 10 días de gestación (eri-10), con la intención de evaluar su capacidad para *trans*-activar al promotor  $\alpha^D$  durante la diferenciación y el desarrollo. Adicionalmente nos interesaba conocer el efecto que la estructura de la cromatina podría tener en esta construcción, por lo mismo también se realizaron las transfecciones de manera estable (Fig. 22 y datos no mostrados). Los

resultados obtenidos muestran que el promotor  $\alpha^D$  no puede ser *trans*-activado en células 6C2 de manera transitoria o estable, consistente con el hecho de que los genes  $\alpha$ -globina muestran niveles basales de expresión en este estado de diferenciación. En células HD3 se observa un pequeño incremento en la actividad del reportero, mientras que en cultivos primarios de eritrocitos de pollo de 10 días (eri-10) se puede observar un incremento en la actividad de 6 veces en presencia del enhancer y de 10 veces con la construcción que contiene al enhancer y a la región de 120 pb (enhancer120). Interesantemente la construcción que solo contiene al la región de 120 pb no presenta ningún efecto en la actividad del reportero.



**Figura 22.** Actividad del enhancer y de la región de 120 pb. **A-** Transfección transitoria en la línea 6C2. **B-** Transfección estable en la línea celular 6C2. **C-** Transfección transitoria en la línea celular HD3. **D-** Transfección transitoria en cultivos primarios de eritrocitos de 10 días de gestación (eri-10).  $\alpha^D$ , promotor  $\alpha^D$  del dominio  $\alpha$ -globina de pollo. **E**, enhancer; 120 pb, región homóloga entre pollo y pato de 120 pb. CAT, reportero de la Cloranfenicol Acetil Transferasa. n, número de veces que se repitió el ensayo de manera independiente.

Estos resultados muestran claramente, que la contribución del fragmento de 120 pb a la actividad del reportero, depende de la presencia del enhancer. Por lo anterior proponemos que las 120 pb son parte integral del enhancer (Fig. 22 “D”, comparar construcciones 1, 2 y 4). El enhancer muestra un efecto en la actividad del gen reportero de manera estado específico, mostrando niveles basales de expresión en pre-

eritroblastos, con un pequeño incremento en eritroblastos, el cual claramente se ve acentuado en eri-10, posiblemente por la expresión de un factor estado específico.

En base a estos resultados surgieron dos preguntas: ¿cuáles son los elementos necesarios para la función reguladora y activadora del enhancer? y ¿cuál es la contribución que la secuencia de 120 pb aporta a la regulación transcripcional vía el enhancer? Para contestar a estas preguntas y con la finalidad de facilitar el análisis, se decidió dividir el estudio en tres partes. En las primeras dos se abordaron las preguntas antes planteadas, en la tercera se determinará la participación del silenciador, sin embargo frecuentemente nos resultó imposible ser estrictos en la división de este análisis.

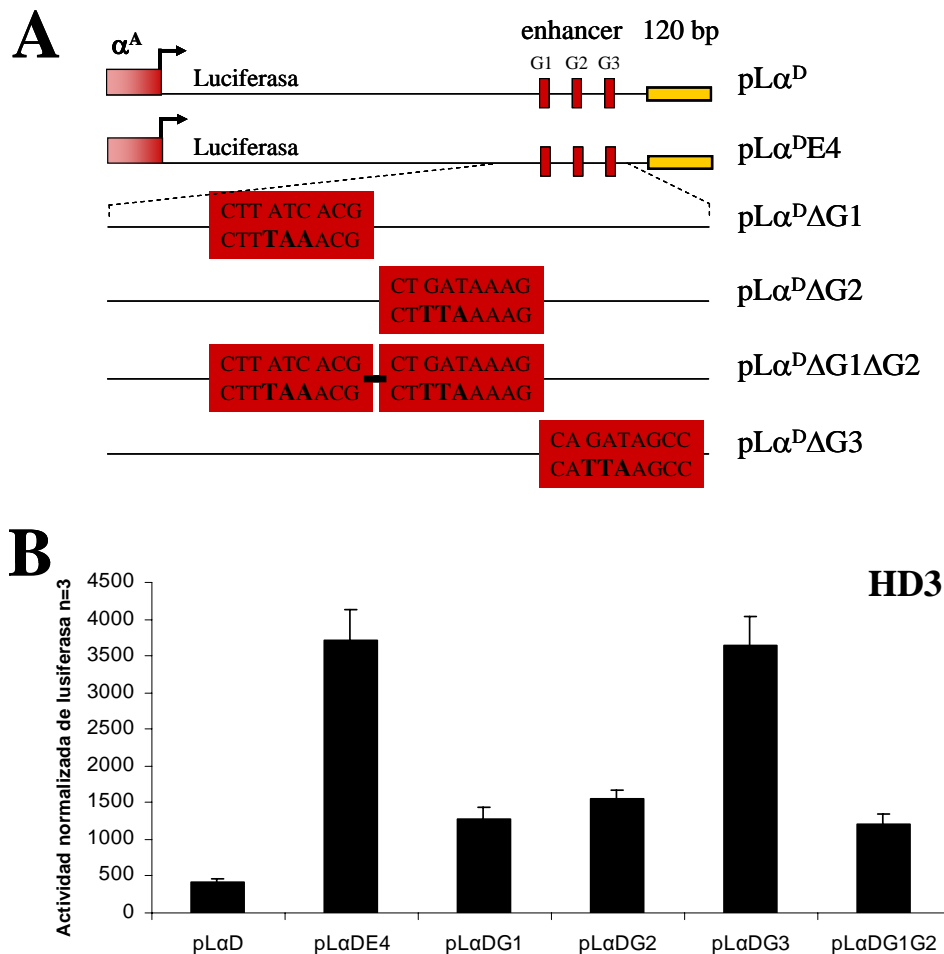
### **Función de los múltiples sitios de interacción al factor GATA-1 en el enhancer**

Con la intención de conocer los elementos de regulación importantes en el elemento enhancer, se generaron mutaciones puntuales de cada uno de los sitios de unión al factor GATA-1<sup>3</sup>. Partiendo de los resultados antes mostrados (Fig. 22, se decidió realizar las mutaciones en el contexto del enhancer y la región homóloga de 120 pb con el gen de luciferasa como reportero (Fig. 23).

Estas construcciones fueron transfectadas transitoriamente en la línea celular HD3. Como se puede observar, la mutación de los sitios G1 y G2 reduce la actividad del reportero en 75% en relación al control (Fig. 23; comparar carriles 3 y 4 con el 2). La mutación en G3 no mostró ningún efecto. La doble mutante G1G2 muestra un efecto similar al observado en las construcciones con las mutaciones por separado. Por lo anterior podemos concluir que los dos primeros sitios de interacción al factor transcripcional GATA-1 (y posiblemente a miembros de esta familia), resultan ser indispensables, reduciendo dramáticamente la actividad del enhancer sobre su promotor. El efecto de estas mutaciones no es aditivo.

---

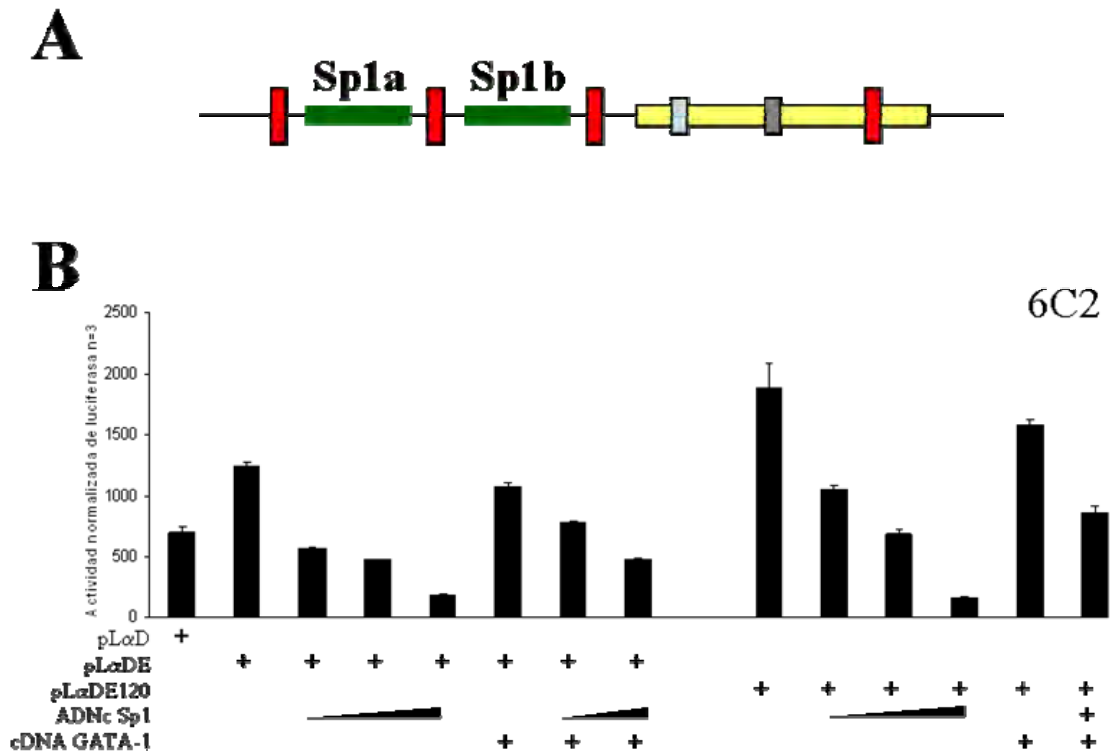
<sup>3</sup> Las mutaciones puntuales de las tres cajas GATA (G1, G2 y G3) y la doble G1G2 (Fig. 23 "A") se realizaron con el paquete de Quick Change Site Directed Mutagenesis (Stratagene) sobre el vector antes descrito pL $\alpha$ <sup>D</sup>E120.



**Figura 23.** Estudio funcional de efecto de las mutaciones puntuales de los sitios de interacción al factor transcripcional GATA-1 en la actividad del enhancer. **A**, Esquema mostrando el diseño de las construcciones y los cambios en las secuencias nucleotídicas utilizadas en este análisis. A la derecha se agregó el nombre de los vectores generados. **B**, Transfecciones transitorias en la línea HD3 de las construcciones mostradas en “A”. G1, G2 y G3, numeración de las tres cajas de interacción a GATA-1 localizadas en el enhancer. Rectángulos rojos horizontales, representación de cada una de las tres cajas de interacción a GATA-1; se muestra en ellas la secuencia particular de interacción a la proteína GATA-1 y en negritas se resalta las bases que fueron intercambiadas para generar las mutaciones en la secuencia de interacción al factor.

El factor transcripcional Sp1 juega un papel importante en la activación de los genes  $\beta$ -globina mediada por el LCR (229). En el enhancer  $\alpha$ -globina, entre las secuencias de unión al factor GATA-1 localizamos dos sitios de unión al factor general de transcripción Sp1 (Fig. 24A). Por lo mismo, nos interesó estudiar su posible participación en la activación mediada por el enhancer  $\alpha$ -globina. Por medio de una serie de co-transfecciones de los ADNc de Sp1 y GATA-1, con las construcciones que incluyen al enhancer y al enhancer120 en la línea celular 6C2, estudiamos su efecto en la regulación del enhancer. Como se puede observar, a diferencia de lo esperado, se

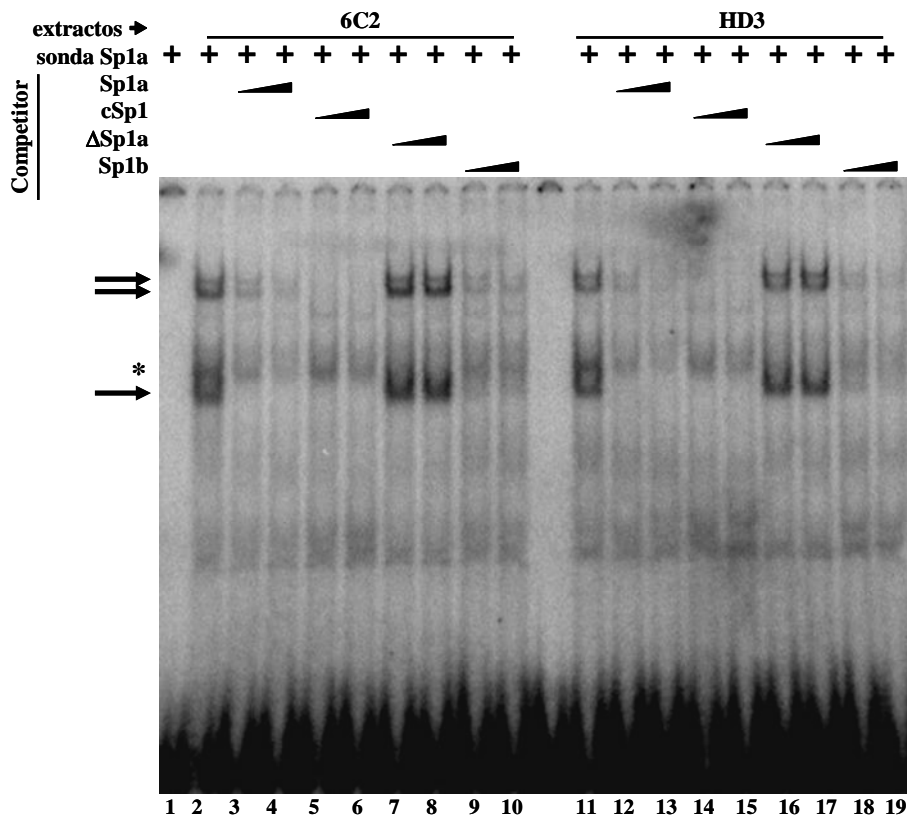
observó un efecto negativo en los niveles de expresión del reportero mediado por el enhancer (Fig. 24). La co-transfección del ADNc de Sp1 en presencia del vector pL $\alpha$ <sup>DE</sup> o pL $\alpha$ <sup>DE120</sup>, genera una disminución en los niveles de expresión a diferencia del efecto positivo que muestra la co-transfección del ADNc de GATA-1. Proponemos que este estudio deberá ser repetido en eritrocitos, donde el efecto negativo podría ser evaluado con mayor claridad.



**Figura 24.** Estudio del papel de la proteína Sp1 en la actividad del enhancer. **A-** Esquema representando la región que comprende al enhancer y las 120 pb, los rectángulos verdes muestran la localización de los sitios de interacción a Sp1 encontrados. Para facilitar su estudio estos fueron nombrados Sp1a y Sp1b **B-** Transfección transitoria en la línea celular 6C2 de los vectores mostrados. Triángulos; representan un incremento en la concentración del vector.

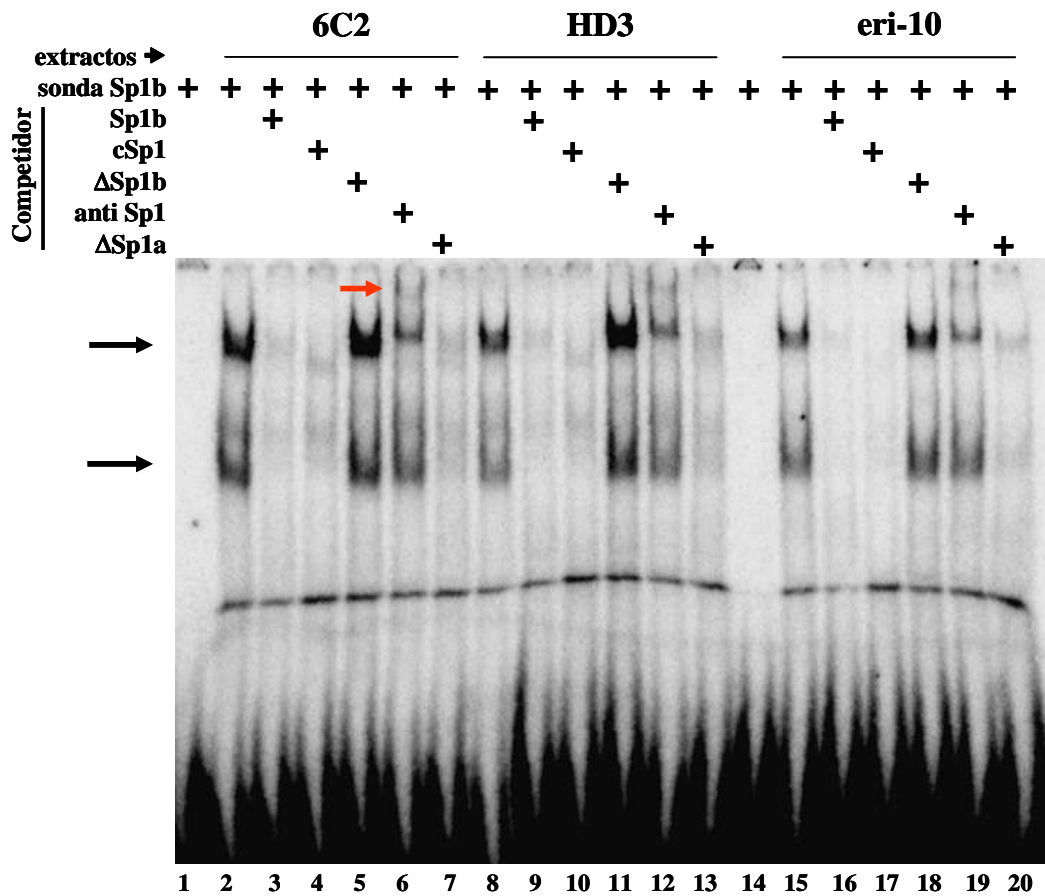
Una vez determinado el efecto de esta proteína en los niveles de actividad mediados por el enhancer, nos interesaba saber ¿con cual de los sitios Sp1 localizados en el enhancer interacciona esta proteína? Para contestar a esta pregunta, se realizaron una serie de ensayos de retardo (Figs. 25 y 26). Como se puede observar la sonda Sp1a muestra al menos tres complejos de retardo (Fig. 25, flechas negras), los cuales son competidos de manera específica con las sondas frías que incluyen a los sitios de interacción para Sp1a ó Sp1b (carriles 3, 4,9 y 10) y con la sonda que contiene la

secuencia consenso para Sp1 (carriles 5 y 6). Además, estos no son competidos por la mutación puntual realizada en este sitio (carriles 7 y 8).



**Figura 25.** Análisis por retardo de las interacciones en la secuencia Sp1a. Gel de retardo en el que se muestra la interacción de al menos 3 complejos de retardo (flechas negras) en pre-eritroblastos (6C2, carril 2) y eritroblastos (HD3, carril 11). En el resto de los carriles se pueden observar las competencias con las sondas frías para Sp1a (auto-competencia), cSp1 (sitio consenso de interacción a Sp1), ΔSp1a (sonda Sp1a con la secuencia de interacción a Sp1 mutada) y Sp1b (sonda que incluye la caja de interacción Sp1b). Asterisco, interacción inespecífica.

El análisis del segundo sitio de interacción a Sp1, nombrado Sp1b, presenta resultados muy similares, excepto por la presencia de solo dos complejos de retardo (Fig. 26, flechas negras). Para comprobar de manera tajante que la proteína Sp1 tiene la capacidad de interactuar con esta secuencia, al menos *in vitro*, se realizó el súper-retardo con el anticuerpo anti-Sp1, el cual genera una banda de súper-retardo (flechas rojas). De estos resultados podemos concluir que la proteína Sp1 interactúa con las secuencias asociadas al enhancer, mostrando una función negativa en la regulación de este.



**Figura 26.** Ensayo de movilidad retardada demostrando las interacciones en la caja Sp1b. Gel de retardo en el que se muestra la interacción de al menos 2 complejos de retardo (flechas negras) en pre-eritroblastos (6C2, carril 2), eritroblastos (HD3, carril 8) y en eri-10 (carril 15). En el resto de los carriles se pueden observar las competencias con las sondas frías para Sp1b (auto-competencia), cSp1 (sitio consenso de interacción a Sp1), ΔSp1b (sonda Sp1b con la secuencia de interacción a Sp1 mutada), Sp1a (sonda que incluye la caja de interacción Sp1a) y el super-retardo obtenido utilizando el anticuerpo para Sp1 (flecha roja).

La caracterización de la región del enhancer nos muestra que esta constituida por tres sitios de interacción al factor transcripcional GATA-1, de los cuales la presencia de los primeros dos resulta indispensable para que este elemento ejerza su actividad de *trans*-activación sobre su promotor blanco (ver discusión). Como se mostró anteriormente, se encontró la presencia de dos sitios de interacción a la proteína Sp1, los cuales se encuentran intercalados entre las cajas GATA-1. A estas cajas interacciona la proteína Sp1, formando varios complejos de retardo, a diferencia de lo esperado, el efecto que ejercen en la actividad del enhancer es negativo. De manera muy interesante, la co-transfección del ADNc de GATA-1 y de Sp1 ejerce efectos contrarios. Este

resultado nos invita a proponer de manera preliminar un modelo, en el cual el balance entre estas dos proteínas podría estar jugando un papel importante en la activación-represión de la actividad del enhancer.

### Caracterización de la secuencia conservada de 120 pb

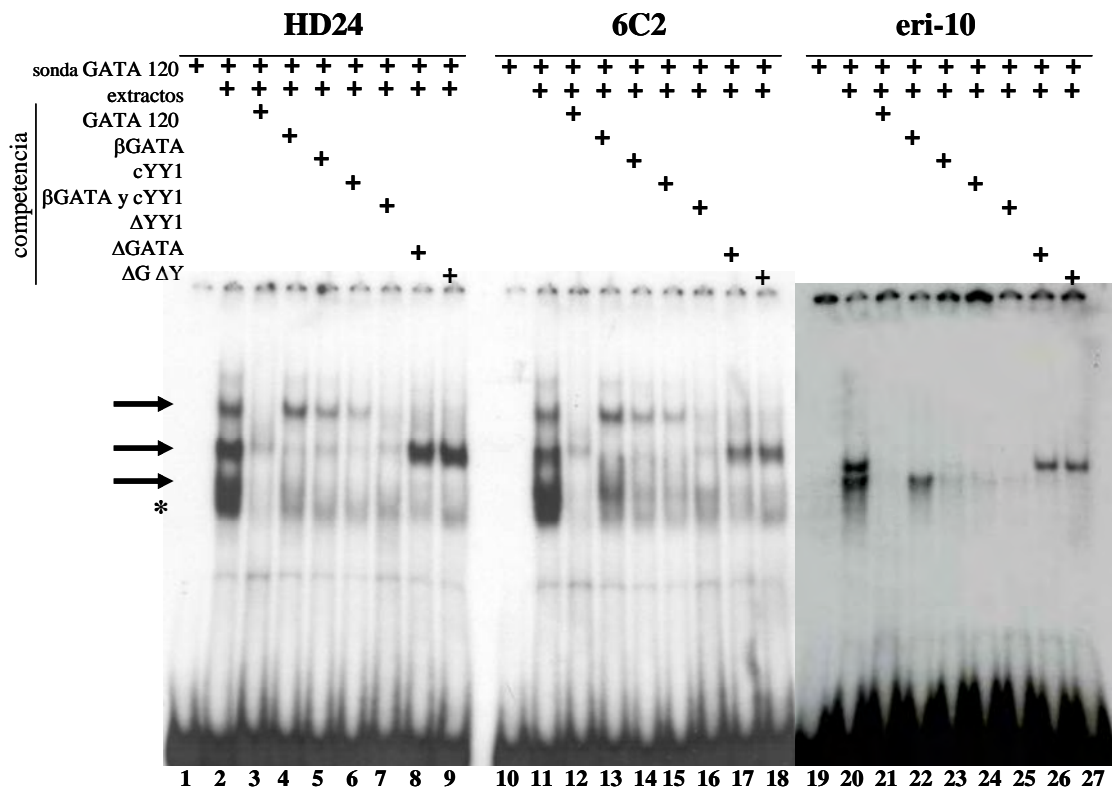
Con la intención de conocer la participación que la secuencia de 120 pb tiene en la actividad del enhancer y con el antecedente de que el sitio de interacción correspondiente al cuarto sitio de unión a GATA-1 generaba dos complejos de retardo (232), se decidió realizar un análisis detallado a nivel de la secuencia nucleotídica. Encontramos que la secuencia de interacción al factor transcripcional GATA-1 se co-localiza con la secuencia de reconocimiento al factor YY1 (Fig. 27). De manera coincidente, esta co-localización también fue descrita en el enhancer  $\epsilon$ -globina; 309).



**Figura 27.** Análisis de la secuencia de interacción a GATA-1 localizada en el fragmento de 120 pb. Arriba, alineamiento entre la región que comprende la cuarta caja de interacción a GATA-1 (GATA120), un motivo de unión a GATA-1/YY1 descrito en el silenciador  $\epsilon$  de humano (69) y la secuencia consenso a YY1 (69, 70). Se resaltan las bases conservadas del motivo de unión a YY1. En rojo se resalta el sitio de interacción a GATA-1. Abajo, se muestran las bases que fueron cambiadas (en minúsculas y en negro) para generar las secuencias con los sitios a GATA-1 ( $\Delta$ GATA) y YY1 ( $\Delta$ YY1) mutados.

Con la intención de definir la interacción y posible función de este factor en la secuencia de 120 pb, se realizó un análisis de interacción proteína-ADN mediante geles de retardo (Fig. 28 y 29). Como se puede observar se encontró la presencia de múltiples complejos de retardo cuando se utilizaron las líneas celulares HD24 y 6C2, donde proponemos en base a las competencias con las sondas frías consenso a GATA-1 y YY1 ( $\beta$ GATA y cYY1, respectivamente), que la primera banda corresponde al factor transcripcional GATA-1, la segunda al factor transcripcional YY1 y la tercera es un posible complejo formado por ambos factores (Fig. 28, en orden ascendente).



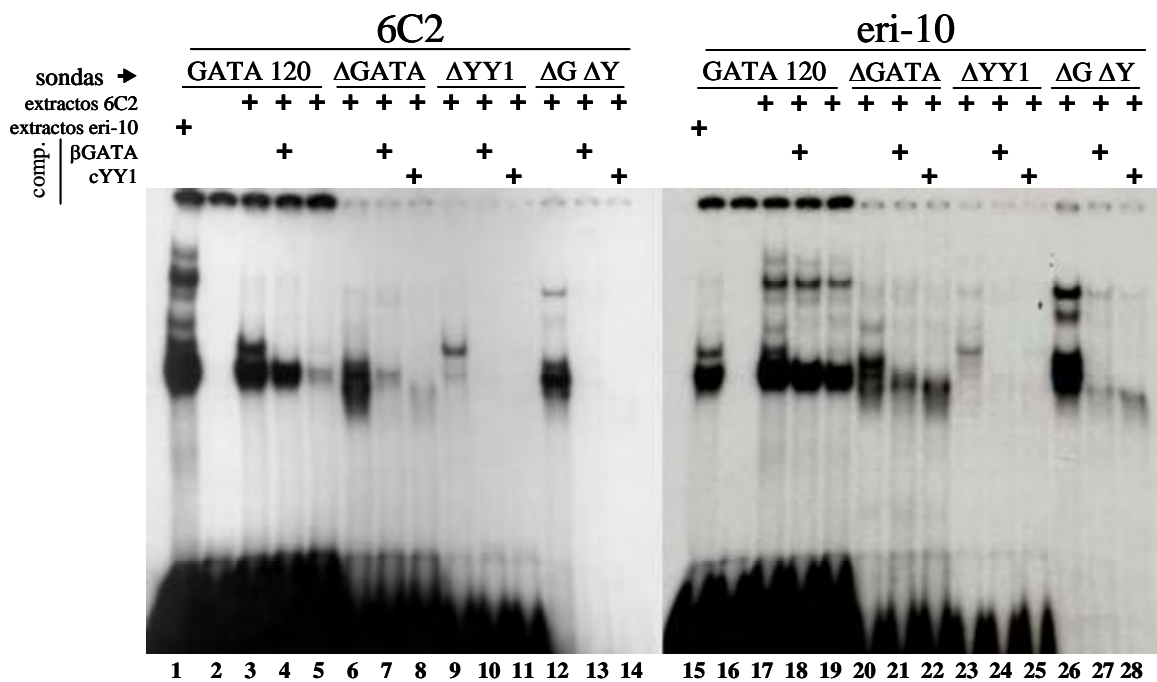


**Figura 28.** Estudio de las interacciones en la región de la cuarta caja GATA-1. Gel de retardo en el que se muestra la formación de al menos tres complejos de retardo (flechas negras), en la línea eritroide-mieloide (HD24, carril 2) y 6C2 (carril 11); en eri-10 (carril 20) se observa la presencia solo de dos complejos de retardo. En el resto de los carriles se pueden observar las competencias con las sondas frías GATA120 (auto-competencia), la competencia con un sitio conocido de interacción a GATA-1 ( $\beta$ GATA-1) del enhancer  $\beta^{A/\epsilon}$  de pollo, con cYY1 (sitio consenso para YY1), con  $\Delta$ YY1 y  $\Delta$ GATA (ver figura 27) y con  $\Delta$ G $\Delta$ Y (doble mutante a GATA y YY1). El asterisco, representa una interacción inespecífica.

En el retardo con eritrocitos, no se observa la banda del complejo y la banda correspondiente a la interacción de YY1 se ve incrementada (Fig. 28, carril 20). De manera consistente, en todos los casos cuando se compete con sondas no radioactivas que contienen la secuencia de unión a GATA-1 (ya sea GATA120,  $\beta$ GATA o  $\Delta$ YY1), de manera consistente y reproducible se compete la banda perteneciente a este factor. De forma similar las bandas pertenecientes a YY1 o al complejo se compiten. En las competiciones con las sondas que incluyen las secuencias que contienen al factor YY1 (GATA120, cYY1 o  $\Delta$ GATA), se observo el mismo comportamiento, sin embargo en este caso la banda perteneciente a GATA-1 se ve menos competida (carriles 8, 9, 17, 18, 26 y 27), mostrando una aparente dependencia el factor YY1 por parte de GATA-1.

Resultados similares se observan cuando marcamos radiactivamente ahora cada una de las sondas que contienen las mutaciones en los sitios GATA-1 y YY1 (Fig. 29).

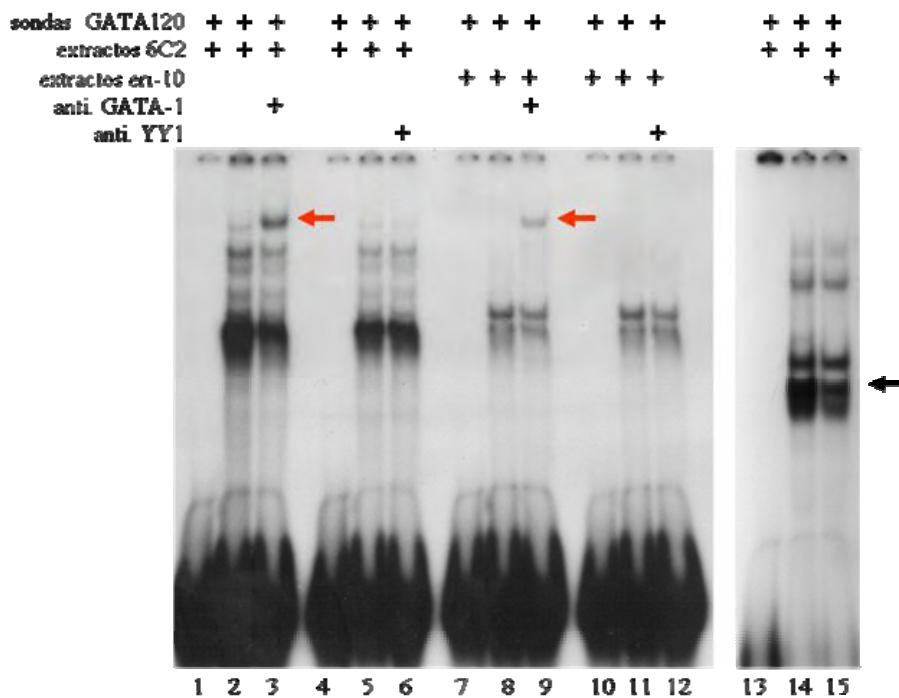
Por lo que sugerimos que los factores GATA-1 y YY1 se encuentran interaccionando en el fragmento de 120 pb de manera cooperativa, donde YY1 parece depender de la unión de GATA-1 y en menor grado GATA-1 de YY1. En un escenario alternativo, proponemos una interacción física entre GATA-1 y YY1. Este último escenario fue comprobado por medio de un ensayo de ChIP re-ChIP, donde confirmamos la coexistencia al menos en el mismo complejo de estas dos proteínas y posteriormente por medio de un ensayo de inmunoprecipitación y “pull-down” se confirmó una interacción física (50, apéndice 2).



**Figura 29.** Estudio *in vitro* de las interacciones en la secuencia de la cuarta caja de unión a GATA-1. Como sondas se utilizaron las secuencias GATA120 (ver figura 27), ΔGATA, ΔYY1 Y ΔGΔY (con las mutaciones puntuales en GATA, YY1 y la doble mutante, respectivamente; ver figura 27). Se pueden observar las competencias con las sondas frías GATA120 (auto-competencia) y con cYY1 (sitio consenso para YY1), con ΔYY1 y ΔGATA (ver figura 27) y con ΔGΔY (doble mutante a GATA y YY1).

Para confirmar la interacción de estas dos proteínas en el fragmento de 120 pb, se realizaron los súper-retardos utilizando los anticuerpos que reconocen a las proteínas GATA-1 y YY1. En el caso de GATA-1, se observó un súper-retardo muy claro tanto en la línea celular 6C2, como en eritrocitos (Fig. 30, flechas rojas). En el caso del factor YY1, el súper-retardo no fue claro, como se puede observar en la figura 30 (flechas

negras) se observó una competición parcial por parte del anticuerpo en 6C2. En eritrocitos, el súper-retardo fue más claro (50, apéndice 2).



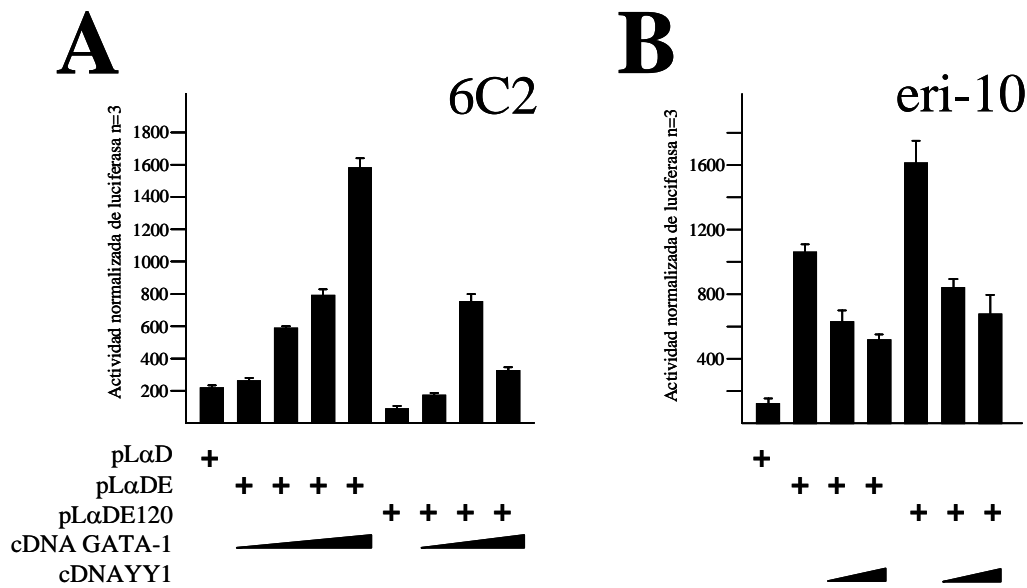
**Figura 30.** Estudio por super-retardo de las interacciones de las proteínas GATA-1 y YY1 con la secuencia correspondiente al cuarto sitio de unión al factor GATA-1. Gel de retardo en el que se muestra la formación de diferentes complejos de super-retardo, utilizando los anticuerpos anti-GATA y anti-Sp1 (flechas rojas). Con el anticuerpo contra YY1 no se observó super-retardo, pero sí una disminución en la intensidad del complejo de retardo (flecha negra).

De este análisis podemos concluir que en el fragmento de 120 pb interaccionan al menos el factor transcripcional GATA-1 y el factor YY1, mostrando estos dos una asociación novedosa aunque hasta el momento su papel funcional seguía siendo una incógnita.

### Estudio funcional de la secuencia conservada de 120 pb

¿Cuál es la función que los factores GATA-1 y YY1 están llevando a cabo con su interacción en la secuencia de las 120 pb? Reportes previos muestran que los niveles de la proteína GATA-1 se encuentran elevados en estadios tempranos de la diferenciación y posteriormente, en eritrocitos terminalmente diferenciados bajan drásticamente (317). En un reporte reciente se mostró que la sobre-expresión del factor transcripcional

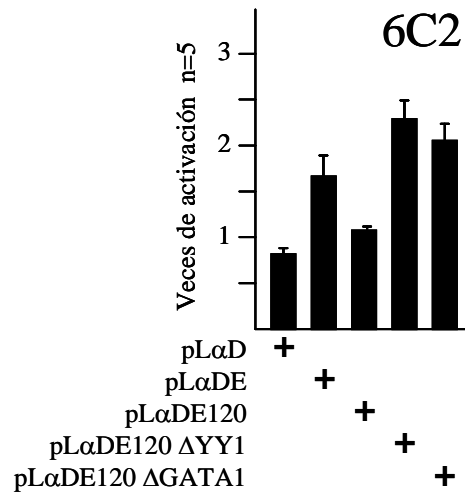
GATA-1 en una línea no-eritroide es capaz de generar patrones de expresión eritroide específicos (140). En base a estos antecedentes, decidimos sobre-expresar a los factores GATA-1 y YY1 en la línea celular 6C2 correspondiente a pre-eritroblastos (Fig. 31). Para medir el efecto de estos factores, se co-transfectaron los vectores de expresión que contienen al enhancer y al enhancer120, ambas construcciones tienen al gen reportero de la luciferasa.



**Figura 31.** Determinación de la función de la secuencia de 120 pb en la actividad del enhancer. **A-** Co-transfección transitoria de las construcciones que contienen al enhancer o al enhancer con el ADNc de GATA-1, en la línea celular 6C2. **B-** Co-transfección transitoria de las construcciones que contienen al enhancer o al enhancer con el ADNc de YY1, en eri-10.

Como era de esperarse, la sobre-expresión de la proteína GATA-1 en el contexto del enhancer incrementó los niveles de reportero (Fig. 31, vector  $\alpha^{DE}$ ), sin embargo y de manera muy interesante, la sobre-expresión en el contexto de la construcción que contiene a la secuencia de 120 pb, mostró niveles reducidos de expresión (Fig. 31, vector  $\alpha^{DE120}$ ). De este resultado surgen 2 interpretaciones, la primera es que la secuencia de 120 pb pueda influir como un regulador negativo. Sin embargo, los resultados obtenidos que se muestran en la figura 22, sugieren que el efecto de esta región en eritroblastos es positivo. Por lo anterior al demostrar que estas secuencias tiene un efecto sobre la actividad del enhancer de manera estado específico y recordando que por sí solo se muestra como un elemento neutro que depende de la

presencia del enhancer, proponemos que se trata de un modulador tiempo y tejido específico de la actividad del enhancer.



**Figura 32.** Determinación de la función de la secuencia de 120 pb vía los sitios de interacción a los factores GATA-1 y YY1, en la actividad del enhancer. Transfección transitoria de las construcciones que contienen la mutación en las secuencias de interacción a GATA (pLαDE120ΔGATA1) y a YY1 (pLαDE120ΔYY1) en el contexto del enhancer120.

Para comprobar esta hipótesis, decidimos ahora sobre-expresar al ADNc de YY1 en eri-10. Como se puede observar en la figura 31, el efecto de la sobre-expresión de YY1 es negativo. Sin embargo a diferencia de lo que esperábamos, el efecto negativo también se observó en la construcción que no contiene la región homóloga de 120 pb, posiblemente por la unión de YY1 a otros sitios o por una posible interacción con el factor GATA-1 (Fig. 31 “B”, co-transfecciones pLα<sup>D</sup>E con el cADN de YY1). Con la intención de comprobar nuestros resultados y confirmar que el efecto observado se debía a las secuencias para GATA y YY1 localizadas en la región de las 120 pb, se realizaron las mutaciones puntuales (ver material y métodos) de las secuencias de interacción a estas proteínas, en el contexto del enhancer y el enhancer120 (Fig. 32). Apoyando nuestros resultados anteriores, se observa la des-represión en la actividad del enhancer.

En base a este resultado y los resultados anteriores, donde se mostró la interacción directa entre las proteínas GATA-1 y YY1 y su función diferencial en la secuencia de 120 pb con un efecto directo en la actividad del enhancer, proponemos que esta región

es parte integral del enhancer y que se encuentra modulando al enhancer de manera tiempo- y tejido-específica.

Atribuimos este efecto a la interacción física que se encontró entre GATA-1 y YY1 y su posible efecto en otra parte del enhancer<sup>4</sup>. Es importante resaltar que la sobreexpresión de YY1 muestra su mayor efecto en la actividad del enhancer, en la construcción que contiene a la secuencia de 120 pb (Fig. 31). En conclusión, el enhancer en presencia del fragmento de 120 pb muestra un efecto negativo vía los sitios de interacción a GATA-1 y YY1 en pre-eritroblastos. Cuando se realiza el mismo estudio en eritrocitos, el efecto de las 120 pb ahora es positivo.

### **Estudio *in vitro* del elemento silenciador**

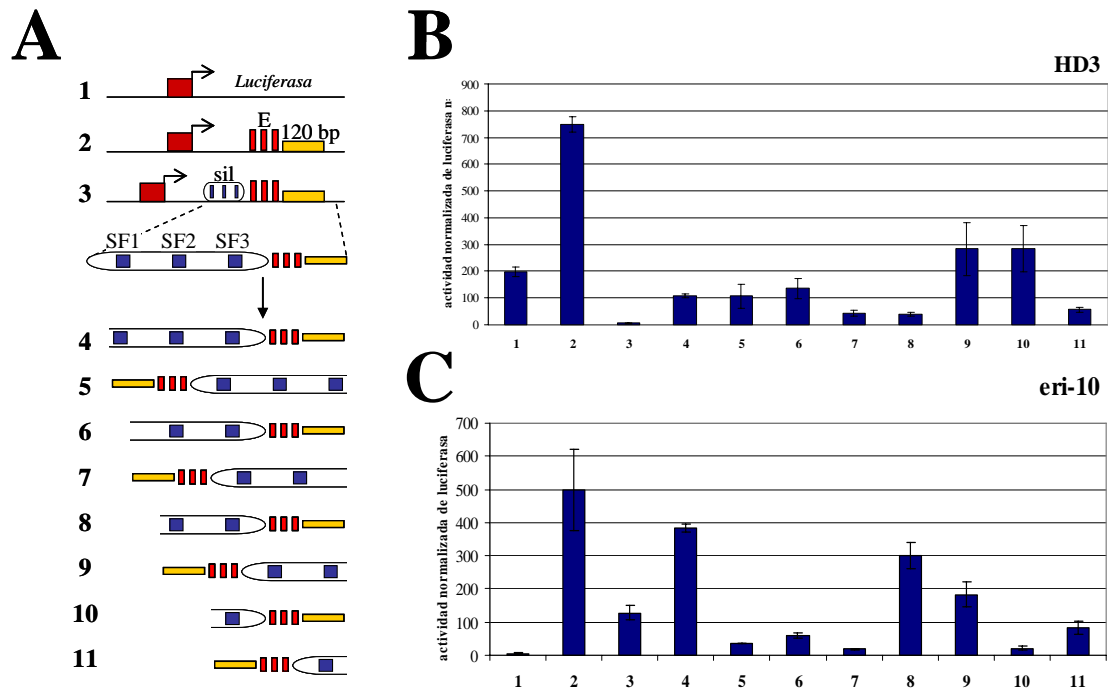
Mediante estudios funcionales y de interacción proteína-ADN, se identificaron tres sitios de interacción a proteínas en el silenciador (SF1-3), en estadios donde los genes globina no se encuentran activos (232). La caracterización de este elemento resulta un punto clave en el entendimiento de los mecanismos de regulación diferencial del dominio  $\alpha$ -globina. El estudio en la expresión de los genes globina mostró un recambio en la expresión de sus genes, caracterizado por el silenciamiento del gen embrionario  $\pi$  y la activación de los genes adultos (Fig. 15). ¿Podría ser el silenciador  $\alpha$ -globina, el elemento de regulación encargado de dirigir o al menos participar en el recambio de la expresión de los genes  $\alpha$ -globina, vía el silenciamiento del gen embrionario  $\pi$ ? ¿Por qué se encuentra localizado a un costado del enhancer?

Mediante un estudio por transfecciones transitorias en las líneas celular HD3 y en cultivos primarios de eritrocitos de pollo de 10 días de gestación (eri-10), se evaluó la capacidad que tiene el silenciador  $\alpha$ -globina, para inhibir la *trans*-activación mediada por el enhancer sobre el promotor  $\alpha^D$  durante la diferenciación y el desarrollo (Fig. 33). Como se puede observar, la construcción que contiene al silenciador reduce la actividad mediada por el enhancer en la línea celular HD3 y en eri-10 (carriles 3 contra 2, respectivamente). También se observó de manera reproducible en eritroblastos (datos no mostrados y Fig. 33 “B”, comparar carril 3 contra 1), que en presencia del silenciador la actividad del reportero mediada por el promotor  $\alpha^D$  se ve afectada. Este resultado podría

---

<sup>4</sup> Esta teoría posteriormente fue comprobada, se encontró una dependencia en la interacción de GATA-1 por parte de YY1, en su interacción con su primer sitio de interacción en el enhancer (Fig. 46, carril 10).

ser explicado a través de tres escenarios distintos: el silenciador tiene su efecto sobre el elemento promotor, sobre el enhancer o sobre ambos. Lo anterior podría darse, dependiendo del estado de diferenciación y/o de la fuerza que el elemento sobre el que se encuentra actuando presente. Esta pregunta será retomada mas adelante.

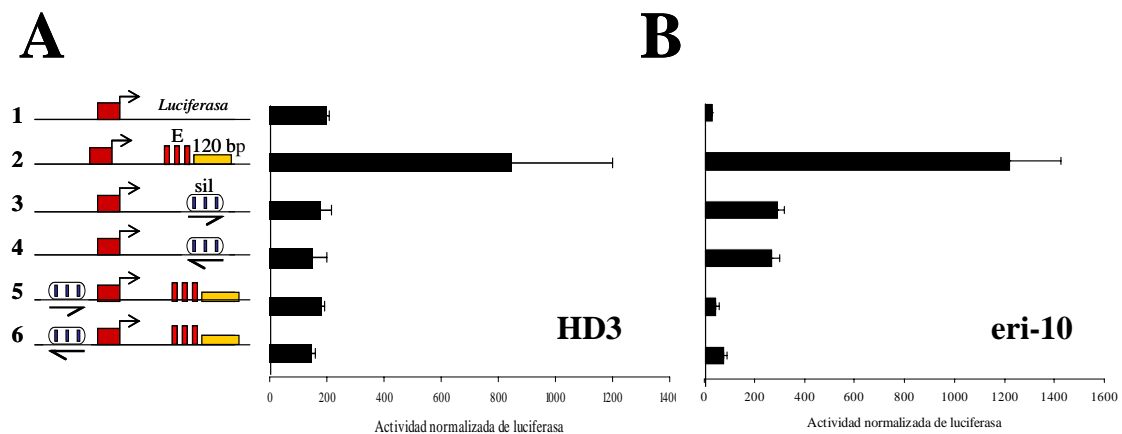


**Figura 33.** Análisis por deleciones consecutivas del silenciador  $\alpha$ -globina. **A-** Esquema representando las diferentes construcciones utilizadas en este estudio. De la construcción 4 a la 11, solo se muestra el fragmento y la orientación del fragmento que fue integrado en la posición en la que originalmente se encontraba el silenciador /enhancer120. **B-** transfección transitoria de las construcciones mostradas en “A” en la línea celular HD3. **C-** transfección transitoria de las construcciones mostradas en “A” en eri-10.

Con el interés de identificar elementos de regulación importantes dentro del silenciador, se realizó un estudio por deleciones consecutivas (Fig. 33). Los resultados no son concluyentes. El sitio denominado SF3 parece ser suficiente para conservar la actividad de silenciamiento. Para comprobar esta teoría se propone realizar las construcciones que eliminan al sitio SF3 (ver perspectivas). Alternativamente se podrían realizar mutaciones puntuales en las regiones SF1-3 del silenciador.

Para entender si el efecto que presenta el silenciador depende de su posición con respecto al enhancer, se realizaron las construcciones que separa al silenciador del enhancer120 (Fig. 34). Se muestra que el efecto del silenciador es sobre el enhancer120

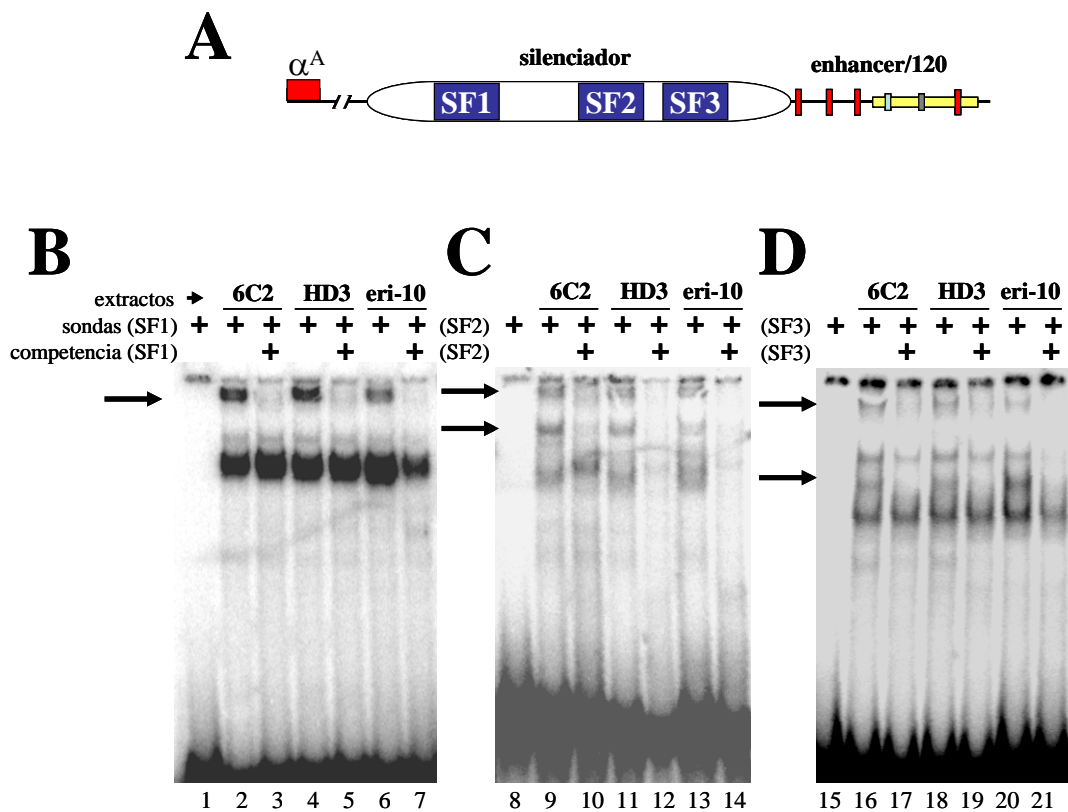
ó algún proceso iniciado desde este elemento. Lo anterior en base a que el silenciador, cuando se encuentra solo (carriles 3 y 4), no reporta una reducción en los niveles de expresión del promotor, incluso se observó un incremento en la actividad. Este incremento se podría deber a que la región definida de manera arbitraria como silenciador contenga parte del enhancer. Esta conclusión de vio reforzada cuando se estudio el efecto del enhancer en el contexto de la cromatina (ver mas adelante).



**Figura 34.** Análisis por transfecciones transitorias de la actividad del silenciador  $\alpha$ -globina. **A-** Esquema de las diferentes construcciones utilizadas en este estudio. La flecha negra colocada por debajo del silenciador muestra la orientación en que este se integro. Transfección transitoria de las construcciones mostradas en la línea celular HD3. **B-** transfección transitoria de las construcciones mostradas en “A” en eri-10.

La localización del silenciador nos parece muy particular, ¿por qué el elemento silenciador se encuentra contiguo al enhancer y entre este último y los genes globina? Solamente por su localización, podríamos pensar en diferentes mecanismos mediante los cuales estaría ejerciendo su efecto sobre el enhancer, todos ellos con un punto en común, vía la interferencia del efecto regulador del enhancer. La inversión la región del silenciador-enhancer con respecto al promotor (datos no publicados), no muestra ningún efecto. Por lo mismo, nos dimos a la tarea de separar a estos elementos. Como se puede observar el la figura 34, el silenciador fuera del contexto del enhancer sigue mostrando el mismo efecto. Este resultado muestra claramente que el silenciador sin importar que se encuentre río arriba o río abajo del promotor o en cualquiera de las orientaciones, sigue ejerciendo su efecto de silenciamiento, corroborando su actividad de silenciador clásico. El estudio de estas construcciones en un contexto estable, es decir en el contexto de la cromatina, nos ayudaría a mejorar su caracterización (ver mas adelante).





**Figura 35.** Estudio de las interacciones proteicas en la región del silenciador  $\alpha$ -globina A, Esquema de la región del silenciador/enhancer120, donde se resaltan las tres cajas SF1, SF2 y SF3 previamente reportadas como los sitios que presentaban interacciones proteicas en el silenciador (50) y los cuales fueron utilizados como sondas radioactivas para este estudio. **B**, Gel de retardo en el que se muestra la formación de un complejo específico. **C**, Gel de retardo en el que se muestra la formación de dos complejos específicos. **D**, Gel de retardo en el que se muestra la formación de dos complejos específicos.

Finalmente, se realizó un estudio de retardo usando extractos nucleares provenientes de distintas líneas eritroides (Fig. 35), utilizando como sondas radioactivas a los sitios SF1, SF2, y SF3 (50). El análisis de retardo muestra la interacción de múltiples y diferentes complejos (Fig. 35) que de manera muy interesante pueden ser competidos entre ellos aunque las sondas a nivel de secuencia no muestren sitios conservados (datos no mostrados).

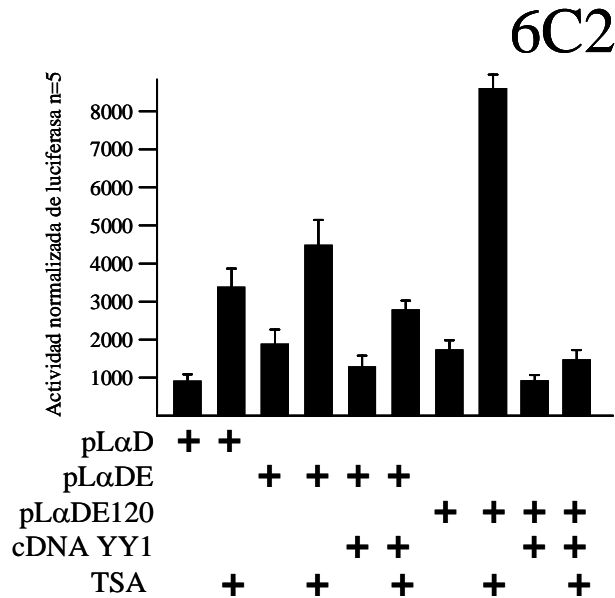
La caracterización de la región del silenciador nos muestra que a esta región se unen varios complejos proteicos, los cuales resulta indispensable caracterizar para entender la función del silenciador. Por los estudios previamente mostrados, podemos concluir que este silenciador puede ser definido como un silenciador clásico, el cual

muestra una actividad sobre la actividad del enhancer a grandes distancias de la región promotora y de manera independiente de su localización y orientación.

### **La conexión con la cromatina**

La estructuración en cromatina del ADN representa el segundo nivel al cual la transcripción puede ser regulada. Para que un factor transcripcional pueda acceder a sus secuencias blanco, se necesita romper con el efecto represor que genera esta. Entre las diferentes MPT de las histonas, la acetilación es de las pocas modificaciones a las que se les ha logrado atribuir un efecto de manera global, coincidiendo la hipo-acetilación con regiones cerradas y represivas de la cromatina y la hiper-acetilación con regiones abiertas y transcripcionalmente activas (268). Uno de los objetivos que nuestro grupo ha abordado en los últimos años, es comprender el papel que la estructura de la cromatina ejerce sobre la actividad transcripcional usando como modelo al dominio  $\alpha$ -globina.

El efecto negativo que genera la estructura de la cromatina, puede ser antagonizado mediante el uso de inhibidores de remodeladores de la cromatina, como la Tricostatina-A (TSA), la cual inhibe a las desacetilasas de histonas (136). Como parte del estudio *in vitro* de los elementos de regulación en el enhancer, mediante el uso de esta droga en transfecciones transitorias en la línea celular 6C2 (Fig. 36), obtuvimos un primer indicio de la posible participación activa de la estructura de la cromatina en la represión transcripcional de esta región.



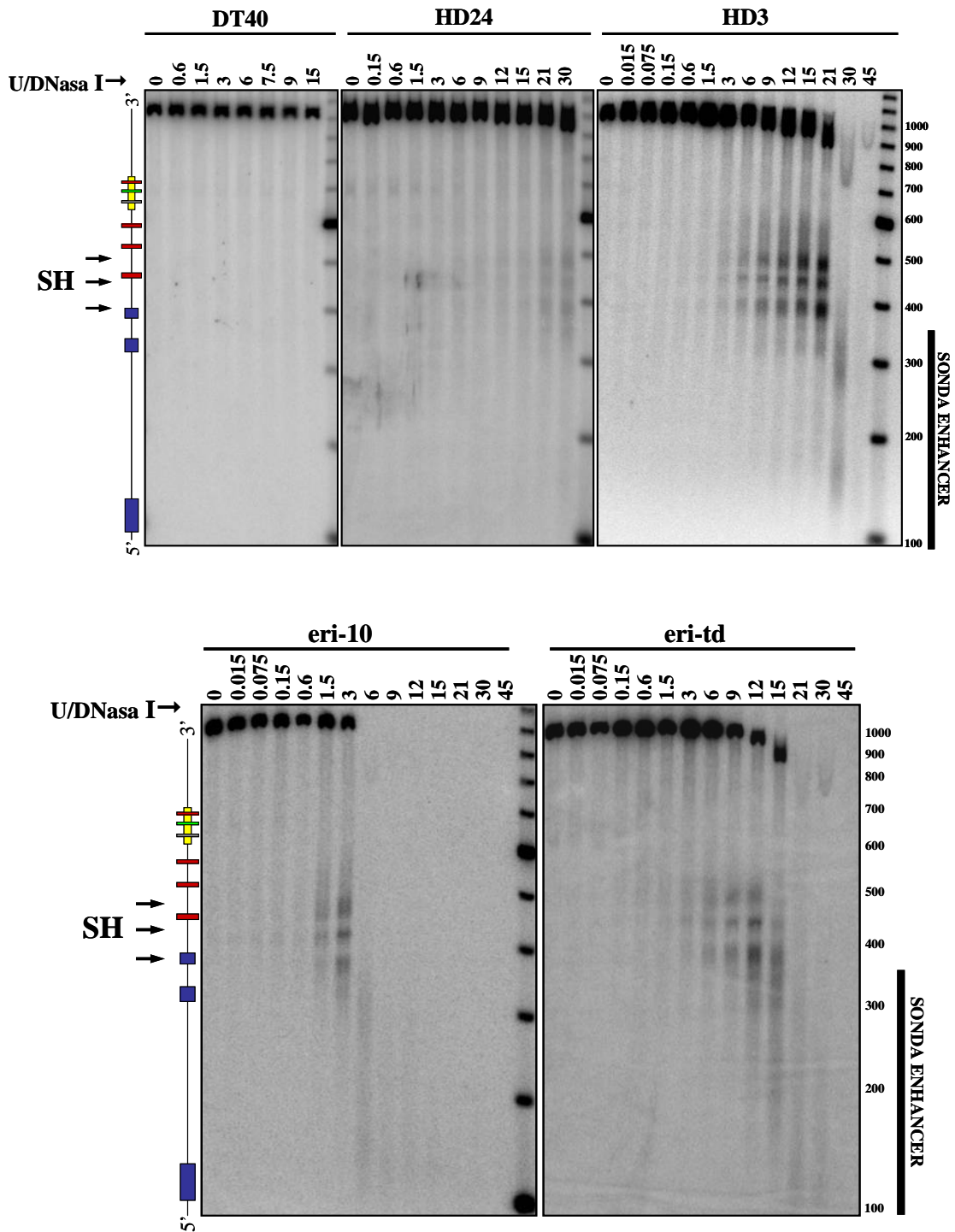
**Figura 36.** Análisis por transfecciones transitorias de la participación de la cromatina en la actividad del enhancer. Transfección transitoria mostrando el efecto positivo del inhibidor de desacetilasas (TSA) en las construcciones mostradas. El efecto estimulador de la actividad del enhancer, se ve reprimido con la co-transfección del ADNc de YY1.

En conclusión el dominio  $\alpha$ -globina de pollo presenta en su extremo 3' un elemento de regulación con actividades antagónicas. Por un lado se encuentra el silenciador al cual interaccionan diferentes factores nucleares. Este elemento de regulación se comporta como un silenciador clásico e interesantemente su actividad es independiente de la posición que presenta de manera endógena contiguo al enhancer. Por el otro lado el elemento enhancer que a su vez esta constituido por dos elementos, el “core” del enhancer el cual presenta tres sitios de unión al factor transcripcional GATA-1, siendo los primeros dos fundamentales para su actividad *in vitro*, y la región de 120 pb, en la cual encontramos la interacción de múltiples complejos proteicos que están modulando la actividad del enhancer de una forma tejido específica. Sin embargo los estudios *in vitro* que hasta este momento presentamos, no nos muestran la contribución que de manera endógena, esta aportando la estructura de la cromatina en su regulación.

## **Caracterización de la estructura de la cromatina en los elementos de regulación**

La regulación transcripcional gira en torno a la estructuración del genoma en cromatina y su efecto regulador sobre la expresión de los genes (99). El ensayo de digestión a la DNasa I ha sido utilizado en los últimos 20 años para localizar elementos de regulación a distancia potencialmente activos y como punto de partida en la determinación de su estructuración en cromatina (48).

Antes de que se realizara este trabajo, no teníamos idea de la posible participación de la cromatina en la regulación del silenciador-enhancer. Por lo mismo y como primera aproximación, se realizó un ensayo de accesibilidad a la DNasa I en núcleos aislados, en diferentes células de origen eritroides y no-eritroides (Fig. 37). Se observó un patrón de corte a la DNasa I eritroide específico (en base a que la línea no eritroide no muestra patrón de corte). La generación de estos sitios de hipersensibilidad (SH) a la DNasa I se genera de manera gradual a través de la diferenciación y el desarrollo, coincidente con los niveles de expresión observados para los genes  $\alpha$ -globina (Fig. 15). En la línea celular eritroide HD3, la accesibilidad a esta enzima es significativamente mayor a la observada en las células HD24. Este resultado sugiere que el enhancer se estructura en cromatina activa, a través de un proceso gradual, que inicia incluso de manera previa a la activación de los genes globina.



**Figura 37.** Análisis de la estructura de la cromatina en la región del silenciador/enhancer, mediante el ensayo de accesibilidad a la endonucleasa DNasa I. Se muestran los southern blots ilustrando la presencia de tres sitios de hipersensibilidad obtenidos al corte por la DNasa I (flechas negras). En la parte superior de cada southern se muestra la línea celular en que se realizó el estudio y las unidades de ADNasa I con que se digirió cada ensayo (representados por cada uno de los carriles).

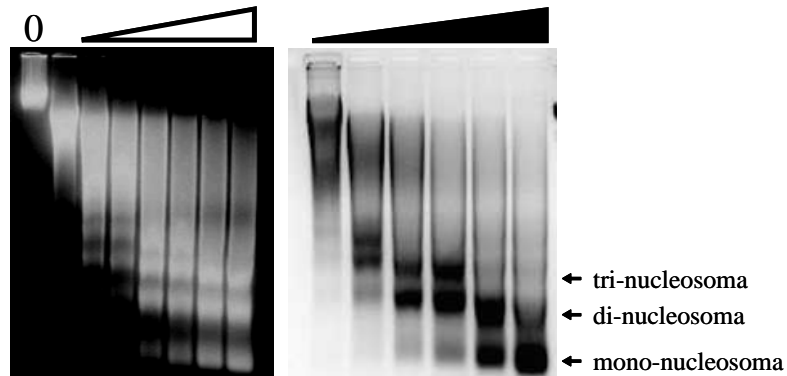
Este proceso gradual de remodelación de la cromatina, se ilustra por el número de unidades de enzima necesarias para obtener un patrón que muestre la misma intensidad de corte a la DNasas I, si comparamos a las líneas celulares HD24 y HD3, se necesitaron 7 veces más unidades (Fig. 37, comparar 21 contra 3 unidades, respectivamente). En eri-10 (estadio donde los genes globina se encuentran activos) se necesitaron 4 veces menos unidades que en células HD3 (Fig. 37, comparar 1.5 contra 6 unidades, respectivamente), es decir que en eritrocitos la estructura de la cromatina se encuentra 14 veces más accesible que en la línea eritroide-mieloide. Finalmente en eritrocitos terminalmente diferenciados (eri-td, cuando los genes globina dejan de ser expresados), la estructura de la cromatina se cierra parcialmente (Fig. 37).

Este resultado nos invita a proponer un modelo de apertura de la cromatina gradual, en el cual el enhancer previo a ser requerido para activar a los genes globina, se encuentra en un estado “semi-activo”, como ya se ha mostrado en algunos trabajos y se ha propuesto para el modelo de pizarra (52, 53). La expresión en tiempo y tejido específico de un factor transcripcional, podrá regular la temporalidad (del inglés “timing”) de expresión de los genes vía el enhancer. Si correlacionamos la expresión, con la estructura de la cromatina del enhancer, encontramos que sus genes se están expresando en el estado coincidente con la mayor apertura de la cromatina y posteriormente se re-estructura parcialmente en una cromatina más compacta, cuando el enhancer parece ya no ser funcional.

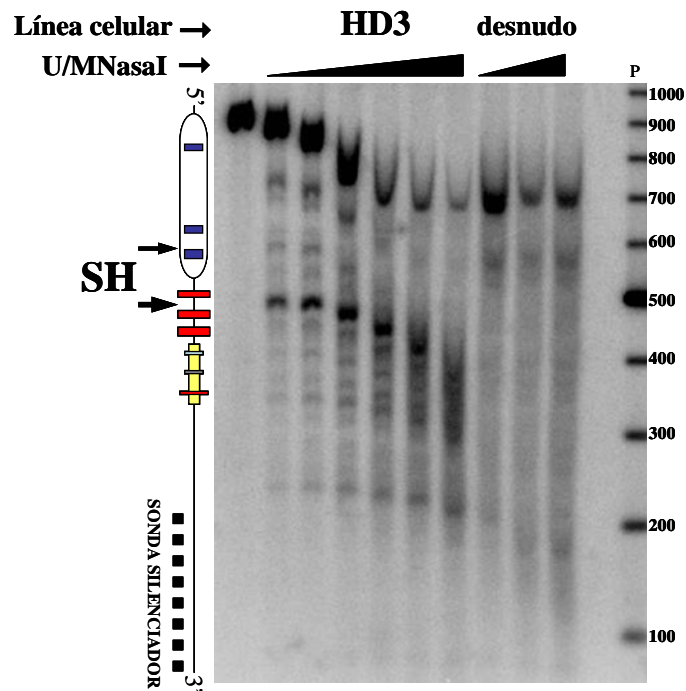
### **Determinación del posicionamiento nucleosomal**

El posicionamiento específico de cada uno de los nucleosomas, podría implicar la liberación u ocultamiento de secuencias de regulación. Con el fin de determinar el posicionamiento nucleosomal en los elementos silenciador y enhancer<sup>120</sup>, se realizaron digestiones nucleares con la endonucleasa micrococcal (MNasa), la cual tiene la

propiedad de digerir en las secuencias inter-nucleosomales, generando patrones de mono-, di-, etc. nucleosomas (Fig. 38).



**Figura 38.** Calibración de las digestiones en núcleos con concentraciones creciente de la MNasa. Geles de agarosa mostrando los patrones de corte obtenidos al digerir núcleos celulares con concentraciones crecientes (triángulos blanco y negro) de la MNasa. “0” representa el control de ADN genómico no digerido por la MNasa.



**Figura 39.** Análisis del posicionamiento nucleosomal en el elemento silenciador-enhancer. Estandarización del análisis del posicionamiento nucleosomal por la digestión de núcleos con la MNasa. Se muestra el southern blot de la hibridación realizada a la línea celular HD3, mostrando la presencia de varias bandas generadas por el corte del ADN por parte de la MNasa en secuencias inter-nucleosomales o incluso zonas libres de nucleosomas. Se resaltan dos sitios de hipersensibilidad (SH, flechas negras) que proponemos deben ser importantes para la actividad del enhancer (ver mas adelante). “desnudo” se refiere al control de digestión con la enzima en ADN genómico, este control nos sirve para descartar bandas inespecíficas.

La digestión de núcleos con este enzima, nos permite por una hibridación indirecta localizar por patrones de bandas, las regiones que se encuentran delimitando a los nucleosomas (Fig. 39, carriles 1-7). Adicionalmente por una digestión *in vitro* de ADN genómico digerido con la misma enzima, podemos descartar que los patrones de corte observados se deban a la preferencia de corte que el ADN por su composición en bases púricas y pirimídicas presente (Fig. 39).

Una vez establecida la técnica y los controles necesarios, se realizó el mismo análisis en las diferentes líneas celulares eritroides y en eri-10 y en eri-td (Fig. 40). Los resultados obtenidos ponen en evidencias un patrón nucleosomal eritroide específico (Fig. 40), caracterizado por la presencia de dos SH (flechas negras) que van incrementando su intensidad a partir de HD24 (donde aparecen), hasta llegar a su máxima intensidad (máxima frecuencia de corte), en el estadio coincidente con la expresión de los genes globina. Este resultado se puede interpretar como que estadísticamente se está viendo favorecido un posicionamiento nucleosomal, aparentemente<sup>5</sup>, en la región localizada entre el silenciador y el enhancer. Un aspecto adicional, es la presencia de un nucleosoma fijo en la región río abajo del enhancer (flechas blancas) y de manera muy particular co-localizando con la secuencia de 120 pb (encargada de modular la actividad del enhancer), aparecen 4 SH (asteriscos) como consecuencia del remodelamiento activo de la cromatina al que se ve sujeta esta región. Este remodelamiento coincide con los estadios donde el enhancer mostró tener su mayor actividad (Fig. 22).

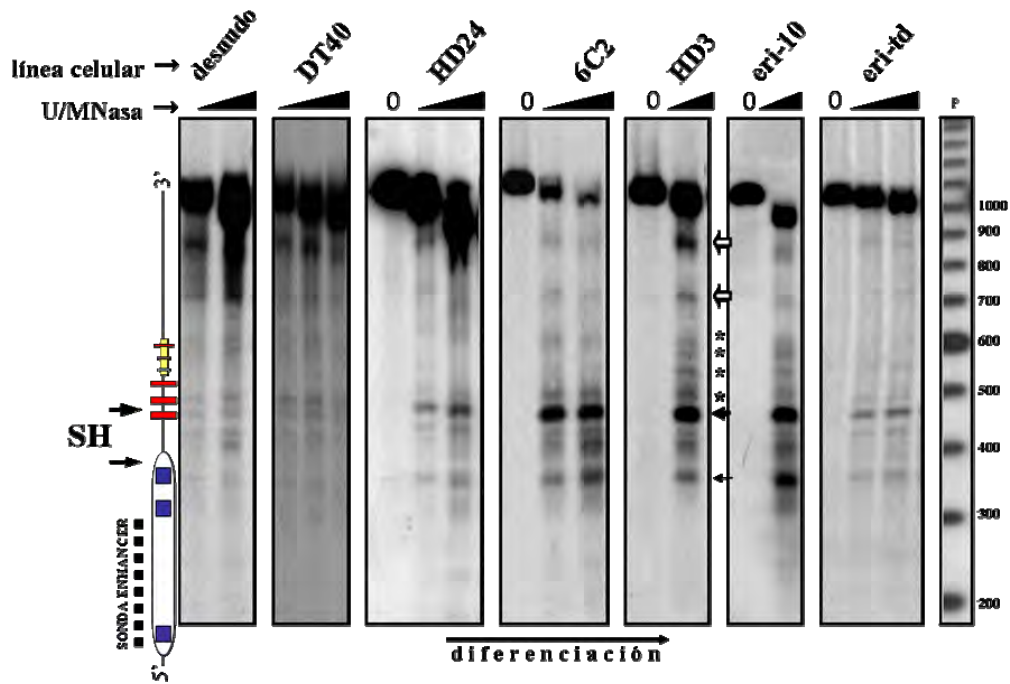
Para corroborar nuestros resultados y poder tener una mayor definición en la región del silenciador, decidimos repetir el análisis del posicionamiento nucleosomal utilizando las mismas membranas, pero ahora con una sonda que incrementa la resolución en la región del silenciador (Fig. 41). Los datos obtenidos mostraron de manera consistente, que la conformación de la cromatina en el enhancer se repite; caracterizado por la aparición de dos SH intermedios (flechas negras) y el remodelamiento en la secuencia de 120 pb. Interesantemente identificamos otro nucleosoma fijo sobre el silenciador (flechas blancas). Este tipo de estructuras cromatínicas enmarcadas por dos nucleosomas ha sido descrita en la literatura para

---

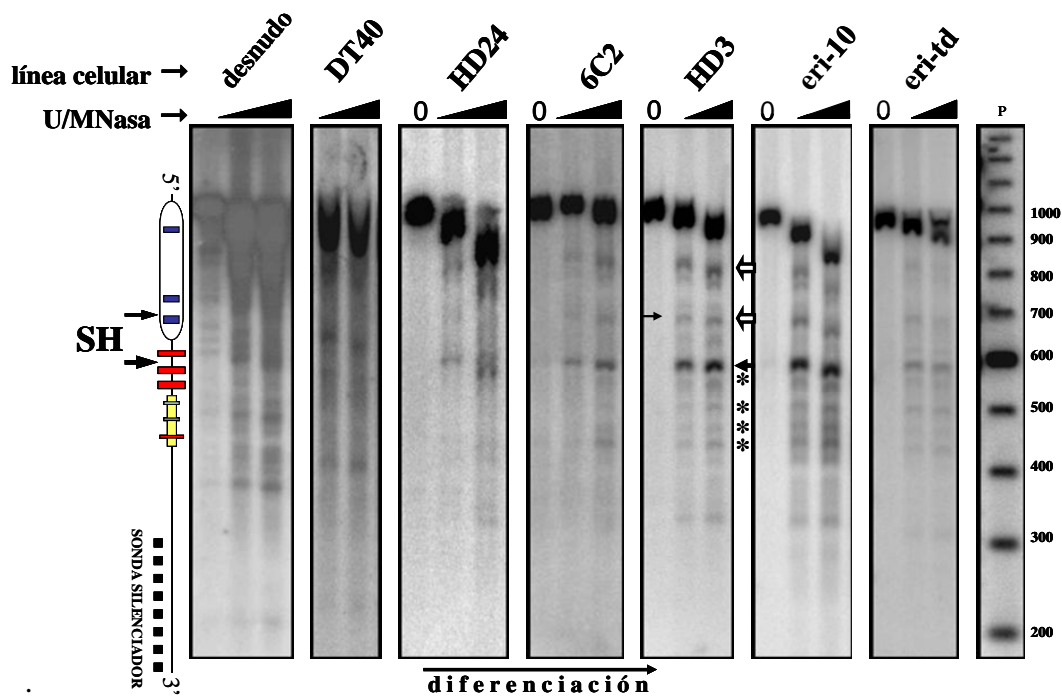
<sup>5</sup> Aparentemente, ya que este estudio tiene una definición de  $\pm 40$  pb.



regiones promotoras (249), que además coincide con zonas libres de nucleosomas (318, 211, 202).

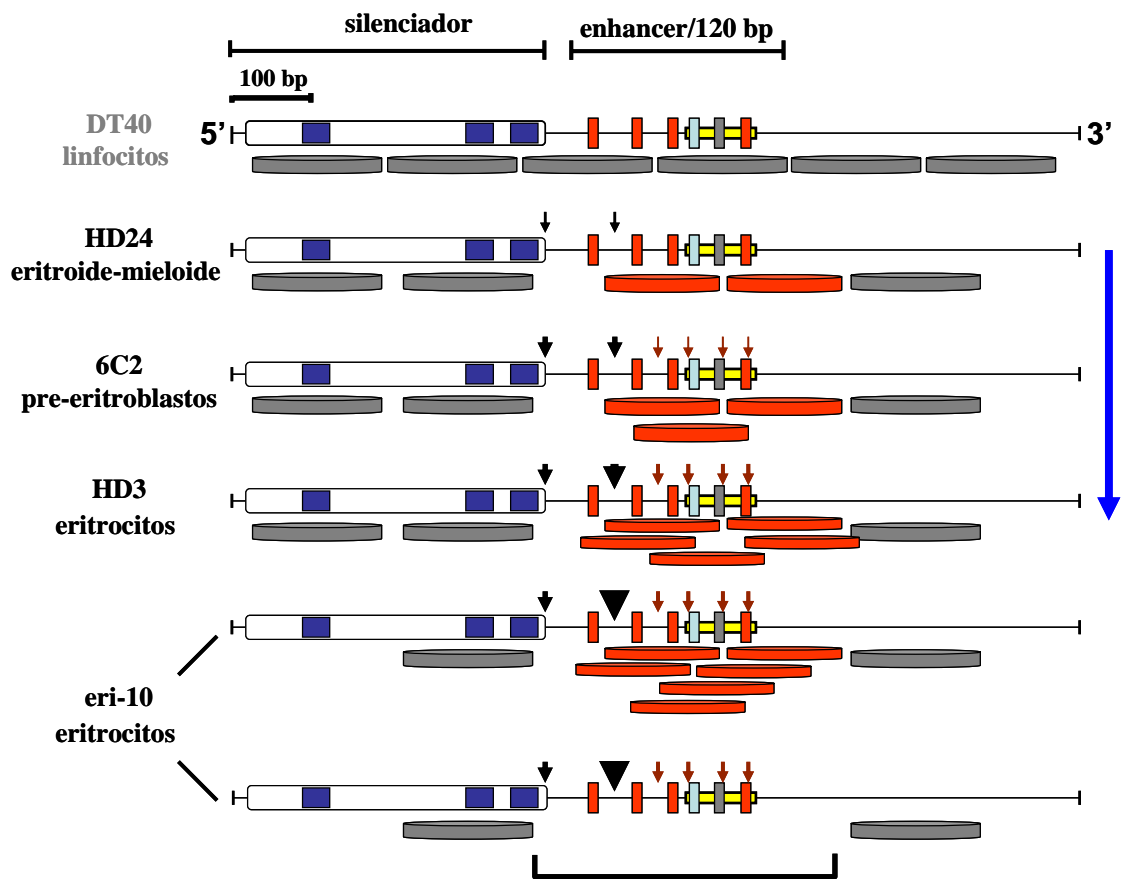


**Figura 40.** Análisis del posicionamiento nucleosomal en el enhancer 120pb. Análisis del posicionamiento nucleosomal en la región del enhancer (sonda enhancer línea negra punteada) en las diferentes líneas celulares que se muestran en la parte superior. Para cada una de estas líneas se realizó un estudio como el que se muestra en la figura 39 (solo se muestran los carriles importantes). Flechas negras SH principales. Asteriscos, remodelamiento en la secuencia de 120 pb. Flechas blancas, presencia de un nucleosoma fijo río abajo del enhancer.



**Figura 41.** Análisis del posicionamiento nucleosomal sobre el silenciador. Re-hibridación de las membranas utilizadas en la figura 40 con la sonda silenciador. Flechas negras SH principales. Asteriscos, remodelamiento en la secuencia de 120 pb. Flechas blancas, presencia de un nucleosoma fijo sobre el silenciador.

Para facilitar la interpretación de los resultados, las membranas fueron cuantificadas con un equipo Storm de Molecular Dynamics para el análisis de muestras marcadas con fósforo y procesadas por el programa ImageQuant 5.2 de (Amersham). El resumen de todas las membranas procesadas con la determinación del posicionamiento y remodelamiento nucleosomal observado, se muestran en la figura 42.



**Figura 42.** Resumen del análisis del posicionamiento nucleosomal en la región del silenciador/enhancer120. Esquema a escala representando cada uno de los estadios donde se analizó el posicionamiento nucleosomal. Cilindros planos grises, nucleosomas fijos. Cilindros planos rojos, nucleosomas activos (en movimiento constante). Flechas negras, sitios de hipersensibilidad (su grosor representa la intensidad de corte). Flecha azul, diferenciación. Corchete, región libre de nucleosomas. El tamaño e intensidad de las flechas, esquematiza el incremento en la intensidad de corte por la MNasa.

En esta figura se esquematiza el patrón eritroide específico, caracterizado por la presencia de dos SH (flechas negras), este patrón eritroide se va formando gradualmente a través de la diferenciación y el desarrollo, significado de un posicionamiento nucleosomal que se está viendo estadísticamente favorecido. También se observan una serie de SH (flechas rojas) coincidentes con la secuencia 120 pb, los cuales son más evidentes en los estadios donde el enhancer se encuentra más activo. Dos son los escenarios posibles para poder explicar el patrón de SH observado: el primero plantea un movimiento activo de los nucleosomas en la región del enhancer, como consecuencia de los diferentes procesos de remodelación y cinéticas de las proteínas que en él interaccionan. Un aspecto muy interesante, es el hecho de que el enhancer se encuentre delimitado por dos nucleosomas estáticos, que parece como si estuvieran enmarcando

las secuencias relevantes para la actividad del enhancer. El segundo sugiere la formación de una región libre de nucleosomas, que de la misma forma estaría flanqueada por dos nucleosomas estáticos.

A partir de estos resultados surgieron dos preguntas ¿el remodelamiento en la estructura de nucleosomas significa la generación de regiones más accesibles en la cromatina, para que los factores transcripcionales puedan reconocer sus secuencias? y ¿cuál es la localización y posible significado de los SHs intermedios observados?

### **Evaluación semicuantitativa de la remodelación de la cromatina**

Para contestar a la primera pregunta, se realizó un análisis *in situ* de accesibilidad a enzimas de restricción a través de la diferenciación eritroide y desarrollo del pollo. En este estudio mediante una digestión enzimática en núcleos aislados y analizando de manera indirecta por southern blot, es posible medir la accesibilidad de manera semicuantitativa<sup>6</sup> de diferentes enzimas de restricción a sus secuencias de reconocimiento en el ADN en su contexto cromatínico. En la figura (Fig. 43), se muestran los controles y la cuantificación que se realizó para cada ensayo, todas las digestiones fueron hibridadas con las sondas silenciador y enhancer, buscando resultados reproducibles y consistentes.

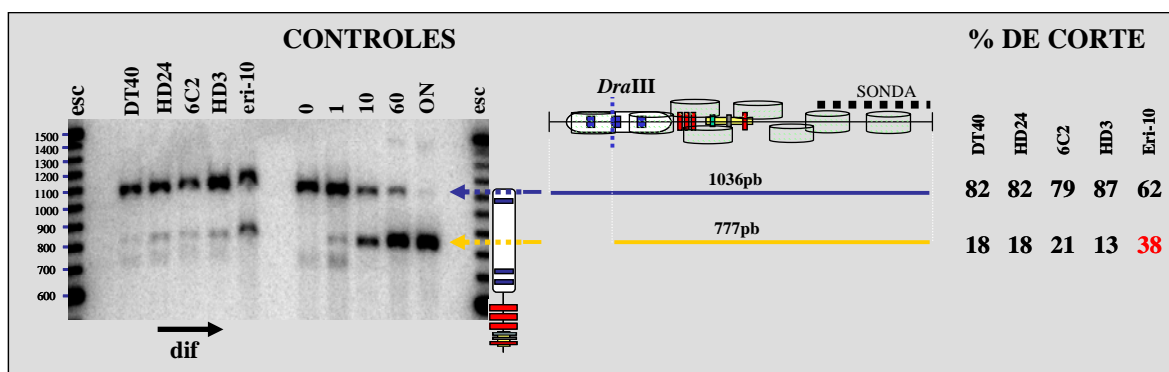
Con la finalidad de analizar la accesibilidad diferencial que presentaba la estructura de la cromatina del silenciador, se realizó una digestión en núcleos aislados con la enzima *DraIII* (Fig. 43, D), la cual presenta un sitio de reconocimiento en la secuencia asociada al silenciador. Como se puede observar, la estructura de la cromatina en el silenciador permanece “cerrada” (ya que permanece inaccesible al corte por las enzimas de restricción). Es hasta el estadio de eritrocitos donde se observa un incremento en la accesibilidad, en una etapa donde los genes globina son transcripcionalmente activados. Este resultado fue confirmado con una digestión con la enzima *AvaI*, la cual presenta dos secuencias de corte en el silenciador (A1 y A2) (Fig. 44).

¿Cual podría ser el significado de que la estructura de la cromatina en el silenciador se encontrara abierta al corte por nucleasas en un estadio donde se supone

---

<sup>6</sup> Por resultados que no muestro, se encontró que este estudio de realizarse por triplicado, es tan exacto que podría tomarse como un estudio cuantitativo.

que el enhancer es elemento que debería estar regulando? Si regresamos a los patrones de expresión en eritrocitos la respuesta parece evidente, el silenciador podría estar participando silenciando al gen embrionario  $\pi$ , permitiendo la activación de los genes adultos por parte del enhancer. Para confirmar esta hipótesis de manera parcial<sup>7</sup>, se realizó un análisis de accesibilidad a enzimas de restricción en núcleos aislados correspondientes a distintas etapas del desarrollo (Fig. 45). Como se puede observar, el silenciador se encuentra en una conformación de la cromatina más abierta y de manera coincidente, con el estadio donde se lleva a cabo el recambio en la expresión de los genes globina. Posteriormente se observa una estructuración en cromatina cerrada, en un proceso que parece pasivo<sup>8</sup>. En conclusión el silenciador pondría ser el responsable del silenciamiento del gen embrionario  $\pi$ .



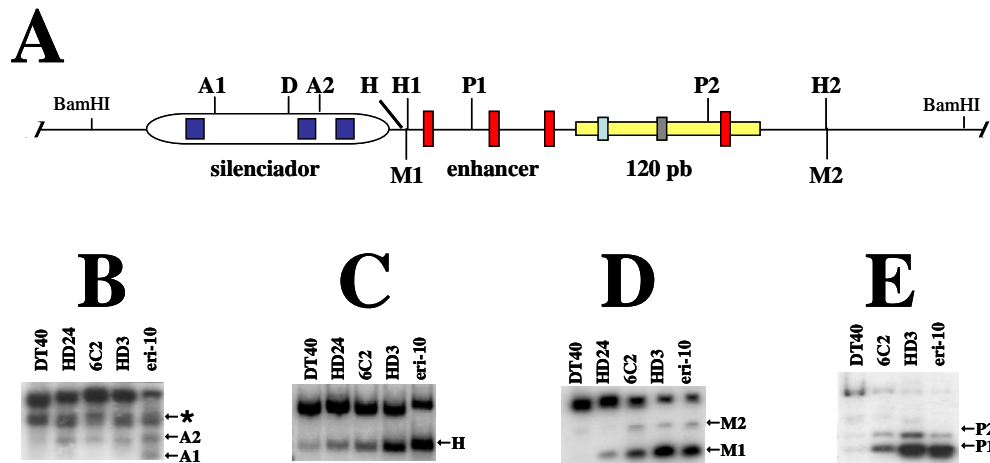
**Figura 43.** Análisis de la accesibilidad diferencial de la cromatina a lo largo de distintas etapas de la diferenciación eritroide. Izquierda, southern blot de la digestión en núcleos que se realizó con la enzima *DraIII* en cada una de las líneas indicadas. Como control se digirió ADN genómico para descartar las bandas que se debieran a una hibridación inespecífica. En medio, esquema del diseño experimental y las bandas esperadas utilizando la sonda silenciador (para cada estudio de accesibilidad se hibridó con ambas sondas, silenciador y enhancer). Derecha, cuantificación con el programa Image Quant 5.2 de las digestiones observadas en el southern de la izquierda.

Para analizar la dinámica en la accesibilidad a la estructura de la cromatina en la región del enhancer<sup>120</sup>, se utilizaron las enzimas *HhaI* (H), *MspI* que presenta dos

<sup>7</sup> Ya que una comprobación contundente necesitaría otro tipo de estudios, como la mutación puntual del silenciador en un sistema donde se pueda inducir a diferenciación y se presente este fenómeno de recambio en la expresión de los genes.

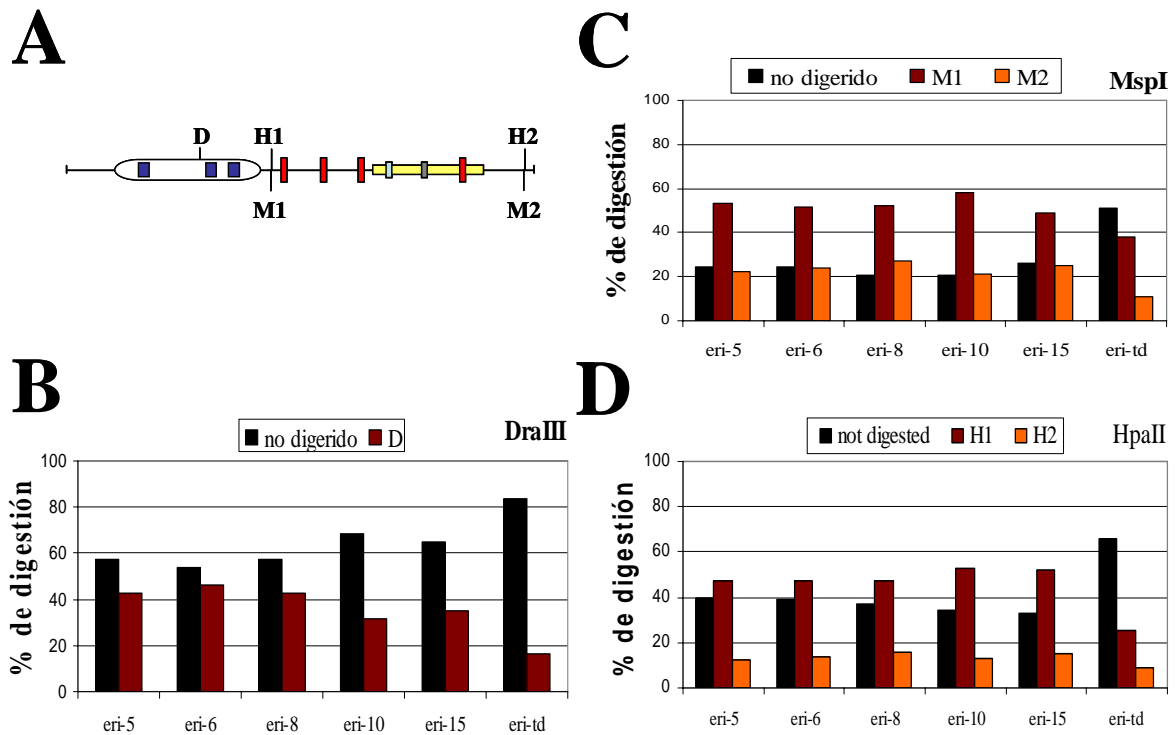
<sup>8</sup> Un mecanismo pasivo se estaría llevando gradualmente de manera ciclo dependiente y sin la utilización de ningún complejo enzimático (como en el caso de la metilación), a diferencia de un proceso activo, donde la participación de proteínas con actividades enzimáticas son las encargadas de llevar a cabo el cambio.

sitios de corte (M1 y M2) y *PstI* con la misma situación (P1 y P2) (Fig. 44). De manera consistente con los resultados obtenidos con los ensayos de DNasa I y MNasa (Fig. 37 y 41), se observa en todos los casos una apertura de la cromatina gradual, caracterizada por una accesibilidad parcial en la línea DT40 y alcanzado su mayor accesibilidad al ADN en eritrocitos.



**Figura 44.** Análisis de la accesibilidad diferencial de la cromatina por parte de distintas enzimas de restricción a través de la diferenciación. **A**, Esquema de la región del silenciador/enhancer120 con la localización de cada una de las enzimas utilizadas en este estudio. **B**, Análisis del silenciador con las enzimas *AvaI* (A1 y A2). Asterisco, localiza una banda obtenida por una hibridación inespecífica. **C-E**, Análisis de la región del enhancer con las enzimas *HhaI* (H), *MspI* (M1 y M2) y *PstI* (P1 y P2).

De la misma manera que se hizo con el silenciador, nos interesó comprender la dinámica que la estructura de la cromatina en términos de su accesibilidad estaba mostrando a través del desarrollo (Fig. 45). Como se puede observar, el enhancer permanecer en una estructura accesible al corte para las enzimas *MspI* y *HpaII*, y es hasta eritrocitos terminalmente diferenciados, donde la estructura de la cromatina adopta una conformación compacta. Como control adicional para estos experimentos, se utilizó la enzima *MspI*, la cual presenta un sitio de restricción localizado río abajo del enhancer, este sitio de restricción se mostró inaccesible al corte por *MspI* (Fig. 43, sitio M2 y datos no mostrados). Por los que podemos concluir que la remodelación de la estructura de la cromatina, genera un incremento en la accesibilidad de manera diferencial.



**Figura 45.** Análisis de la accesibilidad diferencial de la cromatina a través del desarrollo de pollo. **A-** Esquema representando la región del silenciador/enhancer con la localización de cada una de las enzimas utilizadas en este estudio. **B-** Análisis a través del desarrollo con la enzima *DraIII*. Barras negras: ADN no digerido; barras cafés: ADN digerido por *DraIII*. **C-** Análisis a través del desarrollo con la enzima *MspI*. Barras negras: ADN no digerido; Barras cafés: ADN digerido por *MspI* en M1. Barras naranjas: ADN digerido por *MspI* en M2. **D-** Análisis a través del desarrollo con la enzima *HpaII*. Barras negras: ADN no digerido; Barras cafés: ADN digerido por *HpaII* en H1. Barras naranjas: ADN digerido por *HpaII* en H2.

En resumen, estos experimentos muestran una dinámica diferencial entre la estructura local de la cromatina asociada al silenciador y al enhancer, representando los SH localizados en la región intermedia, posiblemente una estructura cromatínica particular entre estos dos elementos. El silenciador se muestra en una estructura de la cromatina compacta a través de la diferenciación y es hasta un estadio localizado alrededor de los 6 días de gestación (coincidente con el recambio de expresión de los genes embrionarios a los adultos), donde se genera posiblemente por algún proceso de remodelación de la estructura de la cromatina, un incremento en la accesibilidad al corte por las enzimas que lo reconocen. Posteriormente, este elemento se vuelve a estructurar en cromatina inaccesible de una manera gradual (y posiblemente pasiva). Por otro lado el enhancer se va estructurando gradualmente en una cromatina accesible a través de la diferenciación, al llegar al estadio de eritrocitos alcanza y se mantiene en su mayor estado de apertura de la cromatina. Finalmente al llegar a eri-td, el enhancer se re-

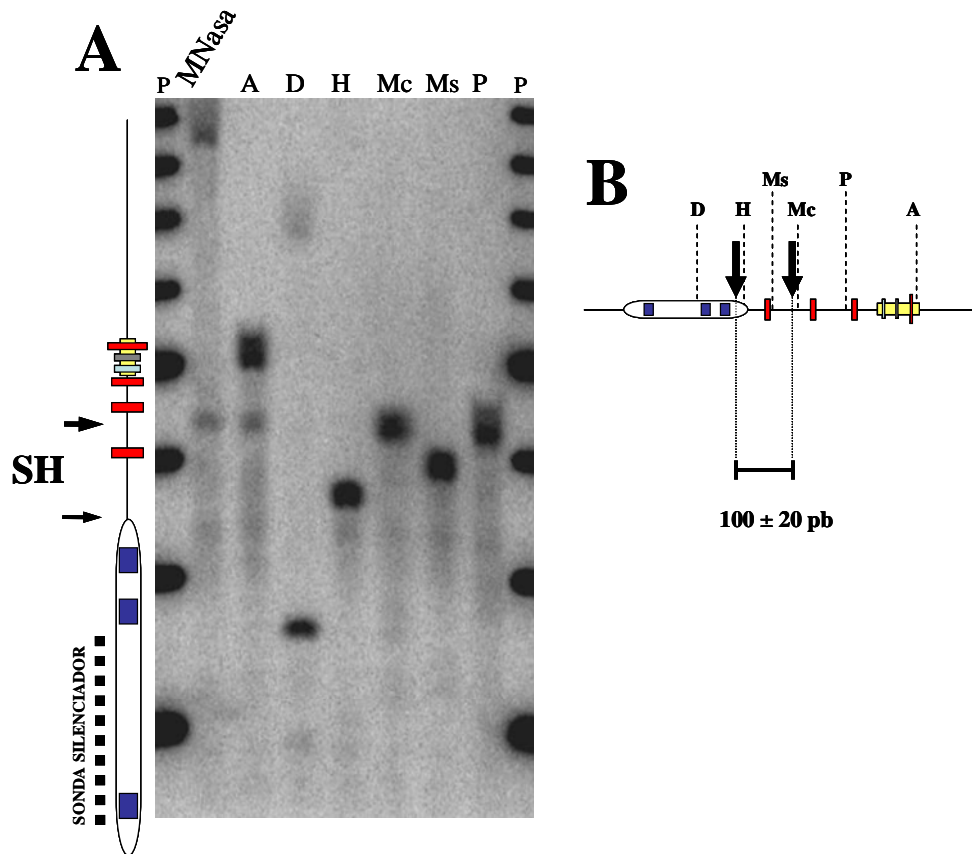
estructura en una cromatina inaccesible. Por lo tanto, el enhancer presenta una mayor ventana de acción, coincidente siempre con la expresión de algún gen de la familia  $\alpha$ -globina.

### **Mapeo de los sitios de hipersensibilidad**

La formación de dos SH en la región intermedia entre el silenciador y el enhancer nos generó un interés particular. Esta región parece estar delimitando al enhancer del silenciador. Alternativamente podría tratarse de un centro de enucleación que atrajera actividades remodeladoras de la cromatina, lo anterior en base a los patrones de digestión observados con la DNasa I, MNasa y enzimas de restricción (Figs. 37, 41, 43 y 45), donde todas coinciden en mostrar a esta región como la primera en la que se llevan a cabo los eventos de remodelación. Para abordar estos posibles escenarios, decidimos determinar la localización precisa y caracterizar la secuencia que se encontraba entre estos dos sitios de corte (Fig. 46).

Por medio de una técnica que combina la digestión *in vivo* de núcleos con la MNasa y la posterior digestión *in vitro* con enzimas de restricción (del ADN previamente digerido con la MNasa), se mapearon los sitios SHs generados por la MNasa (Figs 31 y 34). Sorpresivamente el fragmento de ADN que se encontraba entre estas dos SH, no era lo suficientemente grande para poder contener a un nucleosoma (Figs 30 y 32). El primer SH se localizó entre el silenciador y el enhancer, el segundo SH entre la primera y la segunda caja de unión al factor transcripcional GATA-1, justo entre los sitios que previamente definimos (Fig. 24) necesarios para la actividad del enhancer. Este resultado generó muchas preguntas ¿realmente la región contenida entre estos dos elementos se podría tratar de una región libre de nucleosomas? El sitio contenido entre estas dos hipersensibilidades es un sitio de interacción al factor transcripcional GATA-1 ¿Realmente está interaccionando GATA-1 entre estas dos secuencias? ¿Cuál es el papel que el factor transcripcional GATA-1 podría estar llevando a cabo en la generación de los SH, así como en el posicionamiento nucleosomal del enhancer?

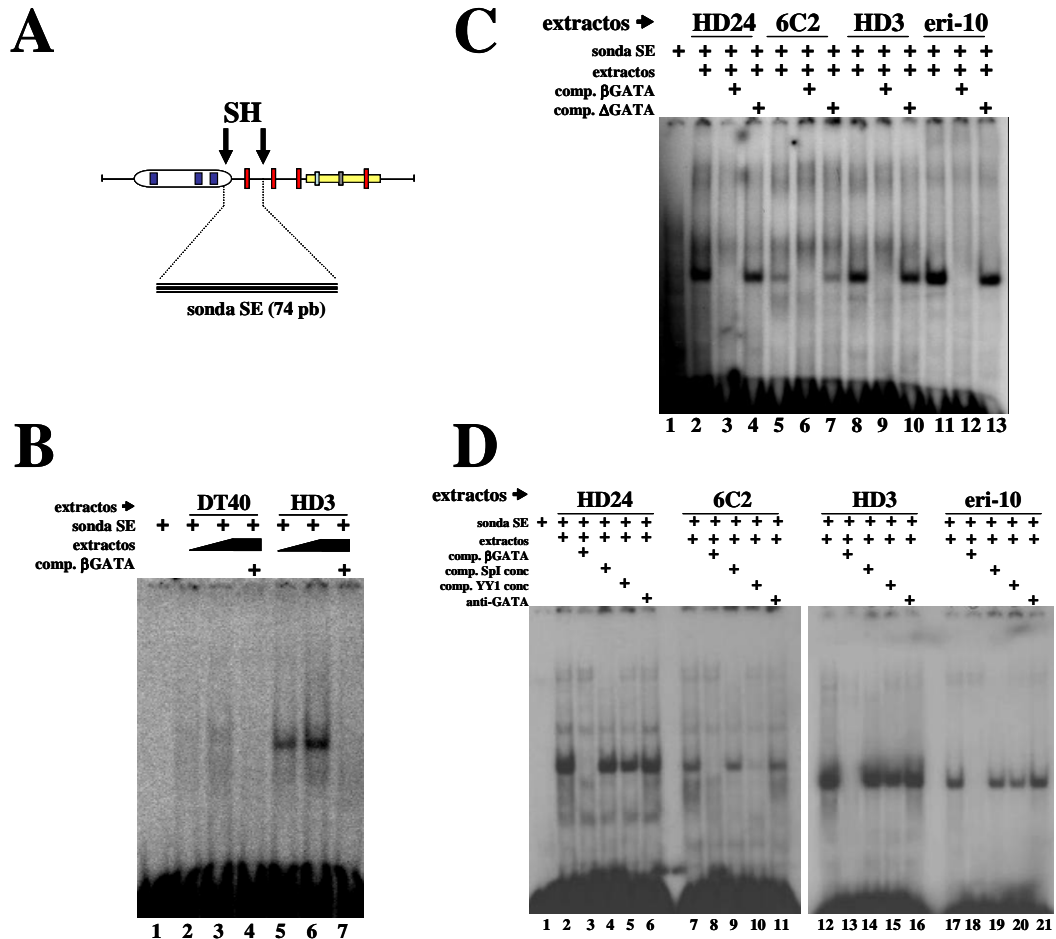




**Figura 46.** Mapeo de los sitios de hipersensibilidad. **A**, Souther blot mostrando el mapeo de los SH (flechas negras) localizados en la región intermedia entre el silenciador y el enhancer, con las enzimas *Ava*I (A), *Dra*III (D), *Hha*I (H), *Msc*I (Mc), *Msp*I (Ms) y *Pst*I (P). **B**, Esquema mostrando el sitio de restricción de las enzimas utilizadas y la determinación de la localización de los SH, así como la distancia que se calculo hay entre ellos.

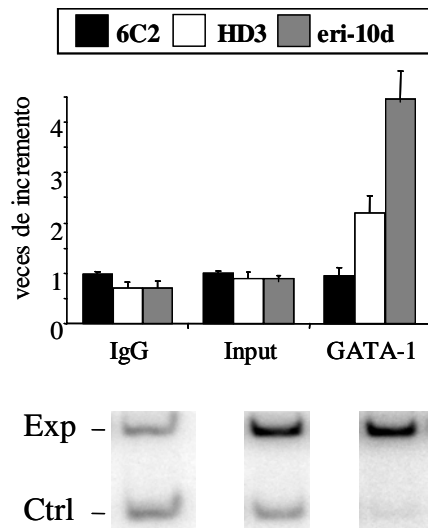
En una primera aproximación para definir las interacciones proteína-ADN que se estaban llevando a cabo en la región comprendida entre los SHs, se realizaron una serie de ensayos de retardo, con los que encontramos que una proteína o complejo, se encontraba interaccionando de manera eritroide específica, posiblemente GATA-1 (Fig. 47 B). Esta interacción es específica para la secuencia de GATA-1, lo anterior lo encontramos al competir de manera positiva con la secuencia consenso para esta proteína, y al encontrar que las secuencias consenso para las proteínas Sp1 o YY1 no competían, así como la secuencia con el sitio de interacción a GATA mutado (Fig. 47 C y D). Sorpresivamente y apoyandonos en los resultados previos (Fig. 31) cuando competimos con la sonda que une al factor YY1 en la línea celular 6C2 (Fig. 47 D, carril 10) se compite el complejo. Este resultado apoya la asociación de YY1 a la región del “core” del enhancer y favorece la interacción entre GATA-1 y YY1, favorece un

modelo en el cual la proteína YY1 podría estar implicada en la modulación de la actividad del enhancer en pre-eritroblastos.



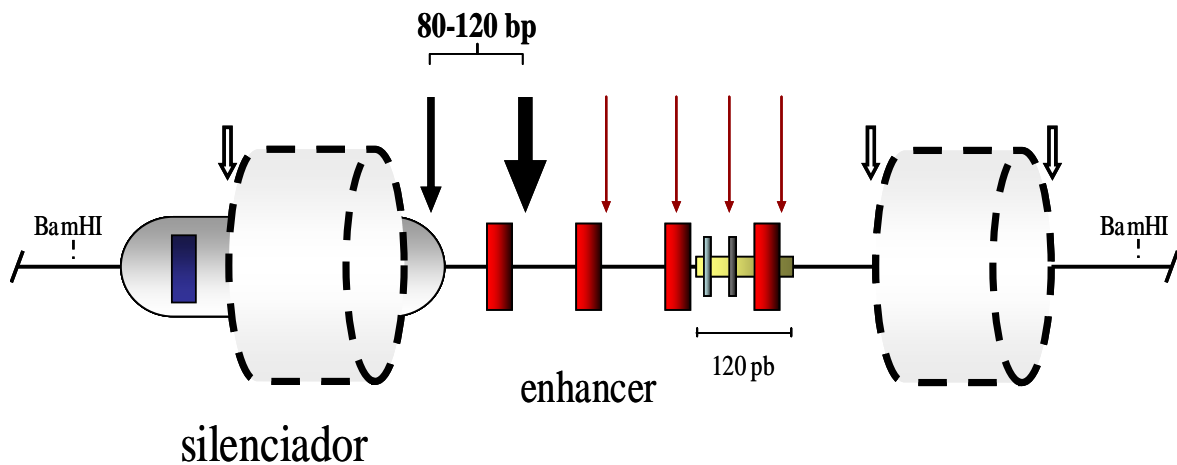
**Figura 47.** Estudio de las interacciones proteicas en la secuencia localizada entre los sitios de hipersensibilidad A- Esquema de la región del silenciador/enhancer, donde se localizan los SH mapeados y la región intermedia que se decidió utilizar como sonda para el análisis por retardo. B- Gel de retardo en el que se muestra la formación de un complejo eritroide específico. C- Gel de retardo en el que se muestra la competencia específica del complejo por la sonda consenso a GATA  $\beta$ GATA y no por la sonda que incluye el sitio G1 mutado (ver fig. 27). D- Gel de retardo en el que se muestra que el complejo no es competido de manera inespecífica por las sondas que incluyen el sitio de interacción a Sp1 (cSp1) o a YY1 (cYY1). También se muestra el ensayo de super-retardo (anti.GATA).

Con la intención de comprobar *in vivo* la interacción del factor transcripcional GATA-1 en la región de “core” del enhancer, se realizó un ChIP semi-cuantitativo con el anticuerpo específico anti-GATA (Fig. 48). Este resultado muestra claramente como el factor GATA-1 se va incorporando gradualmente al enhancer, de manera similar al proceso gradual de remodelación de la estructura de la cromatina en el enhancer.



**Figura 48.** Asociación *in vivo* del factor transcripcional GATA-1 en la región del enhancer  $\alpha$ -globina. Histograma mostrando los resultados normalizados a partir de una PCR-duplex hecha por triplicado y de dos inmunoprecipitaciones independientes. En la parte inferior se muestra un gel representativo del CHIP realizado en eri-10. Exp. Amplificación con los oligos que amplifican en la región del enhancer. Ctrl., representa la amplificación que se utilizó para normalizar, utilizando un par de oligos que amplifican a 3 kb río abajo del enhancer.

Con los resultados mostrados hasta este punto, consideramos tener una primera caracterización de los elementos silenciador y enhancer, suficiente para poder iniciar las preguntas funcionales a nivel de la cromatina. La caracterización a nivel de la cromatina, de la región comprendida en las 1036 pb nos mostró que la remodelación de la cromatina en el enhancer  $\alpha$ -globina es eritroide específica, caracterizada por la presencia de dos SH que aparecen de manera gradual (flechas negras, Fig. 49). El primer SH se localiza entre estos dos elementos y el segunda entre las cajas GATA-1 (G1 y G2), localizando en la región intermedia un sitio de interacción para el factor GATA-1, el cual al ser mutado *in vitro* (así como el sitio G2) genera la pérdida total en la actividad de *trans*-activacional de este elemento. Se demostró la interacción específica del factor transcripcional GATA-1 de manera gradual en el enhancer, coincidente con el patrón de remodelación observado. Para explicar este resultado, proponemos dos modelos en los cuales la estructura de la cromatina se está remodelando de manera gradual, permitiendo la interacción del factor GATA-1. El segundo modelo propone que GATA-1, se encuentra directamente implicada en esta remodelación gradual.

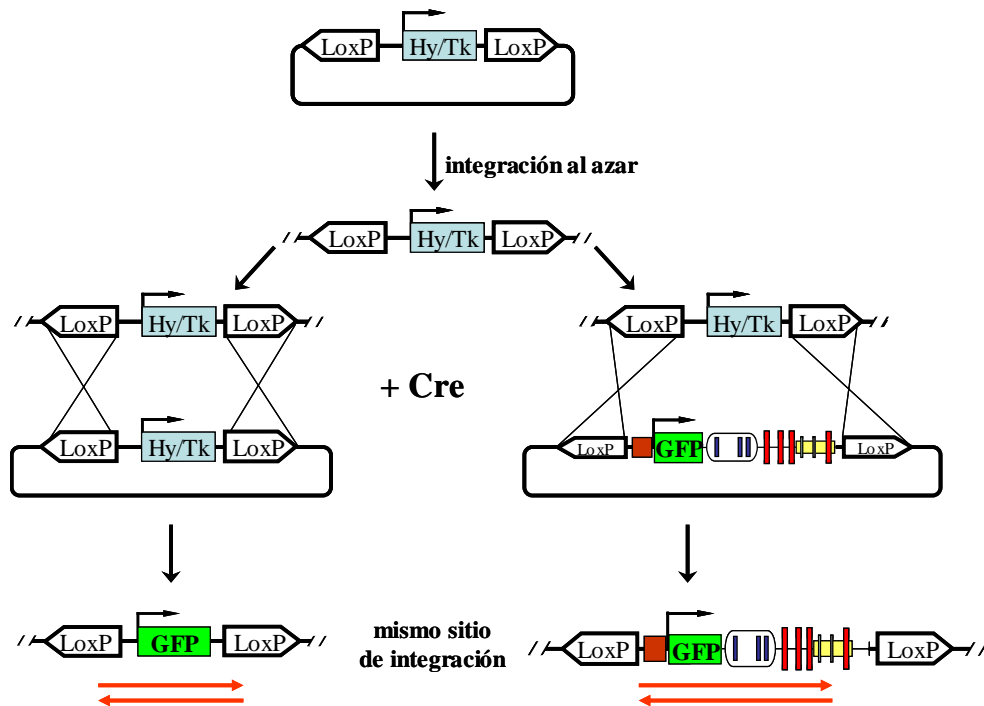


**Figura 49.** Esquema resumen de la caracterización de la estructura de la cromatina determinada en la región silenciador/enhancer. Esquema de la región del silenciador/enhancer, donde se localizan los SH mapeados (flechas negras y rojas) y la aparición de SH constantes (flechas blancas) que muestran la existencia de dos nucleosomas fijos (cilindros punteados), que estarían enmarcando la región enhancer 120.

### **Determinación *in vivo* de la participación de los sitios de unión al factor GATA-1**

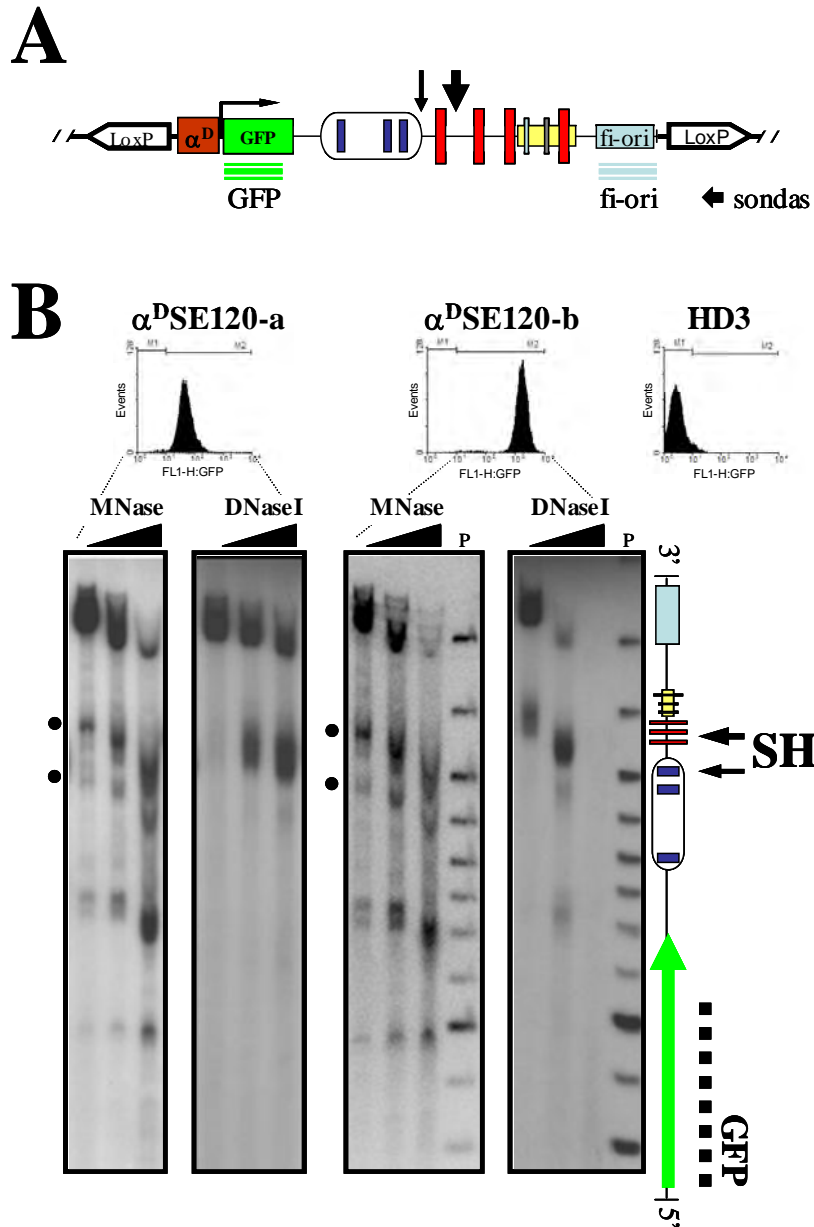
En base a la caracterización de la estructura de la cromatina del enhancer, consideramos que la región intermedia parece ser clave en la regulación de este elemento. La presencia de dos elementos fundamentales de interacción al factor GATA-1, no parece ser mera coincidencia. Por lo mismo decidimos enfocar nuestro estudio en esta región. Como punto de partida se decidió realizar una transfección del ADNc de GATA-1 y estudiar por un ensayo de accesibilidad a la MNasa y a la DNasa I las alteraciones que podría estar generando a nivel de la estructura de la cromatina esta proteína (dato no mostrado), desafortunadamente, los resultados no fueron claros<sup>9</sup>. Por lo mismo decidimos recurrir a un sistema de recombinación sitio específico (Fig. 50).

<sup>9</sup> Ahora me doy cuenta que el principal problema fue el haber realizado el estudio en la línea 6C2 y no a la no eritroide DT40. Lo anterior en base a que en la línea 6C2 los eventos de remodelación GATA dependientes ya se iniciaron y las diferencias obtenidas por una sobre expresión no son cuantificables.



**Figura 50.** Esquema que muestra como funciona el protocolo denominado intercambio de cassette mediado por la recombinasa Cre (RMCE). En la parte superior me muestra el vector que inicialmente fue integrado al azar en la línea celular HD3, este contiene dos sitios de recombinación (loxP), el cassette de resistencia a higromicina (Hy) y el cassette de selección con la timidin cinasa (Tk). Una vez seleccionadas las clonas positivas a higromicina, se genera el recambio de cassette co-transfectando el vector por recombinar (izquierda se muestra el de la GFP, derecha se muestra el utilizado en este estudio) con el ADNc de la recombinasa Cre. La selección de las clonas se realiza por un tratamiento con ganciclovir, el cual en las líneas que no llevaron acabo la recombinación será fosforilado por la Tk y se convertirá en un compuesto toxico para la célula. Este intercambio de cassette tiene la particularidad de permitirnos integrar cualquier construcción siempre en el mismo sitio de integración, es decir, en el mismo entorno cromatínico.

Este sistema tiene la particularidad de permitirnos transfectar de manera estable distintas construcciones en el mismo sitio de interacción. En particular, nos permite comparar diferentes construcciones en el mismo entorno cromatínico, ya que el efecto al que estén sometidas, ya sea positivo o negativo, siempre será el mismo para cada construcción. A su vez y de particular interés para nuestro análisis, nos permite eliminar al marcador de selección, el cual podría alterar los resultados, ya que estos elementos por sí solos podrían estar abriendo la cromatina (ver mas adelante).



**Figura 51.** Validación del modelo de estudio. **A-** Esquema mostrando la construcción que fue integrada. Contiene los sitios de recombinación (*loxP*), el promotor eritroide  $\alpha^D$ , el gen de la proteína verde fluorescente (*GFP*) como reportero, la construcción silenciador/enhancer/120 (en todas las integraciones esta es la una región que varía) y un fragmento del origen de replicación bacteriano *fi-ori*, el cual junto con la región de la *GFP*, fueron utilizadas para diseñar las sondas y poder distinguir la construcción integrada, del dominio endógeno. **B-** Arriba citometría de flujo (FACS) de las líneas analizadas. Se muestra en cada una la región que fue integrada.  $\alpha^DSE120$ , región endógena. HD3 control de expresión negativo. En la parte inferior, análisis del posicionamiento nucleosomal (MNase) y de la generación de SH a la DNase I. Triángulos negros, concentración de la enzima indicada agregada en forma creciente. SH observados (puntos y flechas negras), línea negra punteada, sonda GFP. P, peso molecular de 100 pb. Con los sufijos -a y -b, diferencian clones independientes que contienen la misma construcción.

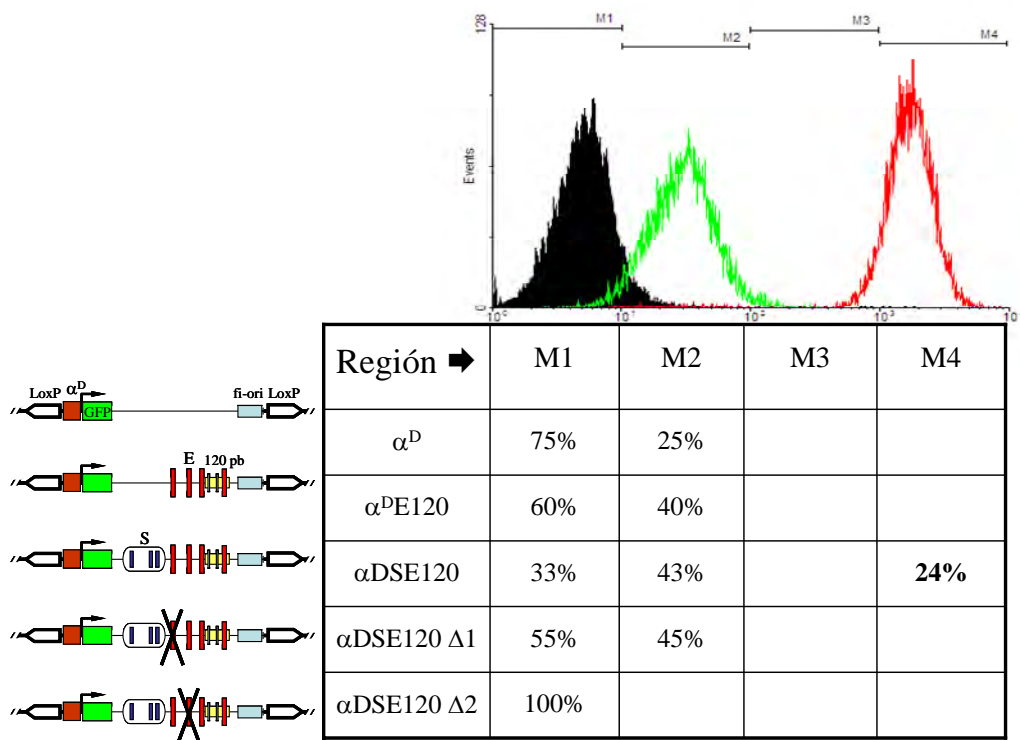
Para validar la utilización de esta técnica, necesitábamos mostrar que se reproducía el patrón de posicionamiento nucleosomal endógeno. Un posicionamiento nucleosomal diferente, no nos permitiría analizar de manera análoga la relación entre la estructura de la cromatina y la actividad del enhancer. En el establecimiento de la técnica, decidimos generar tres líneas estables (nombradas RMCE-HD3-13, -14 y -20) con el sistema de recombinación integrado de manera aleatoria (ver material y métodos). Posteriormente, se generó el intercambio de cassette con la construcción que incluye la región endógena de 1036 pb (pBL1SE4) (Fig. 51). La determinación de los niveles de expresión para las clonas obtenidas, mostró que el enhancer incrementa la actividad de nuestro reportero (Figs. 51 y 52), lo anterior significa que es funcional. Se caracterizó la accesibilidad a la DNasa I y se determinó el posicionamiento nucleosomal por digestiones nucleares con la MNasa (Fig. 51). Demostramos la presencia de un posicionamiento similar al endógeno, caracterizado por dos SH en la región intermedia (Fig. 51, flechas negras) donde de la misma manera que se observó en la determinación del posicionamiento nucleosomal endógeno (Figs. 37 y 40) y ahora resulta de vital importancia resaltar ese resultado para entender los resultados que se mostraran en la figura 53, el SH que se encuentra entre las dos cajas GATA (comparar Figs. 40 y 51, flecha negra mas grande) muestra una mayor intensidad tanto en su forma endógena como en los transgenes. De manera apenas perceptible, se alcanzan a ver los cortes generados en la región de 120 pb (asteriscos) y el nucleosoma localizado en la región del silenciador (flechas blancas)<sup>10</sup>. Finalmente, la obtención de los SH observadas al corte por la DNasa I (flechas negras), terminan por validar nuestro modelo de estudio. En conclusión, la estructuración en cromatina del enhancer  $\alpha$ -globina (a similitud del endógeno) en un contexto cromatínico diferente al endógeno, nos indica que este posicionamiento posiblemente sea necesario para la función de este elemento.

Una vez establecido el sistema, se realizó el intercambio de cassette utilizando las construcciones que presentan la delección del silenciador y las mutaciones puntuales en los sitios de interacción a GATA-1 en las líneas RMCE-HD·13, -14 y -20. El objetivo inicial fue obtener las clonas estables para cada una de las construcciones a analizar, en cada uno de los tres sitios de integración. Desafortunadamente por lo complicado que resultó el sistema, no se logró obtener clonas en los tres sitios de integración para todas

---

<sup>10</sup> El nucleosoma posicionado en la región río abajo del enhancer se observo por la hibridación con la sonda fi-ori la cual hibrida río abajo del enhancer (no mostrados).

las construcciones, sin embargo las clonas obtenidas, fueron suficientes para obtener conclusiones reproducibles.



**Figura 52.** Participación de los sitios de interacción a GATA-1 en la actividad del enhancer. En la parte superior, esquema mostrando las regiones analizadas por citometría de flujo, Izquierda esquema a similitud del presentado en la figura 50 de las construcciones analizadas. Con cruces negras se marcan las cajas GATA-1 que fueron mutadas. Derecha, cuadro mostrando el porcentaje promedio de expresión de todas las clonas analizadas.

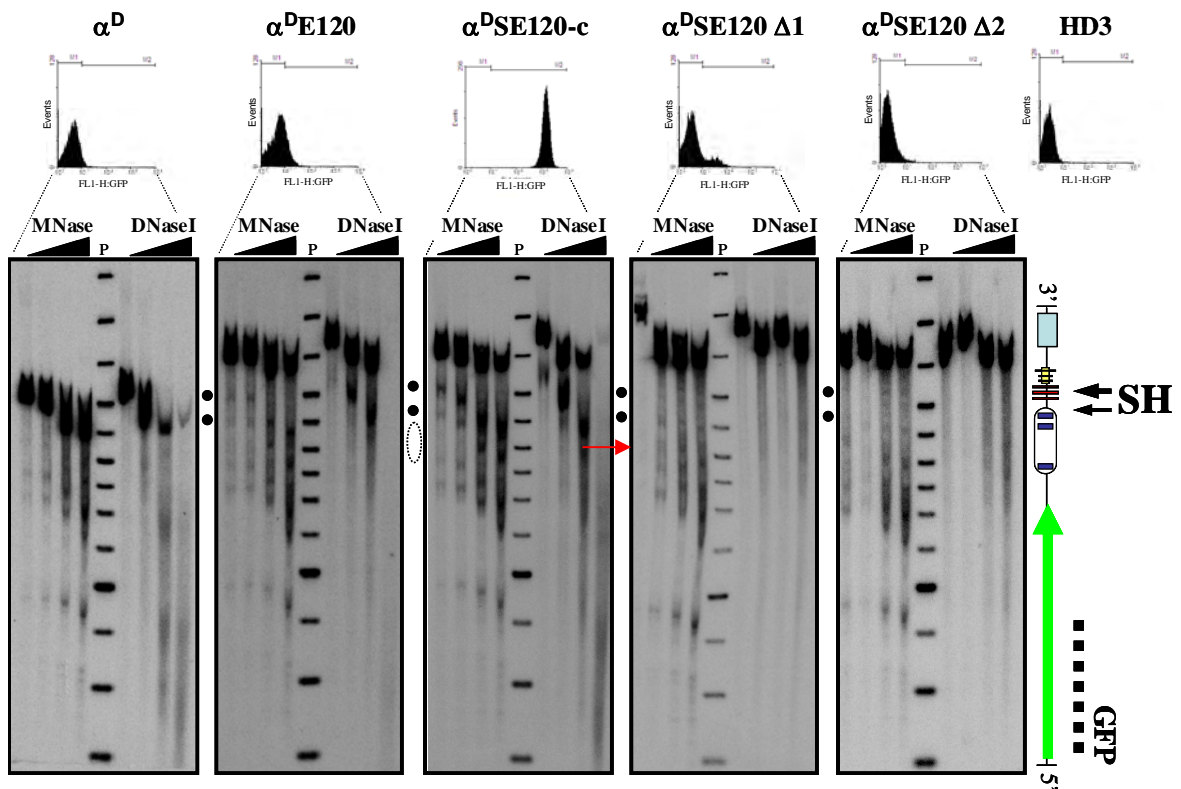
Con el objetivo de reducir al máximo cualquier tendencia en nuestros resultados, se modificó la forma mediante la cual se seleccionaron las clonas positivas. En lugar de separar cada clona por citometría de flujo (FACS), las células que mostraran niveles de expresión de la *GFP* (y generando una tendencia a separar solo clonas positivas a la expresión de *GFP*), se separaron aproximadamente 800 clonas que resistieron el tratamiento con Ganciclovir<sup>11</sup> (ver material y métodos) y vía la técnica de southern blot, se seleccionaron a las clonas que habían realizado el intercambio de cassette de manera exitosa. De todas las clonas obtenidas, se determinó el porcentaje de expresión por citometría de flujo y la recopilación de los resultados se muestra en la figura (Fig. 52). Nos resulta de particular interés el que de todas las clonas obtenidas, las únicas que mostraron altos niveles de expresión son las que contienen la construcción silvestre

<sup>11</sup> El Ganciclovir es un compuesto que al ser fosforilado por la timidin cinasa (la cual esta incluida en el vector de recombinación-figura RMCE-), se vuelve toxico.



(Fig. 52, construcción  $\alpha$ DSE120). Estos resultados apoyan un modelo de “reostat” para el funcionamiento del enhancer.

Adicionalmente, a diferencia de los resultados obtenidos *in vitro* (Fig. 23), el enhancer sin la región del silenciador no es funcional, lo anterior significa que el silenciador en el contexto de la cromatina resulta una parte integral del enhancer. Este resultado será discutido más ampliamente en las conclusiones.



**Figura 53.** Participación de los sitios de interacción a GATA-1 en la estructuración de la cromatina. Arriba citometría de flujo (FACS) de las líneas analizadas. Se muestra en cada una la región que fue integrada.  $\alpha$ D, promotor con el reportero de la verde fluorescente,  $\alpha$ DE120, igual que  $\alpha$ D más la región del enhance,  $\alpha$ DSE120, igual que  $\alpha$ D más la región endógena del silenciador/enhancer120, el sufijo -c se refiere a que esta es la clona c, diferente a las mostradas en la figura 50.  $\alpha$ DSE120  $\Delta$ 1 y  $\alpha$ DSE120  $\Delta$ 2, igual que  $\alpha$ DSE120 pero con los sitios de interacción a GATA-1 G1 y G2 (respectivamente) mutados. HD3 control de expresión negativo. Abajo, análisis del posicionamiento nucleosomal (MNasa) y de la generación de SH a la DNasa I. Triángulos negros, concentración de la enzima indicada agregada en forma creciente. SH observados (puntos y flechas negras), línea negra punteada, sonda GFP. Ovalo punteado, nucleosoma fijo, flecha roja, muestra la banda que se genera cuando se pierde el posicionamiento del nucleosoma. P, peso molecular de 100 pb.

Para determinar el efecto que a nivel de la estructura de la cromatina presentaban la eliminación del silenciador y las mutaciones puntuales de las cajas GATA y así poder determinar la correlación que existe entre la estructura de la cromatina y la actividad de enhancer, se realizaron los ensayos de digestión en núcleos con la MNasa y la DNasa I en las clonas obtenidas por RMCE (Fig. 53). Estos resultados muestran que la construcción endógena ( $\alpha^D$ SE120) presenta niveles elevados de expresión del gen reportero, su estructuración en cromatina asemeja la endógena caracterizada por la presencia de los SH a la DNasa I (flechas negras), los SH a la MNasa (puntos negros) y un nucleosoma bien posicionado en la secuencia del silenciador. Al eliminar la secuencia correspondiente al silenciador (construcción  $\alpha^D$ E120), se observa la pérdida en los niveles de expresión del gen reportero, el posicionamiento nucleosomal característico del enhancer se ve parcialmente alterado, mostrando el 100% de las clonas analizadas que no contienen al silenciador, una disminución importante en la intensidad de los SH (característico) localizados entre los dos sitios GATA-1 del enhancer (comparar las construcciones  $\alpha^D$ E120 y  $\alpha^D$ SE120). También se observa el deslocalizamiento del nucleosoma que se encontraba fijo sobre el silenciador (óvalo blanco). Finalmente, los SH a la DNasa I no se pierde, este resultado lo atribuimos a que el sitio de interacción al factor GATA sigue presente y por lo mismo el factor GATA-1 al seguir interaccionando genera el SH.

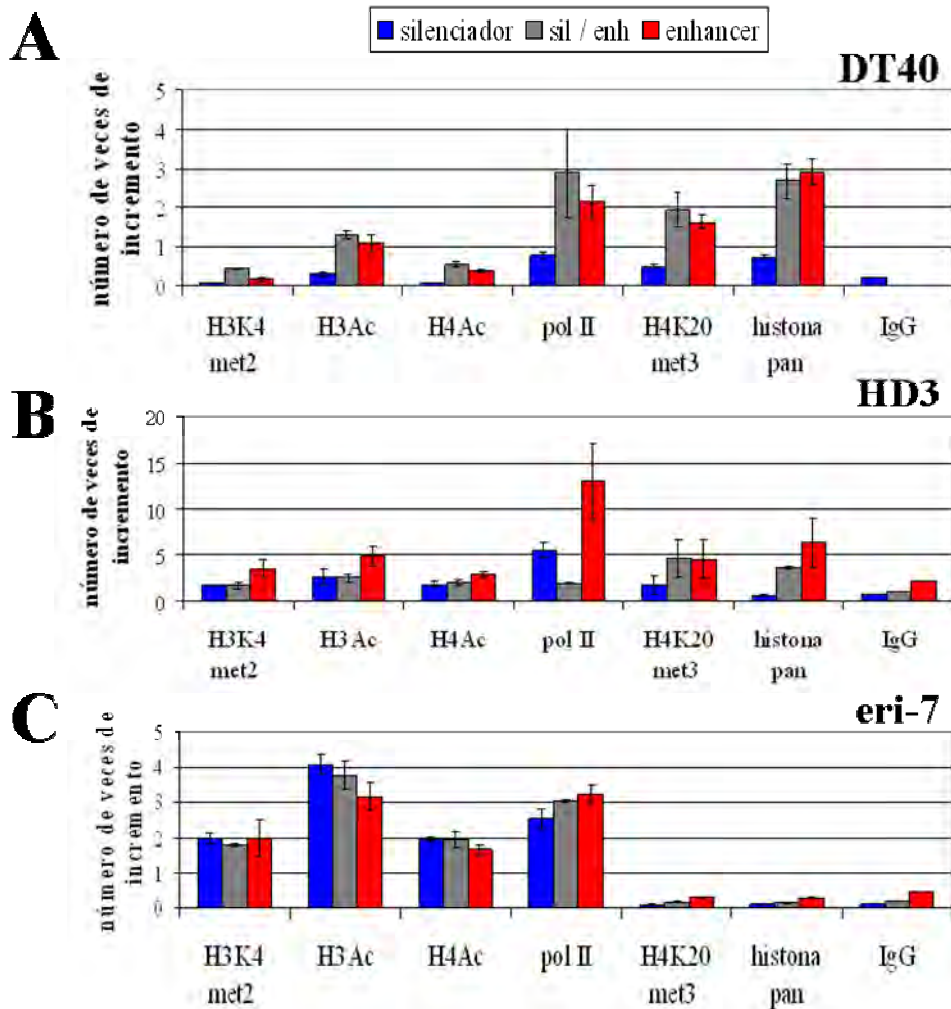
El resultado más relevante, se origina a partir de la mutación de los sitios de interacción con el factor eritroide específico GATA-1 ( $\alpha^D$ SE120 $\Delta$ 1 y  $\alpha^D$ SE120 $\Delta$ 2). En ambos casos se observa la pérdida total en la expresión del gen reportero, la pérdida en el patrón nucleosomal eritroide específico y adquirido a lo largo de la diferenciación, la pérdida total a los SH a la DNasa I y el deslocalizamiento del nucleosoma localizado en el silenciador. Este resultado muestra de manera complementaria con las transfecciones transitorias (Fig. 23). Que el factor transcripcional GATA-1 tiene un papel principal en la función del enhancer, posiblemente reconociendo su secuencia blanco en el complejo de cromatina cerrada. Esta particularidad nuestros resultados sugieren una función de activador y centro de reclutamiento de las actividades necesarias para remodelar la estructura de la cromatina del enhancer. Con una visión más clara de lo que el factor transcripcional GATA-1 podría estar haciendo en la regulación del enhancer, no me parece casual el que la mayoría de los elementos eritroide específicos presenten blancos de interacción a esta proteína.

## **Modificaciones post-traduccionales de las histonas**

El estudio de las modificaciones post-traduccionales de las histonas vía la técnica de ChIP, representa una estrategia experimental indispensable para comprender los procesos que se regulan a nivel de la cromatina. En particular, los patrones de modificaciones de las histonas pueden evidenciar la función o el estado (activo o inactivo) en que se encuentran los elementos analizados, adicionalmente vía esta técnica se puede determinar la interacción de otras proteínas al ADN. El estudio de las MPT de las histonas en la región silenciador enhancer120 (Fig. 54), encontramos en la línea celular HD3 un patrón de modificaciones positivas y negativas mezclado, este tipo de patrones han sido descritos como patrones intermedios (ver conclusiones), adicionalmente podría ser el resultado de una heterogeneidad de expresión en las poblaciones analizadas. En eri-7 encontramos marcas de cromatina activa en concordancia con la actividad transcripcional propuesta para este elemento.

Adicionalmente, en células HD3 encontramos un enriquecimiento importante de la RNA pol II cuando utilizamos los oligonucleotidos que amplifican en la región del enhancer (Fig. 54, en la figura 56 A se muestran los oligonucleotidos utilizados en este estudio), este enriquecimiento resulta coincidente con la región donde se localiza el segundo sitio de interacción a GATA-1, muy cercano a la secuencia de 120 pb donde se encuentran los sitios de interacción a EKLF, NF-E2 y un cuarto sitio de interacción a GATA-1, las últimas dos proteínas han sido previamente involucradas en el reclutamiento de la RNA pol II a los promotores B-globina (118), por el tipo de estudio realizado por nosotros no podemos descartar la participación de estas dos proteínas en el reclutamiento de la RNA pol II (Fig. 54, HD3). Este enriquecimiento baja cuando analizamos el estadio donde el enhancer se encuentra expresando a los genes globina (Fig. 54, eri-7). Este resultado está en concordancia con un modelo de “tracking” para el funcionamiento del enhancer, donde como se ha mostrado anteriormente, cuando el enhancer HS2 activa a su promotor blanco, la polimerasa vía su encarrilamiento en la fibra de ADN, se re-localiza a la región el promotor (166). Como control se utilizó la línea DT40, la cual muestra marcas de cromatina cerrada, entre ellas la presencia del octámero de histonas (histona-pan, ver material y métodos) el cual de manera muy interesante y en apoyo a uno de los escenarios que previamente planteamos, vemos como se va perdiendo conforme avanzamos en la diferenciación. Este resultado está

apoyando un modelo en el cual en el enhancer se está generando una región libre de nucleosomas.

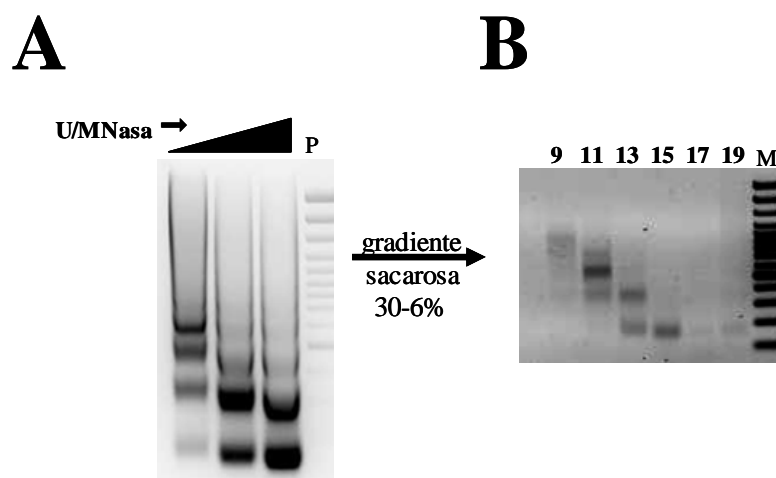


**Figura 54.** Determinación de las modificaciones post-traduccionales de las histonas y determinación de una posible región libre de nucleosomas. Gráficas mostrando el promedio de abundancia relativa de 3 ensayos de amplificación cuantitativa con los oligonucleotidos mostrados en la figura 55A, para las marcas señaladas en la región del silenciador/enhancer 120. H3K4me2 (histona 3 lisina 4 di-metilada), H3ac (histona 3 acetilada), H4ac (histona 4 acetilada), pol II (RNA polimerasa II), H4K20me3 (histona 4 lisina 20 tri.metilada), histon pan (anticuerpo que reconoce cualquiera de las 5 histonas) e IgG (anticuerpo inespecífico IgG de conejo).

### Regiones libres de nucleosomas

Para confirmar si el enhancer se encontraba en una región libre de nucleosomas, se realizó un estudio de protección de templado (Fig. 56 B y C), el cual tiene como fundamento el que a una digestión intensiva con la MNasa, solamente las regiones que

se encuentren protegidas por nucleosomas, no serán digeridas y por lo anterior pueden ser amplificadas por PCR. Como se puede ver en la figura 55, se realizó la digestión intensiva con tres concentraciones distintas de MNasa<sup>12</sup>, se purificó por un gradiente de sacarosa la fracción perteneciente a mono- y di-nucleosomas y se realizó una PCR duplex utilizando una región de heterocromatina como control para normalizar (Fig 56 B). La determinación inicial de la protección que la estructura de la cromatina estaba generando al templado de ADN en la región silenciador y enhancer (Fig. 56), muestra un efecto de pérdida en la protección del templado conforme avanzábamos en la diferenciación.

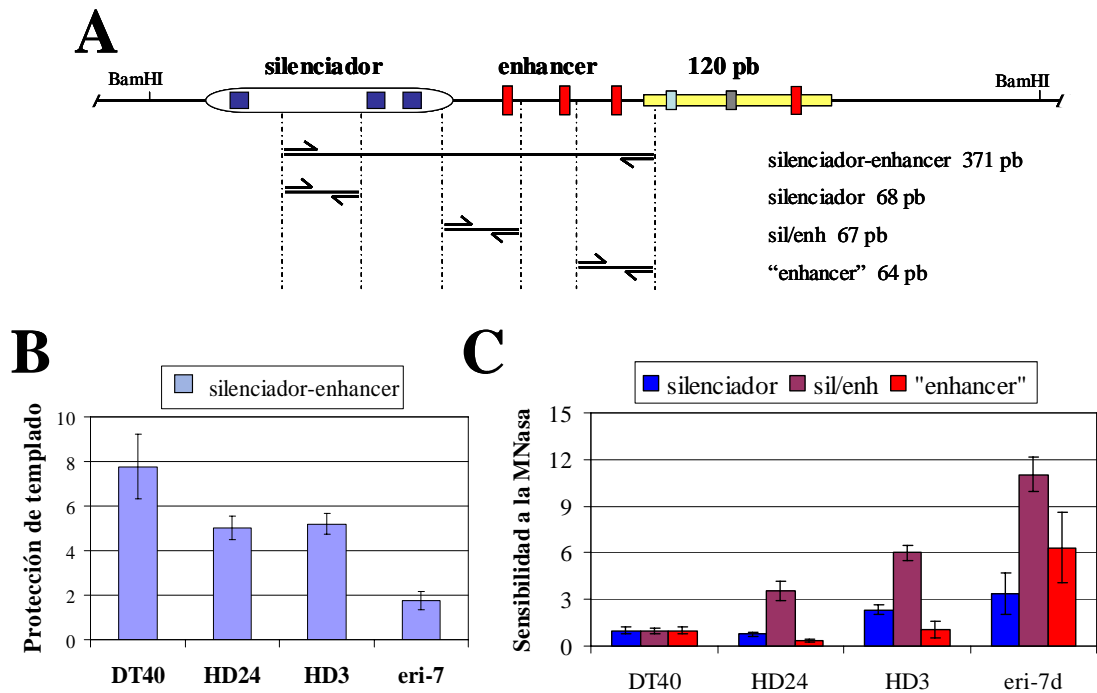


**Figura 55.** Aislamiento de mono- y di-nucleosomas. **A**, Gel de agarosa donde se ejemplifican tres digestiones extensivas con la MNasa. **B**, Una vez digerida la cromatina, se purifican mono- y di-nucleosomas por un gradiente de sacarosa al 30-6%.

Posteriormente, con la intención de determinar si localmente la región silenciador enhancer se encontraba estructurada en nucleosomas o se había perdido la interacción ADN-nucleosomas, se diseñaron tres juegos de oligonucleotidos, los cuales van a amplificar pequeños fragmentos de ADN (de entre 68 y 74 pb) localizados justo al interior de las regiones que teóricamente podrían estar ocupadas por un nucleosoma, por lo mismo si el nucleosoma se encontraba presente, el fragmento de ADN sería protegido, pero si el nucleosoma no se encontrara, el templado sería cortado por la nucleasa. En la figura 56 B se puede observar como en la región localizada entre el enhancer y el silenciador (oligonucleótidos sil/enh) claramente se va incrementando la

<sup>12</sup> Al no conocer en relación al genoma que tan accesible se encuentra la región por analizar, se hacen tres digestiones con diferentes concentraciones de la MNasa, aumentando la probabilidad de que la región de estudio se encuentre en estado de mono- y di-nucleosomas en alguna de ellas.

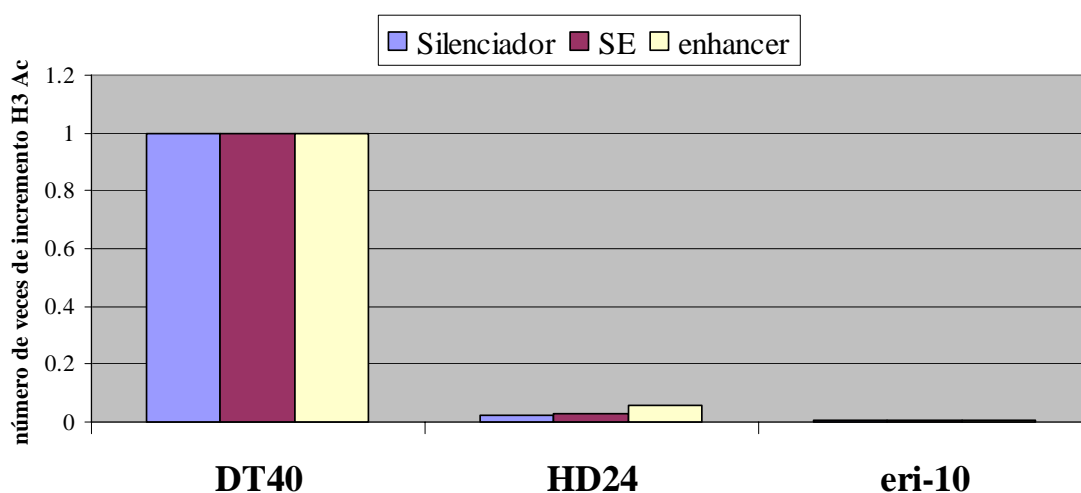
sensibilidad, sinónimo de pérdida de nucleosomas, las regiones del silenciador y el enhancer (oligonucleótidos silenciador y enhancer, respectivamente) muestran una menor sensibilidad al corte de la MNasa, resultado de una mayor estabilidad del complejo ADN-nucleosoma. Este resultado nos invita a proponer un modelo en el cual la región del silenciador/enhancer120, se encuentra estructurado en un complejo ADN-histona poco estable.



**Figura 56.** Determinación de la presencia de nucleosomas en la región del enhancer. **A-** Esquema de la región del silenciador/enhancer con la localización de cada uno de los pares de oligonucleótidos que se utilizaron para este estudio. **B-** La gráfica representa el promedio de tres amplificaciones cuantitativas sobre ADN genómico (oligonucleótidos Sil/enh) purificado de los mono- y di- nucleosomas aislados. Mostrando la protección que muestra la región silenciador/enhancer al corte por la MNasa en 4 estados de diferenciación **C-** La gráfica representa el inverso del promedio de tres amplificaciones cuantitativas sobre ADN genómico (oligos silenciador, SE y enhancer) purificado de los mono- y di-nucleosomas aislados. Mostrando la sensibilidad diferencial que muestra la región silenciador/enhancer al corte por la MNasa en 4 estados de diferenciación

Resulta importante resaltar que en un inicio los experimentos de ChIP se intentaron realizar por PCR en tiempo real, pero mostraron resultados no interpretables, como el que se muestra en la figura 57. En un principio se pensó que existía un problema en el diseño del experimento, ahora este resultado apoya nuestros resultados y un modelo en el cual la región del silenciador/enhancer se está generando una región

desprovista de nucleosomas y por lo mismo el templado al encontrarse desprotegido, es digerido por la nucleasa, reflejado en los bajos niveles de amplificación.



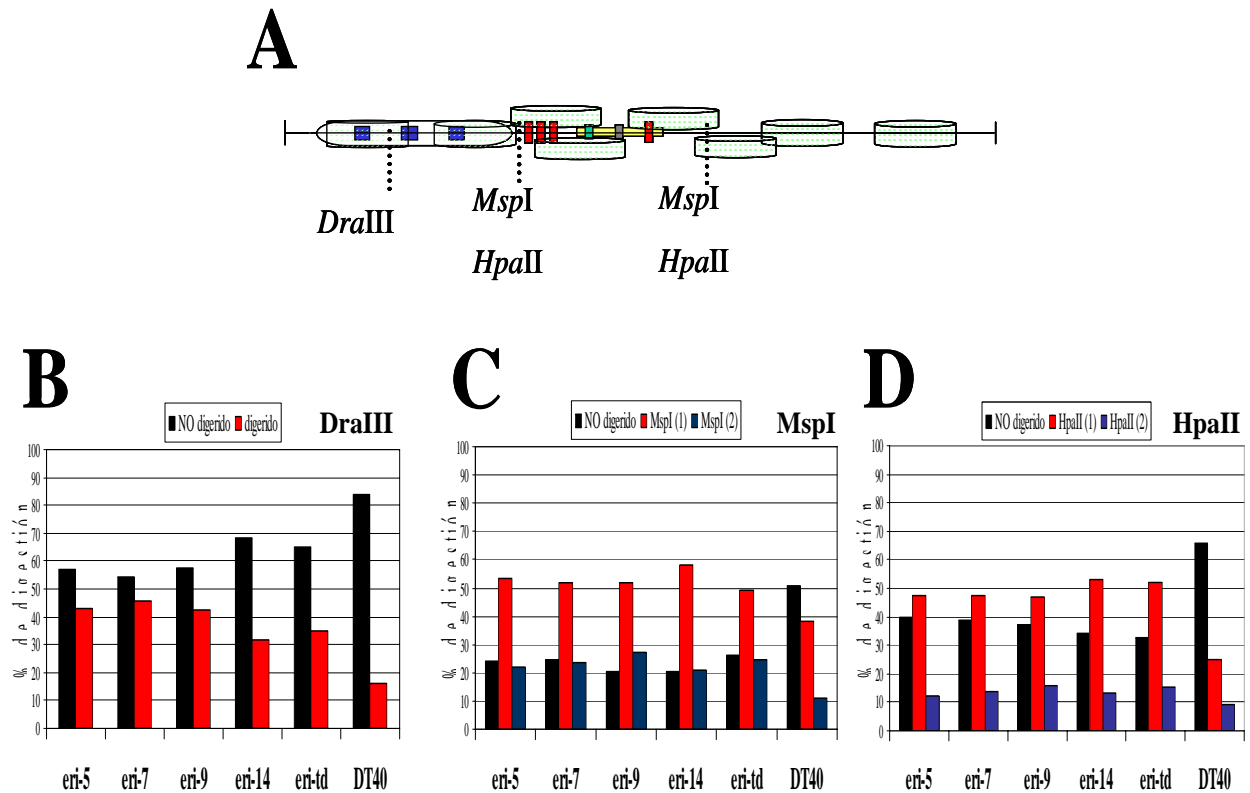
**Figura 57.** Determinación de la acetilación de la histona H3 por PCR en tiempo real en la región silenciador/enhancer 120. Gráficas mostrando la abundancia relativa de la acetilación de la histona H3, determinada por la amplificación por PCR en tiempo sobre mono- y di nucleosomas aslados por gradientes de sacarosa.

La exposición de las secuencias del enhancer al corte por nucleasas se incrementa en los estadios donde los genes globina son activos, consistente con la estructuración en cromatina activa. En estados tardíos del desarrollo de pollo, observamos que la estructura de la cromatina se re-estructura en una cromatina compacta. El estudio de los mecanismos que generan la represión de las secuencias que presenta el enhancer, resulta un tema importante en el entendimiento de los mecanismos implicados en la regulación de un enhancer.

### **Metilación del ADN en el fragmento de 1036 pb**

En el dominio  $\alpha$ -globina se ha reportado que la metilación del ADN estaría jugando un papel importante en la represión del gen embrionario. El enhancer  $\alpha$ -globina muestra un porcentaje del 59% en contenido en Citocinas y Guaninas. Este resultado es muy interesante por que clasifica a la región en que se localizan el silenciador y el enhancer, como una isla CpG. Regiones donde la metilación de di-nucleótidos CpG podría estar participando en estadios tardíos del desarrollo. Para determinar una posible

participación de la metilación del ADN en esta región, se realizaron digestiones *in vitro* del ADN con enzimas metilosensibles (Fig. 58). Esta figura sugiere que el silenciador se estaría metilando a partir del día 14 y el enhancer solo en líneas no eritroides. En un escenario para explicar este resultado propongo, que el silenciador una vez que ya ejerció su actividad, no presentara mayor efecto sobre los genes adultos y es sujeto a metilación.

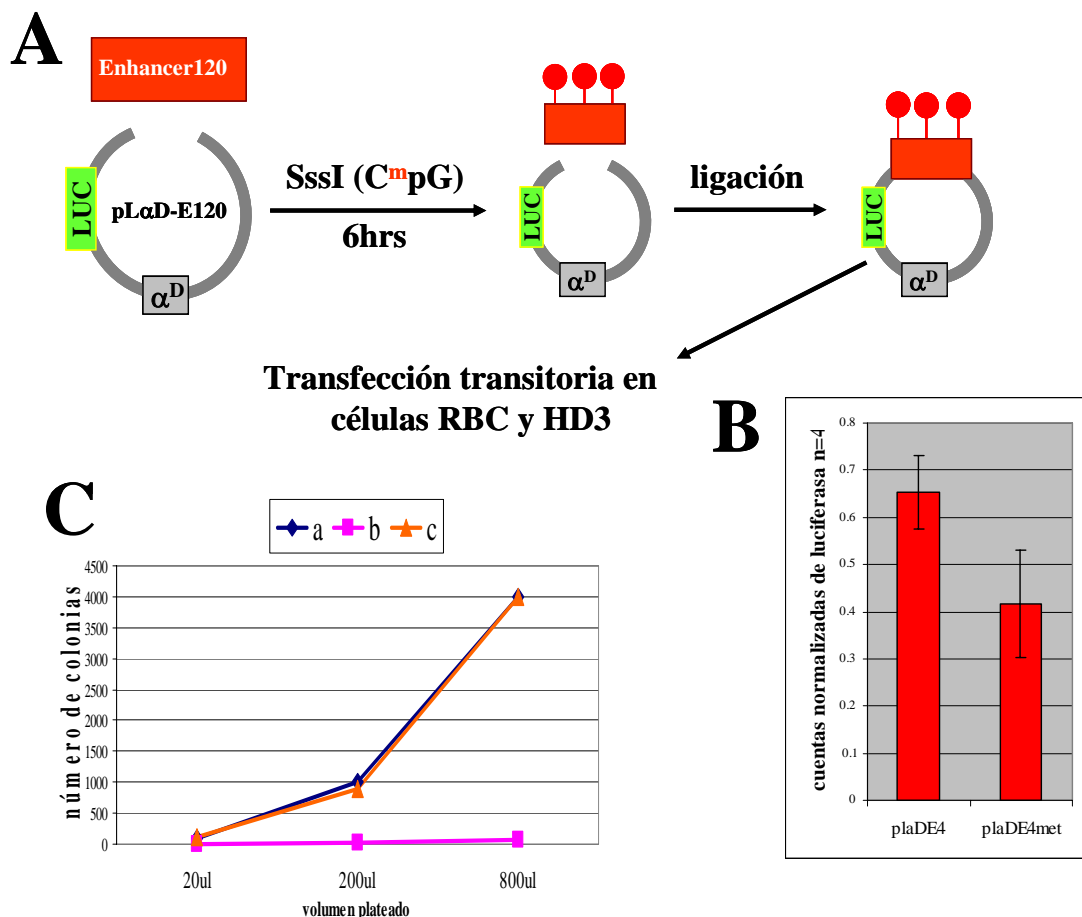


**Figura 58.** Análisis de la metilación en la región del silenciador/enhancer120. **A**, Esquema de la región del silenciador/enhancer con la localización de cada una de las enzimas utilizadas en este estudio. **B**, Análisis a través del desarrollo con la enzima *DraIII*. Barras negras, ADN no digerido. Barras cafés ADN digerido por *DraIII*. DT40 control de digestión. **C**, Análisis a través del desarrollo con la enzima *MspI*. Barras negras, ADN no digerido. Barras rojas ADN digerido por *MspI* en M1. Barras azules ADN digerido en M2. **D**, Análisis a través del desarrollo con la enzima *HpaII*. Barras negras, ADN no digerido. Barras rojas ADN digerido por *HpaII* en H1. Barras azules ADN digerido por *HpaII* en H2.

Para determinar si el enhancer o el silenciador presentaban metilación en sus secuencias, se desarrollo la técnica bisulfito de sodio. Desafortunadamente por motivos ajenos a nosotros no se logró obtener las secuencias modificadas. Por lo mismo



decidimos realizar un estudio *in vitro* para determinar si el enhancer metilado de manera artificial se veía afectado (Fig. 59). Cuando el enhancer es metilado de manera *in vitro*, los niveles del gen reportero bajan entre un 30 y un 50 %, resultado que nos invita a proponer que la metilación del enhancer podría estar participando en mantener al enhancer transcripcionalmente inactivo en un estado en el que ya no va a ser necesitado. Este resultado se ve apoyado por unos estudios que se realizaron en las clonas estables (dato no mostrado), donde las clonas que no presentaban actividad de reportero, las tratamos de reactivar con un tratamiento con TSA sin que mostraran mayor efecto, con la utilización de 5-aza-citidina<sup>13</sup> se reactivaron parcialmente las clonas, apoyando la idea de una posible implicación por parte de la metilación en la actividad de enhancer.

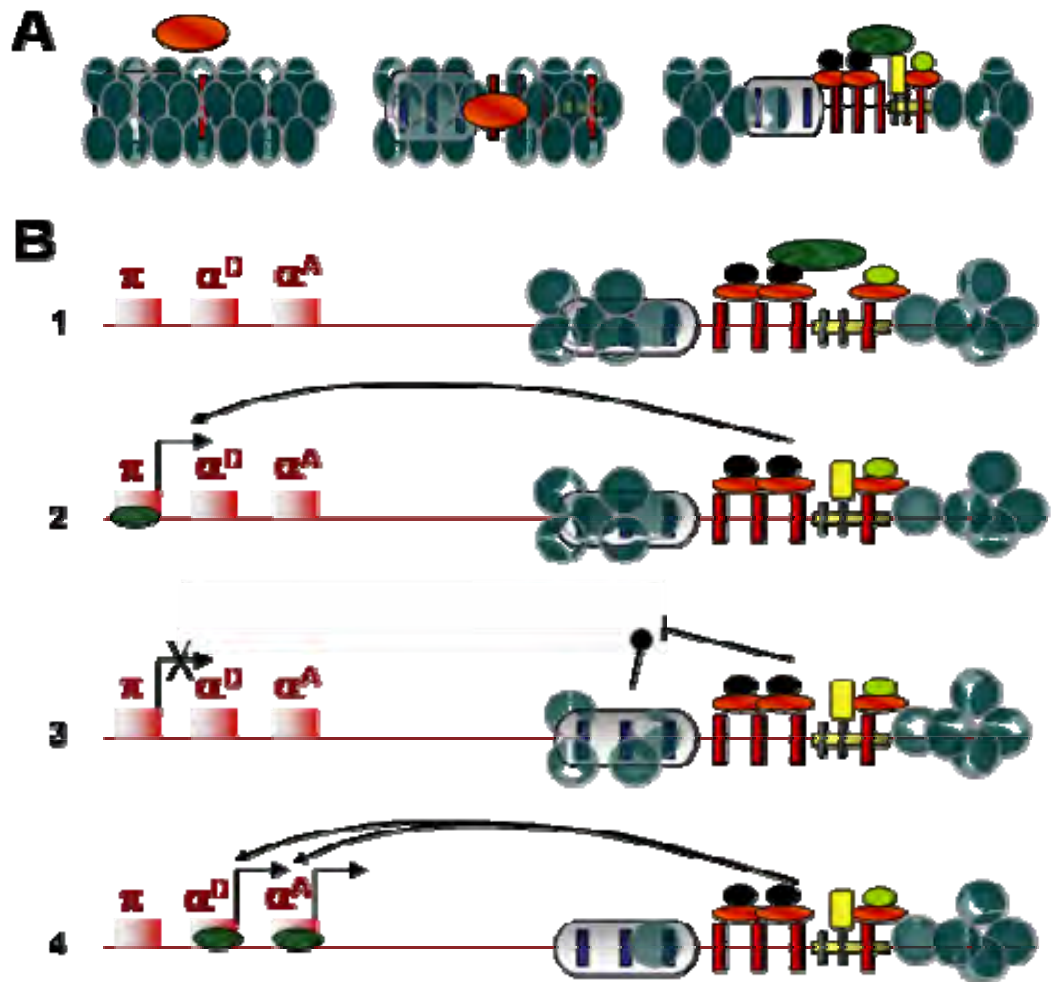


**Figura 59.** Análisis de la metilación en la región del silenciador/enhancer120. **A**, Esquema de la región del silenciador/enhancer120 con la localización de cada una de las enzimas utilizadas en este estudio. **B**, Análisis a través del desarrollo con la enzima *DraIII*. Barras negras, ADN no

<sup>13</sup> LA 5-aza-citidina es un análogo de la citosina que por una modificación química, no puede ser metilada y por lo mismo su uso muestra la pérdida gradual de la metilación.

digerido. Barras cafés ADN digerido por *DraIII*. DT40 control de digestión. C, Análisis a través del desarrollo con la enzima *MspI*. Barras negras, ADN no digerido. Barras rojas ADN digerido por *MspI* en M1. Barras azules ADN digerido en M2.

Con los resultados presentados en esta tesis y apoyados en la literatura, proponemos de manera especulativa, un modelo de funcionamiento del elemento silenciador enhancer (Fig. 60). Proponemos que el factor transcripcional GATA-1 (ya sea GATA-2 o GATA-1, Fig. 60, A), será la primera proteína en interaccionar a la región del enhancer, cuando este elemento aun se encuentra estructurado en cromatina represiva. GATA-1 vía la interacción de sus dedos de zinc en el complejo ADN-histonas, desestabilizará la interacción de la hebra de ADN al octámero de histonas (ver discusión), facilitando el acceso a otras proteínas a sus secuencias de interacción en el enhancer, a factores y co-factores transcripcionales y colaborando en la formación de un arreglo de nucleosomas posicionados con un patrón característico de las líneas eritroides y necesaria para que el enhancer pueda activar a sus genes blanco. Esta interacción se lleva de manera gradual hasta un estado previo a la activación de los genes eritroides (Fig. 60 A y B paso 1). En este momento el enhancer se encuentra estructurado en un estado semi-activo, listo para ser activado en un momento en específico y disparar la cadena de eventos necesarios para lograr la activación de los genes globina (Fig. 60 B paso 2). La activación de los genes globina podría resultar de la expresión de un factor eritroide tiempo-tejido-específico, que debería coincidir con la re-localización de la polimerasa, a los promotores globina. Proponemos que la región de 120 pb se debe encontrar íntimamente ligada a esta activación. Regulando inicialmente al enhancer negativamente vía su interacción con reguladores negativos (entre ellos YY1) y posteriormente con reguladores positivos (entre ellos remodeladores de la estructura de la cromatina). En un estadio tardío, el silenciador romperá la comunicación que existe entre enhancer y el promotor embrionario (Fig. 60 B paso 3), permitiendo ahora los genes adultos, por un modelo de competencia (Fig. 60, B paso 4), interaccionen con este elemento y sean activados. En este momento el silenciador se apagará (significado de que ya no es necesario) y estructurará en cromatina represiva. El enhancer  $\alpha$ -globina regulará la actividad de los genes adultos en estadios avanzados del desarrollo de pollo, hasta llegar a un estado donde ya no se le necesite y también se estructure en cromatina represiva.



**Figura 60.** Modelo de funcionamiento del elemento silenciador enhancer120. **A-** estructuración gradual (de izquierda a derecha) del elemento silenciador enhancer120, en una conformación activa, vía la interacción del factor transcripcional GATA-1. Círculos verdes, nucleosomas; ovalo rojo, GATA-1; ovalo verde oscuro, RNA pol II; círculos negros, verde claro y rectángulo amarillo; maquinaria transcripcional. **B-** Cinética de activación de los genes globina, de arriba abajo (con números negros se marcan cada unos de los pasos propuestos en esta cinética de activación), primero 1.- el elemento silenciador enhancer120 se estructura en cromatina activa, 2.- la expresión de alguna proteína tiempo tejido específica (rectángulo amarillo, por ejemplo p45 del complejo NF-E2) genera la activación del enhancer, la re-localización de la pol II a el promotor embrionario. 3.- el “silenciador interrumpe la comunicación que existe entre el enhancer y el promotor embrionario, para que finalmente 4.- los promotores adultos puedan competir por el enhancer.

## Discusión

Una de las preguntas fundamentales concerniente a la función de un LCR y en este caso en particular del enhancer  $\alpha$ -globina, es el mecanismo mediante el cual estos elementos regulan diferencialmente a sus genes blanco. Una de las hipótesis predice que la estructura de la cromatina, característica de estos elementos, es necesaria para su interacción funcional (1). La determinación de la estructura de la cromatina en la que se encuentra conformado el enhancer  $\alpha$ -globina a través de la diferenciación y el desarrollo y la influencia que ésta podría ejercer sobre su actividad, representan la parte central de este trabajo de tesis y resulta un aspecto importante para lograr entender los mecanismos mediante los cuales regulan a sus genes a distancia.

Para entender la participación que la estructura de la cromatina tiene en la regulación del enhancer  $\alpha$ -globina, se realizó un estudio sistemático de la estructura local de la cromatina en una serie de líneas eritroides que se encuentran arrestadas en diferentes estadios de la diferenciación (HD24, 6C2 y HD3), una serie de cultivos primarios de eritrocitos de pollo y en la línea linfóide DT40. Se realizaron experimentos de corte por la ADNasa I, donde se observó la formación gradual de un SH en la región donde se localizan el silenciador y el enhancer (Fig. 37). Este SH es casi imperceptible en la línea eritroide mielóide (HD24), su sensibilidad se va incrementando gradualmente hasta llegar al estadio de eritrocitos de 10 días de gestación (eri-10) donde la accesibilidad a la DNasa I, alcanza su mayor nivel con un incremento de 20 veces, con respecto al observado en la línea celular HD24. Este resultado coincide con los niveles de expresión y la actividad asociada del enhancer, coincidente con una remodelación gradual de su estructura de la cromatina. En la línea control DT40, no se observó la formación de ningún SH, por lo que podemos concluir que el enhancer  $\alpha$ -globina es eritroide específico.

Este remodelamiento se hace más evidente mediante un análisis por digestión con la endonucleasa micrococcal (MNasa), poniendo en evidencia la aparición inicial de un SH entre las dos primeras secuencias de unión al factor GATA-1 localizadas en el enhancer. En este caso se forma un SH similar al previamente reportado en un estudio realizado *in vitro* (3), pero en este caso en un estudio *in situ*. Una vez más, este sitio no se observa en la línea no eritroide (DT40) y tampoco es consecuencia de una región que por su contenido nucleotídico presente mayor frecuencia de corte (control de digestión sobre ADN genómico). Estos resultados, nos hablan de la generación de un SH entre el

silenciador y el enhancer de manera tejida específica. En este mismo análisis, se determinó la presencia de sitios de corte en la secuencia de 120 pb en las líneas celulares HD3 y en cultivos primarios de eri-10, mostrando una remodelación activa de la región. En conclusión, estos cambios estructurales contribuyen a modular la actividad del enhancer. Lo anterior se ve apoyado por la identificación de interacciones entre diferentes proteínas, en particular entre GATA-1 y YY1 (50). Proponemos que la interacción física de GATA-1 y YY1 en la secuencia de 120 pb del enhancer  $\alpha$ -globina, puede estar asociada a la remodelación de la cromatina. El balance entre la concentración de estas dos proteínas (y posiblemente alguna otra) y los co-factores asociados a éstas, podría ser el factor crítico para que el enhancer pase de un estado semi-activo a uno totalmente activo con consecuencias directas en los niveles de expresión de los genes  $\alpha$ -globina durante las distintas etapas del desarrollo del pollo y la diferenciación eritroide.

Por otra parte, se localizaron dos nucleosomas fijos enmarcando la secuencia asociada al enhancer (Figs. 40-42, 49), estos parecen estar delimitando y tal vez contribuyen en el proceso ligado a la remodelación de la estructura de la cromatina del enhancer. El enhancer parece presentar una estructura cromatínica muy particular, que se ve favorecida en los estadios donde los genes globina se expresan, cuando estos dejan de expresarse, en estadios tardíos del desarrollo (eri-TD) el posicionamiento nucleosomal se ve disminuido, favoreciendo la idea de una estructura cromatínica implicada directamente en la actividad del enhancer.

La remodelación de la estructura de la cromatina en el enhancer, se lleva a cabo en un estadio donde los genes globina aun no son activos transcripcionalmente. Esto implica la preparación del elemento para poder ser activado en un momento preciso del desarrollo. Esta activación se podría dar por la expresión de un factor transcripcional tiempo- y tejido-específico. Como ejemplo, en un estudio en el dominio  $\beta$ -globina de pollo, se encontró que la expresión de la proteína NF-E2p45 fomenta la relocalización del enhancer HS2 a una región de eucromatina y su activación (133). Coincidentemente, el enhancer  $\alpha$ -globina presenta en la secuencia de 120 pb un sitio de unión al factor NF-E2. En conclusión, el enhancer  $\alpha$ -globina presenta una apertura gradual de la cromatina conforme avanza la diferenciación y el desarrollo.

## **El significado de la remodelación del enhancer**

Con la intención de determinar de una forma un poco más cuantitativa si la remodelación de la estructura de la cromatina observada, significaba la apertura local de la cromatina, se realizó un estudio de accesibilidad a enzimas de restricción. Se demostró un incremento gradual a través de la diferenciación en la región completa del enhancer. Los resultados obtenidos nos llevaron a plantear la hipótesis de que la remodelación de la cromatina y el incremento en la accesibilidad a endonucleasas en la región del enhancer  $\alpha$ -globina se encuentra acompañado por una disociación completa de los nucleosomas localizados sobre de él. Nuestros datos sugieren fuertemente dicha disociación, lo cual representa una estrategia regulada posiblemente para alcanzar el mayor grado de apertura del enhancer y su óptima actividad funcional. Ahora bien, la pregunta que surge es ¿Quién o cómo se inicia la remodelación de la cromatina en la región del enhancer y por qué mecanismo los nucleosomas se disocian del ADN?

En cuanto a las secuencias asociadas al silenciador, la accesibilidad a enzimas de restricción es pobre en todos los estadios de diferenciación analizados y es sólo hasta llegar al estadio de eritrocitos donde las secuencias se mostraron expuestas al corte por enzimas de restricción, justo en el lapso de tiempo en que el gen embrionario se ve silenciado (Fig. 15, Guerrero *et al.*, en preparación). En conclusión, los elementos silenciador y enhancer presentan a nivel de su estructura de la cromatina, una dinámica de apertura de la cromatina y exposición de sus secuencias blanco diferente, concordante con su actividad diferencial a lo largo de la diferenciación eritroide y el desarrollo del pollo.

## **Mapeo de los SH y el papel que GATA podría estar realizando en el enhancer $\alpha$ -globina**

A partir de los resultados anteriores fue posible determinar la localización precisa de los SH situados entre el silenciador y el enhancer. Consideramos que este dato nos permitirá determinar qué factores proteicos pudieran estar implicados en la formación de una estructura cromatínica tan característica del silenciador y el enhancer  $\alpha$ -globina en estados eritroides, para así poder comprender la función diferencial de estos elementos. Por un estudio combinado de digestiones nucleares con la MNasa y mapeo por restricciones enzimáticas, se determinó la localización de los SH en un rango de  $\pm 10$

pb. El sitio principal se localiza entre la primera y segunda caja de unión al factor GATA-1 y el segundo entre la segunda y la tercera secuencias conservadas del silenciador (Fig. 49). La región genómica que se encuentra comprendida entre los dos SH, contiene una secuencia de unión al factor transcripcional GATA-1. Mediante un ensayo de retardo y competencia con una secuencia consenso así como con la secuencia G1 mutada y por el ensayo de ChIP, se determinó la presencia del factor transcripcional GATA-1 (Figs.47 y 48). Interesantemente, esta interacción se exhibe de manera gradual, de forma análoga al proceso de remodelación de la estructura de la cromatina asociada al enhancer. Por los datos mostrados en la literatura, parecería ser esta proteína un candidato fuerte para ser la proteína que desencadenará los eventos de remodelación del enhancer (1-3, 140, 172).

La co-transfección del ADNc de GATA-1 de pollo con las construcciones que incorporan al enhancer, mostró la aportación positiva de la proteína GATA-1 a la actividad del enhancer. Se ha reportado que la expresión transcripcional de los miembros de esta familia, se encuentra altamente restringida. Se les ha propuesto como factores con una función decisiva, que estaría controlando la transcripción programada a través del desarrollo (172, 173). GATA-1 ha sido implicada en la generación de estructuras de la cromatina activa en prácticamente todos los promotores eritroides (1-3). Por lo anterior, resultaba de particular interés determinar el papel de GATA-1 en la formación de los SH y por ende, en la actividad regulada del enhancer.

### **Papel de GATA-1 en la formación de SH**

Se ha postulado que los SH observados al corte por la DNasa I como zonas del genoma donde la estructura de la cromatina se encuentra más laxa, favoreciendo la unión de activadores y co-activadores (90), otros grupos se refieren a ellos como regiones libres de nucleosomas (240). Alternativamente se ha propuesto que la interacción cooperativa de factores transcripcionales al complejo ADN-nucleosoma genera SH a la DNasa I (3). Se sugiere que parte de la energía requerida para disociar las interacciones electrostáticas entre el ADN y el nucleosoma, podrían ser compensadas por un nuevo tipo de interacciones entre el ADN y los residuos básicos de ciertos factores de transcripción. El grupo de Felsenfeld (3) mostró la unión *in vitro* del factor transcripcional GATA-1 a un nucleosoma, generando el complejo proteína-ADN-nucleosoma de manera muy estable. La unión de GATA-1 a este nucleosoma genera la

disrupción entre el ADN y el nucleosoma en diferentes regiones, pero no la disrupción total. Cuando analizan el patrón de corte con diferentes enzimas, encuentran un patrón de digestión similar al visto cuando GATA esta unida a ADN libre de nucleosomas. Por lo anterior, GATA-1 podría ser el factor que desencadena la serie de eventos necesarios para que el enhancer se remodele y este listo para ser activado en un tiempo específico.

Por otra parte, en un estudio en células eritroleucémicas de ratón por tranfecciones estables, se demostró que la integridad de los sitios de unión a los factores GATA-1 y NF-E2 es necesaria para la formación de los SH en el LCR  $\beta$ -globina humano (1). Sin embargo quedó abierta la pregunta de si un factor transcripcional de manera individual es suficiente para generar SH (2). Por lo anterior, se decidió determinar la función que el factor transcripcional GATA-1 tiene en la actividad del “core” del enhancer y su posible participación en la generación de los SH. Tomando ventaja de la técnica de RMCE previamente reportada (312), se decidió realizar la mutación puntual de los sitios de interacción al factor GATA-1 (Fig. 53). Con la intención de entender si la formación de los SH observados en el enhancer se están formando estocasticamente por la adición consecutiva de factores transcripcionales según el modelo propuesto por el grupo de Lowrey (1), se decidió deletar la región que comprende al silenciador, la cual por ensayos de retardo muestra la interacción de diferentes factores proteicos.

La integración de la construcción que contiene a los elementos silenciador y al enhancer, en un sitio de integración aleatorio, presentó un posicionamiento nucleosomal similar al endógeno en todas las colonias analizadas. Los SH generados al corte por la DNAsa I también se observaron. Esto significa que la región escogida para realizar el estudio de los elementos silenciador y enhancer contiene todos los elementos necesarios para que este elemento se estructure en cromatina activa y sea funcional.

El análisis de las diferentes construcciones estudiadas por la técnica de RMCE, mostró que la delección de la región del silenciador, a diferencia de lo observado *in vitro*, imposibilita al enhancer para activar al promotor. Se observó una pérdida en la intensidad del SH generado entre los sitios de interacción a GATA-1, en el análisis de posicionamiento nucleosomal (Fig. 53). Este sitio de hipersensibilidad resulta ser el primero en formarse durante el proceso de remodelación del enhancer, interesantemente los SH generados al corte por la DNAsa I no se pierden. Proponemos que la pérdida de las secuencias asociadas al silenciador, causan la pérdida del nucleosoma que se encuentra posicionado sobre él y por lo anterior se observa una pérdida parcial del



posicionamiento nucleosomal. En concordancia, la delección de la región del silenciador mostró la aparición de un SH a la MNasa justo en medio de la región donde se localizó al nucleosoma fijo, mostrando el cambio en el posicionamiento nucleosomal. El hecho de que no se pierda o disminuya el SH generado por la DNasa I en la región del enhancer, proponemos se debe a que al no perderse la integridad de los sitios de interacción al factor GATA-1, este factor puede seguir interaccionando y generando los SH, apoyando un modelo en el cual la formación de los SH no es estocástico.

De los resultados anteriores podemos concluir que la capacidad de formar los SH esta comprendida en el “core” del enhancer y que parte de su capacidad de activar a su promotor blanco, depende de la región que hasta el momento ha sido nombrada como silenciador y que ahora sugerimos es parte integral del enhancer. Lo anterior no parece ser tan ilógico dado que la definición clásica del enhancer asocia a estos elementos de regulación con una organización modular que conlleven tanto funciones positivas como negativas. De manera muy particular, cuando separamos al silenciador del enhancer, por transfecciones transitorias encontramos que la región denominada como el elemento silenciador cumplía con todas las características de un silenciador clásico deberá de tener. Cuando ensayamos al silenciador y al promotor por separado, se observó un pequeño efecto positivo, con este resultado proponemos dos posibles escenarios: Uno, el silenciador es parte integral del enhancer, presentando este último actividades negativas y positivas o dos, el silenciador está mal delimitado.

Si la propiedad de formar SH se encontraba en el core del enhancer, es posible que se debiera a la presencia de los sitios de unión al factor GATA-1, por lo que se realizaron las mutaciones puntuales sobre los tres sitios de interacción a GATA-1 localizados en el “core”. El análisis *in vitro* mostró que la mutación tanto en los sitio G1 o G2 era suficiente para disminuir la actividad de reportero en un 80%, efecto antes no obtenido para este elemento. La tercera secuencia de interacción a GATA-1 no presentó ningún efecto.

El estudio *in situ* de manera relevante y por primera vez se encontró que la presencia de los sitio de SH a la DNasa I dependen de la presencia exclusiva del sitio de unión al factor transcripcional GATA-1. Este resultado se reprodujo en diferentes sitios de integración obteniéndose el mismo resultado; la pérdida total de los SH a la DNasa I y una deslocalización del posicionamiento nucleosomal. Lo anterior significa que el posicionamiento nucleosomal característico, la generación de una cromatina activa y la

actividad del enhancer, dependen en gran medida de los sitios de interacción al factor transcripcional GATA-1. Este resultado, en adición con las múltiples funciones en que esta proteína ha sido implicada, se acopla a la función que un regulador maestro de la hematopoyesis estaría desempeñando.

Del 100% de las clonas analizadas, solo presentaron actividad de reportero aquellas en que el enhancer se encontraba en su forma silvestre. Estos resultados apoyan un modelo progresivo. Este modelo propone que un enhancer podría estar incrementando la actividad del promotor. Alternativamente se propuso el modelo de “prendido o apagado”, en el que el enhancer estaría incrementando la probabilidad de que un promotor se active. Este último escenario fue descartado al no encontrarse ninguna clona que en ausencia del enhancer, presentara niveles elevados de reportero. En base a lo anterior proponemos que el funcionamiento del enhancer  $\alpha$ -globina de pollo, se acopla a un modelo progresivo. Lo anterior significa que está afectando los niveles de la transcripción y no el número de células en que se está expresando el reportero.

### **Las particularidades en la estructura de la cromatina del enhancer $\alpha$ -globina**

En años recientes se ha demostrado que patrones determinados de modificaciones de las histonas pueden dar referencia al tipo de funciones transcripcionales en un sitio determinado del genoma. Con el objetivo de complementar la caracterización a nivel cromatínico del enhancer se determinaron algunas modificaciones post-traduccionales de las histonas asociadas al funcionamiento enhancer 3'  $\alpha$ -globina de pollo. En el estadio de eritroblasto, donde los genes globina no se están expresando se logró determinar una combinatoria tanto de marcas positivas como negativas. Este resultado es relevante ya que en años recientes se han reportado la posibilidad de co-localizar marcas bivalentes en las histonas, donde el balance hacia un estado activo o uno apagado, puede ser rápidamente determinado (313).

En eritrocitos primarios de 7 días (eri-7), se determinó que los genes globina adultos se están expresando (Fig. 15, Guerrero y Recillas-Targa, en proceso). Los datos experimentales reflejan una estructura de la cromatina característica de un dominio abierto. Es importante resaltar que los altos niveles para la RNA pol II encontrados en el enhancer en la línea celular HD3 en eri-7 se pierden. Apoyados en estudios hechos por

el grupo de Ann Dean (90), donde se demostró que en la pérdida de la interacción entre el enhancer HS2 el promotor  $\beta$ -major por la interferencia del “insulator” cHS4, los niveles de RNA pol II se re-localizan a la región del enhancer. Proponemos, de existir entre estos elementos, que en células HD3 no hay interacción, dándole al enhancer un papel de poza de activadores transcripcionales. Cuando los genes globina se activan, los niveles de RNA pol II observados en el enhancer se ven disminuidos ya que la RNA pol II ha sido re-localizada a secuencias promotoras. En conclusión, el enhancer  $\alpha$ -globina apoya un modelo progresivo de acción.

Nos pareció muy extraño el que los niveles de inmunoprecipitación recuperados en el ensayo de ChIP fueran tan bajos (Fig. 54y 57). Al analizar estos datos surgió la hipótesis que estos niveles podrán indicar que el enhancer se encuentra desprovisto de nucleosomas. Para demostrar esta hipótesis decidimos realizar un ensayo de ChIP utilizando un anticuerpo específico que reconoce el “core” de todas las histonas. Con este anticuerpo pudimos demostrar una disminución notable y gradual en los niveles inmunoprecipitados conforme se avanza la diferenciación. Este resultado apoya fuertemente la idea de una disociación de las histonas y consecuentemente la formación de una zona desprovista de nucleosomas. Para confirmar este dato se realizó el estudio de inmunoprecipitación en mono-nucleosomas aislados a partir de gradientes de sacarosa. Este estudio nos mostró que el enhancer está perdiendo su organización cromatínica y adquiere una conformación libre de nucleosomas en el estadio en el cual el enhancer es más activo.

Con base a nuestros resultados proponemos que la estructura genómica del enhancer está desprovista de nucleosomas en el estadio en el cual el enhancer es más activo que coincide con el nivel más elevado de remodelación de su cromatina. Sin embargo, el modelo reportado por Boyes y Felsenfeld de interacción del factor GATA-1 y formación de un SH, podría explicar los mismos resultados (3). Las propiedades que muestra el factor GATA-1, al formar el complejo ADN/nucleosoma/GATA liberando secuencias de ADN que se encontraban interaccionando con las histonas. Para el caso del ChIP, aunque la inmunoprecipitación del core sea positiva, el templado estaría expuesto al corte por nucleasas (en este caso en particular de la MNasa), lo que traería como resultado artefactual la no amplificación del templado. En resumen, concluimos que el enhancer se encuentra estructurado en cromatina con características particulares que ameritan profundizar su estudio desde una perspectiva mecánica.

## **El modelo de funcionamiento del enhancer**

El modelo general para explicar el funcionamiento del elemento silenciador-enhancer plantea que estos dos elementos distales de regulación funcionan en su contexto endógeno como una unidad y no como dos elementos distintos y antagónicos. Este modelo se deriva de los datos presentados en este trabajo de tesis relacionados con la caracterización fina de la estructura de la cromatina en esta región genómica. En base a datos de nuestro laboratorio, la actividad *trans*-activadora del enhancer es además modulada por la secuencia conservada de 120 pb, desde la cual se está dictando, al menos en parte, la regulación diferencial de los genes  $\alpha$ -globina de pollo. Proponemos que al llegar al estadio de 5 a 7 días de desarrollo embrionario, el enhancer desde la región previamente definida como el silenciador, daría la señal para que el gen embrionario se apague y por un modelo de competencia ahora los genes adultos pueden interaccionar con el enhancer (Fig. 60). En nuestro grupo favorecemos la idea de un silenciamiento epigenético (en particular metilación) del gen embrionario (Rincón-Arano et al. datos no publicados).

En cuanto a la estructura de la cromatina, GATA-1, además de sus funciones regulatorias como factor transcripcional, parece tener una contribución crítica en la modulación local de la estructura de la cromatina del enhancer. La relevancia de este hallazgo permite imaginar que GATA-1 y sus factores y co-factores asociados contribuyen en el remodelamiento inicial que conlleva a la activación de este elemento. Es decir, que GATA-1 estaría funcionando como un activador transcripcional, desencadenando la serie de eventos que como consecuencia final activarán la expresión de los genes globina. GATA-1 también se encuentra modulando la actividad del enhancer al interaccionar con una región conservada de 120 pb, localizadas 3' del enhancer. Por lo tanto, resulta interesante analizar en este contexto cuales son los co-factores asociados a GATA-1 y las actividades de remodelaje de la cromatina asociadas a éste. Por lo anterior y en base a los resultados obtenidos por nuestro grupo de trabajo, proponemos que el factor transcripcional GATA-1 juega un papel de regulador maestro.

## Conclusiones

- El silenciador  $\alpha$ -globina se comporta como un silenciador clásico.
- El enhancer  $\alpha$ -globina es funcional en eritroblastos y eritrocitos.
- La región de 120 pb de homología se encuentra modulando la actividad del enhancer  $\alpha$ -globina.
- El factor transcripcional GATA-1 interacciona en la región del enhancer  $\alpha$ -globina y presenta un efecto positivo sobre la actividad transcripcional.
- El factor transcripcional YY-1 interacciona en la región del enhancer  $\alpha$ -globina y presenta un efecto negativo sobre la actividad transcripcional.
- El factor transcripcional Sp1 interacciona en la región del enhancer  $\alpha$ -globina y presenta un efecto negativo sobre la actividad transcripcional.
- El enhancer  $\alpha$ -globina se remodela de un modo gradual a través de la diferenciación.
- El enhancer  $\alpha$ -globina es eritroide específico.
- La remodelación de los nucleosomas en la región del enhancer, genera la exposición de sus secuencias de restricción.
- Los elementos silenciador y enhancer presentan a nivel de la estructura de la cromatina, una dinámica de apertura de la cromatina y exposición de sus secuencias blanco diferente.
- El posicionamiento nucleosomal característico, la generación de una cromatina activa y la actividad del enhancer, dependen de los sitios de interacción al factor transcripcional GATA-1.
- El enhancer se encuentra desprovisto de nucleosomas o al menos en una estructura que al corte por nucleasas parece estarlo.
- El enhancer  $\alpha$ -globina apoya un modelo progresivo de acción del enhancer.

## **Perspectivas**

En el presente trabajo de tesis, se planteó el estudio de dos elementos de regulación localizados en la región 3' no-codificante del dominio  $\alpha$ -globina de pollo con la finalidad de definir la contribución de su estructura de la cromatina en sus funciones regulatorias a lo largo de la diferenciación eritroide y el desarrollo del pollo. Los datos obtenidos que han sido descritos y discutidos anteriormente, abren nuevas y amplias perspectivas para la realización de estudios a futuro. Parte de estas perspectivas se exponen a continuación.

### **El silenciador 3' $\alpha$ -globina de pollo**

La caracterización del silenciador en este trabajo no nos permitió obtener datos contundentes relacionados con su función y mecanismos de acción. Incluso las construcciones analizadas por la técnica de RMCE pusieron en duda su existencia. Mientras no se tenga un dato contundente seguiré refiriendo a este fragmento de ADN como el elemento silenciador.

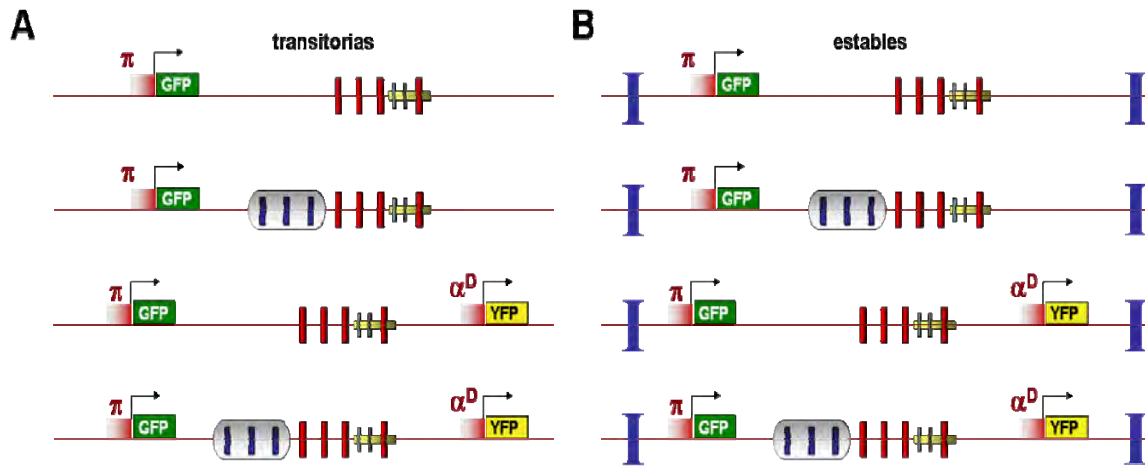
1. A partir de nuestras observaciones consideramos necesario diseñar una nueva serie de experimentos para caracterizar al silenciador. El diseño de estos, deberá tomar en cuenta los resultados obtenidos en el ensayo de recombinación (RMCE, Fig. 50), es decir, las construcciones que fueron generadas para el análisis del silenciador, deberán incluir ahora la región río abajo del enhancer donde localizamos un nucleosoma fijo. Estas construcciones, incluyen una serie de deleciones sucesivas del silenciador, las cuales convendría analizarlas en el contexto de la cromatina por la técnica de RMCE, lo anterior en base a que los resultados obtenidos *in vitro* no concuerdan con los análisis realizados *in situ*, este último estudio al realizarse en el contexto de la cromatina, se encuentra mas cercano a lo que en el contexto endógeno debería de estar ocurriendo.

2. El estudio anterior deberá mostrar con mayor claridad a nivel de secuencia nucleotídica qué elementos tiene un efecto negativo y cuales tienen un efecto positivo sobre la regulación transcripcional. Para continuar con la caracterización del silenciador, propongo, de existir una frontera física entre los elementos negativos y los

elementos positivos que pudieran sostener la existencia del silenciador, que se realice la separación física de estos y se hagan los estudios clásicos de caracterización de un silenciador.

3. Una vez redefinida la región del silenciador (si es que se sigue sosteniendo su existencia), con la intención de comprender si la actividad del silenciador se ejerce sobre el enhancer o las regiones promotoras, realizar las construcciones mostradas en la figura 61 y analizarlas por transfecciones transitorias. Este estudio deberá de realizarse *in vitro*, ya que hasta la fecha no tenemos una línea celular eritroide en la que el silenciador muestre actividad. Las últimas dos construcciones presentan 2 promotores con distintos reporteros, la intención que tiene el diseño de este experimento, es comprender si el silenciador esta actuando sobre el enhancer (o sobre algún proceso iniciado en el) evitando así un incremento en la actividad de reportero de alguno de los dos promotores o a nivel del promotor, donde al seguir el enhancer activo, esperaríamos observar la activación del promotor que no fuera silenciado. Una alternativa para poder estudiar la actividad del silenciador en un contexto estable, sería realizar las construcciones que se muestran en la figura 61 B (de arriba abajo, las primeras dos construcciones) de manera estable en la línea celular HD3 y diferenciar esta línea. Si se observara una actividad de silenciamiento, entonces, valdría la pena analizar las otras construcciones en un contexto estable.

4. El SH que se genera cercano al silenciador (que coincide con el nucleosoma bien posicionado), se incrementa en eritrocitos. Estos resultados sugieren que los eventos de remodelación del silenciador se están llevando a cabo en estados tempranos del desarrollo. Proponemos realizar un ensayo por digestión a la MNasa en cultivos primarios de eritrocitos de 6 y 7 días para buscar un posible cambio en el posicionamiento nucleosomal en la región del silenciador. Coincidentemente, en estos estadios del desarrollo se observó una estructura abierta, accesible a enzimas de restricción. Alternativamente, se podría inducir a diferenciación la línea eritroide HD3 y estudiar el posicionamiento nucleosomal en el silenciador, evaluando la concordancia que existe con respecto a la expresión de los genes globina.



**Figura 61.** Estudio del silenciador. **A**, construcciones propuestas para realizar un estudio por transfecciones transitorias. En este caso proponemos utilizar el reportero de la *GFP*. De arriba abajo, las primeras dos construcciones servirán para comprender si el silenciador en un contexto transitorio, tiene un efecto negativo sobre del promotor embrionario (el efecto que tiene sobre los promotores adultos ya a sido determinado en al menos tres experimentos independientes y realizados por tres personas independientes). El estudio con el segundo juego de construcciones, tiene el objetivo de determinar si el efecto del silenciador inhabilita al enhancer o interrumpe la comunicación que existe entre el enhancer y el promotor o en un segundo escenario, si esta actuando a nivel del promotor. Si estuviera actuando a nivel del promotor, la presencia de un segundo promotor (que contenga a un reportero que exprese en otra longitud de onda, como la *YFP*, de proteína amarilla fluorescente), permitiría observar el efecto del enhancer en los niveles de reportero alterno. Si estuviera actuando a nivel del enhancer, uno esperaría obtener solo niveles basales de reportero **B**, estas construcciones tienen la idea de realizar las mismas preguntas que en “A” pero en un contexto estable, por lo mismo las construcciones deberán estar flanqueadas por “insulators” (I, azul). La última construcción también se podría utilizar en el ensayo de competencia (ver mas adelante).

### **Estructuración en cromatina del enhancer $\alpha$ -globina y la función que el factor transcripcional GATA-1 podría estar desempeñando**

1. Con el objetivo de estudiar los eventos iniciales en la remodelación del enhancer  $\alpha$ -globina y en base a los resultados obtenidos por Lowrey *et al* (2007) donde ven que la expresión del ADNc de GATA-1 en un ambiente no-eritroide, es suficiente para formar los SH endógenos (140), proponemos transfectar el ADNc de GATA-1 en la línea linfóide DT40. Determinar la expresión del mensajero de los genes globina y si estos presentan un incremento en los niveles de mensajero, realizar un análisis a nivel



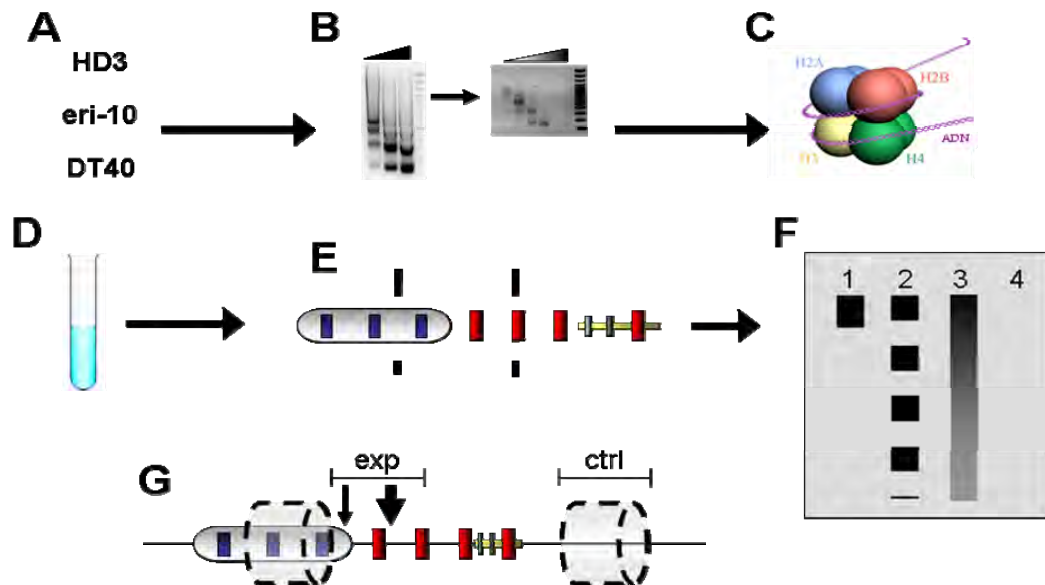
de la estructura de la cromatina para entender los cambios que se están llevando a cabo. La generación de esta herramienta de estudio, nos daría la pauta para realizar preguntas concernientes al establecimiento de una cromatina activa en el dominio  $\alpha$ -globina. Incluso, gracias a que la línea linfoide presenta una mayor frecuencia de recombinación que las líneas eritroides utilizadas por nuestro grupo, se podrían realizar estudios genéticos y de transgénesis.

2. De presentar la transfección transitoria de ADNc un efecto, se podría estudiar el papel *in vivo* de las proteínas implicadas en la regulación del enhancer, generando el “Knockout” de los genes NF-E2 e EKLf y la doble mutante. Estos experimentos tienen la intención final de mostrar si el factor transcripcional GATA-1, tiene la capacidad de iniciar por sí solo, los eventos de remodelación del enhancer  $\alpha$ -globina.

3. Para entender el proceso de inicio de apertura de la cromatina en la región del enhancer  $\alpha$ -globina, podríamos ampliar el estudio de las MPT sobre mono- y di- nucleosomas aislados por gradientes de sacarosa. En paralelo podríamos determinando la interacción de la RNA pol II y la presencia de complejos remodeladores de la cromatina. Este análisis nos ayudaría a entender el estado transcripcional del enhancer a través de la diferenciación y desarrollo del pollo, y a determinar la presencia de factores y co-factores implicados en su regulación.

4. En la línea celular HD3 se encontró un enriquecimiento de la RNA pol II en la región del enhancer, cuando analizamos un estado donde los genes globina se encuentran activos (eri-7), esta, podría estar re-localizando a las secuencias promotoras. Con la intención de incrementar nuestro conocimiento respecto a los mecanismos mediante los cuales el enhancer  $\alpha$ -globina podría estar activando a los genes globina. Propongo confirmar este dato y determinar si la re-localización y por lo tanto el enriquecimiento de la RNA pol II a secuencias promotoras se está llevando a cabo. Este análisis podrá ser complementando si se realiza el mismo estudio en la línea HD3 inducida a diferenciación. Alternativamente y con la finalidad de determinar si el enhancer se ajusta un modelo de “tracking”, se podría estudiar el

enriquecimiento de la RNA pol II en las regiones río arriba y abajo del enhancer, antes, durante y después de la diferenciación.



**Figura 62.** Esquema mostrando los pasos más importantes del ensayo propuesto para determinar la presencia o ausencia de nucleosomas. **A**, partir de la línea celular HD3 y/o de eri-10. La línea DT40 sería un control positivo para evaluar la presencia de nucleosomas. **B**, prepara mono- y di-nucleosomas según la técnica utilizada en esta tesis, con una pequeña modificación, en breve: 1. Obtener núcleos 2. Digerir extensivamente con tres concentraciones distintas de MNasa (triángulo negro), para la reacción (y al mismo tiempo generar un “crosslink” del ADN digerido que siga unido al complejo ADN-histonas) agregando formaldehído al 0.4 % final, incubar 5 minutos a temperatura ambiente. Para asegurarnos de que la reacción enzimática a sido detenida, agregar EDTA 5 mM final. Lavar con PBS-EDTA frío, lisar los núcleos, resuspender en buffer de inmunoprecipitación y realizar el gradiente de sacarosa (triángulo degradado). **C**, realizar un ChIP con el anticuerpo anti histon pan (este anticuerpo inmunoprecipita las 5 histonas). Otro anticuerpo que sería interesante ensayar, sería anti GATA-1. Este anticuerpo al mismo tiempo serviría como control para entender si la región inmunoprecipitada estando o no asociado un nucleosoma (si el anticuerpo anti histonas no inmunoprecipita y el anti GATA-1 si, la conclusión es que no hay nucleosoma). Esferas, histonas. **D**, revertir el “crosslink” y purificar el ADN. **E**, delimitar la región de estudio a fragmentos que se encuentren alrededor de las 200 pb por una digestión enzimática (12 horas mínimo) (para este estudio propongo las enzimas *DraIII* y *PstI*), en caso de ser necesaria una doble digestión, realizarla en el mismo ensayo. En color se muestra la región silenciador enhancer, líneas punteadas representan la región a delimitar. **F**, concentrar la muestra y correrla en un gel al 1.8% (ya que se van a definir fragmentos de menos de 200 pb). La mayor limitante para este experimento será la cantidad de ADN recuperado, por lo mismo el gel de análisis deberá ser lo mas pequeño y con los pozos lo mas chicos que se pueda. Los posibles escenarios a observar son cuatro: 1.- El fragmento se encontraba protegido por un nucleosoma, por lo mismo se observaría una banda parental de alrededor de 200 pb. 2 y 3.- El fragmento de ADN contenía al complejo GATA-nucleosoma, por lo anterior la digestión enzimática generaría cortes al interior del nucleosoma, por lo que es posible que se observen varias bandas (“F2”) o un barrido (“F3”). 4.- La región no se encontraba protegida por un nucleosoma, por lo mismo no habría inmunoprecipitación de esta región en el ChIP anti las histonas. El anticuerpo anti GATA-1 sería un control positivo de inmunoprecipitación, así como la línea DT40. Como sonda radioactiva, se debería utilizar la región completa comprendida entre las enzimas que se utilizaron para delimitar la región. **G**, un control adicional y que al mismo tiempo nos permita

confirmar el resultado en el que proponemos el posicionamiento de un nucleosoma en la región río arriba del enhancer, sería el análisis de esta región. Control (ctrl), región experimental (exp).

5. La presencia de regiones libres de nucleosomas resulta ser un tema muy novedoso e importante para comprender como es que la regulación transcripcional vía elementos de regulación a distancia se lleva a cabo. La hipótesis que propongo es: en la región del enhancer  $\alpha$ -globina tenemos la formación de una estructura muy particular entre GATA-1 y los nucleosomas, es decir, que no se lleva a cabo la disociación total del octámero de histonas.

6. Para profundizar esta caracterización y comprobar mi hipótesis, propongo realizar un ensayo de “footprint” *in vivo* de la región con el que se podría definir con mayor precisión los límites de los nucleosomas posicionados en esa región y la región que deberá contener a GATA-1. Sin embargo aunque la información que este estudio aportaría a la caracterización de la estructura de la cromatina en la región del enhancer sería muy importante, es posible que el resultado fuera engañoso. Como explico en la discusión de mi tesis, por las características muy particulares que se han descrito en la interacción del factor transcripcional GATA-1 con los nucleosomas (3), por ensayos comunes como el ChIP no se podría demostrar si esta región se encuentra libre de nucleosomas. Propongo una técnica que en realidad es una mezcla de otras para determinar la existencia o ausencia de nucleosomas (Fig. 62). Este ensayo tiene la limitante de la cantidad de ADN que se va a recuperar al final, por lo mismo es posible que se tenga que partir de un número grande de células y realizar varios ensayos de ChIP en paralelo para un mismo análisis.

### **La participación de la proteína Sp1, en la actividad del enhancer $\alpha$ -globina**

1. En el dominio  $\beta$ -globina, la proteína Sp1 muestra un efecto positivo en la activación de los genes globina vía el LCR (229). A similitud del dominio  $\alpha$ -globina, los sitios de unión a esta proteína se encuentran en combinación con sitios de unión al factor transcripcional GATA-1, pero para el enhancer  $\alpha$ -globina Sp1 parece mostrar un efecto negativo sobre la actividad del enhancer (Fig. 24). Con el objetivo de tener una primera aproximación sobre su contribución a la actividad del enhancer, se propone la

realización de transfecciones transitorias en la línea eritroide HD3 y en eritrocitos de embriones de 10 días de gestación. Al presentar el enhancer en estos sistemas biológicos un efecto positivo sobre su reportero, resultará más fácil y evidente determinar su participación en la regulación transcripcional vía el enhancer.

2. Para comprender la función que esta proteína deberá estar realizando en un contexto *in vivo*, realizar un ensayo de CHIP contra esta proteína en los diferentes estadios de la diferenciación y desarrollo del pollo. El resultado que este estudio nos muestre, nos permitirá relacionar su presencia con la actividad del enhancer y así poder extrapolar su función.

3. En paralelo para entender por una técnica alterna la función que esta proteína presenta *in vivo* y determinar la dependencia que existe de esta en la regulación de los genes globina vía el enhancer. Proponemos por transfecciones transitorias y estables, utilizando las construcciones previamente reportadas y las construcciones que incluyen la mutación puntual del sitio de interacción a esta proteína, estudiar el efecto de esta proteína en el contexto de la cromatina.

### **Participación de la metilación del ADN en la regulación del enhancer**

1. El elemento silenciador enhancer  $\alpha$ -globina podría ser regulado vía metilación del ADN, considero que se deberían determinar el estado de metilación del ADN de este elemento mediante la técnica de bisulfitos de sodio. De encontrarse metilación *in vivo*, entonces valdría la pena invertir tiempo en los mecanismos mediante los cuales este elemento podría ser metilado y regulado.

### **Papel de la curvatura del ADN en la regulación del enhancer**

1. El efecto que la curvatura del ADN puede tener en la regulación transcripcional ha sido un tema clásico de estudio, en particular en el estudio del ensamblaje del complejo de pre-iniciación. Sin embargo recientemente gracias a nuevos programas y técnicas de análisis, se ha encontrado un incremento en la curvatura del ADN en regiones de regulación conservadas entre especies (314, 315). El enhancer  $\alpha$ -

globina, por los estudios hechos con la MNasa y la DNasa I, parece presentar una curvatura justo en la región que se encuentra entre los dos sitios de hipersensibilidad. Considero que valdría la pena dilucidar la importancia que la curvatura del ADN podría estar teniendo en la regulación transcripcional del enhancer.

### **Modelo de competencia**

1. En base a los resultados realizados en el laboratorio del Dr. Félix Recillas (determinación de la expresión de los genes globina, G Guerrero en preparación), a los resultados que muestran como se remodela la cromatina y a la metilación reportada del gen embrionario  $\pi$ -globina, propongo que el recambio en la expresión de los genes embrionarios a los adultos, se podría estar llevando a cabo por una competencia entre los promotores adultos y el embrionario, por el enhancer  $\alpha$ -globina. El mecanismo implica el silenciamiento del gen embrionario por el silenciador y la posterior metilación de este gen (metilación previamente reportada; 316). Para comprobar este modelo, propongo realizar una construcción que contenga el promotor embrionario y a uno de los promotores adultos, utilizando reporteros distintos como las GFP verde y la amarilla, y localizar entre ellos al enhancer  $\alpha$ -globina (Fig. 61 B, última construcción de arriba abajo). Este estudio debería de ser realizado en transfecciones transitorias (que no incluyan al silenciador) y estables (que incluyan al silenciador). En el caso de las construcciones estables, estas deberán estar preferentemente flanqueada por dos “insulator” que eviten que un segundo enhancer o elementos de regulación a distancia genere datos incorrectos.

2. Finalmente, si se observara una competencia como lo propongo, un segundo experimento podría consistir en mutar al silenciador, obtener las clonas estables en la línea celular HD3 y realizar la diferenciación de la línea. Este experimento me resulta muy interesante, ya que mi hipótesis propone que al momento de llevarse a cabo la diferenciación, la presencia o ausencia del silenciador, podría dictar el promotor que va a ser activado.

## **Material y métodos**

### **Cultivo de tejidos**

La línea celular HD24 es una línea pluripotente eritroide-mieloide, la cual fue transformada con el virus E26 (310, 311) y crecida de acuerdo a previos reportes (111). La línea celular 6C2 corresponde a una línea celular obtenida a partir de médula ósea de pollo y transformada con el virus de la eritroblastosis aviar (AEV), dichas células se consideran como pre-eritroblastos arrestados en estado de CFU-E (Unidades formadoras de colonias eritroides). Esta línea celular se mantuvo en medio  $\alpha$ -MEM (GIBCO) suplementada con 10% (v/v) de suero fetal bovino (SFB), 2% (v/v) de suero de pollo (ChS), 1 mM de Hepes (pH 7.2) y 50 de  $\beta$ -mercaptoetanol. Las células HD3 corresponden a eritroblastos arrestados a través de la infección del virus AEV. Esta línea se mantuvo en medio DMEM (GIBCO) suplementada con 8% de SFB, 2% de ChS. La línea celular linfoide de pollo DT40 (célula hematopoyética no eritroide) fue mantenida en DMEM suplementado con 8% de SFB, 2% de ChS. En todas las líneas celulares utilizadas se complementó el medio con penicilina y estreptomycinina en una concentración final de 500 unidades/ml. Los cultivos primarios de eritrocitos de pollo, fueron obtenidos mediante punción de la vena apical de embriones de pollo incubados durante distintos días y con el fin de obtener eritrocitos terminalmente diferenciados se utilizaron pollos adultos (embriones de *Gallus gallus domesticus*, Alpes, Puebla, México).

### **Transfección transitoria**

Las células HD3 y 6C2 fueron transfectadas transitoriamente por lipofección. Para dicho fin  $2 \times 10^6$  células fueron resuspendidas en 800  $\mu$ l de DMEM sin suero y colocadas en pozos de placas 6 pozos. Por otro lado, 2  $\mu$ g de plásmido de prueba fueron incubados con 200 ng de plásmido control pTK-Renilla (para normalizar la transfección, Promega) en 100  $\mu$ l de DMEM. Así mismo se mantuvo una relación 1:1 de ADN/Liposomas (Lipofectamina 2000, Invitrogen) en 200  $\mu$ l de DMEM. Ambas mezclas fueron incubadas por 10 min a temperatura ambiente, para posteriormente ser mezcladas por goteo e incubadas a temperatura ambiente por 20 min adicionales. Los 200  $\mu$ l finales de la mezcla son agregados a la suspensión celular y, finalmente, se incubaron por 4 h a

37°C. Al terminar la incubación, 2 ml de medio DMEM suplementado con suero fueron agregados a la transfección e incubadas por 48 h a 37°C. La transfección fue analizada por el sistema Dual Luciferasa en un luminómetro 20/20 Turner, de acuerdo con las instrucciones de la compañía (Promega).

Para los ensayos de reactivación con Tricostatina A (TSA) sobre transfecciones transitorias, la suspensión celular transfectada fue dividida a la mitad e incubada por 24 h en presencia o ausencia de TSA.

Para los ensayos de sobre-expresión, los plásmidos con los ADNc de YY1 y GATA-1 fueron co-transfectados utilizando una cantidad máxima de 3 µg, así como ajustando la cantidad de Lipofectamina a usar, conservando la misma relación (1:1). En los ensayos control se utilizó el vector de expresión vacío, usualmente pcADN3.1A.

#### **Ensayo de accesibilidad a la ADNsa I, para las líneas celulares HD24, 6C2, HD3 y DT40**

Este ensayo debe de ser realizado siempre manteniendo las células o núcleos celulares en frío a menos que se indique lo contrario. Se parte de 5 a 6 cajas de 150 mm con las células al 70-80 % de confluencia. Lavar dos veces con 40 ml de PBS (frío). Colectar por centrifugación a 4°C, 5min, 500g (1,200-1,400rpm en la SORBAL). Resuspender lentamente las células en 5 ml de buffer NI (15mM Tris-HCl pH 7.5, sacarosa 300 mM, HEPES 10 mM, KCl 60 mM, MgCl<sub>2</sub>, DTT 1 mM, 4 mM NaCl 5 mM) (según tabla 1).

Número de células	Usar de buffer NI
$1 \times 10^7 - 1 \times 10^8$	2.5 mL
$1 \times 10^9 - 1 \times 10^{10}$	5 mL
$1 \times 10^{11} - 1 \times 10^{12}$	7.5 ml

**Tabla 1.** Relación entre el número de células utilizadas contra el volumen de detergente necesario para obtener núcleos celulares.

Agregar 1 volumen de buffer NI con NP-40 al 1%. Mezclar con pipeta de plástico (fría) en hielo por 60 segundos. Recolectar los núcleos por centrifugación 5 minutos, 500g, 4°C (1,200-1,400rpm en la SORBAL). Lavar 1 vez el pellet con 10ml de buffer NI, resuspendiendo por aspiraciones sucesivas (5 al menos). Centrifugar 5min, 500g, 4°C (1,200-1,400rpm en la SORBAL). Resuspender en el número de ensayos x 100µl de buffer NI. Medir OD's (5µl núcleos en 950µl de buffer stop (10 mM Tris-HCl ph 7.5, EDTA 20 mM y SDS 0.5%); Promedio abs a 260 nm x el factor de dilución = OD's/ml).

Ajustar a 2mM CaCl<sub>2</sub>, dividir en muestras de 300µl y digerir con diferentes concentraciones de ADNsa I (0-10 U/ml + dos muestras para desnudo) por 5 minutos a 25°C. Para la reacción con 300µl de solución de STOP. Agregar 15µl de proteinasaK y dejar 3-6 hr a 50°C. Hacer dos extracciones Fenol:cloroformo:isoamílico (PCIA). RNasa (100µg/ml) e incubar 60 minutos a 37°C. Hacer una extracción. Hacer una extracción cloroformo. Precipitar con 0.33 volúmenes de Amonio 10 M + 1 volumen de isopropanol, dejar ON a -20°C. Centrifugar a 12,000rpm por 30 minutos. (RC-5B, SS-34 rotor). Retirar el EtOH con la p200 y lavar dos veces el pellet con EtOH 70%. Quitar el ETOH con pipeta y secar sobre la mesa de 10 a 20 minutos. Resuspender en 100µl H<sub>2</sub>O y guardar a -20°C. Correr en un gel al 2% agarosa para checar patrón de digestión (de 3 a 4 µg). Digerir con enzimas de restricción y hacer southern con 40 a 70µg.

### **Ensayo de accesibilidad a la DNasa I eritrocitos de pollo**

Recuperar la sangre (20 huevos) en 30 ml PBS 1X, 2mM EDTA (120µl EDTA 0.5M). Centrifugar 5min a 1400rpm a 4°C. Lavar con PBS-EDTA frío. Resuspender en 10ml de PBS (sin EDTA). Centrifugar 5min a 2,000rpm a 4°C. Tomas 3 alícuotas de 10µl y agregar a tres tubos previamente preparados con 1.49ml de buffer de lisis (250µl de TX-100, 2ml NaCl 5 M, llevar a 5ml con H<sub>2</sub>O), mezclar. Centrifugar 2 minutos a 1200rpm. Tomar el sobrenadante y leer a λ 412nm. El promedio de las tres lecturas x 150 (dilución) x 10 (volumen inicial) x 8.56x10<sup>6</sup> (RBC en una OD). Resuspender las células en 5 ml de Buffer NI (según tabla 1). Seguir a partir de este punto el ensayo de accesibilidad a la ADNsa I en líneas celulares.



### **Ensayo de accesibilidad a la MNasa**

Si se va a realizar el estudio en líneas celulares o en eritrocitos de pollo, seguir el protocolo de ensayo de accesibilidad a la DNasa I en líneas celulares o en eritrocitos de pollo respectivamente, hasta el punto donde se cuentan las OD's y a partir de ese punto seguir el siguiente protocolo. Ajustar a 1mM CaCl<sub>2</sub>, dividir en muestras de 300µl y digerir con diferentes concentraciones de MNasa (0, 1, 2, 6, 10, 15, 20, 30 U/ml + dos muestras para desnudo) por 10min. a 25°C. Para la reacción con 300µl de solución de STOP. Agregar 15µl de proteinasa K. Y dejar 3-6hr a 50°C. Hacer dos extracciones PCIA. RNasa (100µg/ml) e incubar 60min. a 37°C. Hacer una extracción PCIA. Hacer una extracción cloroformo. Precipitar con 0.33 volúmenes de Amonio 10 M + 1 volumen de isopropanol, dejar ON a -20°C. Centrifugar a 12,000rpm por 30min. (RC-5B, SS-34 rotor). Retirar el EtOH con la p200 y lavar dos veces el pellet con EtOH 70%. Quitar el ETOH con pipeta y secar sobre la mesa de 10 a 20 min. Resuspender en 100µl H<sub>2</sub>O y guardar a -20°C. Correr en un gel al 2% agarosa para checar patrón de digestión (de 3 a 4 µg). Digerir con enzimas de restricción y hacer southern con 40 µg aproximadamente.

### **Ensayo de accesibilidad a enzimas de restricción**

Si se va a realizar el estudio en líneas celulares o en eritrocitos de pollo, seguir el protocolo de ensayo de accesibilidad a la DNasa I en líneas celulares o en eritrocitos de pollo respectivamente, hasta el punto donde se cuentan las ODs y a partir de ese punto seguir el siguiente protocolo. Utilizar una caja chica (10 ml) por ensayo (al rededor del 50-60% de confluencia, entre 1x10<sup>7</sup> y 1x10<sup>8</sup> de células). Digerir 3,000 OD de núcleos resuspendidos en buffer A (10 mM Tris-HCl, 15 mM NaCl, 60 mM KCl, 0.1 mM EDTA, 5 mM MgCl<sub>2</sub>, 5% glicerol y 1 mM DTT) con 200U de enzima 15minutos a 30°C, en un volumen total de 250 µl. Para la reacción con 300 µl de solución de STOP. Agregar 15 µl de proteinasa K. Y dejar 3-6 hrs a 37°C. Hacer dos extracciones PCIA. Tratamiento con RNasa (100µg/ml) e incubar 60 minutos a 37°C. Hacer una extracción. Hacer una extracción cloroformo. Precipitar con 0.33 volúmenes de Amonio 10 M + 1 volumen de isopropanol, dejar ON a -20°C. Centrifugar a 12,000rpm por 30min. (RC-5B, SS-34 rotor). Retirar el EtOH con la p200 y lavar dos veces el

pellet con EtOH 70%. Quitar el ETOH con pipeta y secar sobre la mesa de 10 a 20 min. Resuspender en 100µl H<sub>2</sub>O y guardar a -20°C. Evaluar la concentración y preparar para el southern.

### **Southern blot**

Antes de transferir el gel, se puede cortar el exceso de este y cortar en el extremo inferior izquierdo un triángulo (rectángulo) para facilitar la orientación. Se recomienda cortar la escalera, ya que en la mayoría de los casos mete ruido. Desnaturalizar en ADN en una solución alcalina (NaOH 0.5M) por 5 minutos. Preparación de la membrana. Cortar una pieza de papel (nylon) aprox. 1mm mas grande que el gel, al igual que dos piezas de papel secante (papel filtro 3mm). Es importante trabajar con guantes y pinzas con punta redonda, de lo contrario la membrana se engrasará y no se humedecerá. Ensamblado del equipo y transferencia de ADN. Colocar una hoja de papel secante en un vidrio (como soporte) permitiendo que los extremos cuelguen. Colocar este dentro de un plato. Llenar el plato hasta el borde del soporte con NaOH 0.5 M, cuando el papel se encuentre perfectamente húmedo, quitar las burbujas utilizando una pipeta como rodillo. Colocar el gel cara abajo sobre el papel, verificando que no queden burbujas entre el papel y el gel. Rodear el gel con egga pack o parafilm, para evitar que las toallas que se pondrán arriba absorban el buffer. Agregar un poco de NaHO 0.5 M (sin que queden burbujas) sobre el gel y colocar la membrana previamente humedecida de tal manera que las marcas coincidan y la membrana cubra los bordes del gel. Sujetar un extremo con un dedo y permitir que se coloque gradualmente. Mojar dos piezas de papel secante en NaOH 0.5 M y colocarlas sobre la membrana, quitando las burbujas con una pipeta. Colocar a continuación de 5-8 cm de servilletas, de tal forma que no sobrepasen los bordes del papel absorbente, y a continuación un vidrio y 400 g de peso. El peso debe ser suficiente de tal manera que la membrana y el gel se encuentren en contacto, pero evitando que la compresión exprima el gel y rehidrate la membrana, impidiendo la buena incorporación del ADN. Permitir que la transferencia del gel se lleve a cabo de 10-15 horas. Quitar las servilletas y los papeles secantes. Marcar los carriles. Poner la membrana en SSC 2X de 1 a 3 minutos, guardar a 4°C o marcar.

Desarmar y pasar la membrana a SSC2X (200ml), de 20 seg a 1min. Meter a un tubo en el horno de hibridación previamente calentado a 65°C, con solución de pre-

hibridación (SSC 20X 6 ml, Denhart 2 ml, SDS 10X 2 ml, 100µl de esperma de salmón y H<sub>2</sub>O 11ml). Pre-hibridar 1 hora a 65°C con agitación constante. Agregar la sonda previamente marcada y determinado su porcentaje de incorporación, desnaturalizarla justo antes de agregarla (5 min a 95°C, 3 min en hielo). Dejar así de 12 a 24 horas.

Lavar y exponer.

### **Lavado de films**

Enjuagar con 50 ml (SSC 2X, SDS 0.1%) precalentada a 65°C. Lavar con 50 ml (SSC 2X, SDS 0.1%) precalentada a 65°C por 45 minutos. Enjuagar con 50 ml (SSC 0.15X, SDS 0.1%) precalentada a 65°C. Lavar con 50 ml (SSC 0.15X, SDS 0.1%) precalentada a 65°C por 45 minutos. Envolver en egapack y poner a exponer.

### **Geles de retardo**

Llenar la tabla conforme a la forma en que se van a realizar los retardos (si se va a competir, agregar un anticuerpo etc.). Agregar el agua, 10 µl de buffer GS 2X (40 mM HEPES pH 7.9, 100 mM KCl, 100 mM MgCl<sub>2</sub>, 10% glicerol y 1 mM DTT) + 1 µl de poli-Idc. Agregar los extractos proteicos. Incubar 15 minutos en hielo y 15 minutos a temperatura ambiente. Agregar las sondas sin marcar radiactivamente con las que se va a competir o los anticuerpos. Incubar 15 minutos en hielo y 15 minutos a temperatura ambiente. Agregar la sonda marcada. Incubar 15 minutos en hielo y 15 minutos a temperatura ambiente. Correr en un gel de archilamida.

### **Immunoprecipitación de la cromatina (ChIP)**

Añadir a las células a una concentración final del 1%, incubarlas a temperatura ambiente por 10 minutos. Añadir glicina sólida a una concentración final de 125 mM y dejar agitando las cajas suavemente por 5 minutos. Lavar 2 veces las células utilizando 10 ml de PBS frío. Resuspender las células en buffer de lisis (5 mM PIPES, pH 8.0, 85 mM de KCl, 0.5% Nonidet P-40, Inhibidores de proteasas) por 10 minutos e incubar en hielo. Centrifugar a 4500 rpm por 5 minutos a 4°C para obtener un pellet de núcleos, descartar el sobrenadante y resuspender el pellet en buffer de lisis para núcleos (50 mM de Tris – Cl, pH 8.1, 10 mM de EDTA, 1% SDS, Inhibidores de proteasas; 250 µl por

condición de anticuerpo) en incubar 10 minutos en hielo. A éste último agregar PMSF (Fenilmetilsulfonilfluorida) a una concentración de 100 µl por ml de Buffer de núcleos (PMSF, 10 U/ml). Para fragmentar el DNA, se podrá recurrir a la sonicación o digestiones prolongadas como con la endonucleasa MNasa (esta última digestión se utilizo en la obtención de mono nucleosomas). Obtener fragmentos de entre 300 a 600pb (Mantener en hielo durante el procedimiento de sonicación, el tiempo y número de pulsos dependerán del sonicador, tipo celular y extensión del crosslinking. Microcentrifugar la muestras a 4°C por 10 minutos a 14,000 rpm. Recuperar el sobrenadante y transferirlo a un nuevo tubo. Dividir las muestras (utilizando  $2 \times 10^7$  células por condición) ajustar a un volumen final de 200 a 600 µl. Agregar 4 µg (1 µg por µl) de anticuerpo específico dependiendo de la muestra, e incubar en una plataforma rotatoria a 4°C overnight.

Lavar las perlas (si las especificaciones del producto lo indican). Se sacan las con buffer IP (0.01% de SDS, 1.1% de tritón X100, 1.2 mM de EDTA, 16.7 mM Tris – Cl, pH 8.1, 167 mM de NaCl) muestras y el rotor a Room Temp. Por 5 horas, y transcurrido el tiempo se procede a añadir cuidadosamente a cada inmunoprecipitación 20 µl de perlas. Se deja rotando overnight a 4°C. Centrifugar 2 min a 3000 rpm y guardar una alícuota del sobrenadante para el “unbound”. Lavar 1X con 1-1.5 ml de RIPA (50 mM de Tris pH 8.0, 150 mM de NaCl, 0.1% de SDS, 0.5% de deoxycholate, 1% de NP – 40, 1 mM de EDTA), 2X High SALT (50 mM Tris pH 8.0, 500 mM de NaCl, 0.1% de SDS, 0.5% de deoxycholate 1% de NP – 40, 1 mM de EDTA), 1X LiCl (50 mM de Tris pH 8.0, 1 mM de EDTA, 250 mM de LiCl, 1% de NP – 40, 0.5% de deoxycholate) y 2X con TE buffer (pH 8) en el cuarto frío con incubaciones de 10 minutos en agitador rotatorio y cada vez quedarnos con las perlas. Añadir 400 µl de TE + 0.5 µl de un stock de 10mg/ml de RNasa A por muestra e incubar 1 hora a 37°C, añadir agua hasta llegar al tope del tubo. Incubar las muestras a 67°C por 5 horas para revertir el entrecruzamiento (se puede dejar toda la noche). Centrifugar a 3000 rpm por 2 minutos y recuperar el sobrenadante en un tubo nuevo. Agregar 60 µg de proteinasa K a 45°C por 2 horas. Agregar 300 µls de TE y hacer 2 fenol:cloroformo y una vez con cloroformo. Precipitar con 30 µls de NaCl 5M y 5 a 20 µg de glicógeno por muestra y con etanol al 100% (“Overnight” a – 20°C). Centrifugar las muestras a máxima velocidad por 20 minutos a 4°C. Dar un lavado con ETOH al 70%, y dejar secar los pellets. Resuspender en agua inyectable y analizar por PCR para cuantificar. Estos

generalmente se resuspenden en 30  $\mu$ l y se diluyen 300 veces más, usando de 2 a 3  $\mu$ l por reacción de PCR.

### **Anticuerpos**

Los anticuerpos utilizados son:  $\alpha$ -GATA-1 (N-200),  $\alpha$ -YY1,  $\alpha$ -RNA Pol II,  $\alpha$ -IgG de ratón-Peroxidasa de rábano (HRP) (Santa Cruz Biotechnology). Los anticuerpos policlonales para H3 y H4 acetiladas fueron obtenidos de Upstate. El Dr. Thomas Jenuwein (IMP, Viena, Austria) proporcionó anticuerpos específicos para las siguientes modificaciones: trimetilación de la lisina 9 de la histona H3 y trimetilación de la lisina 20 de la histona H4. El Dr. Mario Zurita proporciono el Anticuerpo para de ratón anti-Histon pan fue de Chemicon.

### **Obtención de extractos proteicos**

Se trabaja con 3 cajas de células al 80-90% (entre 40 y 50x10<sup>6</sup> células). Se lavaron 2 veces con PBS frío. Se resuspenden en 500 $\mu$ l de buffer RIPA con sal (1M Tris pH 8.0, 1M Hepes pH 8.0, 0.5M EDTA, 5M NaCl, NP40, PMSF 1M), se incuban 1.3 horas en hielo. Se centrifugan a 12,000rpm X 15 minutos a 4°C, se conserva el sobrenadante. Se agrega un volumen de buffer RIPA sin sal (1M Tris pH 8.0, 1M Hepes pH 8.0, 0.5M EDTA, NP40, PMSF 1M, glicerol 5%). Calcular concentración.

### **Recombinación sitio específica**

Se utiliza un sistema de "Cre recombination-mediated cassette exchange (Cre-RMCE)". Este requiere la generación de una construcción integrada en copia sencilla que exprese a la higromicin B fosfotransferasa y a la timidin kinasa del herpes simplex virus (HyTK). Estos marcadores favorecen doble selección, confiriendo resistencia a la higromocina B y sensibilidad al glanciclovir, respectivamente.

Diez días previos a la electroporación, seleccionar las células que expresan HyTK en DMEN que contenga 750  $\mu$ g por ml de higromicina B, para asegurarse de que todas las células expresen la fusión HyTK. Electroporar o transforfar con lipofectamina. Preparar el mix de electroporación mezclando 25  $\mu$ g del plásmido por integrar + 15  $\mu$ g del plásmido que contiene a la Cre y 100  $\mu$ g del acarreador. Precipitar por EtOH y lavar

con EtOH 70%. Resuspender el ADN en 40  $\mu$ l de solución 2 (DMEM. Dulbecco's Modified Eagle Medium suplementado con 10% de suero fetal) en condiciones de esterilidad. Llevar las células a  $5 \times 10^6$  en la solución 2 a temp ambiente. Mezclar las células con el ADN y transferirlos a una cubeta de electroporación. Electroporar. Seleccionar (ver mas adelante).

### **Transfección con lipofectamina**

Disolver los plasmidos por utilizar en  $\mu$ l de DMEM sin suero. Disolver 6  $\mu$ l de reactivo pluss en los 100  $\mu$ l de DMEM anteriores. Disolver 4  $\mu$ l de lipofectamina en 100  $\mu$ l de DMEM sin suero. Incubarlos 15 min a TA. Mezclarlos por goteo (la lipofectamina a los plasmados). Incubar 30 min a TA. Lavar las células por transfectar con D-MEM sin suero y resuspenderlas a  $5 \times 10^5$  en 800  $\mu$ l en D-MEM sin suero. Agregar la mezcla de lipofectamina + plasmidos por goteo a las células. Incubar 4 hr en la incubadora. Llevar a 4 mL con medio complementado. 3 días mas tarde comenzar la selección con Glanciclovir (50 ng/mL conc. Final (stock 2.5 mg/mL).

### **Selección**

Tres días después de la electroporación o transfección por lipofectamina, agregar glanciclovir a una concentración de 10  $\mu$ M por mililitro para seleccionar en contra de aquellas células que continúen expresando HyTK. Para generar colonias aisladas, determinar el número de células vivas contando con azul de Trypan. Platear en una caja de 96 pozos a una densidad de una célula por pozo. Después de una semana expandir a cajas de 24 pozos. Posteriormente a una caja de 6 pozos y analizar ADN genómico.

### **Gradiente de sacarosa para la preparación de mono-, di- y tri- nucleosomas.**

Crece 6 cajas de 150cm. Preparar y digerir núcleos según protocolo (utilizando 2 conc. diferentes –tomar una alícuota de cada una para analizar al final- de MNasa las cuales se juntaran al final). Parar agregando EDTA a una concentración final de 0.5m.

Centrifugar a 25,000 rpm por 5min (retener sobrenadante S1 -100  $\mu$ l para análisis-).

Resuspender los núcleos en buffer de lisis + 0.25mM EDTA (500 $\mu$ l) incubar 15 min a 4°C. Pasar a travez de una jeringa del 20 4 veces y luego a través de una jeringa del 22 4

veces. Centrifugar a 10 K por 10 min (retener sobrenadante S2-100  $\mu$ l para análisis-). Combinar S1 y S2 y llevar a 150mM NaCl con NaCl 5M. Ej. S1 + S2 = 1400  $\mu$ l + 42  $\mu$ l 5M NaCl. Incubar 20 min a 4°C. Centrifugar 15 min a 10 K y guardar "overnight" a 4°C.

### **Preparación del gradiente de sucrosa**

Una noche antes preparar un gradiente de sacarosa del 30% al 6% en (80 mM NaCl, 10 mM Tris-HCl pH 7.4 y EDTA 3 mM) guardar a 4°C. Agregar 1.2 mL de la digestión con la MNasa encima del gradiente de sucrosa. Centrifugar 1.45 min. A 38K, 4°C en la SW40. Colectar fracciones de 22 gotas / fracción (aprox 1 mL por fracción). Guardar a -80°C. Para el análisis de fracciones. Colectar 100  $\mu$ l de cada fracción. Agregar 100  $\mu$ l de NaCl 0.12 M en TE pH 8. Agregar 1  $\mu$ l (10mg/ $\mu$ l) de RNasa e incubar 30 min a 37°C. Agregar 10  $\mu$ l (10mg/ $\mu$ l) de proteinasa K he incubar 30 min a 37°C. Extracción fenol cloroformo (200 $\mu$ l). Extracción cloroformo (200  $\mu$ l). Precipitar con 5  $\mu$ l de NaCl 5M, + 500  $\mu$ l de EtOH. Dejar "overnight" a 4°C. Centrifugar 45 minutos y lavar con 500  $\mu$ l de EtOH 80%. Checar concentración y analizar en un gel TBE 2%.

# GATA-1 Modulates the Chromatin Structure and Activity of the Chicken $\alpha$ -Globin 3' Enhancer<sup>∇†</sup>

Martín Escamilla-Del-Arenal and Félix Recillas-Targa\*

*Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Apartado Postal 70-242, México D.F. 04510, México*

Received 28 May 2007/Returned for modification 31 August 2007/Accepted 23 October 2007

**Long-distance regulatory elements and local chromatin structure are critical for proper regulation of gene expression. Here we characterize the chromatin conformation of the chicken  $\alpha$ -globin silencer-enhancer elements located 3' of the domain. We found a characteristic and erythrocyte-specific structure between the previously defined silencer and the enhancer, defined by two nuclease hypersensitive sites, which appear when the enhancer is active during erythroid differentiation. Fine mapping of these sites demonstrates the absence of a positioned nucleosome and the association of GATA-1. Functional analyses of episomal vectors, as well as stably integrated constructs, revealed that GATA-1 plays a major role in defining both the chromatin structure and the enhancer activity. We detected a progressive enrichment of histone acetylation on critical enhancer nuclear factor binding sites, in correlation with the formation of an apparent nucleosome-free region. On the basis of these results, we propose that the local chromatin structure of the chicken  $\alpha$ -globin enhancer plays a central role in its capacity to differentially regulate  $\alpha$ -globin gene expression during erythroid differentiation and development.**

In recent years, the relevance of promoter elements has been outlined based on the varied promoter architectures required for regulatory specificity (28). A large amount of data has been generated describing how the basal transcriptional machinery is incorporated into promoters and how proximal elements are required for full and specific activity (17). However, a subset of genes requires long-distance regulatory elements for their developmental timing and tissue-specific gene expression (9, 39). Locus control regions (LCRs), enhancers, silencers, and insulators represent some of the elements exerting remote regulation (5, 10, 26). Such regulatory elements require characteristic chromatin structures, and the great majority of these regulatory elements and chromatin components are usually associated with nuclease hypersensitive sites (HSS) (6, 39). Despite progress in our understanding of promoters and long-distance regulatory elements, the mechanisms by which enhancers control gene expression are poorly understood, particularly in terms of how their own chromatin structure modulates their activity.

In early work, we identified a silencer-enhancer element located at the chicken 3' side of the  $\alpha$ -globin domain, around 400 bp downstream of the adult  $\alpha^A$  gene (Fig. 1A) (16, 32, 33). We have adopted this enhancer element as a model system to investigate the differential regulation of chicken  $\alpha$ -globin gene expression during erythroid differentiation and development. Based on recent data, we have proposed that the 3'-side en-

hancer is able to modulate its own function through the binding of GATA-1 and YY1 (Fig. 1B) (33) and the associated chromatin remodeling machinery. Much less is known about the molecular features of the chicken  $\alpha$ -globin silencer. What we have found until now is that the silencer is located side by side with the enhancer and that such location may give particular regulatory properties to both elements. The silencer was functionally defined in transient transfections as capable of reducing the activity of the heterologous simian virus 40 promoter, a weak promoter in erythroid cells (31). In addition, we know that at least three unidentified nuclear factors bind in vitro to it (31).

To better understand the function of the chicken 3' silencer-enhancer elements, here we studied the chromatin organization of this element during chicken development and erythroid differentiation. We hypothesized that the silencer-enhancer activity could be regulated by the interplay between key transcription factors and the acquisition of differential conformation of its chromatin structure. DNase I accessibility and nucleosomal mapping revealed an erythrocyte-specific conformation, based on the generation of two characteristic HSSs. In situ restriction assays confirmed a differential chromatin conformation of the silencer-enhancer elements during erythroid differentiation and development. High-resolution mapping showed that the HSSs are located between the silencer-enhancer elements, colocalizing with two GATA-1 binding sites. Mutational analysis in vivo demonstrated that the two most 5' GATA-1 binding sites regulate the enhancer chromatin conformation and are responsible for its optimal and regulated activity. These experiments further showed that in an integrated context, the silencer element previously defined in transient assays is a functional part of the enhancer. The chromatin remodeling process that occurs over the enhancer is consistent with an enrichment of histone acetylation followed by the local formation of an apparent nucleosome-free region, coincident

\* Corresponding author. Mailing address: Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Apartado Postal 70-242, México D.F. 04510, México. Phone: (52 55) 56 22 56 74. Fax: (52 55) 56 22 56 30. E-mail: frecilla@ifc.unam.mx.

† Supplemental material for this article may be found at <http://mcb.asm.org/>.

<sup>∇</sup> Published ahead of print on 5 November 2007.



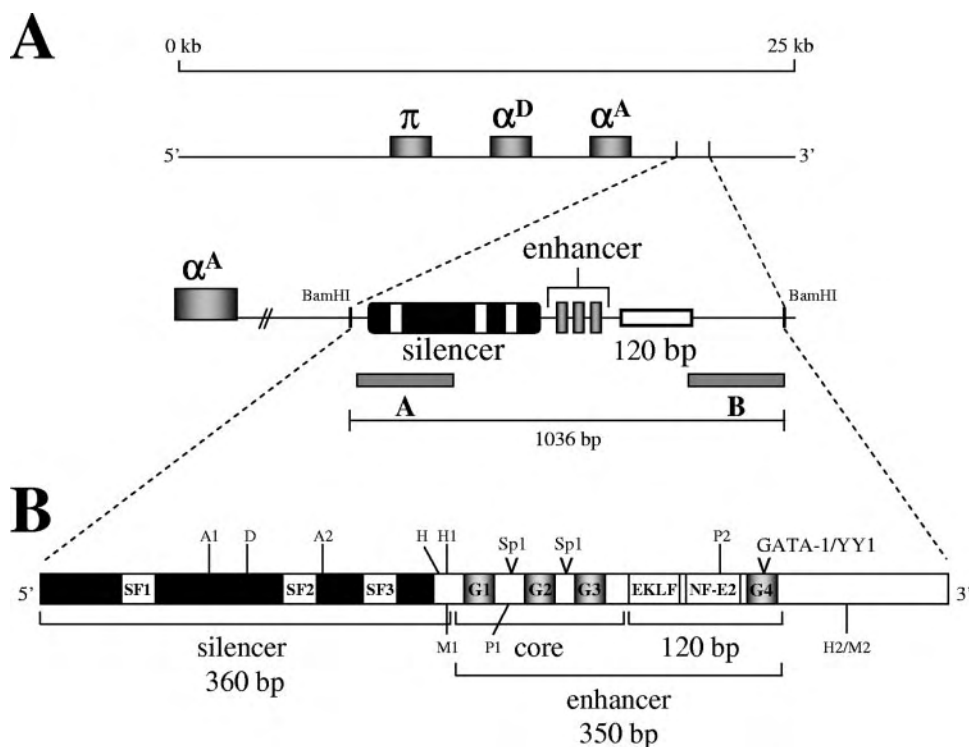


FIG. 1. Scheme of the chicken  $\alpha$ -globin gene domain, the silencer-enhancer, and the detailed distribution of nuclear factor binding sites. (A) The silencer-enhancer regulatory elements are located 397 bp downstream of the adult  $\alpha^A$  globin gene. These elements are included in a 1,036-bp BamHI genomic DNA fragment. A and B probes used in Southern blotting are shown (Fig. 2 and 3; also see Fig. S1 and S2 in the supplemental material). (B) Detailed scheme of the silencer-enhancer DNA-binding sites and key restriction enzyme-cutting sequences. The silencer possesses at least three binding sites for nuclear factors defined by in vitro footprinting (SF1 to SF3) (31). The identity of these nuclear factors remains to be determined. The initial characterization of the enhancer described the presence of three GATA-1 binding sites named the core enhancer (16). The core enhancer GATA-1 sites are named G1, G2, and G3. Sequence alignment of the duck and chicken homologous sequences revealed a highly conserved 120-bp DNA fragment with binding sites for Sp1/EKLF, NF-E2, a fourth GATA-1 (G4), and YY1 (33).

with the stage of maximal enhancer activity. Our results demonstrate that GATA-1 is mainly responsible for the erythrocyte-specific chromatin conformation and activity of the chicken  $\alpha$ -globin 3' enhancer.

#### MATERIALS AND METHODS

**Cell culture.** All cell lines are derived from chicken (*Gallus gallus*). HD24 cells, which are pluripotent erythroid-myeloid cells transformed by the E26 virus, were grown as previously reported (3, 4, 30). 6C2 cells, corresponding to erythroblastosis virus-transformed bone marrow cells arrested at the CFU-E stage, are considered to be a preerythroblast line (3, 4). 6C2 cells, as well as the avian erythroblastosis virus-transformed and temperature-sensitive erythroblast line LSCCHD3 (herein referred to as HD3), were grown as previously described (6, 33). The DT40 lymphoid cell line was grown in DMEM supplemented with 50  $\mu$ M 2-mercaptoethanol, 10% (vol/vol) fetal calf serum, 5% (vol/vol) chicken serum, and 10% (vol/vol) tryptose phosphate broth (13).

**Plasmid constructs.** DNA subfragments containing the chicken  $\alpha$ -globin 3' silencer-enhancer was PCR amplified from 10-day-old embryonic erythrocyte (10d-RBC) genomic DNA of *G. gallus* (Alpes, Puebla, México) by use of the following primers: ASIL1-PacI (5'-CGCGCTTAATTAAGATCCCATGCCTACTCCTACC-3'), A7-AscI (5'-CAAGTGGCGCGCCTGCAGCAGGTTGAGCAGACC-3'), and AE1-NheI (5'-CGCGCGCTAGCGCGCAGGGTGAAGCTGTGCTG-3'). All DNA fragments were cloned into pG $\alpha$ P3 (kindly provided by Héctor Rincón-Arango) containing the *GFP* reporter gene under control of the  $\alpha^D$  promoter. Mutations of the three GATA-1 binding sites present in the core enhancer (Fig. 1B; also see Fig. 7 below; also see Fig. S5 in the supplemental material) were generated in the context of the 1,036-bp silencer-enhancer construct by use of the QuikChange site-directed mutagenesis kit (Stratagene) with the following primers: Mutgata01U (5'-GGCCCCGGCTCTTTAAACGGCC

AGCAGCAGGG-3'), Mutgata01D (5'-CCCTGCTGCTGGCCGTTTAAAGAGCCGGGGGCC-3'), Mutgata02U (5'-GGGGGCTGCAGGTGGCTTTAAAGAGAGCTGACAGGC-3'), Mutgata02D (5'-GCCTGCAGCTCTTTAAAGCCACCTGCAGCCCC-3'), Mutgata03U (5'-GGGACGTGGGCAGCATTAAAGCCTCGGGTGGG-3'), and Mutgata03D (5'-CCCACCGAGGCTTAATGCTGCCACGTCCC-3'). All the DNA fragments were NaeI-EcoRI digested and subcloned into the SmaI-EcoRI sites of the pL1-HYTK-L2 vector, replacing the cytomegalovirus HYTK gene (14). All recombinant plasmids were sequenced employing the primer ASIL1-NheI (5'-CGCGGATCCCATGCCTACTCCTACCCTATG-3').

**DNase I, MNase, and restriction endonuclease nuclear digestions.** For DNase I digestion,  $1 \times 10^9$  cells were washed twice in cold phosphate-buffered saline and resuspended in 2.5 ml of NI buffer (15 mM Tris-HCl [pH 7.5], 300 mM sucrose, 10 mM HEPES, 60 mM KCl, 4 mM MgCl<sub>2</sub>, 1 mM dithiothreitol, 5 mM NaCl), and 2.5 ml of lysis buffer (NI buffer plus 0.5% NP-40) was then added to the isolated nuclei and incubated for 1 min. Nucleus integrity was monitored during this incubation, and the nuclei were then pelleted and washed once with NI buffer and resuspended in 100  $\mu$ l of NI buffer. Five microliters of the resuspended nuclei was mixed with stop buffer (10 mM Tris-HCl [pH 7.5], 20 mM EDTA, 0.5% sodium dodecyl sulfate), and optical densities (OD) were determined ( $A_{260} \times 200$ ). The volume corresponding to 3,000 OD units was adjusted to 300  $\mu$ l with NI buffer solution in a final concentration of 2 mM CaCl<sub>2</sub>, and nuclei were digested with 0, 0.1, 0.5, 1, 3, 5, and 15 U of DNase I (Worthington) for 5 min at 25°C. The reaction was stopped with 300  $\mu$ l of stop buffer. Samples were treated with 150  $\mu$ g of proteinase K at 55°C for 6 h, phenol-chloroform extracted twice, and incubated with 50  $\mu$ g of RNase A at 37°C for 1 h. The samples were phenol-chloroform extracted once more and DNA was precipitated with isopropanol. For micrococcal nuclease (MNase) digestions, we followed the same protocol as for DNase I with two modifications. Nuclei were adjusted to 1 mM of CaCl<sub>2</sub> and digested with increasing concentrations of

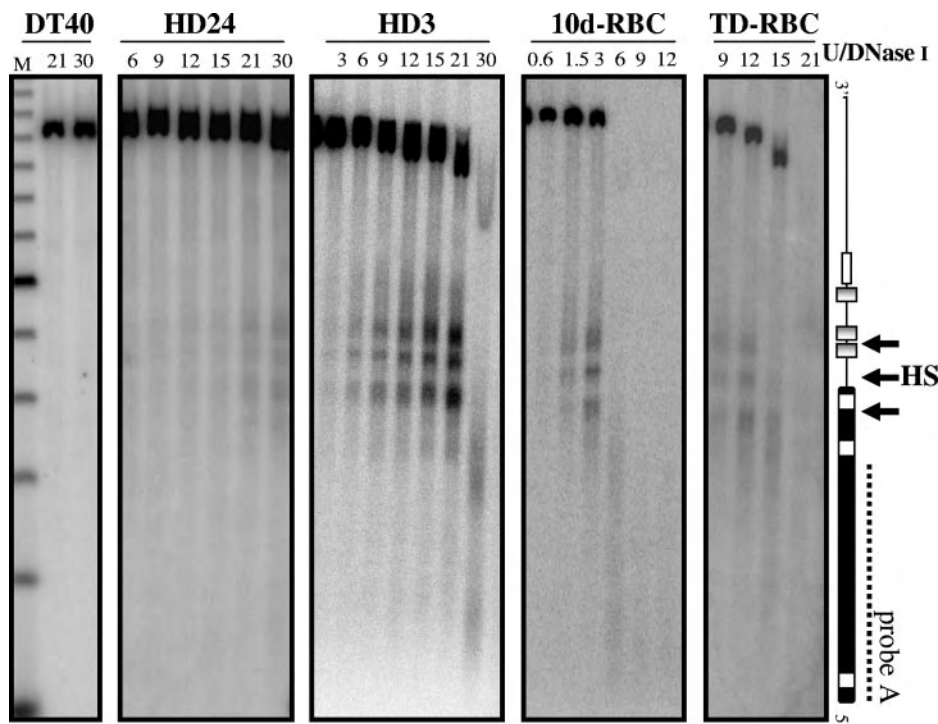


FIG. 2. Chromatin structure over the chicken 3'-side  $\alpha$ -globin silencer-enhancer. Southern blots showing the presence of DNase I-HSs. Notice the appearance of three DNase I-HSs over the core enhancer and the transition zone between the silencer and the enhancer, from HD24 cells to TD-RBCs (vertical arrows). In 10d-RBCs, the appearance of the HSs is almost 1 order of magnitude higher than that seen for HD3 digested nuclei (compare lanes corresponding to 6 and 1.5 units of DNase I). DT40 cells show only the higher concentration of enzyme to evidence the absence of HSs. M, 100-bp molecular mass marker.

MNase (Worthington). For restriction endonuclease digestions, we followed the same protocol as for DNase I with minor modifications. Nuclei were resuspended in buffer A (10 mM Tris-HCl, 15 mM NaCl, 60 mM KCl, 0.1 mM EDTA, 5 mM MgCl<sub>2</sub>, 5% [vol/vol] glycerol, 1 mM dithiothreitol). Nuclear concentrations were adjusted to 3,000 OD in 300  $\mu$ l buffer A and digested with 200 U of the corresponding restriction enzyme for 15 min at 30°C.

**Isolation of mono- and dinucleosomes.** Mono- and dinucleosomes were purified as described previously (18, 27). Three independent aliquots of isolated nuclei were digested with increasing concentrations of MNase. All MNase digestions were incubated for 25 min at 25°C. The reactions were stopped by adding EDTA (pH 8.0) to a final concentration of 10 mM. The three digests were combined and mono- and dinucleosomes were separated on a 6 to 30% sucrose gradient (18). For the MNase sensitivity assay, duplex PCR was performed with DNA associated with mono- and dinucleosomes insolated by sucrose gradient. The normalization was done using the *FR* gene sequence (primer 5.613 in reference 27). Genomic DNA from 10d-RBCs was used to determine the difference in sensitivity enrichment (*n*-fold) between genomic DNA and MNase-digested input DNA.

**ChIP.** The chromatin immunoprecipitation (ChIP) assay was done as previously reported (35) employing 2  $\mu$ g of acetylated H3 (acH3), acH4, and H3K4me2 (Upstate),  $\alpha$ -RNA polymerase II (Pol II), anti-GATA-1 (H-200), and  $\alpha$ -immunoglobulin G (Santa Cruz Biotechnology) antibodies. H4K20me3 antibody was kindly provided by Thomas Jenuwein and mouse antihistone, pan-monoclonal antibody recognizing histones H1, H3, H4, H2A, and H2B from Chemicon International (MAB3422) was kindly provided by Mario Zurita (24). Immunoprecipitated DNA was analyzed by PCR using primers specific for the silencer-enhancer (SE120) region with the primers RealTSil-U (5'-TGTGGCC CTCATAGGACTTCC-3'), RealTSil-D (5'-GACCCACAAGTGGCACATA G-3'), RealTSE-U (5'-GGATGAGCGAGGGTGAAGCT-3'), and RealTSE-D (5'-GCTGTGGCCGTGATAAGACG-3') and the enhancer (E4) RealTenh-U (5'-CAGGCTCTCCTCCAGTCACG-3') and RealTenh-D (5'-T ACCCACCCGAGGCTATCTGC-3'). For semiquantitative ChIP assay, duplex PCR was performed as described previously (35). Control primers for normalization were selected from the chicken  $\beta$ -globin-adjacent 16 kb of condensed chromatin and the *FR* gene HSA sequence corresponding to an open chromatin

region in the chicken genome (see primers 10.35 and 5.613 in reference 27). For GATA-1 ChIP assays, normalization primers were selected around 3 kb downstream of the enhancer, the 3' noncoding region of the chicken  $\alpha$ -globin domain and the primer sequence area and were as follows: Ctrl3' $\alpha$ F (5'-CCGTTTCAA ACCAACCTACTGGACT-3') and Ctrl3' $\alpha$ R (5'-GCTGTACGCTTCAGCTCA ATATCAG-3') (16, 33).

**RMCE.** Recombinase-mediated cassette exchange (RMCE) was basically performed as described previously (14). The pL1-HYTK-L2 vector was stably transfected in HD3 cells, and three independent clones were isolated (RMCE-HD3-13, -14, and -20). Transgene integrity was confirmed by Southern blotting (clones kindly provided by Héctor Rinón-Arano). Newly generated plasmids were co-transfected with the pBS185 plasmid, which contains the Cre recombinase (cytomegalovirus Cre; Clontech). Three rounds of hygromycin selection, each for a period of 3 to 5 days, to allow cell recovery, were performed. Stably transfected cells were selected based on Southern blotting and PCR from a pool of 700 hygromycin-resistant individual colonies.

RESULTS

**Erythrocyte-specific formation of DNase I-HSs defining the silencer-enhancer transition.** To assess the contribution of the local chromatin structure to the activity of the 3' chicken  $\alpha$ -globin silencer-enhancer (Fig. 1), we initially characterized its chromatin organization during erythroid differentiation and development. We took advantage of transformed chicken erythroid cell lines representing different developmental stages: the most primitive are HD24 cells, representing a myeloid-erythroid stage, whereas the most mature are HD3 cells, corresponding to erythroblasts. To complete the differentiation spectrum, we isolated nuclei from 10d-RBCs as well as terminally differentiated erythrocytes (TD-RBCs) (Fig. 2). Based on

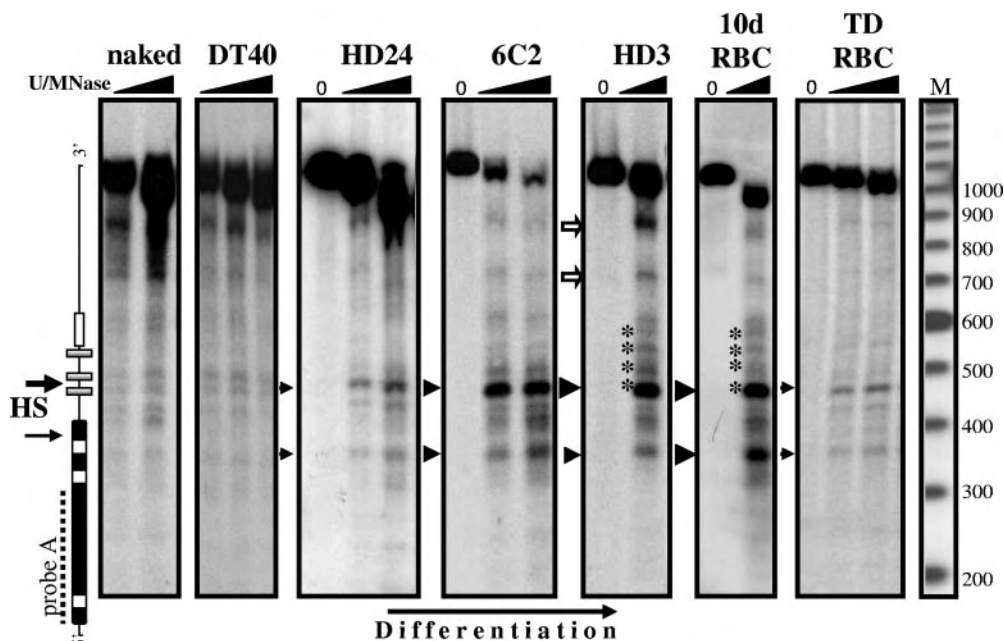


FIG. 3. Nucleosome positioning determined by MNase digestion. The patterns of MNase digestion of naked DNA and of DT40 nuclei are similar. Interestingly, at the earliest stages of erythroid differentiation there are two MNase-HSs, the increasing intensity of which reaches the highest level in 10d-RBCs (black arrows). These two MNase-HSs are found between the silencer and the enhancer, and the most 3' HS is more intense (large horizontal arrow). Starting with HD3 cells, when the enhancer begins to be active, there is the appearance of several bands (asterisks) 3' to the second MNase-HS over the 120-bp DNA fragment (open white box in the vertical scheme), suggestive of an increase in nucleosome mobility or nucleosome exclusion, consistent with a more open chromatin organization (asterisks). The open horizontal arrows indicate the digestion pattern created by the constant positioning of a nucleosome downstream to the enhancer.

the identification of DNase I-HSs by DNase I digestions, we observed the progressive appearance of a cluster of at least three DNase I-HSs located between the enhancer and the silencer (Fig. 2). The nonerythroid DT40 cell line showed no DNase I-HS formation. Consistent with this change in chromatin structure, we have previously demonstrated that the enhancer is inactive in HD24 and 6C2 cells, is modestly active in HD3 cells, and reaches the highest levels of *trans*-activation potential in 10d-RBCs (33). These results demonstrate a localized and progressive increment of DNase I sensitivity over the silencer-enhancer.

**Chromatin structure of the silencer-enhancer based on differential positioning of a nucleosome.** To further characterize the chromatin structure of these regulatory elements and determine nucleosome positions over the silencer-enhancer elements, we carried out MNase digestions (Fig. 3). We incorporated two additional controls: naked (genomic) DNA (see Fig. S1 in the supplemental material) and chicken preerythroblast 6C2 cells, considered to be differentiation intermediates between the HD24 and HD3 cell lines (Fig. 3) (30). Two prominent erythrocyte-specific MNase-HSs are clearly observed with hypersensitivity progressively increasing in intensity in a differentiation stage-specific manner (Fig. 3), as observed for the DNase I-HSs (Fig. 2). The 5' MNase-HS is located between the silencer and enhancer, and the most 3' MNase-HS, which is consistently more intense, appears to colocalize between the first and second GATA-1 binding sites (Fig. 3). The distance between these two MNase-HSs is apparently not sufficient to accommodate a nucleosome (data not shown). We found that additional bands appear for HD3 cells and 10d-

RBCs, suggesting different nucleosome positions over the enhancer representative of an average of templates (Fig. 3). This chromatin remodeling affects sequences corresponding to the 120-bp DNA fragment previously shown to be responsible for modulation of the enhancer activity (Fig. 1B and 3) (33). For TD-RBCs, we observed a general decrease in MNase sensitivity consistent with the progressive silencing of the  $\alpha$ -globin gene expression (Fig. 3; TD-RBCs and data not shown).

Utilizing Southern hybridization with probe B, we then determined the nucleosome positioning pattern over the silencer element, revealing stable positioning of one nucleosome over the silencer (Fig. 3; also see Fig. S2 in the supplemental material). Notice that another well-positioned nucleosome is also observed downstream from the 120-bp DNA fragment (Fig. 3).

In conclusion, the DNase I and MNase assays support the fact that the silencer-enhancer chromatin structure is gradually remodeled in a stage-specific manner, reaching maximal chromatin accessibility when the enhancer is fully active (HD3 cells and 10d-RBCs) (see Fig. S3 in the supplemental material).

**Evaluation of the chromatin template accessibility during erythroid differentiation and development.** Chromatin remodeling *in vivo* can be detected by an increase in the restriction enzyme accessibility of nucleosome assembled DNA in preparations of isolated nuclei (1, 41). Restriction enzyme accessibility assays were used to determine if particular nucleosomes over the silencer-enhancer elements were subjected to chromatin remodeling (Fig. 4). Nuclei from cells representative of the distinct erythroid stages and convenient restriction sites were chosen, based on the predicted nucleosome positioning.

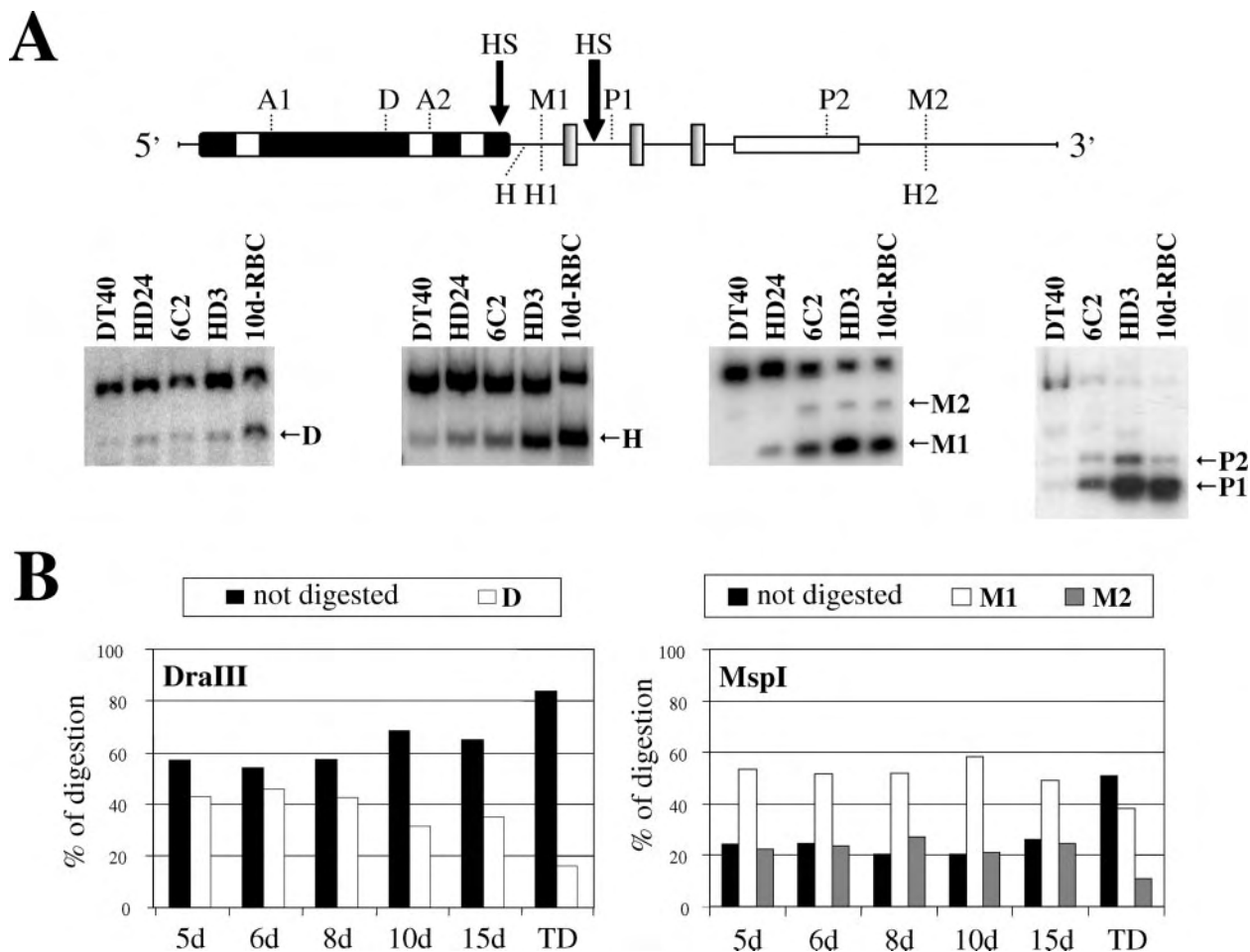


FIG. 4. Analysis of chromatin accessibility by in situ restriction enzyme assays in the course of erythroid differentiation and development. (A) Scheme representing the  $\alpha$ -globin domain with the location of restriction enzymes used. D, DraIII; M1 and M2, MspI; A1 and A2, AvaII; P1 and P2, PstI; H1 and H2, HpaII; and H, HhaI. Vertical arrows indicate the positions of the two MNase-HSs. In situ restriction enzyme accessibility assays were performed with isolated nuclei. For DT40 cells, the region is in a repressive conformation, as shown using all the restriction enzymes. It is important to outline the progressive accessibility observed over the course of erythroid differentiation. In all cases, genomic DNA was BamHI digested and Southern blotting performed with probes from the 5' and 3' sides of the silencer-enhancer (data not shown). (B) Evaluation of local chromatin accessibility during chicken development. Using DraIII (D), there is a decay in enzyme digestion, suggesting a more accessible and relaxed chromatin conformation over the silencer during early stages of development. In contrast, for MspI (M1), accessibility is continuous during development. In TD-RBCs, there is a notorious increment on nondigested material in both cases. The results were quantitated using a phosphorimager and are presented as the percentage of digestion relative to that seen for the input DNA.

We found that the silencer element is poorly accessible to DraIII in DT40, HD24, 6C2, and HD3 cells (Fig. 4A and data not shown). In contrast, there is a clear accessibility increase in 10d-RBCs, suggesting a change in the local chromatin structure (Fig. 4A and B). Based on these results, we propose that the silencer is remodeled until late stages of erythroid differentiation. In an analysis of the region between the silencer and enhancer through MspI and HhaI digestion and over the enhancer through PstI (P1 and P2) (Fig. 4B), we detected a gradual increase of enzyme accessibility consistent with the DNase I- and MNase-HS results (Fig. 4B). Interestingly, the second and most 3' MspI site (M2) located outside of the enhancer shows a constant and weak accessibility in 6C2 and HD3 cells and 10d-RBCs, suggesting that the chromatin remodeling is restricted solely to the enhancer sequence (Fig. 4 and data not shown).

To better understand the differential chromatin remodeling

on the silencer-enhancer region in an in vivo context, we surveyed restriction enzyme accessibility during chicken development by incubating DraIII and MspI restriction enzymes on isolated nuclei from embryonic primary red blood cells (Fig. 4B). The silencer chromatin, which is digested by DraIII, is more accessible during embryonic developmental days 5 to 8 (Fig. 4B), despite the possible well-positioned nucleosome (Fig. 4A). In contrast, restricted accessibility is concomitant with the switch from embryonic to adult gene expression (Fig. 4B, 10-d to 15d-RBCs). When we analyzed the region between the silencer and enhancer through MspI digestion, we found a constant degree of accessibility coincident with the fact that adult genes are transcribed (Fig. 4B, M1 site). Similar results were obtained using the HpaII restriction enzyme (data not shown).

In summary, in situ restriction enzyme accessibility assays

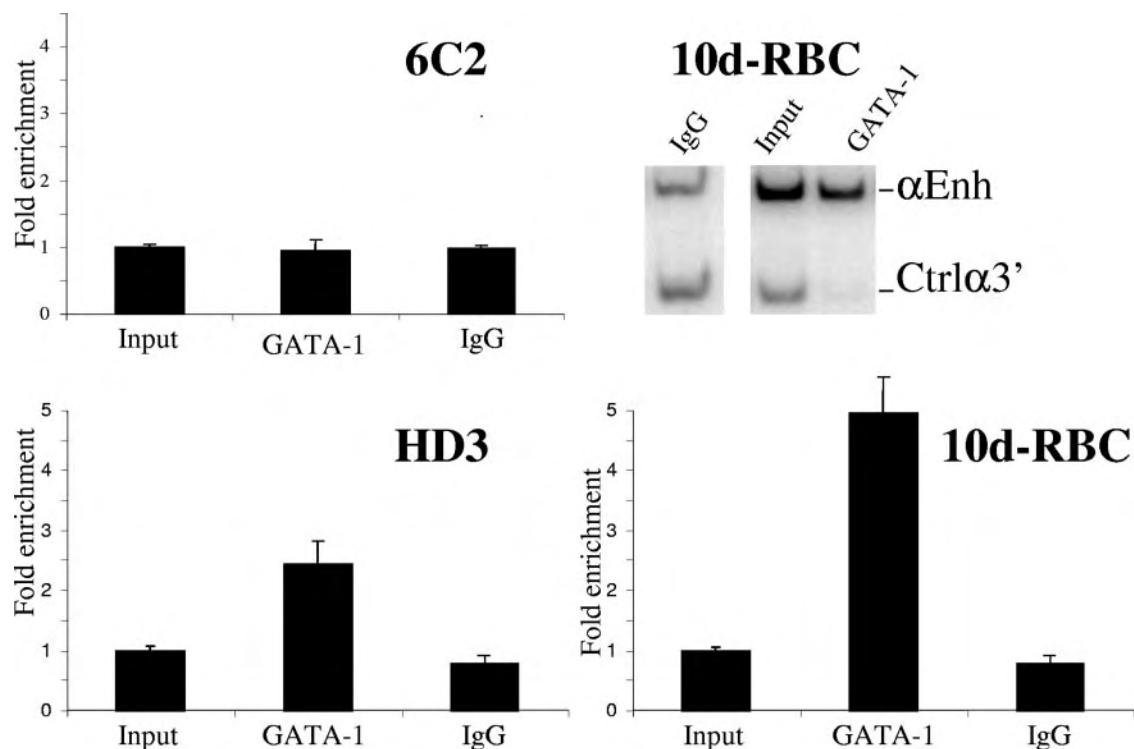


FIG. 5. In vivo interaction of the transcriptional factor GATA-1 with the chicken 3' enhancer  $\alpha$ -globin. The histograms summarize the results of semiquantitative duplex PCR performed in triplicate from two independent ChIPs. A representative gel of 10d-RBC ChIP is shown.  $\alpha$ Enh represents the PCR amplification products covering the silencer-enhancer sequences and Ctrl $\alpha$ 3' represents the corresponding normalization product from a distal 3' noncoding sequence located 3 kb downstream of the enhancer (16). IgG, immunoglobulin G.

confirmed the progressive changes in chromatin conformation. These results are in agreement with our initial hypothesis that predicts that a fully active enhancer is subjected to extensive local chromatin remodeling.

**Mapping of the erythrocyte-specific MNase I-HSs and in vivo differential binding of GATA-1.** The appearance of two prominent erythrocyte-specific MNase-HSs prompted us to map more precisely their position in relation to the silencer-enhancer sequences. By Southern blotting and different restriction enzyme digestions, we determined that the 5' MNase HSs is located over the silencer and the most prominent 3' MNase-HS was found between the first and second GATA-1 binding sites of the core enhancer encompassing a DNA fragment of  $100 \pm 20$  bp (Fig. 4A and data not shown). Furthermore, we determined the in vitro GATA-1 interaction in this central motif (data not shown). To demonstrate GATA-1 association at the enhancer in vivo, we performed a semiquantitative ChIP assay (Fig. 5). Interestingly, we found a gradual incorporation of GATA-1 to the enhancer, in which the highest amount of factor coincides with the peak of enhancer activity. These results suggest a possible correlation between GATA-1 association and the degree of enhancer chromatin remodeling and activity.

**GATA-1 binding is critical for the enhancer chromatin structure conformation.** Based on these observations and the known contribution of GATA-1 to enhancer activity, we determined by use of point mutations that the interactions of GATA-1 at the first and second GATA-1 binding sequences

are critical for enhancer activity in transient transfection assays (data not shown). To determine whether GATA-1 binding sites in the core of the 3'-side chicken  $\alpha$ -globin enhancer contribute in vivo to the establishment of its chromatin organization, we used a RMCE strategy, which allows comparison of the activities of different constructs within the same genomic insertion site (13, 14). First, we generated three independent stable clones with a single copy of the pL1-HYTK-L2 vector. Clones were generated in HD3 cells and named RMCE-HD3-13, -14, and -20. We then constructed a new series of vectors, including an intact silencer-enhancer with and without mutated GATA-1 sites and the adult  $\alpha^D$  globin gene promoter driving *GFP* gene expression (Fig. 6A). A portion of the pEGF-1 plasmid was left on the 3' side to allow hybridization of a specific probe (Fig. 6A). Inverted LoxP sequences (Cre recombinase target sites) were placed flanking the  $\alpha^D$  promoter, the *EGFP* reporter gene, and the silencer-enhancer element to induce cassette exchange through Cre recombinase transient expression. Nonrecombinant clones were discarded by negative selection with ganciclovir (14). It is important to outline that such a system allowed us to eliminate the influence of selection marker genes on stably integrated constructs. Once we established the RMCE experimental system, we were able to assess the contribution of the GATA-1 sites of the core enhancer to the local chromatin structure, as well as the enhancer-mediated transcriptional activation.

With this experimental system, we were able to recapitulate the endogenous chromatin structure pattern of the 3'-side

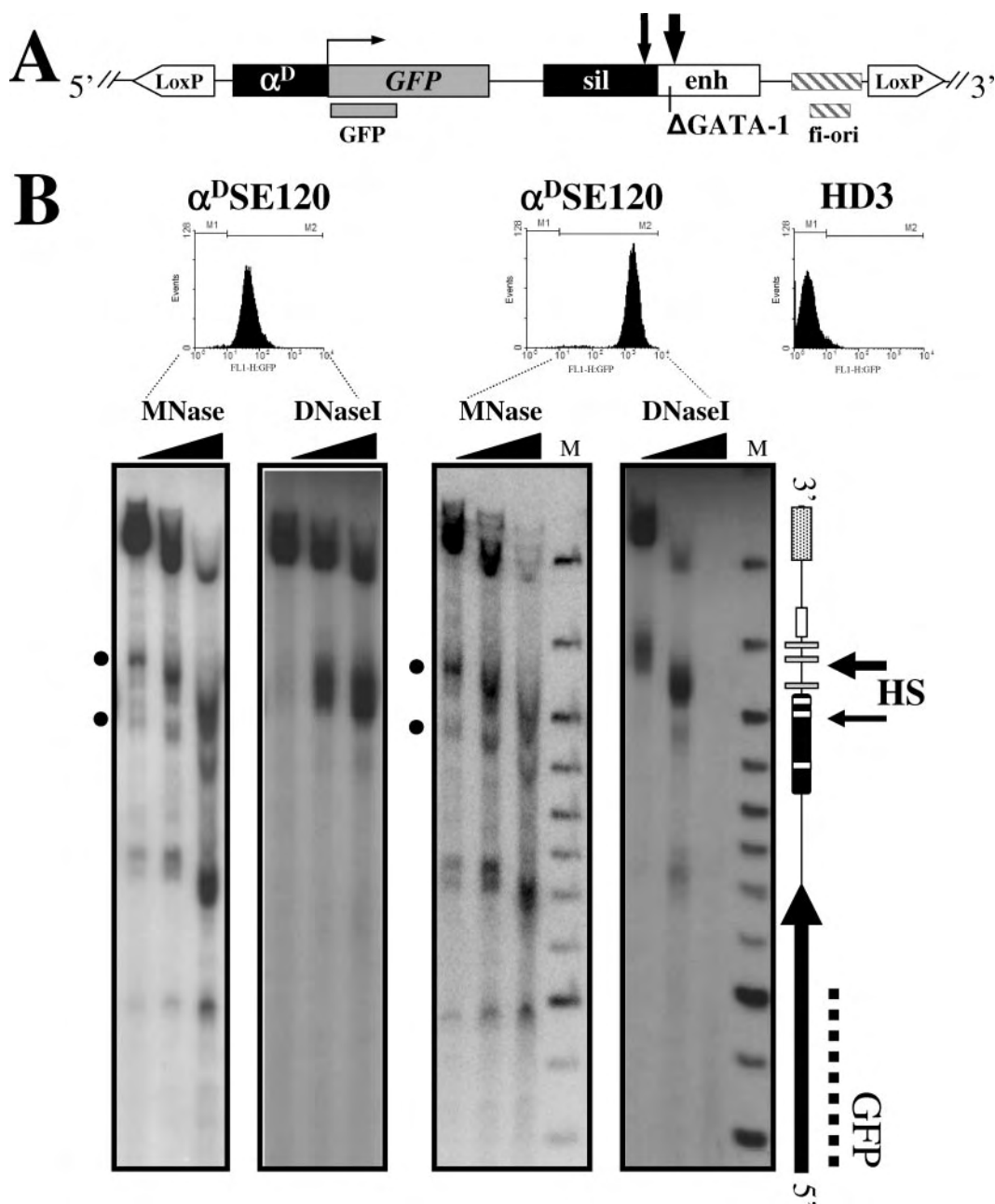


FIG. 6. Reconstitution of the chromatin structure at the silencer-enhancer in stably transfected HD3 cell lines. (A) Transgene flanked by opposing LoxP sequence-specific homologous recombination sites. The specific probes used for Southern blot hybridization are shown (GFP and fi-ori). Arrows represent the HS determined. (B) RMCE was used to direct the insertion of different constructs into the same integration RMCE-HD3-13 site (in the absence of a selectable marker). MNase- and DNase I-HSs assays were performed with individual clones, showing recapitulation of the endogenous chromatin structure over the silencer-enhancer. *EGFP* expression was evaluated by fluorescent flow cytometry. It is important to note that in the presence of the silencer ( $\alpha^D$ SE120), there is a remarkable enhancement of gene expression and the reconstitution of the endogenous chromatin structure. Black dots indicate the positions of the two MNase-HSs described previously in the endogenous context.

$\alpha$ -globin silencer-enhancer elements (compare Fig. 6B with Fig. 3). Unexpectedly, and contrary to our previously published in vitro results (31), when the silencer is incorporated to the construct we found a robust enhancement of the *GFP* expression (Fig. 7, top; compare pL $\alpha^D$ E120 with pL $\alpha^D$ SE120). Based on these results, we incorporated in this study a construct lacking the silencer (Fig. 7,  $\alpha^D$ E120). In the absence of the silencer, we observed a reduction in *GFP* gene expression

along with changes in chromatin structure (Fig. 7; compare  $\alpha^D$ E120 with  $\alpha^D$ SE120). The strong 3' MNase-HS is diminished, whereas the DNase I-HS remains. We attribute the permanence of this DNase I-HS to the presence of GATA-1 binding sites. These data are consistently reproducible in the context of the two other integration sites (see Fig. S4 and S5 in the supplemental material). In contrast to our earlier results, the silencer is not behaving as a negative regulator in the

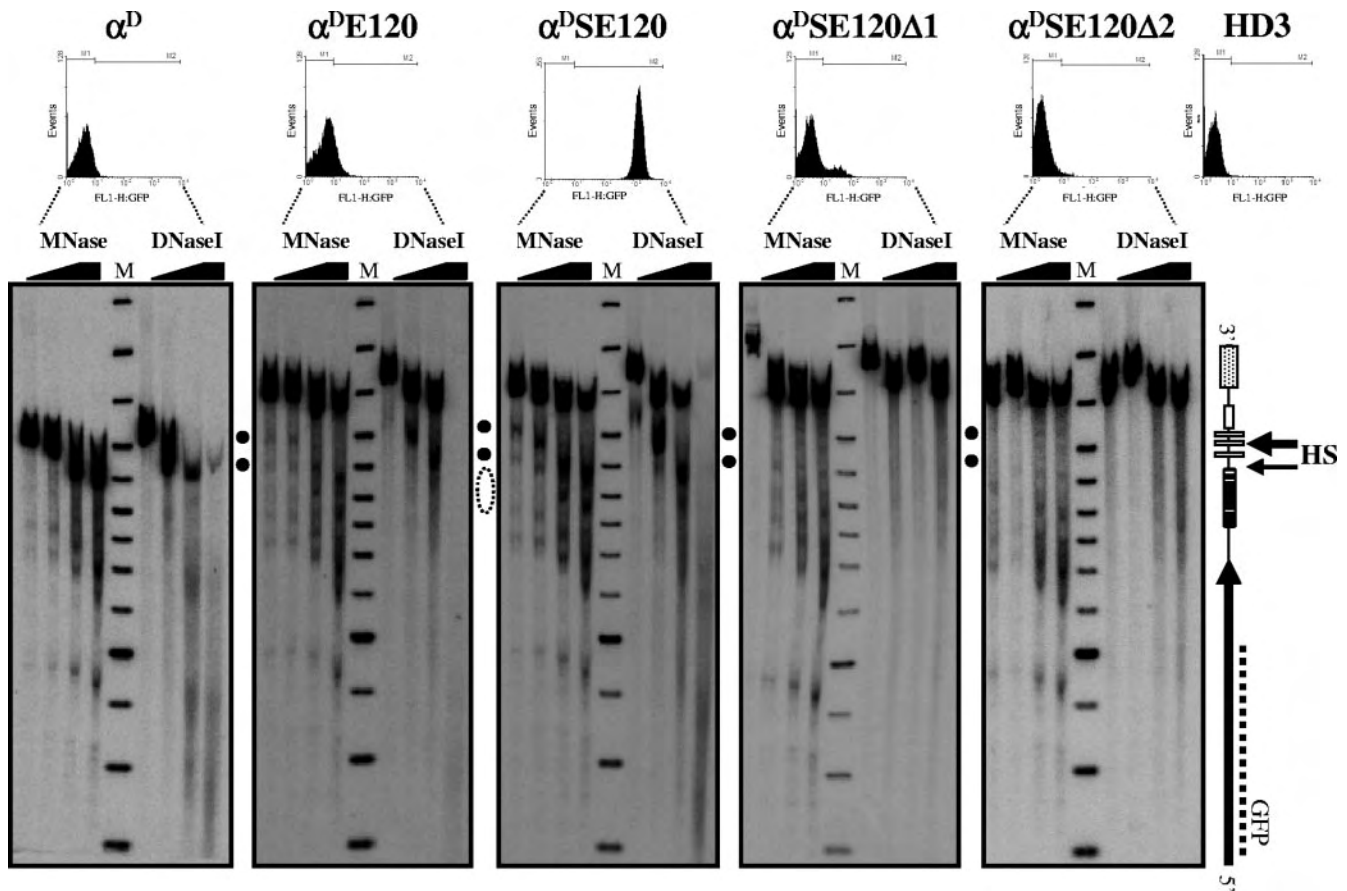


FIG. 7. Mutational analysis of the core enhancer GATA-1 binding sites and their effect on gene expression and chromatin structure. The transcriptional activity of the construct containing the  $\alpha^D$  promoter ( $\alpha^D$ ),  $\alpha^D$  plus the enhancer (core enhancer and the 120-bp DNA fragment; shown as  $\alpha^D E120$ ), and the silencer-enhancer at the RMCE-HD3-13 integration site was tested by fluorescence flow cytometry. DNase I- and MNase-HSs assays are shown in the lower panel. In addition, mutations of the first and second GATA-1 binding sites at the core enhancer were analyzed ( $\alpha^D SE120 \Delta G1$  and  $\alpha^D SE120 \Delta G2$ ). Black circles show the positions of the two HSs. The dotted oval shows the positioning of the static nucleosome that is lost in the GATA-1 point mutation clones.

integrated context. Instead, it seems to be an integral component of the enhancer.

Analysis of constructs harboring GATA-1 point mutations revealed that the first ( $\Delta G1$ ) and second ( $\Delta G2$ ) GATA-1 binding sites are critical for enhancer activity and, most importantly, that their mutation caused the complete loss of DNase I-HSs (Fig. 7; also see Fig. S5 in the supplemental material). In addition, no reporter gene expression was detected from either construct by fluorescence-activated cell sorting analysis (Fig. 7). With respect to the nucleosome positioning, the general chromatin structure over the region was completely altered, with the loss of the 3' MNase-HS, characteristic of all the erythroid cells tested. These observations are reproducible in the context of the RMCE-HD3-14 and -20 integration sites (see Fig. S4 and S5 in the supplemental material; also data not shown).

In summary, we have been able to recreate the native chromatin structure pattern of the silencer-enhancer in distinct integration sites in HD3 cells. Furthermore, the previously defined silencer seems to contribute to the overall enhancer activity, possibly through the stable positioning of a nucleosome (see Fig. S1 to S3 in the supplemental material). But

most importantly, the first and second GATA-1 binding sequences are responsible for the acquisition of a chromatin structure permissive for enhancer activity.

**Histone posttranslational modifications over the  $\alpha$ -globin 3'-side regulatory element.** Covalent histone modifications, particularly histone acetylation and methylation, regulate chromatin structure and function in intimate association with transcription factors. We therefore performed a series of semi-quantitative ChIP assays to address the dynamics of histone modification on the complete  $\alpha$ -globin 3' enhancer during erythroid differentiation (Fig. 8). Chromatin from DT40 and HD3 cells and 7d-RBCs was immunoprecipitated using specific antibodies against H3K4me2, acH3, and acH4 as representatives of open chromatin and H4K20me3 as a representative of a repressive modification (Fig. 8A). Furthermore, we also surveyed the region for the *in vivo* association of RNA Pol II and verified the presence of nucleosomes employing a panmonoclonal antibody which is able to recognize all five histones (Fig. 8A). Interestingly, for HD3 cells we found both activating and repressive marks (Fig. 8A). We also found RNA Pol II present predominantly on the 3' of the enhancer (Fig. 8A). For 7d-RBCs, in which embryonic  $\pi$  gene expression switches to adult

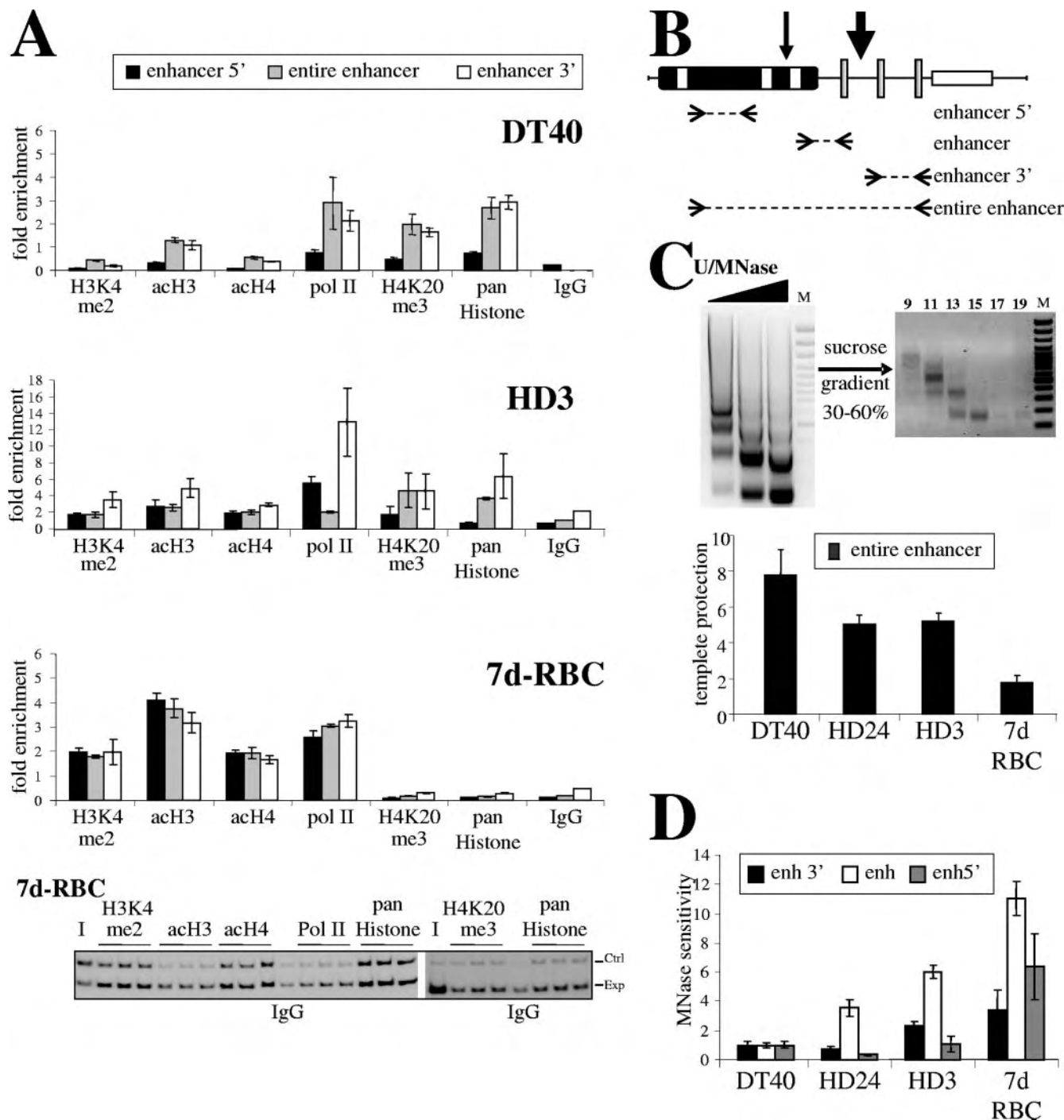


FIG. 8. Histone modification analysis by ChIP assays and identification of a potential nucleosome-free region over the enhancer. (A) The histograms summarize the results of quantitative duplex PCR performed in triplicate from two independent immunoprecipitations. We amplified the 5' part of the enhancer (enhancer 5') and the 3' part (enhancer 3') as well as the entire enhancer (see scheme in panel B). A representative gel of 7d-RBC ChIP is shown, with each PCR performed in triplicate. Exp represents the PCR amplification products of the entire enhancer and Ctrl represents the corresponding normalization primers. For normalization, the central part of the 16 kb of condensed chromatin (for open marks), located upstream to the chicken  $\beta$ -globin domain, and the *FR* gene HSA sequences (for repressive marks) were used. (B) Localization of the primers used as shown in panels A and C. (C) Purification of mono- and dinucleosomes by use of a sucrose gradient. Numbers represent the sucrose fractions. DNA associated with mono- and dinucleosomes was used as the template for PCR amplification to determine the degree of template protection as a measure of nucleosome depletion. The normalization was done as described for panel A. The difference ( $n$ -fold) between genomic DNA and MNase-digested input DNA was determined. (D) Duplex PCR was performed with DNA associated with mono- and dinucleosomes isolated by sucrose gradient treatment to determine the MNase sensitivity as a measure of nucleosome depletion. The normalization was done as described for panel A. The difference ( $n$ -fold) between genomic DNA and MNase-digested input DNA was determined. Data are presented as increase ( $n$ -fold) in MNase sensitivity with respect to that seen for DT40.



$\alpha^D$  and  $\alpha^A$  gene expression (16), we found enrichment of open chromatin marks, in agreement with our previous observations. In addition, we observed a redistribution of RNA Pol II, which was equally present across the enhancer. We note that these data are still consistent with the recruitment of RNA Pol II and components of the preinitiation complex to distal regulatory elements (22, 39, 40). Unexpectedly, the general levels of immunoprecipitated chromatin were low in 7d-RBCs in comparison to what was seen for DT40 and HD3 cells, suggesting a diminishment or even absence of nucleosomes in the regions studied, as supported by the pan-histone monoclonal antibody ChIP. For DT40 nonerythroid cells, we observed the absence of open chromatin marks, a modest presence of RNA Pol II, enrichment of histone H4K20me3, and the presence of histones (Fig. 8A). Overall, these results are in agreement with the progressive changes in the chromatin structure of the enhancer and its associated activity.

To further assess the observed nucleosome depletion, MNase-digested chromatin was purified and mono- and dinucleosomes were isolated by use of a sucrose gradient (Fig. 8C). Duplex PCR was used for normalization, with primers covering the entire enhancer and the chicken *FR* gene HSA (27) (Fig. 8B; also see Materials and Methods). These results were clearer when we analyzed individual parts of the enhancer with primers used in the ChIP analysis (Fig. 8A and C) with the chicken *FR* gene HSA as a control for duplex PCR normalization (27). The entire enhancer region shows a gradual increase in MNase sensitivity (Fig. 2 to 5), but higher levels of sensitivity are associated with the central region, for which we demonstrated a GATA-1-dependent chromatin remodeling (Fig. 8D; data are presented as enrichment [*n*-fold] of MNase sensitivity [see Materials and Methods]).

Together, the ChIP results support a progressive chromatin remodeling during erythroid differentiation. Furthermore, in 7d-RBCs, a developmental stage in which the enhancer is more active, there is an apparent nucleosomal depletion over the 3'-side  $\alpha$ -globin enhancer.

## DISCUSSION

The contribution of the local chromatin structure to the enhancer's method of action is not well understood. Our hypothesis predicts that characteristic chromatin structures over enhancers and LCRs are needed for their function and long-range interaction with promoters. In the present work, we assessed the role of the chromatin structure and GATA-1-dependent remodeling on the activity of the 3'  $\alpha$ -globin silencer-enhancer during erythroid differentiation. A systematic characterization of the chromatin structure of this approximately 1-kb element established the erythrocyte-specific appearance of two characteristic HSs located over the enhancer and the transition point between the silencer and the enhancer. In the later erythroid differentiation stages, the silencer and the enhancer acquire the most accessible (relaxed) chromatin conformation. We also observed chromatin remodeling over a 120-bp region previously proposed to be responsible for the modulation of the enhancer activity in the stages during which the enhancer is active and the adult  $\alpha$ -globin genes are expressed (33). Fine mapping of the two HSs showed that a GATA-1 binding site is located between them. These chroma-

tin-based studies demonstrated that the first and second GATA-1 binding sequences are critical for the proper chromatin remodeling of the silencer-enhancer elements and for enhancer activity. Unexpectedly and in contrast to our previous transient transfection results, in the integrated context, the silencer contributes to the enhancer activity. Finally, at the stage in which the enhancer is most active, we observed the formation of an apparent nucleosome-free zone. Together, our results demonstrate that local chromatin structure plays essential role in enhancer activity and that GATA-1 is fundamental for the formation of a characteristic erythrocyte-specific chromatin conformation and for regulated enhancer function. GATA-1 and GATA-2 are expressed in megakaryocytic and erythroid lineages, playing crucial roles in cell maturation and differentiation (8, 29). GATA-1 possesses a fundamental function in the control of transcriptional programs during development (11, 12) and interacts with a plethora of transcription factors, such as YY1 (33) and FOG-1 (8). GATA-1, like several other transcriptional factors, is able to induce transcriptional activation or repression in erythroid cells (34). Moreover, it is well documented that GATA-1 interacts stably with chromatin at promoters and enhancers and functions at the level of chromatin structure (2, 20, 37) and local chromatin remodeling (6, 7, 25). Indeed, in a reconstituted experimental system, GATA-1 perturbs the nucleosome, generating an apparently nucleosome-free DNA-GATA-1 complex responsible for HS formation (7).

An interesting aspect of our results is the observation of a nucleosome positioned over the previously defined silencer element, which we now find contributes to enhancer activity (see Fig. S1 to S3 in the supplemental material). It is possible that the nucleosome over this region represents a signal for the regulated attraction of GATA-1 and associated chromatin remodeling of the  $\alpha$ -globin enhancer. We believe that GATA-1 binding to the first motif and the generation of an apparent nucleosome-free region depends, at least in part, on such nucleosome positioning. This does not seem so unlikely based on the growing list of examples in which specific and strategically located nucleosomes contribute to the formation of nucleosome-free zones, in particular, over promoter elements (15, 42). In the same way, when we deleted the silencer or mutated the GATA-1 binding site, we observed the delocalization of a putative nucleosome just 5' of the first HS. The delocalization of the nucleosome by GATA-1 point mutation supports a model in which GATA-1 is working as a nucleation center of chromatin remodeling activities.

There has been an increased interest in the existence of a genomic positioning code for nucleosomes (1, 21, 36). It has been proposed that the degree of nucleosome depletion correlates with the level of gene expression and that nucleosome-free regions found for active promoters could also be a novel feature of enhancers (19). Our data are consistent with these observations. There is also an apparent correlation between the maximal degree of nucleosome depletion, the highest 3'-side  $\alpha$ -globin enhancer activity, and local chromatin remodeling (Fig. 3; also see Fig. S3 in the supplemental material). Recent elegant observations from others describe GATA-1 and GATA-2 binding along the mouse  $\alpha$ -globin domain during hematopoiesis that, in collaboration with FOG-1, contributes to regulated long-range intrachromosomal interactions (2, 37,

38). Based on our results, we propose that such regulation is highly dependent on the local chromatin structure of the enhancer. A detailed survey of histone marks and transcription factors and cofactors has suggested that each HS in the human  $\beta$ -globin LCR has distinct structural and functional features (23). These results are accompanied by the demonstration of nucleosome exclusion, and an inverse correlation between GATA-1 binding and nucleosome loss seems to emerge (23). In the case of the chicken  $\alpha$ -globin enhancer, one possible scenario is that the enhancer is behaving like an LCR in terms of the diversity of its genomic organization, chromatin structure, and activity.

In conclusion, we propose a tripartite composition for the regulatory element located at the 3' end of the chicken  $\alpha$ -globin domain. This element possesses a modular organization and has an activity strictly regulated with erythroid differentiation and development. The central component is tightly dependent on GATA-1 *trans*-acting potential, chromatin remodeling, and nucleosome displacement properties. The second, initially called a silencer element, is associated with a well-positioned nucleosome, and the third component is the 120-bp DNA fragment previously demonstrated to modulate the enhancer activity (33). In summary, the GATA-1 binding sites of the  $\alpha$ -globin enhancer are fundamental for the generation of an active chromatin structure and the modulation of the enhancer function in an erythroid environment. At this point, the challenge is to integrate and define how these activities operate *in vivo* and in the context of the  $\alpha$ -globin differential gene expression and chromatin domain organization.

#### ACKNOWLEDGMENTS

We thank Ann Dean, Gary Felsenfeld, David J. Clark, Rodolfo Ghirlando, Mayra Furlan-Magaril, Ernesto Soto-Reyes, and Paul Delgado-Olguín for helpful comments on the manuscript. We thank Georgina Guerrero for her excellent technical assistance and members of the Félix Recillas-Targa laboratory for stimulating scientific discussions and suggestions. We also thank the technical assistance of Héctor Rincón-Arango, Cristina Aranda, Estela García Conzález, and Itzel Escobedo Ávila. We thank L. Ongay, G. Codiz, and M. Mora from the Unidad de Biología Molecular from the Instituto de Fisiología Celular, UNAM, for DNA sequencing and fluorescence-activated cell sorting facility.

This work was supported by grants from the Dirección General de Asuntos del Personal Académico-UNAM (IN203200, IX230104, IN209403, and IN214407), Consejo Nacional de Ciencia y Tecnología, CONACyT (33863-N, 42653-Q, and 58767), and the Third World Academy of Sciences (TWAS, grant 01-055 RG/BIO/LA). M.E.-D.-A was a fellowship recipient from CONACyT.

#### REFERENCES

- Almer, A., H. Rudolph, A. Hinnen, and W. Horz. 1986. Removal of positioned nucleosomes from the yeast PHO5 promoter upon PHO5 induction releases additional upstream activity DNA elements. *EMBO J.* **5**:2689–2696.
- Anguita, E., J. Hughes, C. Heyworth, G. A. Blobel, W. G. Wood, and D. R. Higgs. 2004. Globin gene activation during haemopoiesis is driven by complexes nucleated by GATA-1 and GATA-2. *EMBO J.* **23**:2841–2852.
- Beug, H., A. von Kirchbach, G. Doderlein, J. F. Conscience, and T. Graf. 1979. Chicken hematopoietic cells transformed by seven strains of defective avian leukemia viruses display three distinct phenotypes of differentiation. *Cell* **18**:375–390.
- Beug, H., S. Palmieri, C. Freudenstein, H. Zentgraf, and T. Graf. 1982. Hormone-dependent terminal differentiation *in vitro* of chicken erythroleukemia cells transformed by ts mutants of avian erythroblastosis virus. *Cell* **28**:907–919.
- Blackwood, E. M., and J. T. Kadonaga. 1998. Going the distance: a current view of enhancer action. *Science* **281**:60–63.
- Boyes, J., and G. Felsenfeld. 1996. Tissue-specific factors additively increase the probability of the all-or-none formation of a hypersensitive site. *EMBO J.* **15**:2496–2507.
- Boyes, J., J. Omichinski, D. Clark, M. Pikaart, and G. Felsenfeld. 1998. Perturbation of nucleosome structure by the erythroid transition factor GATA-1. *J. Mol. Biol.* **279**:529–544.
- Bresnick, E. H., M. L. Matowicz, S. Pal, and K. D. Johnson. 2005. Developmental control via GATA factor interplay at chromatin domains. *J. Cell. Physiol.* **205**:1–9.
- Bulger, M., and M. Groudine. 1999. Looping versus linking: towards a model for long distance gene activation. *Genes Dev.* **13**:2465–2477.
- Chakalova, L., E. Debrand, J. A. Mitchell, C. S. Osborne, and P. Fraser. 2005. Replication and transcription: shaping the landscape of the genome. *Nat. Rev. Genet.* **6**:669–677.
- Chen, E. S., M. Yanagida, and K. Takahashi. 2003. Does a GATA factor make the bed for centromeric nucleosomes? *Cell Cycle* **2**:277–278.
- Cirillo, L. A., F. R. Lin, I. Cuesta, D. Friedman, M. Jarnik, and K. S. Zaret. 2002. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* **9**:279–289.
- Dieken, E. S., E. M. Epner, S. Fiering, R. E. Fournier, and M. Groudine. 1996. Efficient modification of human chromosomal alleles using recombination-proficient chicken/human microcell hybrids. *Nat. Genet.* **12**:174–182.
- Feng, Y. Q., J. Seibler, R. Alani, A. Eisen, K. A. Westerman, P. Leboulch, S. Fiering, and E. E. Bouhassira. 1999. Site-specific chromosomal integration in mammalian cells: highly efficient CRE recombinase-mediated cassette exchange. *J. Mol. Biol.* **292**:779–785.
- Gal-Yam, E. N., S. Jeong, A. Tamay, G. Egger, A. S. Lee, and P. A. Jones. 2006. Constitutive nucleosome depletion and ordered factor assembly at the *GRP78* promoter revealed by single molecule footprinting. *PLoS Genet.* **2**:1451–1463.
- Guerrero, G., P. Delgado-Olguín, M. Escamilla-Del-Arenal, M. Furlan-Magaril, E. Rebollar, I. A. De La Rosa-Velázquez, E. Soto-Reyes, H. Rincón-Arango, C. Valdes-Quezada, V. Valadez-Graham, and F. Recillas-Targa. 2006. Globin genes transcriptional switching, chromatin structure and linked lessons to epigenetics in cancer: a comparative overview. *Comp. Biochem. Physiol. A* **147**:750–760.
- Hahn, S. 2004. Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* **11**:394–402.
- Hebbes, T. R., A. L. Clayton, A. W. Thorne, and C. Crane-Robinson. 1994. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken  $\beta$ -globin chromosomal domain. *EMBO J.* **13**:1823–1830.
- Heintzman, N., R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. B. Hawkins, L. O. Barrera, S. V. Calcar, K. A. Ching, W. Wang, Z. Weng, R. D. Green, G. E. Crawford, and B. Ren. 2007. Distinct and predictive chromatin signature of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39**:311–318.
- Im, H., J. A. Grass, K. D. Johnson, S. Kim, M. E. Bayer, A. N. Imbalzano, J. J. Bieker, and E. H. Bresnick. 2005. Chromatin domain activation via GATA-1 utilization of a small subset of dispersed GATA motifs within a broad chromosomal region. *Proc. Natl. Acad. Sci. USA* **102**:17065–17070.
- Ioshikhes, I. P., I. Albert, S. J. Zanton, and B. F. Pugh. 2006. Nucleosome positions predicted through comparative genomics. *Nat. Genet.* **10**:1210–1214.
- Johnson, K. D., J. A. Grass, M. E. Boyer, C. M. Kiekhafer, G. A. Blobel, M. J. Weiss, and E. H. Bresnick. 2002. Cooperative activities of hematopoietic regulators recruit RNA polymerase II to a tissue-specific chromatin domain. *Proc. Natl. Acad. Sci. USA* **99**:11760–11765.
- Kim, A., S. H. Song, M. Brand, and A. Dean. 2007. Nucleosome and transcription activator antagonism at human  $\beta$ -globin locus control region DNase I hypersensitive sites. *Nucleic Acids Res.* **35**:5831–5838.
- Kutney, S. N., R. Hong, T. Macfarlan, and D. Chakravarti. 2004. A signaling role of histone-binding proteins and INHAT subunits pp32 and set/TAF- $\beta$  in intergenic chromatin hypoacetylation and transcriptional repression. *J. Biol. Chem.* **279**:30850–30855.
- Layon, M. E., C. J. Ackley, R. J. West, and C. H. Lowrey. 2007. Expression of GATA-1 in a non-hematopoietic cell line induces  $\beta$ -globin locus control region chromatin structure remodeling and an erythroid pattern of gene expression. *J. Mol. Biol.* **366**:737–744.
- Ling, J. Q., T. Li, J. F. Hu, T. H. Va, H. L. Chen, X. W. Qiu, A. M. Cherry, and A. R. Hoffman. 2006. CTCF mediates interchromosomal colocalization between *Igf2/H19* and *Wsb1/Nf1*. *Science* **312**:269–272.
- Litt, M. D., M. Simpson, F. Recillas-Targa, M. N. Prioleau, and G. Felsenfeld. 2001. Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *EMBO J.* **20**:2224–2235.
- Mellor, J. 2005. The dynamics of chromatin remodeling at promoters. *Mol. Cell* **19**:147–157.
- Morceau, F., M. Schnekenburger, M. Dicato, and M. Diederich. 2004. GATA-1: friends, brothers and coworkers. *Ann. N. Y. Acad. Sci.* **1030**:537–554.
- Prioleau, M. N., P. Nony, M. Simpson, and G. Felsenfeld. 1999. An insulator element and condensed chromatin region separate the chicken  $\beta$ -globin locus from an independently regulated erythroid-specific folate receptor gene. *EMBO J.* **18**:4035–4048.
- Recillas-Targa, F., C. V. de Moura Gallo, M. Huesca, K. Scherrer, and L.

- Marcaud. 1993. Silencer and enhancer elements located at the 3'-side of the chicken and duck  $\alpha$ -globin-encoding gene domains. *Gene* **129**:229–237.
32. Recillas-Targa, F. 2000. The chicken  $\alpha$ - and  $\beta$ -globin domains and their chromatin organization. *Cell. Mol. Biol. Lett.* **5**:451–467.
33. Rincón-Arango, H., V. Valadez-Graham, G. Guerrero, M. Escamilla-Del-Arenal, and F. Recillas-Targa. 2005. YY1 and GATA-1 interaction modulate the chicken 3'-side  $\alpha$ -globin enhancer activity. *J. Mol. Biol.* **349**:961–975.
34. Rodríguez, P., E. Bonte, J. Krijgsveld, K. E. Kolodziej, B. Guyot, A. J. R. Heck, P. Vyas, E. deBoer, F. Grosveld, and J. Strouboulis. 2005. GATA-1 forms distinct activating and repressive complexes in erythroid cells. *EMBO J.* **24**:2354–2366.
35. Sawado, T., K. Igarashi, and M. Groudine. 2001. Activation of  $\beta$ -major globin gene transcription is associated with recruitment of NF-E2 to the  $\beta$ -globin LCR and gene promoter. *Proc. Natl. Acad. Sci. USA* **98**:10226–10231.
36. Segal, E., Y. Fondufe-Mittendorf, L. Chen, A. Thastrom, Y. Field, I. K. Moore, J.-P. Z. Wang, and J. Widom. 2006. A genomic code for nucleosome positioning. *Nature* **442**:772–778.
37. Vakoc, C. R., D. L. Letting, N. Gheldof, T. Sawado, M. A. Bender, M. Groudine, M. J. Weiss, J. Dekker, and G. A. Blobel. 2005. Proximity among distant regulatory elements at the  $\beta$ -globin locus requires GATA-1 and FOG-1. *Mol. Cell* **17**:453–462.
38. Vernimmen, D., M. De Gobbi, J. A., Sloane-Stanley, W. G. Wood, and D. R. Higgs. 2007. Long-range chromosomal interactions regulate the timing of the transition between paused and active gene expression. *EMBO J.* **26**:2041–2051.
39. West, A. G., and P. Fraser. 2005. Remote control of gene transcription. *Hum. Mol. Genet.* **14**:R101–R111.
40. Zhao, H., and A. Dean. 2004. An insulator block spreading of histone acetylation and interferes with RNA polymerase II transfer between an enhancer and gene. *Nucleic Acids Res.* **32**:4903–4919.
41. Zhou, J., C. M. Chau, Z. Deng, R. Shiekhatter, M.-P. Spindler, A. Schepers, and P. M. Lieberan. 2005. Cell cycle regulation of chromatin at an origin of DNA replication. *EMBO J.* **24**:1406–1417.
42. Zou, Y., Q. Yu, and X. Bi. 2006. Asymmetric positioning of nucleosomes and directional establishment of transcriptionally silent chromatin by *Saccharomyces cerevisiae* silencers. *Mol. Cell. Biol.* **26**:7806–7819.

## YY1 and GATA-1 Interaction Modulate the Chicken 3'-Side $\alpha$ -Globin Enhancer Activity

Héctor Rincón-Arano<sup>†</sup>, Viviana Valadez-Graham<sup>†</sup>, Georgina Guerrero Martín Escamilla-Del-Arenal and Félix Recillas-Targa\*

*Instituto de Fisiología Celular  
Departamento de Genética  
Molecular, Universidad  
Nacional Autónoma de México  
Apartado Postal 70-242, México  
D.F. 04510, México*

Studying the chicken  $\alpha$ -globin domain as a model system of gene regulation, we have previously identified contiguous silencer-enhancer elements located on the 3'-side of the domain. To better characterize the enhancer we performed a systematic functional analysis to define its expression influence range and the ubiquitous and stage-specific transcriptional regulators interacting with this control element. In contrast to previous reports, we found that, in addition to a core element that includes three GATA-1 binding sites, the enhancer incorporates a 120 base-pair DNA fragment where EKLF, NF-E2 and a fourth GATA-1 factor could interact. Functional experiments demonstrate that the enhancer activity over the adult  $\alpha^D$  promoter is differentially regulated. We found that the transcriptional factor Ying Yang 1 (YY1) binds to the 120 base-pair DNA fragment and its effect over the enhancer activity is GATA-1-dependent. In addition, we characterize a novel physical interaction between GATA-1 and YY1 that influences the enhancer function. Experiments using a histone deacetylation inhibitor indicate that, in pre-erythroblasts, the enhancer down-regulation could be influenced by a closed chromatin conformation. Our observations show that the originally defined enhancer possesses a more complex composition than previously assumed. We propose that its activity is modulated through differential nuclear factor interactions and chromatin modifications at distinct erythroid stages.

© 2005 Elsevier Ltd. All rights reserved.

\*Corresponding author

**Keywords:** globin genes; enhancer; YY1; GATA-1; erythroid differentiation

### Introduction

Eukaryotic gene expression requires precise coordination between tissue and stage-specific regulatory factors in the nuclear and chromatin structural environments.<sup>1,2</sup> Promoters and long distance regulatory elements are only partly responsible for such regulation. The globin genes have been an attractive research topic over the years, particularly the chicken  $\alpha$  and  $\beta$ -globin gene clusters because they are subjected to sophisticated

transcriptional regulation, presumably achieved through diverse regulatory elements including enhancers, silencers, locus control regions (LCRs) and insulators.<sup>1,3–6</sup> The study of gene families is based on differential gene expression with physico-chemical properties specific to each developmental stage.<sup>4,5</sup>

Locus control regions and enhancers are thought to regulate gene expression *via* one of two basic models: the binary or the progressive model.<sup>7</sup> Recently, two new experimental strategies demonstrated that the mechanism of action of the mammalian  $\beta$ -globin LCR involves chromatin loop formation with multiple physical contacts in close proximity to individual promoters and the DNase I hypersensitive sites that are located at long range.<sup>2,8–11</sup>

As in the human and mouse  $\alpha$ -globin loci, the chicken  $\alpha$ -globin domain is located in a constitutively open chromatin configuration.<sup>3,4</sup> This lack of domain shielding, like that seen in the chicken  $\beta$ -globin domain, particularly by specialized

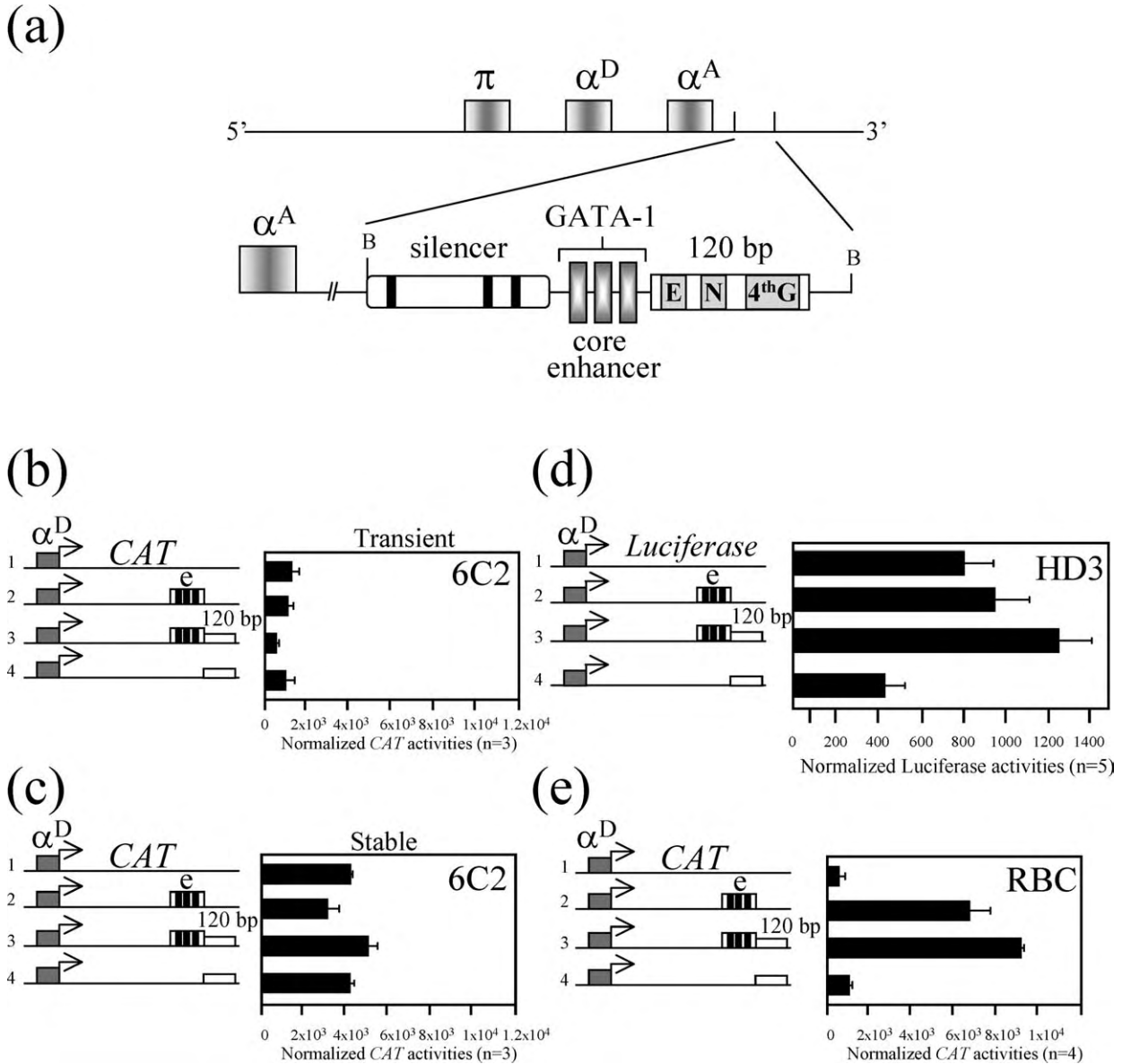
<sup>†</sup> H.R.-A. and V.V.-G. contributed equally to this work.

Abbreviations used: LCR, locus control region; EKLF, erythroid Krüppel-like factor; YY1, Ying Yang 1 nuclear factor; FACS, fluorescence activated cell sorting; CAT, chloramphenicol acetyltransferase; RBC, red blood cells; NF-E2, nuclear factor erythroid 2; GST, glutathione S-transferase; ChIP, chromatin immunoprecipitation; TSA, trichostatin-A.

E-mail address of the corresponding author: frecilla@ifc.unam.mx

chromatin structures like insulators, suggests a requirement for tighter control in developmental regulation of gene expression.<sup>3-5</sup> Here, we are particularly interested in the functional characterization of the 3'-side chicken  $\alpha$ -globin enhancer during distinct erythroid differentiation stages.

Previous work has identified silencer-enhancer elements in the 3'-side of the chicken  $\alpha$ -globin locus, around 450 base-pairs (bp) downstream of the adult  $\alpha^A$  globin gene.<sup>12,13</sup> Knezetic & Felsenfeld identified three GATA-family binding sites, as sufficient for the enhancer activity.<sup>12</sup> Alignment of duck and



**Figure 1.** Scheme of the chicken  $\alpha$ -globin gene domain and functional characterization of the 3' chicken  $\alpha$ -globin enhancer. (a) The silencer-enhancer control elements are located around 450 bp downstream of the adult  $\alpha^A$  gene. B represents BamHI restriction sites that liberate a 1036 bp DNA fragment. The enhancer is located adjacent to the silencer where at least three DNA-binding sequences for unknown factors have been previously defined (vertical black boxes, SF1, SF2 and SF3).<sup>13,15</sup> The first characterization of the enhancer determined the presence of three GATA-1 binding sites core enhancer.<sup>12,13</sup> From a sequence alignment of the duck and chicken homologous sequences, and DNA-protein interaction experiments we identified an adjacent, highly conserved 120 bp DNA fragment with binding sites for Sp1/EKLF (E), NF-E2 (N) and GATA-1 (4<sup>th</sup>G).<sup>13,14</sup> (b) Plasmids containing different enhancer sub-fragments were transiently transfected using the CAT reporter gene under the control of the chicken adult  $\alpha^D$  promoter; e, core enhancer (containing three GATA-1 sites); 120 bp, the 120 bp DNA fragment (open box). Plasmids were transfected into the transformed chicken pre-erythroblast 6C2 cell line. (c) The same constructs were stably co-transfected with a plasmid containing the *Neo* resistance gene and pools of stable lines were analyzed. Transfection efficiencies were normalized to protein concentrations. (d) The regulatory elements were transferred to a luciferase reporter vector and transiently transfected into HD3 erythroblast cells. (e) Ten day-old chicken embryonic erythrocytes (RBC) were transiently transfected with the same plasmids as in (a). Data shown in histograms represent the average of at least three independent experiments and error bars show the SEM here and in the following transfections.

chicken homologous sequences enabled us to identify a highly conserved 120 bp DNA fragment, downstream of the three GATA-1 motifs. This DNA fragment includes binding sequences for Sp1 family, NF-E2 (a leucine-zipper family member) and GATA-family of transcription factors (Figure 1(a)).<sup>13,14</sup>

To further characterize the biological and functional significance of this highly conserved DNA fragment we addressed two main questions. The first was to determine the contribution of the 120 bp DNA fragment to the full enhancer activity. The second was to define which nuclear factors bind to this region and their effect in the enhancer function during distinct erythroid differentiation stages. Using the chicken  $\alpha^D$  adult globin promoter we were able to demonstrate that the 120 bp DNA fragment may be an integral component of the enhancer. Interestingly, by itself the 120 bp DNA fragment does not possess any enhancer activity. Functional characterization of the enhancer demonstrated that its activity depends on the erythroid stage. DNA-protein and protein-protein *in vivo* interactions allowed us to identify a novel association of the Ying Yang 1 (YY1) nuclear factor in the proximity of the fourth GATA-1 site, contacting physically the erythroid-specific GATA-1 factor. We propose that YY1, with the contribution of other nuclear factors and co-factors, is able to modulate the chicken  $\alpha$ -globin 3'-side enhancer activity.

## Results

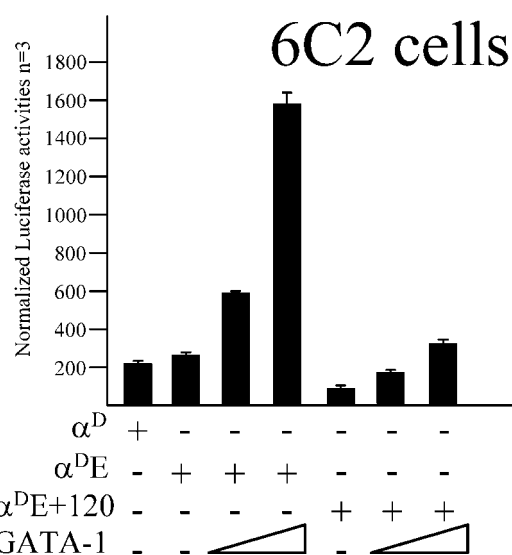
Previous work demonstrated the interaction of the erythroid specific GATA-1 factor with three DNA consensus sequences at the core enhancer (Figure 1(a)).<sup>12-14</sup> Using nuclear extracts from a virally transformed chicken erythroid cell line we have obtained by *in vitro* footprinting the first overview of the entire enhancer that includes the core element and the 120 bp DNA fragment (data not shown). We observed a large protected area covering around 250 bp including three GATA-1 sites and other DNA sequences in addition to the 120 bp DNA fragment (Figure 1(a)). A large complex is formed at the 120 bp DNA fragment that includes the binding sites for EKLF, NF-E2 and GATA-family members (Figure 1(a)).<sup>13,14</sup> This result suggested to us that the chicken  $\alpha$ -globin enhancer may comprise not only GATA-1 binding sites, but also sites for other erythroid and non-erythroid factors.<sup>12</sup> Interestingly, there is a similarity between the number and types of factors interacting with the human  $\alpha$ -globin -40HS upstream regulatory element and the 3'-side chicken  $\alpha$ -globin enhancer.<sup>15,16</sup>

### Functional activity of the enhancer in distinct erythroid stages

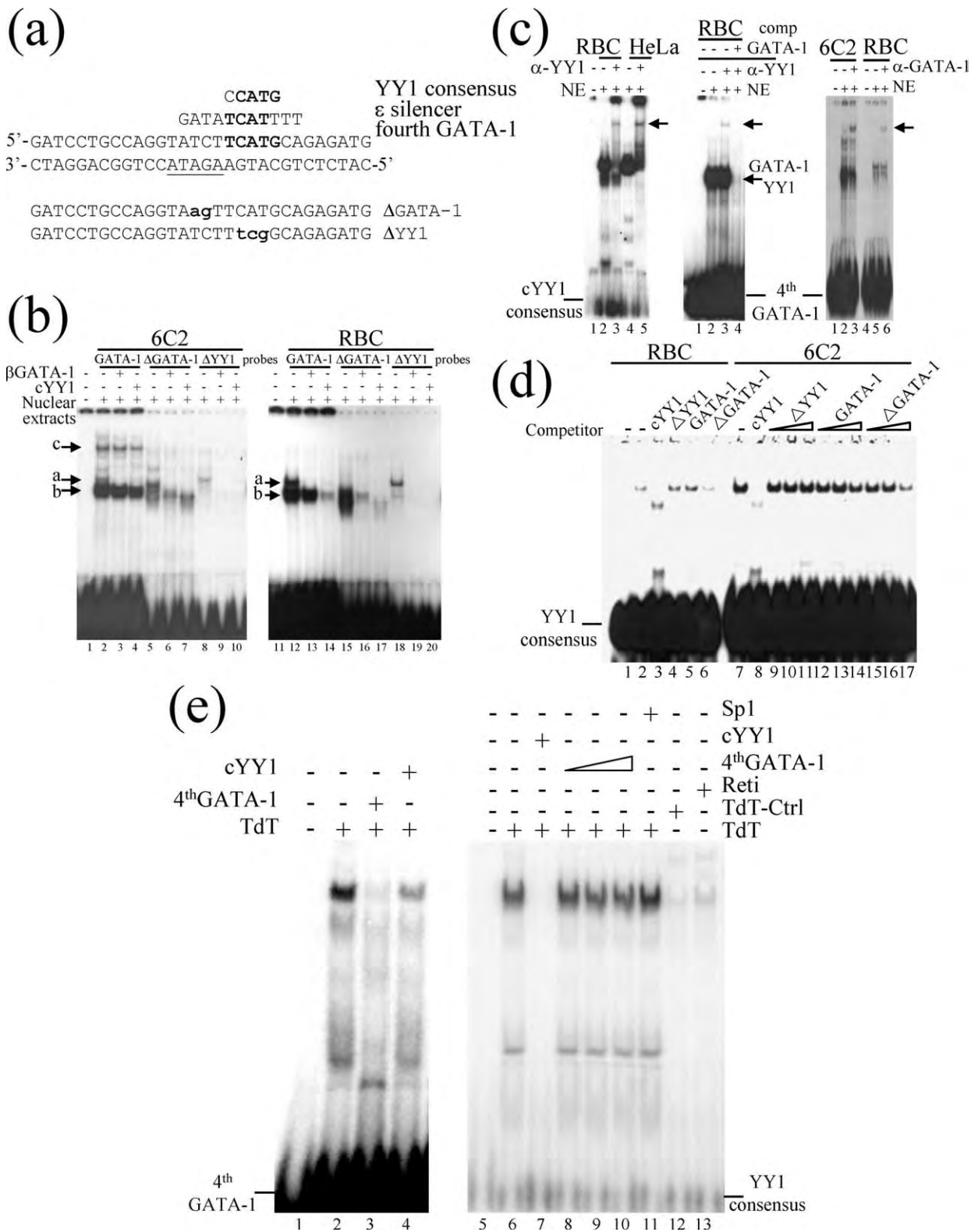
To study the activity of the chicken  $\alpha$ -globin

enhancer at different developmental stages, we used a panel of cell lines and primary cells. The chicken erythroid 6C2 cell line that is arrested at the erythroid colony forming unit erythrocyte stage (CFU-E), is considered as a pre-erythroblast.<sup>17-20</sup> To analyze more advanced erythroid stages we decided to take advantage of the AEV-transformed HD3 cell line that corresponds to an erythroblast and ten day-old circulating embryonic erythrocytes (RBC).<sup>20-22</sup> These cell lines and primary cultures have been used extensively as model systems in numerous studies to examine erythroid-specific gene regulation and chromatin structure in early developmental stages.<sup>17-22</sup>

The effect of the  $\alpha$ -globin enhancer was first analyzed on the chicken adult  $\alpha^D$  promoter driving the expression of the CAT reporter gene in the chicken pre-erythroblast 6C2 cell line (Figure 1(b); construct 1). The enhancer was tested in different combinations in transient transfections and no significant activity in any construct was observed (Figure 1(b)). To exclude the possibility that low transfection efficiency was responsible for this result, we linked the  $\alpha^D$  promoter alone or the  $\alpha^D$  promoter in the presence of the entire enhancer to the green fluorescent protein gene (*GFP*). Transfection efficiencies were evaluated by fluorescence cytometry (FACS) using the viral CMV promoter as a positive control (data not shown). High levels of *GFP* expression were obtained with the control plasmid, leading us to conclude that the lack of activity of the test constructs is not due to the



**Figure 2.** Trans-activation of the chicken 3'-side  $\alpha$ -globin enhancer by GATA-1 over-expression. The chicken GATA-1 cDNA under the control of the CMV promoter was co-transfected with increasing amounts of plasmid (2  $\mu$ g and 4  $\mu$ g), in the presence of the core enhancer composed of the three GATA-1 sites (Figure 1(a);  $\alpha^{DE}$ ) and with the full enhancer that includes the 120 bp DNA fragment ( $\alpha^{DE}+120$ ). Co-transfections were done in 6C2 cells and transcriptional activation was measured with the dual-luciferase assay.



**Figure 3.** DNA-protein interactions at the fourth GATA-1 binding site located on the 120 bp DNA fragment. (a) DNA sequence of the fourth GATA-1 binding site. On the top is shown sequence alignment with the GATA-1/YY1 motif found in the human  $\epsilon$  silencer<sup>25</sup> and one of the YY1 consensus motifs. On the bottom, shown in lower case are the mutations that were introduced. (b) Gel-shift using the fourth GATA-1 sequence (GATA-1), and mutants thereof ( $\Delta$ GATA-1,  $\Delta$ YY1 and see (a)) as labeled probes. The competitors used were a known GATA-1 binding sequence ( $\beta$ GATA-1) from the chicken  $\beta^A/\epsilon$  enhancer<sup>29</sup> and a consensus YY1 sequence (cYY1).<sup>25,26</sup> (c) On the left panel, the consensus YY1 sequence was used as a probe with RBC and HeLa cell nuclear extracts in a supershift assay employing a polyclonal YY1 antibody ( $\alpha$ -YY1). The supershift complexes are shown with an arrow. On the central panel, the fourth GATA-1 motif was radioactively labeled and a supershift is seen using RBC nuclear extracts. To be able to show the supershift complex the

transfection efficiency. We then asked whether the same set of constructs will give different results in the chromatin environment of stably transfected 6C2 cells (Figure 1(c)). Cells were co-transfected with the promoter-less pEGFP-1 plasmid carrying the neomycin resistance cassette (*Neo<sup>r</sup>*). Consistent with the data from transient transfection experiments, we found that the 3'-side  $\alpha$ -globin enhancer is not functional in this cell line (Figure 1(c)). In summary, in a pre-erythroblast transformed cell line the 3'-side enhancer has no influence over the  $\alpha^D$  promoter, consistent with the fact that chicken  $\alpha$ -globin genes are expressed at a basal level in 6C2 cells (G.G. & F.R.-T., unpublished observations). Interestingly, there is considerable transcription activation of a reporter gene by the chicken adult  $\beta^A$  promoter and the  $\beta^A/\epsilon$  enhancer in 6C2 cells, suggesting that the pools of transcription factors are not limiting.<sup>23</sup>

### The 3'-side $\alpha$ -globin enhancer activity in transformed chicken erythroblast HD3 cells

To test the enhancer activity in a more mature erythroblast stage, we transfected the same constructs into the HD3 avian transformed cell line that corresponds to the erythroblast stage.<sup>21,22</sup> For these experiments we used the luciferase assay, which is more sensitive. We observed only a modest increase in luciferase activity for the core and the entire enhancer, the latter of which also contains the 120 bp DNA fragment, in particular when the activity of those constructs is compared to the promoter alone (Figure 1(d); compare constructs 1, 2 and 3). When the 120 bp DNA fragment is tested alone there is even a reduction in the reporter activity (Figure 1(d); compare constructs 1 and 4). We believe that this reduction is not significant since the overall promoter activity is low. Similar results were obtained with cell lines stably expressing these constructs (data not shown). These results demonstrate that the 3'-side chicken  $\alpha$ -globin enhancer shows a modest activity in HD3 cells.

### Transient transfections in ten day-old chicken embryonic erythrocytes

The same constructs tested above were then transiently transfected into chicken ten day-old embryonic erythrocytes (RBC). In this context, a robust activation by the core enhancer and the core enhancer incorporating the 120 bp DNA fragment is

seen (Figure 1(e); constructs 2 and 3). Consistent with our previous observations, the 120 bp DNA fragment alone is not able to give any significant activation to the reporter gene. Together, these results support the idea that the 3'-side chicken  $\alpha$ -globin enhancer is not only composed of the core three GATA-1 sites as previously defined, but in addition suggested that the 120 bp DNA fragment and its associated nuclear factors contribute, in an erythroblast stage manner, to the activity of the enhancer.

### Trans-activation of the 3'-side $\alpha$ -globin enhancer in 6C2 cells by GATA-1

Based on published results showing that transcription and protein levels of GATA-1 are elevated in the early stages of erythroblast differentiation, and that they decrease drastically when red cells are terminally differentiated,<sup>23</sup> we decided to over-express the chicken GATA-1 cDNA in the pre-erythroblast 6C2 cells where the enhancer is not active (see Figure 1(b) and (c)).<sup>23,24</sup> As expected, increased concentrations of chicken GATA-1 cDNA are able to stimulate reporter activity (Figure 2). Interestingly, a clear difference in the *trans*-activation capacity of GATA-1 is observed between the core enhancer, composed of three GATA-1 binding sites ( $\alpha^{DE}$ ), and the full enhancer ( $\alpha^{DE} + 120$ ) formed by the core enhancer and the 120 bp DNA fragment. Higher reporter activity is reached only when the core enhancer is present in the construct. In contrast, when the full enhancer is tested there is just partial *trans*-activation. These results are not dependent on GATA-1 effect over the  $\alpha^D$  promoter alone, since increasing amounts of GATA-1 cDNA do not *trans*-activate the reporter gene expression in 6C2 cells (data not shown). Furthermore, it is important to notice that the incorporation of the 120 bp DNA fragment is sufficient to abate enhancer activity by half (Figure 2; compare  $\alpha^{DE}$  to  $\alpha^{DE} + 120$ ). We interpret these results to indicate that the 120 bp element seems to contain regulatory components that can modulate the enhancer activity.

### The Ying Yang 1 factor interacts with the 120 bp DNA fragment of the enhancer

Previous observations in terminally differentiated erythroblasts showed two retarded complexes when the GATA-1 site located in the 120 bp DNA

---

film was over-exposed, therefore the two retarded bands corresponding to GATA-1 and YY1 are not clearly defined. The GATA-1/YY1 complex and the supershift were efficiently and specifically self-competed with 100-fold molar excess of cold GATA-1 sequence (lane 4). On the right panel, supershift assay using an antibody against GATA-1 ( $\alpha$ -GATA-1) and the fourth GATA-1 was used as radioactive probe. (d) To confirm the interaction of the YY1 factor with the fourth GATA-1 sequence, we performed a gel shift assay using consensus YY1 sequence (cYY1) as labeled probe.<sup>26,27</sup> Cold competitions are shown where cYY1 corresponds to the self-competition and the other sequence competitors have already been described in (b). (e) Gel-shift using YY1 protein obtained by *in vitro* transcription/translation (TdT) using the fourth GATA-1 (4<sup>th</sup>GATA-1) and YY1 consensus sequences as labeled probes. Reti, represents reticulocyte lysate and TdT-Ctrl, is a TdT reaction with the empty vector, and a consensus Sp1 sequence is used as non-specific competitor.



fragment is used as a probe in a gel-shift assay.<sup>15</sup> These data suggest the interaction of an additional factor in close proximity or overlapping the GATA-1 motif on the 120 bp DNA fragment (Figure 1(a)). Detailed sequence analysis revealed the presence of the TCATG motif adjacent to the fourth GATA-1 site in the 120 bp DNA fragment (Figure 3(a)). The TCAT sequence has been previously identified in the human  $\epsilon$ -globin silencer contacting the multi-functional Ying Yang 1 (YY1) factor adjacent to a GATA-1 binding site.<sup>25,26</sup> A more recent report defines the same TCAT motif as a YY1 binding site in the murine  $\beta$ -interferon promoter with activating and repressing effects on transcription.<sup>27</sup>

This evidence prompted us to ask if YY1 could interact with the 120 bp DNA fragment, in proximity to the fourth GATA-1 binding site (Figure 3(a)). We used the fourth GATA-1, a mutated fourth GATA-1 site ( $\Delta$ GATA-1) and a mutated YY1 site in the fourth GATA-1 sequence ( $\Delta$ YY1) as labeled probes in a gel-shift assay (Figure 3(b) and (d)). The retarded complexes were competed with a known GATA-1 sequence from the chicken  $\beta$ -globin enhancer ( $\beta^A/\epsilon$ ) and with a YY1 consensus sequence.<sup>27,28</sup> We observe a different pattern of retarded complexes using 6C2 and RBCs as source of nuclear extracts. The complexity of the interaction is exemplified by  $\Delta$ GATA-1 and  $\Delta$ YY1 probes. Using the  $\Delta$ GATA-1 probe there is a clear reduction of the b complex corresponding to the GATA-1 factor (Figure 3(b), lanes 5–7 and 15–17). In contrast, using  $\Delta$ YY1 we practically lose all interactions (Figure 3(b), lanes 8–10 and 18–20). We interpret these data to indicate interdependency between GATA-1 and YY1 binding to the fourth GATA-1 sequence. Further detailed characterization is needed to understand the hierarchy and functional importance of those interactions. It is particularly relevant to the present work that multiple retarded complexes are observed using 6C2 cell extracts (Figure 3(b), complexes a, b and c).

To gain additional support for the interaction of YY1 with the 120 bp DNA fragment, we performed a supershift assay using a polyclonal antibody against human YY1 (Figure 3(c)). Since we did not know if the YY1 antibody recognizes this factor in chicken erythroid nuclear extracts, we first incubated the YY1 antibody with labeled YY1 consensus sequence in the presence of ten day-old chicken embryonic erythrocytes or human HeLa nuclear extracts (Figure 3(c)). A clear supershift complex is seen in each case (Figure 3(c); left panel lanes 3 and 5). We next performed the experiment using the GATA-1 sequence of the 120 bp DNA fragment and RBC as source of nuclear extracts. A specific supershift is seen when the YY1 antibody is used, strongly supporting the presence of YY1 on this sequence (Figure 3(c); central panel lane 3). With RBC extract, the retarded complex is not so clearly affected by YY1 antibody, probably because in the retarded complex GATA-1 remains bound (Figure 3(c); central panel lane 3). Interestingly, when 6C2 cell nuclear extracts were used no clear supershift

was detected even though the intensity of the main complex is reduced, suggesting that the YY1 antibody recognizes some component of the complex (data not shown).

A supershift experiment was then carried out using antibodies against GATA-1 (Figure 3(c); right panel). Incubating the GATA-1 antibody with 6C2 cell nuclear extracts revealed the presence of a supershift complex. In addition the intensity of all the bands diminished, suggesting a nucleating role for GATA-1 which probably facilitates multiple protein–protein and protein–DNA interactions (Figure 3(c); right panel lane 3). In contrast, in RBC, GATA-1 is clearly supershifted and the retarded complexes generated by GATA-1 and YY1 factors are reduced (Figure 3(c); right panel lane 6).

Next, we carried out an experiment employing the YY1 consensus sequence as a probe, with RBC and 6C2 cells as sources of nuclear extracts. Specific competitors containing mutations of the fourth GATA-1 sequence and the novel YY1 interacting sequence identified at the fourth GATA-1 site were employed (Figure 3(d)). Consistent with the above results, YY1 interaction with the consensus sequence is only competed when the GATA-1 sequence is mutated leaving access to the YY1 binding sequence. Interestingly, when the wild-type GATA-1 of the 120 bp DNA sequence is used as a cold competitor, there is no competition in RBC nuclear extracts (Figure 3(d); lane 5). In contrast, in 6C2 cells using increasing concentrations of cold competitors, slight competition is seen (Figure 3(d); lanes 13 and 14). This result shows that YY1 cannot bind or partially interact with its recognition sequence when GATA-1 is present in RBC. However, in 6C2 cells partial competition may reflect a difference in GATA-1 and YY1 retarded complex formation with other nuclear factors. Finally, further confirmation of YY1 interaction with the fourth GATA-1 sequence is evident from competition with the GATA-1 mutant sequence (Figure 3(d); lanes 6 and 17). Previous reports have shown that GATA-1 mRNA and protein levels vary during development, peaking around four days of chicken embryonic development then decreasing gradually after day 8.<sup>23</sup> To our knowledge no similar data are available for YY1 in chicken cells. To address this issue, we performed Western blots using HD24 (one of the most primitive avian erythroid cell lines), 6C2, HD3, RBC and the human erythroleukemic K562 cells as source of protein extracts (data not shown). We found no significant variation in relative YY1 abundance in the chicken cell-types tested. Thus, gel-shift results do not reflect variations in the relative concentrations of YY1 in the extracts.

Finally, to confirm the *in vitro* interaction of YY1 to the fourth GATA-1 sequence we carried out an *in vitro* transcription and translation experiment using the full-length YY1 cDNA as template (Figure 3(e)). Using the fourth GATA-1 sequence as probe we observed a single retarded band corresponding

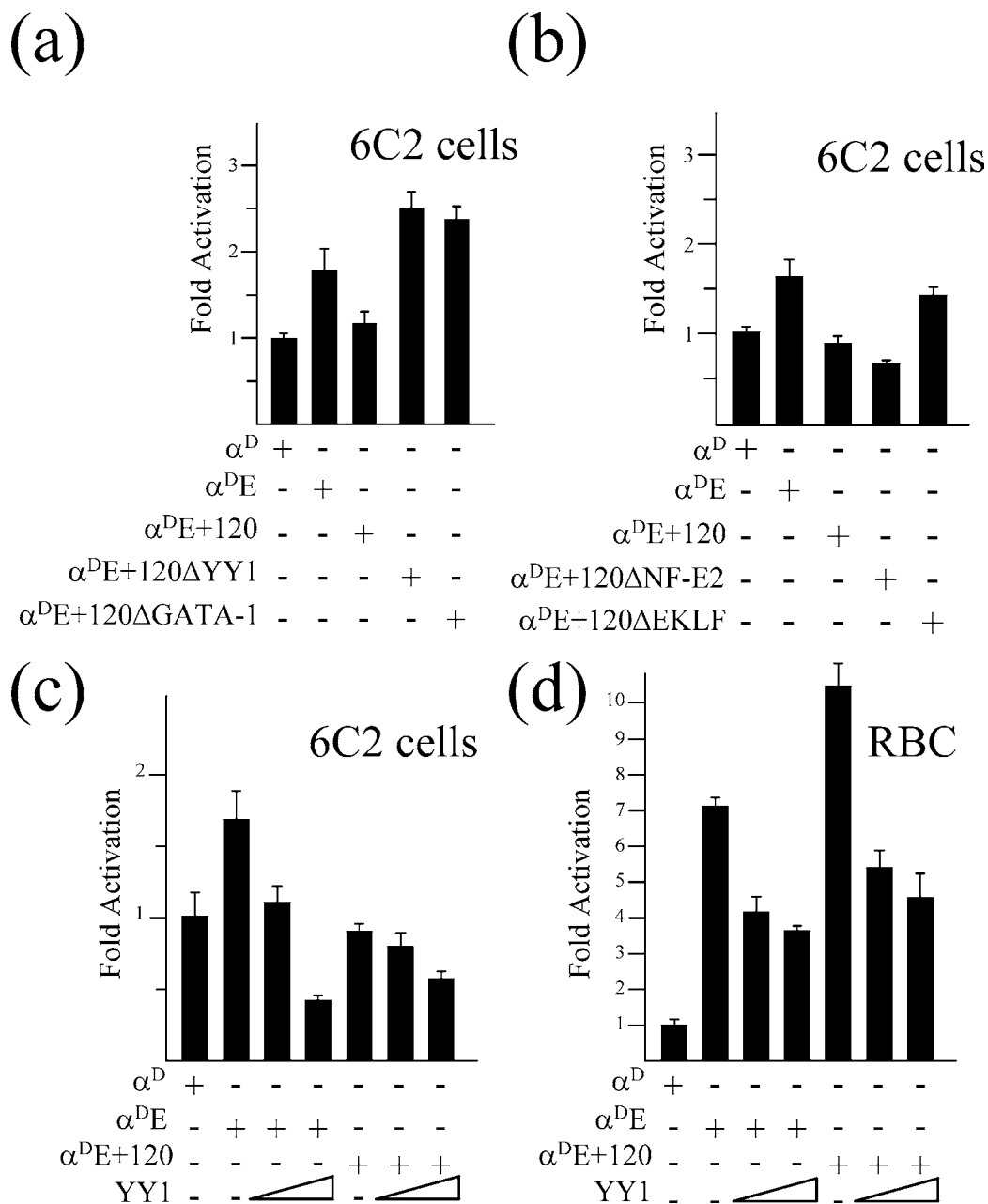
to YY1 that is competed with the unlabeled fourth GATA-1 sequence and YY1 consensus motif (Figure 3(e); lanes 3 and 4). As control, we used the YY1 consensus sequence as labeled probe demonstrating the specificity of the *in vitro* transcribed and translation YY1 binding.

In summary, there is differential nuclear factor interaction at the fourth GATA-1 sequence of the chicken 3'-side  $\alpha$ -globin enhancer. The composition of the complexes formed varies in the two erythroid

stages tested, in particular in 6C2 cells where the lack of activity of the full enhancer was observed.

**The 120 bp DNA fragment modulates the chicken  $\alpha$ -globin enhancer activity**

To further understanding of the contribution of the 120 bp DNA fragment to the overall enhancer activity, we tested point mutations of the YY1 and GATA-1 motifs, and we carried out a series of



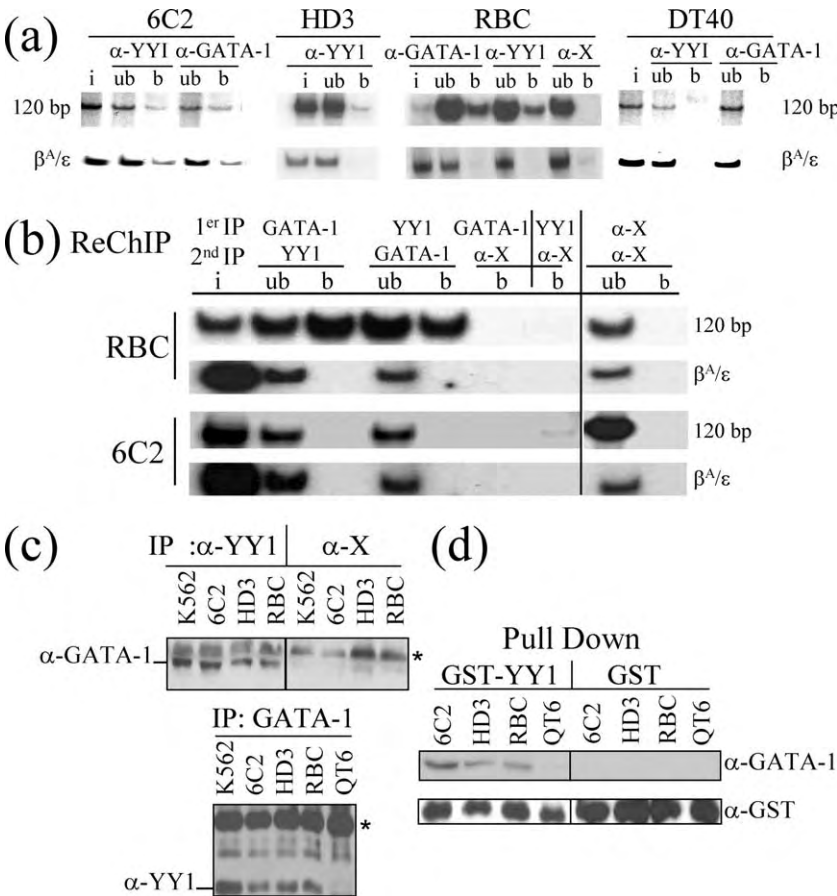
**Figure 4.** Mutant analysis and *trans*-effect of YY1 cDNA on the chicken 3'-side chicken  $\alpha$ -globin enhancer in 6C2 cells and RBCs. (a) Transient transfections of the chicken  $\alpha^D$  promoter, the core enhancer ( $\alpha^{DE}$ ), the entire enhancer ( $\alpha^{DE+120}$ ), and mutants thereof ( $\alpha^{DE+120\Delta YY1}$ ,  $\alpha^{DE+120\Delta GATA-1}$ ) were performed. (b) In the context of the full enhancer ( $\alpha^{DE+120}$ ), NF-E2 and EKLF mutants located in the 120 bp DNA fragment were tested. (c) Human YY1 cDNA (YY1) was co-transfected with 1  $\mu$ g and 2  $\mu$ g of plasmid, in 6C2 cells, in the context of the core enhancer and the full enhancer. (d) The same experiment as in (c) was now performed in ten day-old chicken RBCs. Results are shown as fold activation and represent at least five independent experiments and error bars show the SEM.

*trans*-activation experiments in which we over-expressed human YY1 (Figure 4). We first transiently transfected into 6C2 cells the  $\alpha^D$  promoter and the core enhancer ( $\alpha^{DE}$ ), the complete enhancer ( $\alpha^{DE}+120$ ) and the complete enhancer with the fourth GATA-1 and YY1 sites mutated ( $\alpha^{DE}+120\Delta YY1$  and  $\alpha^{DE}+120\Delta GATA-1$ ). Importantly, the YY1 site mutation relieves part of the repressive effect previously attributed to the 120 bp DNA fragment (Figures 2 and 4(a)). Surprisingly, when the fourth GATA-1 and the adjacent YY1 sites located in the 120 bp DNA fragment were mutated we found a clear de-repression (Figures 2 and 4(a); constructs  $\alpha^{DE}+120\Delta YY1$  and  $\alpha^{DE}+120\Delta GATA-1$ ). This result suggests that YY1, in conjunction with GATA-1, exerts a negative influence over the enhancer in 6C2 cells.

In the same context, we next evaluated the effect of EKLF and NF-E2 mutations (Figures 1(a) and 4(b)). Interestingly, absence of the NF-E2 binding site accentuates a reduction of the reporter activity favoring a positive but partial role for NF-E2 in 6C2 cells. The same results were obtained with stably transfected chicken erythroblast transformed HD3 cells (data not shown). The EKLF binding site

mutation does not seem to significantly affect the enhancer activity in 6C2 cells. In conclusion, EKLF and NF-E2 do not seem to participate in enhancer modulation in 6C2 cells, but they certainly contribute to positive enhancer function at more advanced erythroid stages when the globin genes are very actively transcribed.

To confirm YY1 contribution to negative modulation of enhancer activity we co-transfected the human YY1 cDNA (Figure 4(c) and (d)). Interestingly, in both 6C2 and RBC cells, over-expression of YY1 induces *trans*-repression of the reporter gene. For 6C2 cells the result is consistent with the suggested negative role for YY1 (Figure 4(c)). These results are not dependent on the effect of YY1 on the  $\alpha^D$  promoter (data not shown). Unexpectedly, the YY1 repressive effect is also observed in RBC, a more advanced erythroid stage compared to 6C2 cells, favoring YY1 modulation of the enhancer (Figure 4(d)). This effect is seen both in the presence of the core enhancer ( $\alpha^{DE}$ ) and the complete enhancer ( $\alpha^{DE}+120$ ). This result could be explained by the presence of an uncharacterized YY1 binding site in the core enhancer or YY1 interaction with GATA-1 factors



**Figure 5.** *In vitro* and *in vivo* interaction of YY1 and GATA-1 with the 120 bp DNA fragment of the 3'-side chicken  $\alpha$ -globin enhancer. (a) Chromatin immunoprecipitation assays using two different sets of primers which amplify the chicken  $\alpha$ -globin enhancer 120 bp DNA fragment (120 bp) generating a 169 bp amplification product and the chicken  $\beta^A/\epsilon$  globin enhancer ( $\beta^A/\epsilon$ ) generating a 301 bp amplification product. Input (i), bound (b) and unbound (ub) amplification products are shown. In all cases the linear range of PCR amplification was previously established (data not shown). An irrelevant antibody ( $\alpha$ -X) was used as a control. Erythroid (6C2 and HD3) and non-erythroid cells (DT40) were used in these experiments. This is a representative result from at least two independent experiments. (b) To confirm an interaction between GATA-1 and YY1 at the 3'-side enhancer we performed ReChIP experiments. We show a representative result from two independent experiments on each cell-type. (c) Immunoprecipitation (IP) using YY1 and non-specific ( $\alpha$ -X) antibodies (upper panel) and

GATA-1 (lower panel). The immunoprecipitated material was subjected to Western blot employing anti-GATA-1 antibody ( $\alpha$ -GATA-1; upper panel) and anti-YY1 antibody ( $\alpha$ -YY1; lower panel), respectively. \*, Non-specific band that serves as loading control. (d) Pull-down assay using the GST-YY1 fusion protein and  $\alpha$ -GATA-1 antibody for detection. As a control, GST protein alone was detected (lower panel).

located on the core enhancer. Alternatively, a physical interaction between GATA-1 and YY1 could explain why the core enhancer could still be *trans*-repressed when high levels of YY1 are induced (Figure 4(d), and see below).

In summary, we have shown that the GATA-1 and YY1 binding sites located in the 120 bp DNA fragment are able to modulate the chicken 3'-side  $\alpha$ -globin enhancer in an erythroid stage-dependent way (Figures 1, 2 and 4).

### ***In vivo* co-localization and interaction of GATA-1 and YY1 with the chicken $\alpha$ -globin enhancer**

To confirm the *in vivo* YY1 localization within the fourth GATA-1 sequence, chromatin immunoprecipitation experiments were done (Figure 5(a)). We used 6C2, HD3 and RBC as a source of erythroid cells and the chicken lymphoma cell line DT40 as a source of non-erythroid cells. We verified the *in vivo* interaction of YY1 and GATA-1 with primers amplifying the 120 bp DNA fragment (Figure 5(a)). No amplification products are seen when an irrelevant antibody ( $\alpha$ -X: anti-negative IgG) or when non-erythroid DT40 cells are used as the source of chromatin. Interestingly, we observed an amplification product from the YY1 and GATA-1 immunoprecipitated samples using control primers which amplify a region of the chicken  $\beta^A/\epsilon$  enhancer.<sup>29</sup> This suggests the possibility that YY1 is also interacting with the  $\beta$ -globin enhancer, which has not been reported previously, although there is a putative YY1 site in the  $\beta^A/\epsilon$  enhancer defined by TRANSFAC analysis (data not shown).

In order to demonstrate that GATA-1 and YY1 co-localize, *in vivo*, specifically at the fourth GATA-1 sequence of the 3'-side  $\alpha$ -globin enhancer, we carried out a ReChIP experiment (Figure 5(b)). As shown in Figure 5(b), GATA-1 or YY1 antibodies were used for the first round of immunoprecipitation and the order reversed for the second round of immunoprecipitation. We are able to specifically enrich the 120 bp DNA fragment using RBC chromatin. Unexpectedly, using 6C2 cell chromatin, we are not able to see any enrichment of this fragment when the second round of immunoprecipitation was performed, even though our *in vitro* and *in vivo* DNA-protein experiment supports such an interaction. Possibly, the multi-peptide contacts formed on 6C2 cell chromatin hinder or mask access of the antibodies to their epitopes in the ReChIP assay. An alternative explanation is variation in the local and/or relative abundance of the nuclear factors involved. These results confirm GATA-1/YY1 presence *in vivo* at the 3'-side chicken  $\alpha$ -globin enhancer.

Since GATA-1 and YY1 are present simultaneously on the 120 bp DNA fragment, we asked if there is a direct interaction between these factors. To address this, we performed co-immunoprecipitation experiments using antibodies against YY1, GATA-1 and an irrelevant antibody (Figure 5(c)). In addition to the chicken 6C2, HD3 and RBC cells we

used human erythroleukemic K562 cells and the non-erythroid, GATA-1 non-expressing quail fibroblast transformed QT6 cell line.<sup>30,31</sup> Interaction between GATA-1 and YY1 was seen in erythroid cells but not when non-specific antibody or QT6 cells were employed (Figure 5(c)). This interaction is thus not restricted to avian cells, since similar results were obtained with K562 cells. To confirm GATA-1 and YY1 contact we generated a GST-YY1 fusion protein for *in vitro* pull-down assay (Figure 5(d)). Cell lysates from erythroid and non-erythroid cells were incubated with the GST alone and the GST-YY1 fusion protein. GATA-1 was detected only in erythroid cells, demonstrating a physical interaction with YY1.

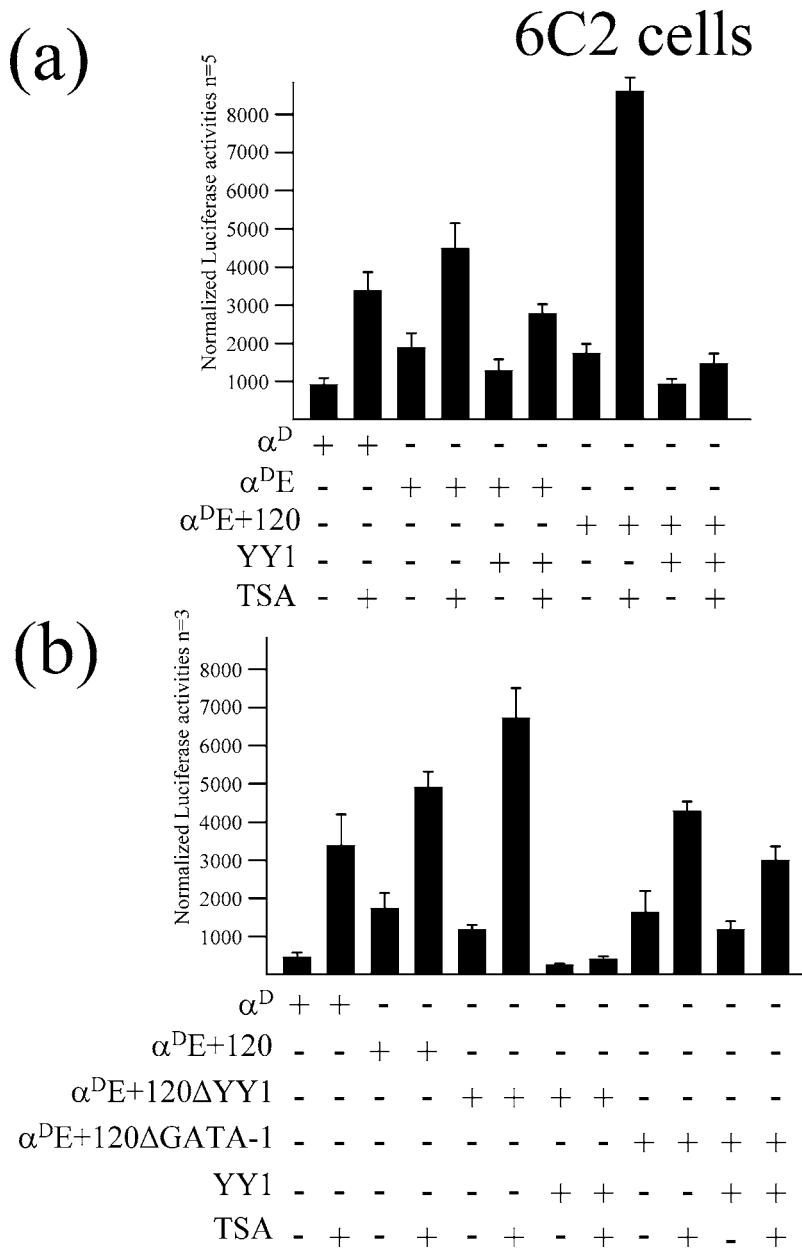
In summary, both functional and *in vitro* and *in vivo* DNA-protein interaction experiments strongly argue in favor of a close relationship between GATA-1 and YY1. For the first time and as a central result of our investigation, a physical interaction between GATA-1 and YY1 is demonstrated with direct consequences on the modulation of the chicken  $\alpha$ -globin enhancer activity.

### **Histone deacetylation modulates the 3'-side chicken $\alpha$ -globin enhancer activity in pre-erythroblasts**

It is well established that histone hyperacetylation is associated with transcriptionally active chromatin, whereas histone hypoacetylation is correlated with transcriptionally inactive chromatin.<sup>32</sup> In addition, it is recently more generally accepted that transiently transfected plasmids acquire some degree of chromatinization and that nucleosomes can assemble on them.<sup>33,34</sup> Furthermore, it has been established that YY1 has a bimodal function with respect to histone acetylation, recruiting either histone acetylases or histone deacetylases.<sup>26,35,36</sup> Based on this evidence, we asked if the negative modulation of the enhancer activity seen in the pre-erythroblast 6C2 cells is due in part to the action of histone deacetylases.

We incubated the 6C2 cells with trichostatin-A (TSA), and determined the influence of this histone deacetylase inhibitor on the activity of different reporter constructs (Figure 6). We tested the  $\alpha^D$  promoter alone, the core enhancer ( $\alpha^{DE}$ ), the entire enhancer ( $\alpha^{DE}+120$ ) and the effect of co-transfection of YY1 cDNA (Figure 6(a)). In all cases, including the  $\alpha^D$  promoter alone, we observe an increase in activity after incubation with TSA. Interestingly, when the entire enhancer ( $\alpha^{DE}+120$ ) is tested we obtain the highest activation levels, suggesting that factor(s) responsible for recruitment of histone deacetylation activity could be located in the 120 bp DNA fragment. Over-expression of YY1 has little effect in the absence of TSA, but almost completely attenuates the stimulatory effect of TSA treatment on the full enhancer.

That the highest level of TSA activation is seen when the entire enhancer is transfected reinforces the hypothesis that the fourth GATA-1 and YY1



**Figure 6.** The chicken 3'-side  $\alpha$ -globin enhancer activity is stimulated in the presence of the histone deacetylation inhibitor TSA in pre-erythroblasts. (a) The chicken adult  $\alpha^D$  promoter alone ( $\alpha^D$ ), the promoter with the core enhancer ( $\alpha^{DE}$ ) and the entire enhancer ( $\alpha^{DE+120}$ ) were assayed in the presence or absence of 2.5 ng/ml of the histone deacetylase inhibitor TSA. In addition 2  $\mu$ g of the human YY1 cDNA was co-transfected with and without TSA.<sup>33,34</sup> All the constructs were transiently transfected in chicken transformed pre-erythroblast 6C2 cells and transcriptional activation was measured with the dual-luciferase assay. (b) *Trans*-activation effect of the human YY1 cDNA (YY1) in the context of the entire enhancer ( $\alpha^{DE+120}$ ) and the mutated YY1 and the fourth GATA-1 sequences ( $\alpha^{DE+120\Delta GATA-1}$  and  $\alpha^{DE+120\Delta YY1}$ ), with and without TSA.

binding sites participate in enhancer modulation in 6C2 cells (Figure 6(a);  $\alpha^{DE+120}$ ). Therefore, we performed the experiment using the mutant YY1 and mutant GATA-1 sites in the context of the full enhancer ( $\alpha^{DE+120\Delta YY1}$  and  $\alpha^{DE+120\Delta GATA-1}$ ). Interestingly, when either the YY1 or GATA-1 sites are mutated, there is still a significant activation in the presence of TSA (Figure 6(b)). Over-expression of YY1 in the absence of TSA strongly inhibited the YY1 mutant enhancer construct, and the stimulatory effect of TSA was completely abrogated in the presence of YY1. In contrast, YY1 expression had little effect on the GATA-1 mutant enhancer, and TSA could still activate it in the presence of over-expressed YY1.

These results emphasize that over-expression of YY1, although modestly deleterious to the function

of the full enhancer, strongly antagonizes the stimulatory effect of TSA. One possibility is that YY1 is involved in the recruitment of deacetylases to the 120 bp enhancer fragment. Alternatively, it is possible, since we have not directly demonstrated histone acetylation changes at the enhancer, that the stimulatory effect of TSA is due to acetylation of YY1 or GATA-1, or to a more indirect mechanism.<sup>37,38</sup> These possibilities are not mutually exclusive. The results obtained with the histone deacetylase inhibitor TSA lead us to propose that part of the negative modulation of the 3'-side enhancer is due to the action of histone deacetylases, possibly by generating a closed chromatin conformation as a way to control activity of the enhancer over distinct erythroid-specific stages.

## Discussion

Differential regulation of gene expression is at the center of many biological processes, including hemoglobinopathies. The globin group of genes exemplify, from the perspective of chromatin structure and erythroid gene expression, an excellent model for understanding mechanisms governing tissue and developmental-stage specific gene regulation. In this investigation we carried out a detailed study of the chicken  $\alpha$ -globin 3'-side enhancer in widely used cell lines representing different stages of erythroid differentiation.<sup>17-22</sup> Deletional and mutational analysis show that in addition to the core  $\alpha$ -globin enhancer, there is a 120 bp DNA fragment that is an integral part of the enhancer and is important for the highest level of activation in ten day-old embryonic chicken erythrocytes. Our studies reveal the unexpected presence of the YY1 transcription factor in close proximity to GATA-1; their physical contact and their relationship seems, in part, to be responsible for the modulation of enhancer activity at distinct erythroid stages. Finally, experiments using the histone deacetylase inhibitor TSA prompt us to suggest that part of the modulation of the enhancer in pre-erythroblasts could be due to histone acetylation and formation of a closed chromatin conformation.

We have previously established that the 3'-side chicken  $\alpha$ -globin enhancer contains binding sites for a classical set of nuclear factors: GATA-1, NF-E2, EKLF, Sp1 and now YY1 (Figure 1(a)).<sup>14,15,39</sup> The recognition sequence for GATA-1 is found in globin promoters, enhancers and LCRs. Usually two or more binding sites are clustered together<sup>31</sup> and in some cases the sites are found in tandem, apparently increasing their affinity.<sup>31,38</sup> Our *trans*-activation experiments and *in vitro* and *in vivo* DNA-protein interaction results support the importance of the 120 bp DNA fragment immediately downstream of the characterized enhancer to overall enhancer function. Combinatorial complex formation on these sequences involving YY1 and GATA-1 contributes to the enhancer function in distinct erythroid stages (Figures 2, 4 and 5). GATA-1 is able to interact with the histone acetyltransferase CBP *in vitro* and *in vivo*, strongly supporting the influence of GATA-1 in chromatin remodeling.<sup>40-42</sup> GATA-1 also partners with a co-factor, FOG-1, to positively and negatively modulate its response during erythroid and megakaryocyte differentiation.<sup>43</sup> Recently, it has been demonstrated that GATA-1 and FOG contribute topologically to long distance enhancer-promoter communication favoring loop formation.<sup>11</sup>

Another pathway involves the GATA-1 homolog GATA-2, which is the predominant GATA species at earlier stages of differentiation.<sup>44-46</sup> We speculate that GATA-2 and not GATA-1 may interact at the 3'-side  $\alpha$ -globin enhancer in the pre-erythroblast stage, when the globin locus is prepared, at the level of chromatin domain structure, to initiate regulated

globin gene expression.<sup>18,20,40</sup> This proposal is consistent with the nucleation attributes of GATA-2 and GATA-1, recently reported by Higgs and collaborators.<sup>47</sup> Chromatin immunoprecipitation experiments, together with mouse  $\alpha$ -globin gene expression data during the haematopoietic program demonstrated that first GATA-2 and later GATA-1, in association with FOG-1, are responsible for the nucleation of multiple interactions with other transcription factors and co-factors.<sup>47</sup> We suggest that GATA-1 is also responsible for nucleating several interactions, in particular with YY1, with consequences for the chicken  $\alpha$ -globin enhancer activity in early stages of erythropoiesis and on its own chromatin structure (M.E.-D.-A. & F.R.-T., unpublished observations).

An unexpected observation was the presence of the Ying Yang 1 nuclear factor, both *in vivo* and *in vitro*, close to the GATA-1 site in the 120 bp DNA fragment of the chicken 3'-side  $\alpha$ -globin enhancer (Figures 3 and 5). YY1 is involved in positive and negative activation based on co-factor recruitment, including histone acetyltransferases like CBP and p300, and with histone deacetylases like HDAC1, HDAC2 and HDAC3.<sup>26,48</sup> Furthermore, YY1 itself can be acetylated by both p300 and PCAF acetyltransferases, and deacetylated by HDAC1 and HDAC2 modifying its affinities for DNA or co-factors.<sup>37</sup> Our results demonstrate that YY1 and GATA-1 physically interact and simultaneously occupy, *in vivo*, the chicken  $\alpha$ -globin enhancer.

A complementary view of the negative modulation of YY1 in 6C2 cells is that GATA-1 and YY1 are competing for their DNA-binding site in this erythroid stage, and that YY1 is displacing GATA-1 interaction because of overlapping binding sites. Alternatively, the relative concentration of GATA-1 or its affinity for its binding site at the pre-erythroblast stage can play a role. Possibly, post-translational modifications of YY1 and GATA-1 affect their binding affinities. It has been reported that GATA-1 acetylation is known to increase its affinity for DNA.<sup>38</sup> It will be appealing to determine whether covalent modifications of these two transcription factors demonstrate a functional relationship.

Interestingly, in recent publications the relevance and the relationship between YY1 and members of the Polycomb group of proteins have been outlined.<sup>49-51</sup> Our results using TSA showed partial reactivation of the full enhancer, and when we over-expressed YY1 we found an increase in the repression effect. Possibly, in addition to the involvement of HDACs in modulating enhancer activity at early erythroid stages, members of the Polycomb family of repressive proteins, in particular EZH2, through H3 lysine 27 trimethylation, could contribute to the negative modulation of the enhancer in 6C2 cells.

At later erythroid stages, GATA-1 and YY1 appear to be interacting exclusively at the fourth GATA-1 motif (Figure 3). In ten day-old chicken erythrocytes we are able to demonstrate *in vivo* the

interaction of YY1 and GATA-1 in the 3'-side enhancer, which correlates temporally with our finding that the enhancer is fully active at this erythroid stage. We favor the explanation that, in contrast to the incorporation of other proteins seen in 6C2 cells, in RBCs the activity of the enhancer is more dependent on GATA-1. Interestingly the contribution of YY1 in chicken erythrocytes remains to be explained. One possibility is that, in erythrocytes, YY1 switches its function and cooperates with GATA-1 to activate  $\alpha$ -globin gene transcription through the enhancer. Another possibility could involve the complementary incorporation of erythroid and non-erythroid factors such as NF-E2, EKLF or Sp1. Thus YY1 plasticity in terms of its contribution to enhancer activity modulation is appealing since YY1 can confer the capacity to shift the balance between gene activation and repression.

The major findings presented here are that the 3'-side chicken  $\alpha$ -globin enhancer is differentially activated during distinct erythroid stages, and the discovery of YY1/GATA-1 relationship that may form the basis of the plasticity of the enhancer function. We are now in a position to ask novel questions concerning the detailed mechanism of action of these regulatory elements and the way in which the enhancer differentially modulates the chicken  $\alpha$ -globin genes in distinct erythroid differentiation stages and during development. The distribution of DNA-binding proteins over the region predicts a differential combinatorial action of those factors influenced by polypeptide concentrations, competition for DNA-binding motifs, and chromatin remodeling activities. We are currently analyzing the chromatin structure over the silencer-enhancer to further determine the contribution of chromatin remodeling events to the silencer and enhancer activities.

## Materials and Methods

### Cell culture

The 6C2 cell line corresponds to an avian erythroblastosis transformed chicken bone marrow cell line arrested at the CFU-E stage and considered to be pre-erythroblast.<sup>21,22</sup> 6C2 cells were grown in minimal essential medium ( $\alpha$ -MEM; GIBCO) supplemented with 10% (v/v) fetal calf serum, 2% (v/v) chicken serum, 1 mM Hepes (pH 7.2) and 50  $\mu$ M 2-mercaptoethanol. The avian erythroblastosis virus-transformed and temperature-sensitive chicken erythroblast cell line LSCCHD3 (HD3) and Japanese quail fibroblasts (QT6 cells) were grown at 37 °C in DMEM medium (GIBCO) supplemented with 8% fetal calf serum, 2% chicken serum and antibiotics.<sup>21,22,30</sup> DT40 lymphoid cell line was grown in DMEM supplemented with 50  $\mu$ M 2-mercaptoethanol, 10% fetal calf serum, 5% chicken serum, and 10% tryptose phosphate broth. HeLa cells were grown in DMEM supplemented with 10% fetal calf serum and human K562 cells in IMEM and 10% fetal calf serum.

### Plasmid constructs

DNA sub-fragments containing the chicken  $\alpha$ -globin 3'-side enhancer<sup>13,29</sup> were PCR amplified from ten day-old embryonic erythrocytes from *Gallus gallus* (Alpes, Puebla, México) using genomic DNA previously fractionated by EcoRI enzymatic digestion, with the following primers: AE1, 5'-CGCGGATCCGCGCAGGGTGAAGCTGTGCTG-3'; A120P4, 5'-CGCGGATCCGACCGTGGTAA CCCCAGGCCG-3'; A120P3, 5'-CGCGGATCCGACGTG GGCAGCAGATAGCCTCG-3'. All synthetic oligonucleotides contain a BamHI restriction site. Mutation of the EKLF, NF-E2, fourth GATA-1 site and double mutant NF-E2/GATA-1 site were generated in the context of the full enhancer, i.e. including the original three GATA-1 binding sites.<sup>12,13</sup> The two-step PCR overlapping method was used to generate two and three base-pair changes using the following primers (with their corresponding complementary strands): MABeklf02, 5'-GCAGATA GCCTCGtTtGGTAAAACCTGC-3'; MABnfe02, 5'-CCT GCCAGGTGCTagGTAAGTGCAGTCTG-3'; and MAB-gata02, 5'-GATCCTGCCAGGTaagTTCATGCAGAGATC-3'. All the mutated DNA fragments were directly cloned into the pCAT $\alpha^D$  reporter vector<sup>52</sup> and the DNA was sequenced using the primer 5'-GGGGAGGTGTGG GAGGTTT-3'. We generated the pG $\alpha^D$  vector by cloning the  $\alpha^D$  globin promoter<sup>53</sup> into the BglIII site of the pGL3-basic poly-linker (Promega). All the DNA sub-fragments to be tested were sub-cloned into the BamHI site downstream of the firefly luciferase reporter gene.

### Transient and stable transfections

Ten day-old chicken embryonic erythrocyte transient transfections were performed essentially as described.<sup>17,53,54</sup> Typically, 6  $\mu$ g of the pC $\alpha$ 3' derived plasmids and 6  $\mu$ g of pTk- $\beta$ gal were co-transfected. CAT activity assays were performed as described<sup>17</sup> and the  $\beta$ -galactosidase activity for normalization was obtained using a  $\beta$ -galactosidase kit (Promega). Dual luciferase assay transfections and measurements were carried out according to manufacturer's instructions (Promega) in a TD-20/20 Turner BioSystems luminometer. For 6C2 and HD3 transformed cell lines stable transfections we used the pE $\alpha$ D plasmids, which contain the green fluorescent protein as the reporter gene and the neomycin-resistance cassette (pEGFP-1, CLONTECH). After selection, stable pools were split in three and neomycin-resistant clones were analyzed by fluorescent cytometry. For *trans*-activation experiments 2  $\mu$ g of test plasmids were co-transfected with 200 ng of *Renilla* (pRL-TK, Promega) and 2  $\mu$ g and 4  $\mu$ g of chicken GATA-1 cDNA (kindly provided by Cecilia Trainor) or human YY1 cDNA (kindly provided by Kenneth Walsh and Tamar Aprahamian). Plasmids were transiently transfected with Lipofectamine (Invitrogen) in a 1:1 ratio. For TSA reactivation experiments we first established the TSA concentration using 2.5 ng/ml, 5 ng/ml, 10 ng/ml, 15 ng/ml and 20 ng/ml of TSA and selected 2.5 ng/ml as the optimum concentration for 6C2 cells (data not shown). In addition, we tested the *Renilla* vectors with different promoters (CMV, TK and SV40) to verify that TSA had no effect on their activity (data not shown). Thus  $1 \times 10^6$  cells were co-transfected and 24 hours post-transfection were split in two parts. One part was incubated with TSA for 24 hours and luciferase activity was measured 24 hours after TSA incubation.

### Electrophoretic mobility shift analysis

Gel-shift assays were carried out essentially as described by Valadez-Graham *et al.*<sup>53</sup> For supershift reactions, anti-YY1 polyclonal antibody (H-414; Santa Cruz Biotechnology) was pre-incubated with nuclear extracts overnight with agitation at 4 °C. The following oligonucleotides were used in different mobility shift experiments (in all cases only one of the complementary chains is shown): A1G (fourth GATA-1), 5'-CGCGGATC CTGCCAGGTATCTTCATGCAGAGATG-3'; MABgata02 oligonucleotide was used as the mutant fourth GATA-1 site ( $\Delta$ GATA-1); AYY1A ( $\Delta$ YY1), 5'-GATTGCCAGGTA TCTTtcgGCAGAGATG-3'; Sp1 (Sp1 consensus), 5'-ATT CGATCGGGGCGGGGGCGAGC-3', YY1A (YY1 consensus), 5'-CGCTCCGCGGCCATCTTGGCGGCTGGT-3'<sup>27</sup> and cGATA-1 (GATA-1 consensus), 5'-CACGCGGATAA GATAAGGCCGGGCT-3'.<sup>29</sup>

### Chromatin immunoprecipitation

The ChIP assay was performed as described by Farnham and collaborators with slight modifications.<sup>55</sup> The following antibodies were used: goat anti-mouse IgG (Santa Cruz Biotechnology) as a negative control and 10  $\mu$ g of anti-GATA-1 (H-200; Santa Cruz Biotechnology) and anti-YY1 (H-414; Santa Cruz Biotechnology). Radioactive PCR reactions were then performed. Two pairs of oligonucleotides were designed. The first one amplifies a 169 bp DNA fragment of the chicken  $\alpha$ -globin 3'-side enhancer containing the 120 bp element, and the oligonucleotide sequences were as follows: A120P3, 5'-GAC GTGGGCAGCAGATAGCCTCG-3' and A120P4, 5'-GCC GGACCCCAATGGTGCCAG-3'. The second pair of oligonucleotides, used as a control, amplifies a 301 bp sequence containing a portion of the chicken  $\beta$ -globin enhancer, named  $\beta^A/\epsilon$  (110–558 bp after the poly(A) signal of the  $\beta^A$  gene):<sup>29</sup> 0026 5'-CTGGGTGGGGGCA GGT-3', 0027 5'-CTGCTTTTGCTGCCCTGTG-3'. PCR reactions were performed at an annealing temperature of 58.5 °C for 28–30 cycles, run on a 6% (w/v) polyacrylamide gel at 150 V in 0.5 $\times$  TBE. Gels were dried and exposed to X-ray film (Kodak Bio-Max) for 24 to 72 hours and films were developed. ReChIP experiments were performed as described.<sup>56</sup>

### Co-immunoprecipitation and pull-down assays

Cells were lysed in RIPA buffer (150 mM NaCl, 5 mM EDTA, 50 mM Hepes (pH 7.5), 0.5% (w/v) sodium deoxycholate, 1% Nonidet P-40 (NP-40), 10 mM 2-mercaptoethanol) containing 2 mM phenylmethylsulfonyl fluoride (PMSF), 50  $\mu$ g/ml of aprotinin A, 25  $\mu$ g/ml of leupeptin, and 25  $\mu$ g/ml of pepstatin, for 15 minutes at 4 °C. Cell lysates were then cleared by centrifugation at 16,000 g for five minutes and kept cold on ice. Aliquots of 100  $\mu$ l of lysate were incubated with 2  $\mu$ g of specific antibodies. Immunoprecipitates were resolved by sodium dodecyl sulfate (SDS)/12% polyacrylamide gel electrophoresis (PAGE) and the proteins transferred to polyvinylidene fluoride membranes (Immobilon-P; Millipore, Bedford, MA). Membranes were incubated in blocking buffer (1% (w/v) BSA, 5% (w/v) non-fat dry milk, 0.1% (v/v) Tween 20 in TBS) overnight at room temperature. Membranes were subsequently probed with the corresponding antibody in blocking buffer for one hour at room temperature: anti-GATA-1 polyclonal antibody at 1:1000. Membranes were six times washed with TBS+

0.1% Tween 20, for five minutes each time, and incubated with a 1/15,000 dilution of HRP-conjugated goat anti-rabbit IgG for one hour at room temperature. After six more washes with TBS +0.1% Tween 20, antibody-reactive proteins were detected using a chemiluminescence substrate (Pierce, Rockford, IL) according to the manufacturer instructions.

For pull-down assay, full-length chicken YY1 cDNA (kindly provided by Guangchao Sui and Yang Shi) was cloned into the pGEX4T-1 vector and expressed as a GST fusion protein in *Escherichia coli* BL21 (pRIL). The fusion protein was immobilized to Sepharose-glutathione beads and incubated with 500  $\mu$ g of cell lysates for six hours at 4 °C. Interacting proteins were resolved in a 10% SDS-PAGE and transferred to PDVF membranes. Protein detection was performed as in the immunoprecipitation assay.

### In vitro transcription/translation

The full-length chicken YY1 cDNA was *in vitro* transcribed and translated using the TnT reticulocyte lysate-coupled *in vitro* transcription/translation system (Promega).

### Acknowledgements

We are particularly indebted to Ann Dean and Catherine M. Farrell for critical reading of the manuscript, suggestions and constant support. We also thank Mark Groudine, Tomoyuki Sawado, Jessica Halow, and Paul Delgado-Olguín for critical reading of the manuscript. We thank Ana María López Colomé for use of the egg incubator facility. This work was supported by the Dirección General de Asuntos del Personal Académico-UNAM (IN203200 and IN209403), the Consejo Nacional de Ciencia y Tecnología, CONACyT (33863-N and 42653-Q), the Third World Academy of Sciences (TWAS, grant 01-055 RG/BIO/LA), and Fundación Miguel Alemán, A.C.

### References

1. Bulger, M. & Groudine, M. (1999). Looping *versus* linking: toward a model for long-distance gene activation. *Genes Dev.* **13**, 2465–2477.
2. Kosak, S. T. & Groudine, M. (2004). Form follows function: the genomic organization of cellular differentiation. *Genes Dev.* **18**, 1371–1384.
3. Recillas-Targa, F. & Razin, S. V. (2001). Chromatin domains and regulation of gene expression: familiar and enigmatic clusters of chicken globin genes. *Crit. Rev. Eukaryot. Gene Expr.* **11**, 227–242.
4. Razin, S. V., Farrell, C. M. & Recillas-Targa, F. (2003). Genomic domains and regulatory elements operating at the domain level. *Int. Rev. Cytol.* **226**, 63–125.
5. Recillas-Targa, F. (2000). The chicken  $\alpha$ - and  $\beta$ -globin gene domains and their chromatin organization. *Cell. Mol. Biol. Letters*, **5**, 451–467.
6. Burgess-Beusse, B., Farrell, C., Gaszner, M., Litt, M., Mutskov, V., Recillas-Targa, F. *et al.* (2002). The



- insulation of genes from external enhancers and silencing chromatin. *Proc. Natl Acad. Sci. USA*, **99**, 16433–16437.
7. Blackwood, E. M. & Kadonaga, J. T. (1998). Going the distance: a current view of enhancer action. *Science*, **281**, 60–63.
  8. Carter, D., Chakalova, L., Osborne, C. S., Dai, Y.-F. & Fraser, P. (2002). Long-range chromatin regulatory interactions *in vivo*. *Nature Genet.* **32**, 1–4.
  9. Patrinos, G. P., de Krom, M., de Boer, E., Langeveld, A., Iman, A. M., Strouboulis, J. *et al.* (2004). Multiple interactions between regulatory regions are required to stabilize and active chromatin hub. *Genes Dev.* **18**, 1495–1509.
  10. Tolhuis, B., Palstra, R.-J., Splinter, E., Grosveld, F. & de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active  $\beta$ -globin locus. *Mol. Cell*, **10**, 1453–1465.
  11. Vakoc, C. R., Letting, D. L., Gheldof, D., Sawado, T., Bender, M. A., Groudine, M. *et al.* (2005). Proximity among distant regulatory elements at the  $\beta$ -globin locus requires GATA-1 and FOG-1. *Mol. Cell*, **17**, 453–462.
  12. Knezetic, J. A. & Felsenfeld, G. (1989). Identification and characterization of a chicken  $\alpha$ -globin enhancer. *Mol. Cell. Biol.* **9**, 893–901.
  13. Recillas Targa, F., De Moura Gallo, C. V., Huesca, M., Scherrer, K. & Marcaud, L. (1993). Silencer and enhancer elements located at the 3'-side of the chicken and duck  $\alpha$ -globin-encoding gene domains. *Gene*, **129**, 229–237.
  14. Recillas Targa, F., De Moura Gallo, C. V. & Scherrer, K. (1993). Analysis of the distribution of protein binding DNA motives in the vicinity of the 3'-side chicken  $\alpha$ -globin enhancer. *Biochem. Biophys. Res. Commun.* **190**, 1163–1172.
  15. Recillas Targa, F., Razin, S. V., De Moura Gallo, C. V., Marcaud, L. & Scherrer, K. (1995). Silencer and enhancer elements and the framing structures of the chicken  $\alpha$ -globin gene domain. In *Molecular Biology of Hemoglobin Switching* (Stamatoyannopoulos, G., ed.), pp. 203–215, Intercept, Andover.
  16. Higgs, D. R., Sharpe, J. A., Goudron, G., Craddock, C. F., Vyas, P., Picketts, D. J. *et al.* (1995). Silencer and enhancer elements and the framing structures of the chicken  $\alpha$ -globin gene domain. In *Molecular Biology of Hemoglobin Switching* (Stamatoyannopoulos, G., ed.), pp. 165–177, Intercept, Andover.
  17. Recillas-Targa, F., Bell, A. C. & Felsenfeld, G. (1999). Positional enhancer-blocking activity of the chicken  $\beta$ -globin insulator in transiently transfected cells. *Proc. Natl Acad. Sci. USA*, **96**, 14354–14359.
  18. Prioleau, M.-N., Nony, P., Simpson, M. & Felsenfeld, G. (1999). An insulator element and condensed chromatin region separate the chicken  $\beta$ -globin locus from an independently regulated erythroid-specific folate receptor gene. *EMBO J.* **18**, 4035–4048.
  19. Trainor, C. D., Evans, T. D. & Felsenfeld, G. (1995). Negative regulation of chicken GATA-1 promoter activity mediated by a hormone response element. *Mol. Endocrinol.* **9**, 1135–1146.
  20. Litt, M. D., Simpson, M., Recillas-Targa, F., Prioleau, M.-N. & Felsenfeld, G. (2001). Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *EMBO J.* **20**, 2224–2235.
  21. Beug, H., von Kirchbach, A., Döderlein, G., Conscience, J.-F. & Graf, T. (1979). Chicken hematopoietic cells transformed by seven strains of defective avian leukemia viruses display three distinct phenotypes of differentiation. *Cell*, **18**, 375–390.
  22. Beug, H., Palmieri, S., Freudenstein, C., Zentgraf, H. & Graf, T. (1982). Hormone-dependent terminal differentiation *in vitro* of chicken erythroleukemia cells transformed by ts mutants of avian erythroblastosis virus. *Cell*, **28**, 907–919.
  23. Minie, M. E., Kimura, T. & Felsenfeld, G. (1992). The developmental switch in embryonic  $\rho$ -globin expression is correlated with erythroid lineage-specific differences in transcription factor levels. *Development*, **115**, 1149–1164.
  24. Pikaart, M. J., Recillas-Targa, F. & Felsenfeld, G. (1998). Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev.* **12**, 2852–2862.
  25. Raich, N., Clegg, C. H., Grofti, J., Romeo, P. H. & Stamatoyannopoulos, G. (1995). GATA-1 and YY1 are developmental repressors of the human epsilon-globin gene. *EMBO J.* **14**, 801–809.
  26. Thomas, M. J. & Seto, E. (1999). Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key? *Gene*, **236**, 197–208.
  27. Weill, L., Shestakova, E. & Bonnefoy, E. (2003). Transcription factor YY1 binds to the murine beta interferon promoter and regulates its transcription capacity with a dual activator/repressor role. *J. Virol.* **77**, 2903–2914.
  28. Brown, J. L., Mucci, D., Whiteley, M., Dirksen, M.-L. & Kassis, J. A. (1998). The *Drosophila* Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. *Mol. Cell*, **7**, 1057–1064.
  29. Reitman, M. & Felsenfeld, G. (1988). Mutational analysis of the chicken  $\beta$ -globin enhancer reveals two positive-acting domains. *Proc. Natl Acad. Sci. USA*, **85**, 6267–6271.
  30. Moscovici, C., Moscovici, M. G., Jimenez, H., Lai, M. M., Hayman, M. J. & Vogt, P. K. (1977). Continuous tissue culture cell lines derived from chemically induced tumors of Japanese quail. *Cell*, **11**, 95–103.
  31. Trainor, C. D., Ghirlando, R. & Simpson, M. A. (2000). GATA zinc finger interactions modulate DNA binding and transactivation. *J. Biol. Chem.* **275**, 28157–28166.
  32. Felsenfeld, G. & Groudine, M. (2003). Controlling the double helix. *Nature*, **421**, 448–453.
  33. Jeong, S. & Stein, A. (1994). Micrococcal nuclease digestion of nuclei reveals extended nucleosome ladders having anomalous DNA lengths for chromatin assembled on non-replicating plasmids in transfected cells. *Nucl. Acids Res.* **22**, 370–375.
  34. Krumm, A., Madisen, L., Yang, X.-J., Goodman, R., Nakatani, Y. & Groudine, M. (1998). Long-distance transcriptional enhancement by the histone acetyltransferase PCAF. *Proc. Natl Acad. Sci. USA*, **95**, 13501–13506.
  35. Hake, S. B., Xiao, A. & Allis, C. D. (2004). Linking the epigenetic “language” of covalent histone modifications to cancer. *Br. J. Cancer*, **90**, 761–769.
  36. Huang, N. E., Lin, C.-H., Lin, Y.-S. & Yu, W. C. Y. (2003). Modulation of YY1 activity by SAP30. *Biochem. Biophys. Res. Commun.* **306**, 267–275.
  37. Yao, Y. L., Yang, W. M. & Seto, E. (2001). Regulation of transcription factor YY1 by acetylation and deacetylation. *Mol. Cell. Biol.* **21**, 5979–5991.

38. Boyes, J., Byfield, P., Nakatani, Y. & Ogryzko, V. (1998). Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature*, **396**, 594–598.
39. Merika, M. & Orkin, S. H. (1995). Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Krüppel family of proteins Sp1 and EKLf. *Mol. Cell. Biol.* **15**, 2437–2447.
40. Boyes, J. & Felsenfeld, G. (1996). Tissue-specific factors additively increase the probability of the all-or-none formation of a hypersensitive site. *EMBO J.* **15**, 2496–2507.
41. Blobel, G. A., Kakajima, T., Eckner, R., Montminy, R. & Orkin, S. H. (1998). CREB-binding protein (CBP) cooperates with transcription factor GATA-1 and is required for erythroid differentiation. *Proc. Natl Acad. Sci. USA*, **95**, 2061–2066.
42. Letting, D. L., Rakowski, C., Weiss, M. J. & Blobel, G. A. (2003). Formation of a tissue-specific histone acetylation pattern by the hematopoietic transcription factor GATA-1. *Mol. Cell. Biol.* **23**, 1334–1340.
43. Tsang, A. P., Visvader, J. E., Turner, C. A., Fujiwara, Y., Yu, C., Weiss, M. J. *et al.* (1997). FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell*, **90**, 109–119.
44. Ikonomi, P., Noguchi, C. T., Miller, W., Kassahun, H., Hardisson, R. & Schechter, A. N. (2000). Levels of GATA-1/GATA-2 transcription factors modulate expression of embryonic and fetal hemoglobins. *Gene*, **261**, 277–287.
45. Leonard, M. W., Brice, M., Engel, J. D. & Papayannopoulou, T. (1993). Dynamics of GATA transcription factor expression during erythroid differentiation. *Blood*, **82**, 1–9.
46. Leonard, M. W., Lim, K. C. & Engel, J. D. (1993). Expression of the chicken GATA factor family during early erythroid development and differentiation. *Development*, **119**, 519–531.
47. Anguita, E., Hughes, J., Heyworth, C., Blobel, G. A., Wood, W. G. & Higgs, D. R. (2004). Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2. *EMBO J.* **23**, 2841–2852.
48. Lee, J. S., Galvin, K. M., See, R. H., Eckner, R., Livingston, D., Moran, E. & Shi, Y. (1995). Relief of YY1 transcriptional repression by adenovirus E1A is mediated by E1A-associated protein p300. *Genes Dev.* **9**, 1188–1198.
49. Atchison, L., Ghias, A., Wilkinson, F., Bonini, N. & Atchison, M. L. (2003). Transcription factor YY1 functions as a PcG protein *in vivo*. *EMBO J.* **22**, 1347–1358.
50. Srinivasan, L. & Atchison, M. L. (2004). YY1 DNA binding and PcG recruitment requires CtBP. *Genes Dev.* **18**, 2596–2601.
51. Caretti, G., Di Padova, M., Micales, B., Lyons, G. E. & Sartorelli, V. (2004). The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. *Genes Dev.* **18**, 2627–2638.
52. Razin, S. V., Shen, K., Ioudinkova, E. & Scherrer, K. (1999). Functional analysis of DNA sequences located within a cluster of DNase I hypersensitive sites colocalizing with a MAR element at the upstream border of the chicken  $\alpha$ -globin gene domain. *J. Cell. Biochem.* **74**, 38–49.
53. Valadez-Graham, V., Razin, S. V. & Recillas-Targa, F. (2004). CTCF-dependent enhancer blockers at the upstream region of the chicken  $\alpha$ -globin gene domain. *Nucl. Acids Res.* **32**, 1354–1362.
54. Mason, M. M., Grasso, J. A., Gavrilova, O. & Reitman, M. (1996). Identification of functional elements of the chicken  $\epsilon$ -globin promoter involved in stage-specific interaction with the  $\beta/\epsilon$  enhancer. *J. Biol. Chem.* **271**, 25459–25467.
55. Weinmann, A. S., Yan, P. S., Oberley, M. J., Huang, T. H. & Farnham, P. J. (2002). Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev.* **16**, 235–244.
56. Hatzis, P. & Talinidis, I. (2002). Dynamics of enhancer-promoter communication during differentiation-induced gene activation. *Mol. Cell*, **10**, 1467–1477.

*Edited by J. Karn*

(Received 7 February 2005; received in revised form 11 April 2005; accepted 19 April 2005)



Review

# Globin genes transcriptional switching, chromatin structure and linked lessons to epigenetics in cancer: A comparative overview <sup>☆</sup>

Georgina Guerrero, Paul Delgado-Olguín, Martín Escamilla-Del-Arenal, Mayra Furlan-Magaril, Eria Rebollar, Inti A. De La Rosa-Velázquez, Ernesto Soto-Reyes, Héctor Rincón-Arango, Christian Valdes-Quezada, Viviana Valadez-Graham, Félix Recillas-Targa <sup>\*</sup>

*Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Apartado Postal 70-242, México, D.F., 04510, Mexico*

Received 6 April 2006; received in revised form 14 September 2006; accepted 22 October 2006  
Available online 11 November 2006

## Abstract

At the present time research situates differential regulation of gene expression in an increasingly complex scenario based on interplay between genetic and epigenetic information networks, which need to be highly coordinated. Here we describe in a comparative way relevant concepts and models derived from studies on the chicken  $\alpha$ - and  $\beta$ -globin group of genes. We discuss models for globin switching and mechanisms for coordinated transcriptional activation. A comparative overview of globin genes chromatin structure, based on their genomic domain organization and epigenetic components is presented. We argue that the results of those studies and their integrative interpretation may contribute to our understanding of epigenetic abnormalities, from  $\beta$ -thalassemias to human cancer. Finally we discuss the interdependency of genetic–epigenetic components and the need of their mutual consideration in order to visualize the regulation of gene expression in a more natural context and consequently better understand cell differentiation, development and cancer.

© 2006 Elsevier Inc. All rights reserved.

*Keywords:* Chromatin; Chromatin domain; Enhancer; Epigenetic; Insulator; Globin genes; Silencer; Transcription

## Contents

1. Introduction . . . . .	751
2. The $\alpha$ - and $\beta$ -globin group of genes: generalities . . . . .	751
3. The chicken $\alpha$ -globin domain and its regulatory components . . . . .	751
4. The chicken $\beta$ -globin domain and regulatory components . . . . .	755
5. The chicken $\alpha$ -globin gene domain and its chromatin organization . . . . .	756
6. The chicken $\beta$ -globin domain and its chromatin organization . . . . .	756
7. Epigenetics and tumour suppressor gene promoters . . . . .	757
8. Conclusions and prospects . . . . .	758
Acknowledgments . . . . .	758
References . . . . .	758

<sup>☆</sup> This paper is part of the 3rd special issue of CBP dedicated to The Face of Latin American Comparative Biochemistry and Physiology organized by Marcelo Hermes-Lima (Brazil) and co-edited by Carlos Navas (Brazil), Rene Belebony (Brazil), Rodrigo Stabeli (Brazil), Tania Zenteno-Savín (Mexico) and the editors of CBP. This issue is dedicated to the memory of two exceptional men, Peter L. Lutz, one of the pioneers of comparative and integrative physiology, and Cicero Lima, journalist, science lover and Hermes-Lima's dad.

<sup>\*</sup> Corresponding author. Tel.: +52 55 56 22 56 74; fax: +52 55 56 22 56 30.

*E-mail address:* [frecilla@ifc.unam.mx](mailto:frecilla@ifc.unam.mx) (F. Recillas-Targa).

## 1. Introduction

Cell lineage commitment in metazoans is established early during development by specific gene expression networks determining highly specific cell differentiation patterns that are mitotically propagated. For several years we have been interested in the relationship between gene regulation and chromatin structure. Epigenetic processes recently emerged as responsible for temporal- and tissue-specific gene expression, not only during cell differentiation and development but also in cancer.

Over the years, comparative genetics had contributed to the understanding of gene organization and conservation among different species (Kosak and Groudine, 2004; Wasserman and Sanderlin, 2004), leading to visualization of a genetic landscape in different organisms. Today, such a view seems to be limited without considering two critical components: the extended non-coding or intergenic sequences and the genome structured into chromatin, i.e., the epigenome. Consequently, to attain an overview of how cellular processes like DNA replication, recombination, DNA repair and differential gene regulation among others occur inside the cell and particularly into the nucleus, we clearly need an integrative genome–epigenome approach (Cremer and Cremer, 2001; Chakalova et al., 2005). Thus a comparative and multidisciplinary approach is nowadays needed to address different genomic processes that could lead us not only to better understand cell physiology but also to propose novel and more effective biomedical strategies to improve human health.

In the present manuscript we present a comparative overview of the  $\alpha$ - and  $\beta$ -globin family of genes, with particular emphasis on the regulation of chicken globin domains, their chromatin organization as well as the epigenetic switching mechanisms occurring during cell differentiation and developmental gene expression. Finally, our knowledge about epigenetic mechanisms of aberrant silencing is linked to tumour suppressor gene promoters in cancer.

## 2. The $\alpha$ - and $\beta$ -globin group of genes: generalities

Since their genomic sequence isolation, globin genes have been one of the most studied loci contributing to the generation of a long list of new concepts and revealing regulatory mechanisms, which are now been extrapolated to other genomic loci in diverse organisms. Part of those observations allowed the introduction of the transcriptional active domain concept that was previously anticipated from cytological evidences, in particular the differential chromosomal band staining (Cremer and Cremer, 2001; Razin et al., 2003). For more than 15 years we have been interested in the study of the chicken  $\alpha$ - and  $\beta$ -globin group of genes (Farache et al., 1990; Recillas-Targa, 2000; Burgess-Beusse et al., 2002; Rincón-Arano et al., 2005). Those investigations have constantly been busted by the large amount of data generated on human and mouse genomes (Stamatoyannopoulos, 2005; Higgs et al., 2005). Even with such accumulation of data and progress, the chicken  $\alpha$ - and  $\beta$ -globin loci remain as an apparent endless source of new information

contributing to novel views for differential gene expression and chromatin conformation at the domain scale, allowing constant proposals for new models.

As in other organisms, the chicken  $\alpha$ - and  $\beta$ -globin domains differentially and coordinately express their genes, which encode distinct types of hemoglobins with variations of oxygen affinity determining their physiological role during development. Each set of genes is grouped as a cluster, which are located in distinct chromosomes, the minichromosome 14 for the  $\alpha$ -globin group of genes and the chromosome 1 for the  $\beta$ -globin domain in the chicken.

## 3. The chicken $\alpha$ -globin domain and its regulatory components

The chicken  $\alpha$ -globin domain is contained in a genomic region of around 40 kilo bases (kb) with an embryonic  $\pi$  gene and two adult  $\alpha^D$  and  $\alpha^A$  genes that are differentially regulated during erythroid differentiation and development (Fig. 1; Recillas-Targa, 2000; Recillas-Targa and Razin, 2001). A series of erythroid-specific and constitutive DNase I hypersensitive sites (DHS) have been identified all along the domain, including the intergenic DHS and those corresponding to individual regulatory elements like promoters, enhancers and structural chromatin components (Fig. 1). Individual promoters contain the classical set of nuclear factor binding sites present in the great majority of other globin genes (Stamatoyannopoulos, 2005; Dean, 2006). Interestingly, a recent publication demonstrated that Sox6 is involved specifically in the  $\epsilon$ -globin gene silencing in definitive erythropoiesis at the mouse  $\beta$ -globin locus (Yi et al., 2006). We hypothesize that Sox6 may contribute, through chromatin remodeling, to the autonomous silencing of the embryonic  $\pi$  gene when adult genes are needed to

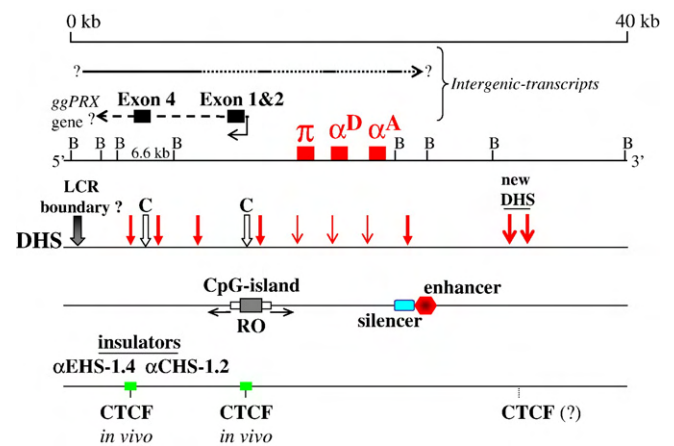


Fig. 1. The chicken  $\alpha$ -globin domain regulatory and chromatin structure elements. The embryonic  $\pi$  gene, and adult  $\alpha^D$  and  $\alpha^A$  genes are shown. *Bam*HI restriction enzyme recognition sites are indicated (B). Intergenic transcripts are presented, although their starting and ending points are not clearly known. Part to the *ggPRX* gene transcript I represented. Distribution of DNase I hypersensitive sites (DHS) is indicated by vertical arrows, with alternate presence of two constitutive (C) sites on the domain 5'-side. Regulatory elements upstream and downstream of the  $\alpha$ -globin genes are shown. The two newly defined insulators are located at the domain 5'-side ( $\alpha$ EHS-1.4 and  $\alpha$ CHS-1.2) and the *in vivo* CTCF binding is indicated.

be expressed. This is an aspect that we are currently investigating in our laboratory. Thus, apparent minor differences in terms of the set of nuclear factors interacting at individual promoters may contribute, in part, to the differential  $\alpha$ -globin gene expression.

A set of hematopoietic transcription factors seems to be recurrently present in the great majority of globin *cis*-regulatory elements; such set includes GATA-1, GATA-2, EKLF, NF-E2 and PU.1 among other erythroid and ubiquitous nuclear factors (Huo and Zhang, 2005). The two main erythroid transcription factors are GATA-1 and GATA-2, which correspond to zinc finger nuclear factors. GATA-1 is required for erythroid precursor cells maturation and interacts with all globin regulatory elements with positive and negative effects, mediated by an increasing number of co-factor interactions, over gene expression. In particular, GATA-1 interacts with a bivalent co-factor named *Friend Of GATA* (FOG) that is able to recruit histone acetylases (HAT) or histone deacetylases (HDAC) (Cantor and Orkin, 2005). FOG-1 co-repressor effect is mediated through its interaction with the ATP-dependent remodeling complex NuRD (Hong et al., 2005). More recently, it has been proposed that GATA-1 interaction with FOG-1 is needed for the formation of a tissue-specific chromatin loop that facilitates long-range enhancer function (Vakoc et al., 2005). Furthermore, GATA-1 can interact with the myeloid and lymphoid specific factor PU.1 with (positive or negative) effects on erythroid and myeloid differentiation (Stopka et al., 2005; Liew et al., 2006). Data from our laboratory demonstrated GATA-1 physical contact with the bivalent transcription factor YY1 that contributes to the negative modulation of the chicken 3'-side  $\alpha$ -globin enhancer (Rincón-Arano et al., 2005). In addition, Boyes et al. (1998) demonstrated that GATA-1 can be acetylated by CBP/p300 HAT, increasing its DNA binding affinity with stimulatory GATA-1-dependent transcription effects. In summary, GATA-1 is a critical erythroid factor with multifaceted regulatory behavior. GATA-2 can also be acetylated but in contrast to GATA-1, it regulates early erythropoiesis in immature progenitors, and when associated with co-factors plays repressive or activating roles on transcription regulation (Huo and Zhang, 2005).

The erythroid krüppel-like factor (EKLF) influences chromatin organization of globin loci during differentiation by recruiting HATs and HDACs. EKLF binds the CACCC sequence motif present in enhancer and LCR elements, and is expressed in erythroid, megakaryocytic and mast cells (Huo and Zhang, 2005). A recent publication showed the isolation and characterization of the chicken EKLF gene that is expressed in the posterior primitive streak, which gives rise to hematopoietic cells, then in the blood islands and in circulating blood cells (Chervenak et al., 2006).

NF-E2 is a heterodimer that belongs to the small Maf proteins family composed by the ubiquitous NF-E2p18 (also known as MafK) and the hematopoietic-specific subunit NF-E2p45 (Andrews et al., 1993). NF-E2 is a leucine-zipper transcriptional activator that binds the consensus sequence TGCTGA(G/C)TCA found in erythroid-specific regulatory elements (Huo and Zhang, 2005). NF-E2 exerts its activator function by recruit-

ment of chromatin remodeling complexes and participates in long-range transfer of RNA polymerase II to globin promoters (Johnson et al., 2001; Huo and Zhang, 2005). More recently, NF-E2p45 subunit has been found to be sumoylated *in vivo* in human erythroid K562 cells (Shyu et al., 2005). NF-E2p45 sumoylation increases NF-E2 *trans*-activation potential together with its DNA binding affinity besides favoring relocation of NF-E2p18 and NF-E2p45 subunits inside the nucleus (Francastel et al., 2001), with consequences on erythroid maturation (Shyu et al., 2005).

In conclusion, erythroid transcription factors interact with tissue-specific and ubiquitous proteins including chromatin modifiers, which in turn could favor the establishment of particular chromatin configurations at specific regulatory elements, affecting epigenetic erythroid gene regulation. However the challenge now is to integrate this information with the aim to understand gene regulation in the context of epigenome and nuclear dynamics (Chakalova et al., 2005; West and Fraser, 2005).

Promoters and silencer–enhancer elements had been identified at the chicken 3'-side of  $\alpha$ -globin locus, located around 450 base pairs (bp) downstream of the adult  $\alpha^A$  gene (Fig. 1; Recillas Targa et al., 1993; Rincón-Arano et al., 2005). The 3'-side  $\alpha$ -globin enhancer is a strong regulatory element capable to activate individual promoters in a regulated way even though the precise mechanisms are under investigation in our group. The enhancer possesses a modest similarity with the chicken  $\beta^A/\epsilon$  enhancer of the  $\beta$ -globin domain (Reitman and Felsenfeld, 1988), however recent observations from our group determined that the  $\alpha$ -globin enhancer is more complex than initially thought (Fig. 2A and C; Rincón-Arano et al., 2005). *In vitro* DNase I footprinting experiments demonstrated multiple interactions of nuclear factors with a DNA fragment of around 400 bp (Fig. 2A and C). Furthermore, we identified a relatively long DNA fragment of 120 bp highly conserved between duck and chicken that turned out to include a fourth GATA-1, in addition to NF-E2 and EKLF binding motifs (Fig. 2B; Recillas Targa et al., 1993). Other factors like Sp1 also interact with such region; in addition, we recently demonstrated the presence of the bivalent YY1 nuclear factor (Fig. 2C; Rincón-Arano et al., 2005). We have also shown that the enhancer activity could be modulated in early erythroid stages, at least to some extent, by the action of YY1 and its associated co-factors, probably through affecting the local chromatin structure (Rincón-Arano et al., 2005). In addition, co-immunoprecipitation and pull-down experiments demonstrated a novel physical interaction between GATA-1 and YY1 that influences enhancer action. Point mutations of YY1 and GATA-1 motifs were tested in transient transfections. The results confirmed that GATA-1 and YY1 association is in part responsible for the down-regulation of the enhancer function in early erythroid stages (Rincón-Arano et al., 2005).

Thus we demonstrated that the 3'-side chicken  $\alpha$ -globin enhancer is differentially activated during distinct erythroid differentiation stages and that YY1-GATA-1 interaction contributes to the plasticity of the enhancer function. Furthermore, the potential combinatorial distribution of DNA binding proteins

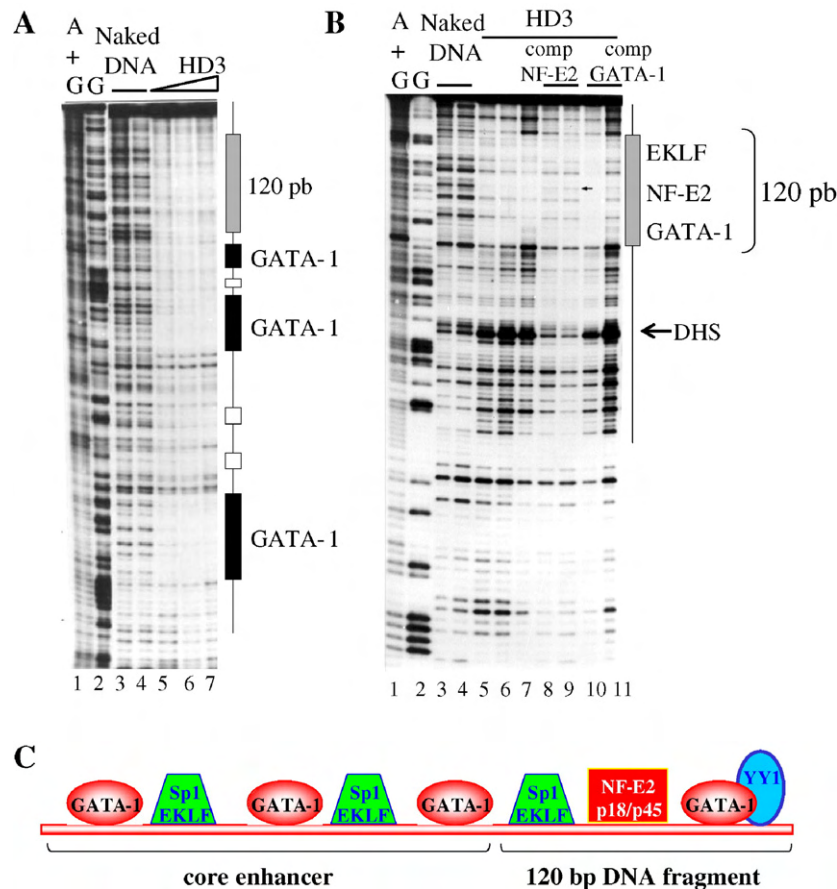


Fig. 2. *In vitro* DNase I footprinting and scheme of the DNA-protein interactions over the 3' chicken  $\alpha$ -globin enhancer. (A) *Hin*P11–*Hin*I (285 bp) DNA fragment containing the full enhancer was radioactive labeled, incubated with increasing concentrations of chicken HD3 cells nuclear extracts and *in vitro* digested with DNase I nuclease. Several footprints are seen in addition to the corresponding GATA-1 protected sequences. In particular, several GGTGG sequences and two Sp1-like motifs are also protected (open boxes). Notice a large footprinted area initiating after the third GATA-1 motif that includes the 120 bp DNA fragment. (B) Characterization of DNA–protein interactions over the 120 bp DNA fragment with a large footprint of around 60 bp (see lanes 5, 6 and 10). We carried out competitions with known sequences that bind NF-E2 from the human porphobilinogen deaminase erythroid promoter and GATA-1 from the chicken  $\beta$ -globin enhancer. Notice the formation of an intense DHS, 40 bp downstream of the protected area. When competing with two concentrations of cold NF-E2 binding sequence, we observe the re-appearance of a band located just in the middle of the NF-E2 recognition sequence with the 120 bp DNA fragment (see small arrow on lanes 8 and 9). Interestingly, this competition causes the loss of the DHS DNA conformation that creates a strong DHS. For GATA-1 competition we observe the re-appearance of the majority of the bands seen in the naked DNA digestion (see lanes 10 and 11). (C) Summary of the defined nuclear factor interactions to the chicken  $\alpha$ -globin 3'-side enhancer.

over the region predicts a differential action of erythroid and non-erythroid transcription factors influenced by polypeptide concentrations, competition for DNA binding motifs and chromatin remodeling activities.

Another regulatory component that recurrently appears in the globin loci is the locus control region or LCR that in conjunction with enhancers and promoters participates in specific gene expression switching (Dean, 2006). LCRs were initially defined as long-distance regulatory elements composed by a set of DNase I hypersensitive sites that influence gene expression and perform a dominant and tissue-specific chromatin opening activity at a genomic domain scale (Bulger and Groudine, 1999; Dean, 2006). From the functional point of view an LCR can be described as a regulatory element that positively drives the expression of a linked transgene dependently on the copy-number and genomic integration site (Grosveld et al., 1987). Among several mechanistic models, recent investigations demonstrated that the  $\beta$ -globin LCR is actively involved in chromosomal loop formation (Bulger and Groudine, 1999;

Chakalova et al., 2005; West and Fraser, 2005). Groudine and collaborators showed that the LCR seems to participate in domain relocalization, which favors transcriptional activation through placing the domain in an active transcription factory inside of the nucleus (Chakalova et al., 2005; Ragozy et al., 2006). Therefore, the LCR seems to play a topological role on tissue-specific genomic domain formation, suggesting its involvement in a domain scale regulatory hierarchy fine-tuning the activities of enhancers and promoters, leading to gene activation or repression.

In comparison with other globin domains, including the chicken  $\beta$ -globin locus, no locus control region (LCR) has been identified in the chicken  $\alpha$ -globin domain (Dean, 2006). At this point, the possibility of the presence of an LCR or other enhancer(s) cannot be discarded. Their existence is under investigation and we have found a novel group of DHS, with apparent erythroid specificity, on the non-coding 3'-side of the domain (Fig. 1). On the other side, a chromatin transition was found by Higgs and collaborators around 20 kb upstream of the

embryonic  $\pi$  gene, where the presence of an LCR had been theoretically predicted (Anguita et al., 2001).

Much less is known about the molecular features of the chicken 3'  $\alpha$ -globin silencer. What we have found until now is that the silencer is located side-by-side to the enhancer and that such location may give particular regulatory properties to both elements. We believe that their activity might be interconnected by means of their chromatin organization. We know that at least *in vitro* three unidentified nuclear factors interact with the silencer element (Recillas Targa et al., 1993). Then, we hypothesize that the silencer could be functional during primitive differentiation events or at some developmental stage when  $\alpha$ -globin genes need to be differentially silenced (see below).

To better understand the function of this dual regulatory element, we have systematically characterized its local chromatin structure. Briefly, micrococcal nuclease and restriction enzyme accessibility experiments suggested that the chromatin at the silencer, where two nucleosomes are positioned, presents an open conformation in early stages of development, and becomes progressively “closed” as development progresses (Escamilla-Del-Arenal and Recillas-Targa, in preparation). Several possible scenarios arise from these results; the silencer could directly interfere with the embryonic gene promoter function *in vivo* or it acts over the enhancer controlling its regulatory functions. Alternatively, the so-called “silencer” may be part of the enhancer and collaborate with its activity in certain circumstances. In conclusion the enhancer and silencer could establish complex functional interactions, with potential implications in chromatin structure and differential gene expression of the  $\alpha$ -globin domain, which still wait to be unraveled.

Over the years the globin gene domains in distinct organisms have demonstrated to be a source of new data that founded novel and unsuspected concepts. Of particular interest is the existence of intergenic transcripts called by Scherrer “giant” or “Full domain transcripts” that correspond to long pre-mRNA that include non-coding and coding sequences (for review Scherrer, 2003). Long time ignored, non-coding transcripts were suggested to contribute to loop formation at a domain scale with a primary topological role that derives in the protection, storage and stability of transcripts, besides processing of those long pre-mRNA's (Scherrer, 2003). Such pioneer vision initiated in the 60s was more recently supported by novel proposals for their functions (Ashe et al., 1997; Gribnau et al., 2000; Razin et al., 2004; Haussecker and Proudfoot, 2005). Today, globin intergenic transcripts are thought to facilitate the sub-domain chromatin opening at globin loci (Gribnau et al., 2000) or mediate RNAi-dependent chromatin silencing at a domain scale (Haussecker and Proudfoot, 2005).

Another aspect of the  $\alpha$ -globin domain that we have been concerned lately in collaboration with Sergey Razin, that seems to be related to the presence of intergenic transcripts, is the identification of a silencer element co-localizing with a CpG-island situated in the 5' non-coding region of the chicken  $\alpha$ -globin domain (Fig. 1; Klochkov et al., 2006). Such silencer is CTCF-dependent and based on several evidences, we propose that this novel regulatory element may regulate the expression of *ggPRX*

housekeeping gene, which begins around 4 kb upstream of the embryonic  $\pi$  gene and is transcribed on the opposite direction of the globin genes. The sequence analyzed corresponds to a CpG-island contacted by CTCF in erythroid cells but not in chicken lymphoid cells, where CTCF binding site and surrounding sequence is hypermethylated (Fig. 3; Klochkov et al., 2006).

At this point a possible model is that the silencer element is active only in erythroid cells where the CTCF recognition site is not methylated, in contrast to lymphoid and non-erythroid cells (Fig. 3). Activation of the silencer in erythroid cells is likely to miss-regulate *ggPRX* gene transcription leading to a situation in which the promoter of this gene may be influenced by the putative upstream LCR, which is needed to activate the  $\alpha$ -globin genes and regulate domain chromatin conformation (Fig. 3). In non-erythroid cells hypermethylation of the area adjacent to the *ggPRX* gene promoter could induce the inactivation of CTCF blocking activity in addition to the inactivation of globin *cis*-regulatory elements allowing full expression of the *ggPRX* gene (Fig. 3). Alternatively, we do not discard the possibility that *ggPRX* transcript participates in some way in the establishment of domain chromatin structure.

What seems very attractive is the fact that upstream non-coding sequences of the chicken  $\alpha$ -globin domain contain intergenic transcripts (Gribnau et al., 2000; Andersen and Panning, 2003; Razin et al., 2004; Bernstein and Allis, 2005). Preliminary data from our group suggest that in distinct stages of development double-stranded non-coding RNAs are synthesized. If such transcripts contribute not only to gene expression regulation but also to chromatin domain organization through regulated recruitment of chromatin remodeling activities is currently under investigation (Gribnau et al., 2000; Haussecker and Proudfoot, 2005). Interestingly, two recent publications support the role of Polycomb and Trithorax protein family

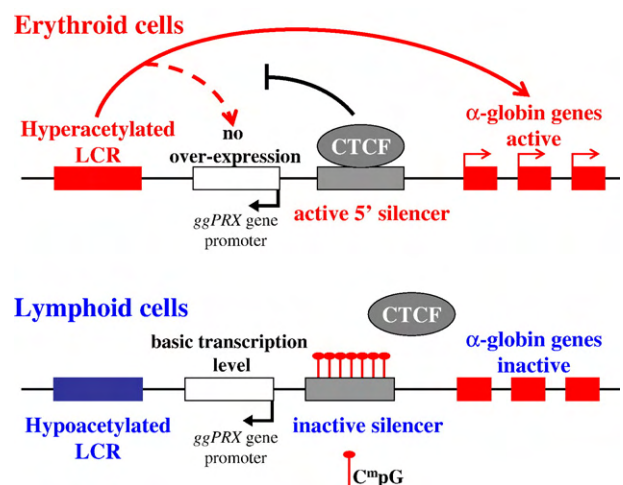


Fig. 3. Model for action of the chicken  $\alpha$ -globin gene 5' silencer. On erythroid cells, CTCF is able to bind *in vivo* to the *ggPRX* gene promoter upstream element (Klochkov et al., 2006), located around 4 kb upstream of the embryonic  $\pi$  gene. The silencer would attenuate or bypass the action of the LCR over the *ggPRX* gene allowing LCR specific regulation of  $\alpha$ -globin genes. In non-erythroid cells, hypermethylation and changes in chromatin conformation will allow normal *ggPRX* gene expression.

members in the establishment of close or open, respectively, chromatin conformations through non-coding RNA intermediates (Sanchez-Elsner et al., 2006; Grimaud et al., 2006).

In summary, the  $\alpha$ -globin locus is apparently much more complex in terms of its differential gene regulation than previously anticipated. Moreover, there are elements that are partially defined and their *in vivo* contributions remain to be elucidated.

#### 4. The chicken $\beta$ -globin domain and regulatory components

As mentioned, the chicken  $\beta$ -globin group of genes is located in the chromosome 1, and contains one embryonic  $\epsilon$  gene and a fetal  $\rho$  gene that are expressed in primitive erythroid lineages (Fig. 4). In definitive erythroid cells two adult genes,  $\beta^A$  and  $\beta^H$  are differentially expressed during development. The  $\beta$ -globin locus possess two main regulatory elements: the  $\beta^A/\epsilon$  enhancer, located between the adult  $\beta^A$  gene and the  $\epsilon$  embryonic gene promoter, and the upstream LCR composed by three erythroid specific DHS (Fig. 4). Even though individually each element has been extensively characterized, the overall participation of these elements in the  $\beta$ -globin expression regulation during development and erythroid differentiation remains unclear. The  $\beta^A/\epsilon$  enhancer was one of the first regulatory elements demonstrated to activate one gene over another by promoter access competition (Foley and Engel, 1992). Competition between stage-specific transcription factors for promoter elements dictates the developmental activation of the adult  $\beta^A$  or the embryonic  $\epsilon$  gene through enhancer action, since this element is located at an equivalent distance between both

genes (Foley and Engel, 1992). In addition the  $\beta^A/\epsilon$  enhancer and the upstream chicken  $\beta$ -globin LCR drive copy-number-dependent expression of a reporter gene in transgenic mice (Reitman et al., 1990).

As mentioned before the  $\beta$ -globin LCR, containing well-characterized DHS, in collaboration with the  $\beta^A/\epsilon$  enhancer is able to drive position-independent gene expression in transgenic mice (Reitman et al., 1990; Abruzzo and Reitman, 1994; Mason et al., 1995). Individually, cHS2 and cHS3 showed significant enhancer activity while cHS1 did not. In particular, cHS2 has been studied in detail and it has been demonstrated that its activity mainly depends on GATA-1, with an apparent contribution of Sp1 and the erythroid EKLf factor (Abruzzo and Reitman, 1994). Functional experiments showed that individually or in conjunction with the other regulatory element fragments, the 5' DHS drive different enhancer and position-independent expression patterns when tested in transgenic mice (Mason et al., 1995). When the entire locus is analyzed, the copy number-dependent expression is maintained with an additional 6-fold increase in expression (Mason et al., 1995). Interestingly, transgenic mice carrying the entire chicken  $\beta$ -globin domain express the globin genes with the proper chicken developmental program.

All this experimental evidence supports the functionality of the chicken  $\beta$ -globin domain LCR element, which requires coordinated action of the four upstream hypersensitive sites and the  $\beta^A/\epsilon$  enhancer. At the 3'-side of the  $\beta$ -globin domain there is no evidence of the existence of a regulatory element equivalent to the 3'-side  $\alpha$ -globin silencer–enhancer, demonstrating a clear difference between both globin domains that are located in distinct chromosomes of the chicken genome. On the other hand, the presence of an LCR in the  $\alpha$ -globin locus remains as an unsolved question.

Epigenetic globin regulation involves DNA methylation. DNA methylation has been demonstrated to occur in avian  $\alpha$ -globin gene promoters during development (Singal et al., 2002a,b). In particular, the chicken embryonic  $\pi$  gene promoter *de novo* methylation correlates with gene silencing at 8 days of development (Singal et al., 2002b). In other words, in 5 days old chicken embryos the embryonic  $\pi$  gene is active and hypomethylated. Embryonic gene promoter methylation is consistent with the association of methyl-CpG-binding proteins and histone deacetylation in 11 days old erythrocytes where the embryonic gene is already silenced (Singal et al., 2002b). These results suggested us that sequential and coordinated epigenetic mechanisms should normally be activated for the so-called autonomous silencing of the embryonic  $\alpha$ -globin  $\pi$  gene that is followed by adult gene expression (Raich et al., 1990). We speculate that the initial step towards  $\pi$  gene silencing could involve the dissociation of the transcription factors regulating the promoter, like GATA-1 and Sp1, and/or the incorporation of a nuclear factor recruiting co-repressors. Those co-repressors would attract chromatin remodelers like histone deacetylases and methylases or even DNA methylation activities. We propose DNA methylation as the last epigenetic repressive component to create the autonomous and in some way irreversible, silencing of  $\pi$  gene when its expression is anymore needed

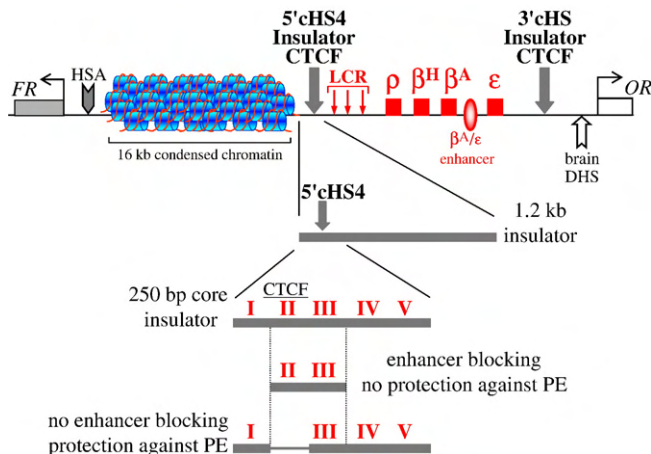


Fig. 4. The chicken  $\beta$ -globin domain and neighbor genomic loci. An erythroid-specific LCR and the strong  $\beta^A/\epsilon$  enhancer are responsible for regulation of the  $\beta$ -globin domain. Domain boundaries (insulators 5'cHS4 and 3'cHS) as well as *in vitro* and *in vivo* CTCF association sites have been determined. The 5'cHS4 constitutive DHS has insulator properties, a CTCF-dependent enhancer-blocking activity, and protects a transgene against position effects and silencing. Five protein-binding sites (I to V) are schematically represented under the map of the locus. Experiments demonstrated that the two cHS4 properties are separable. The constitutive 3'cHS site also binds CTCF but has only enhancer-blocking activity. The neighbor genomic loci are composed by 16 kb of highly condensed chromatin followed by an erythroid specific folate receptor gene (*FR*) apparently regulated by a DHS named HSA. On the 3'-side, a brain specific olfactory receptor gene (*OR*) constitute the next domain.



during development. Addressing these possibilities in detail is in need since the generated knowledge could be extrapolated to for example early epigenetic programming of stem cells and epigenetic miss-regulation of tumour suppressor genes or oncogenes in cancer (see below).

### 5. The chicken $\alpha$ -globin gene domain and its chromatin organization

If we compare the chicken  $\alpha$ -globin with the  $\beta$ -globin domain, the first aspect to be mentioned is that the  $\alpha$ -globin domain in chicken and other organisms is situated in a genomic region of constitutive open chromatin configuration (Recillas-Targa, 2000; Flint et al., 2001; Razin et al., 2003; Higgs et al., 2005). Based on the presence of all the regulatory elements described and the presence of intergenic transcripts (not characterized at the  $\beta$ -globin cluster) it is not surprising that the  $\alpha$ -globin domain is situated in a more relaxed chromatin environment in comparison to the  $\beta$ -globin domain, which is clearly shielded and its chromatin structure modulated (see below). In addition to the presence of non-globin genes over the domain, the mouse and human  $\alpha$ -globin locus are located in relative close proximity to telomeres (Flint et al., 2001; Higgs et al., 2005). All this evidences support the need of a constitutively open chromatin conformation in the  $\alpha$ -globin domain in order to counteract a possible silencing effect from the telomeres.

We have recently characterized a group of three DHS located 11 kb upstream the embryonic  $\pi$  gene in the  $\alpha$ -globin domain (Fig. 1). We have found that the two most distal DHS have characteristics of a chromatin insulator element (Valadez-Graham et al., 2004). The third hypersensitive site apparently corresponds to a silencer element. Interestingly, both elements named  $\alpha$ EHS-1.4 and  $\alpha$ CHS-1.2 possess enhancer-blocking activity and for the  $\alpha$ EHS-1.4 such activity is CTCF-dependent (Valadez-Graham et al., 2004). Furthermore, we have demonstrated *in vitro* and *in vivo* CTCF binding to the  $\alpha$ EHS-1.4 element (Valadez-Graham et al., 2004). We are now testing the second property of insulators, which is the capacity to protect a transgene against chromosomal position effects (Pikaart et al., 1998; Recillas-Targa et al., 2002, 2004), of these enhancer blockers. But the relevant question is if these elements represent some sort of domain boundary. Apparently this is not the case since Higgs and collaborators had found a sharp histone hyperacetylation transition around 6 kb further upstream from the CTCF binding site of the  $\alpha$ EHS-1.4 insulator (Fig. 1; Anguita et al., 2001). Then, at this point the role of these novel insulators in the regulation and chromatin structure of the chicken  $\alpha$ -globin gene domain it is not known. Two proposals arise that in first instance do not seem mutually exclusive. First, the enhancer blocking activities may contribute to intergenic transcripts regulation, avoiding undesired crossover among regulatory signals or blocking more upstream non-globin elements. CTCF, which is distributed all over the domain, could participate in topological organization through mediating differential loop formation as shown for the *Igf2/H19* imprinted loci (Reik et al., 2004; Chakalova et al., 2005). Such model is further supported

by CTCF multimerization capacity and interaction with the nuclear matrix components (Valadez-Graham et al., 2004; Yusufzai et al., 2004; Yusufzai and Felsenfeld, 2004).

Concerning the domain 3'-side chromatin structure, much less is known. We have recently identified three DHS downstream of the adult  $\alpha^A$  gene in proximity to the previously described CR1-repeats (Fig. 1; Farache et al., 1990). In addition, over that genomic region sequences associated to the nuclear matrix, topoisomerase binding consensus sequences and AT-rich DNA segments have also been found (Farache et al., 1990). We are currently interested in defining the functional properties of those sequences, post-translational histone modifications and whether those sequences correspond to the 3'-side boundary of the domain.

Thus, the  $\alpha$ -globin domain incorporates multiple elements related to chromatin structure regulation, which have not been integrated in a functional context. What seems most important is to address the relationship between domain chromatin architecture and developmental  $\alpha$ -globin gene expression. Such complex relationship may involve epigenetic silencing of embryonic or adult genes through at least histone deacetylases and DNA methylation.

### 6. The chicken $\beta$ -globin domain and its chromatin organization

Historically, Felsenfeld and collaborators invested many years of intense investigations leading to the establishment of the chicken  $\beta$ -globin locus as a paradigmatic model system (Burgess-Beusse et al., 2002). We previously mentioned the  $\beta$ -globin locus, containing three erythroid specific DHSs. In fact, there is a fourth DHS, known as cHS4, which is constitutive and clearly delineates the 5' boundary of the domain (Litt et al., 2001a). Interestingly, the 3'-side of the chicken  $\beta$ -globin locus is also defined by the presence of a 3'cHS DHS (Saitoh et al., 2000; Litt et al., 2001a). Such DHS distribution correlates with high levels of histone hyperacetylation, in particular in stages where the  $\beta$ -globin genes are expressed in erythroid cells (Litt et al., 2001a,b). However, the genomic sequence upstream of 5' cHS4 and downstream of the 3'cHS are marked by a drastic drop in DNase I sensitivity and the incorporation of histone repressive marks (Litt et al., 2001b). Moreover, the chromatin upstream of 5' cHS4 site remains in a constitutively condensed or heterochromatin configuration along 16 kb followed by an erythroid folate receptor gene as the next genomic domain (Fig. 4; Litt et al., 2001a; Burgess-Beusse et al., 2002). On the other side of the  $\beta$ -globin domain, downstream to the 3'cHS, an odorant receptor gene with an apparent brain specific gene expression pattern has been found, which constitutes the downstream neighboring genomic domain (Fig. 4; Saitoh et al., 2000). All together, these observations showed a well defined transition from open to a closed chromatin organization upstream and downstream of each constitutive DHS flanking the chicken  $\beta$ -globin domain, supporting the concept of genome organization into chromatin domains (Razin et al., 2003). Hence, these data strongly suggests that insulators and/or boundaries may be contributing to epigenome organization by

delimiting chromatin structure transition areas (Burgess-Beusse et al., 2002; West and Fraser, 2005).

Pioneer work from Felsenfeld research group defined the 5' boundary of the chicken  $\beta$ -globin locus as a 1.2 kb DNA fragment, which possess the two classical properties of insulators: 1) positional enhancer-blocking and 2) protection of a transgene against chromosomal position effects (Burgess-Beusse et al., 2002). Interestingly, both boundary elements (the 5' cHS4 and the 3'cHS), function as enhancer blockers in a CTCF-dependent manner (Burgess-Beusse et al., 2002). Concerning the second property of insulators, initially two copies of the 1.2 kb DNA fragment harboring the 5' cHS4 where shown to protect against chromosomal position effects in distinct cell types and organisms (Pikaart et al., 1998; Recillas-Targa et al., 2002, 2004). Further characterization demonstrated that the minimal element required corresponds to two copies of a 250 bp DNA fragment including five previously defined sequences named footprint FI to FIV, which are recognized by nuclear factors (Fig. 4; Recillas-Targa et al., 2002). Surprisingly, systematic deletion analysis demonstrated that CTCF is not required for protection against position effect activity by the 5' cHS4 insulator (Recillas-Targa et al., 2002). These observations suggest a cohabitation of the two insulator properties in the 5' cHS4 fragment. Further investigations from Felsenfeld group have defined with much more detail the features of the minimal core insulator element, and found that instead of being functionally bipartite, it corresponds to a tripartite insulator element. The chicken  $\beta$ -globin 5' cHS4 insulator is able to: 1) block enhancer promoter interaction in a CTCF-dependent manner, 2) protect a transgene against chromosomal position effect through the active recruitment of proteins favoring histone covalent modifications like histone acetylation and methylation (mediated by the following transcription factors: USF1/2, SET7/9, PRMT1, p300 and CBP) and 3) block epigenetic silencing by DNA methylation probably mediated by BGP factor (West et al., 2004; Huang et al., 2005; Felsenfeld personal communication).

Most recent investigations from our laboratory demonstrated that the core chicken 5' cHS4  $\beta$ -globin insulator is also able to protect a transgene against telomeric position effects. Such a result supports the protective role of the 5' cHS4 insulator even in the presence of one of the most dominant and strong source of epigenetic silencing, confirming the virtues of this insulator (Rincón-Arano and Recillas-Targa, in preparation).

On the other hand, the 3'cHS chromatin insulator is not able to protect a transgene against chromosomal position effects (Recillas-Targa et al., 2002). Therefore, the 3'cHS insulator maintains only one of the two insulator properties. Furthermore, there is a noticeable absence of the histone marks that constitutively seen at the 5' cHS4 insulator as shown by chromatin immunoprecipitation assays (Litt et al., 2001a). These observations suggest that protection against position effects depends on the genomic organization of the  $\beta$ -globin loci and neighbor domains (Fig. 3). Downstream of the 3' domain boundary, a condensed chromatin region equivalent to the 16 kb described at the 5'-side is absent. Moreover, the adjacent gene is expressed exclusively in brain (Felsenfeld personal communication). Thus, we can argue that in the case of the 3'cHS insulator

only enhancer-blocking activity (if there is any *in vivo*) is needed. In contrast, the 5'cHS4 boundary needs both activities, in particular the one desirable to stop a theoretical spreading of repressive chromatin (Fig. 4).

Therefore, specific chromatin modifications can be correlated with differentiation and developmental changes over the chicken  $\beta$ -globin locus and its neighboring genomic domains. The significance of the insulator elements surrounding the  $\beta$ -globin domain and their possible role in establishing and/or maintaining these specific chromatin modifications throughout the domain remain to be clarified.

## 7. Epigenetics and tumour suppressor gene promoters

The comparative view of the  $\alpha$ - and  $\beta$ -globin gene clusters has been illustrated to confirm the relevance of epigenetic regulation during erythroid differentiation and development. Such aspects motivated us to explore some of the epigenetic regulatory phenomena that occur on tumour suppressor genes in normal and pathological conditions. In particular, our main interest is to understand how tumour suppressor gene promoters, commonly corresponding to CpG-islands, are in normal conditions protected against aberrant DNA methylation (Antequera, 2003; Herman and Baylin, 2003). One more time, a comparative approach was followed to analyze two distinct tumour suppressor gene promoters: the human *p53* and retinoblastoma (*Rb*) gene promoters. Retinoblastoma gene promoter represents a classical CpG-island with 27 CpGs in its core element (400 bp). In contrast, the human *p53* gene promoter only has 16 CpGs and strictly should not be considered as a CpG-island. Previous observations from our group and others clearly demonstrated that both promoters can be inactivated by epigenetic mechanisms (Ohtani-Fujita et al., 1997; Stirzaker et al., 1997; Aguirre et al., 2003; Sidhu et al., 2005). In our laboratory we have recently demonstrated that in normal conditions *p53* gene expression is maintained in basal levels through epigenetic mechanisms involving the establishment of repressive, and apparently reversible, chromatin marks upon *p53* transcription is needed. In the case of *Rb* gene promoter, we have discovered a novel transcriptional regulator, the 11-zinc finger multivalent CTCF factor (Ohlsson et al., 2001; Klenova and Ohlsson, 2005). We have found that CTCF contributes positively to the *Rb* gene expression regulation; moreover, preliminary data supports a protective role of CTCF against DNA methylation (Recillas-Targa et al., 2006). Thus, comparatively these two tumour suppressor gene promoters are regulated, in part, through epigenetic mechanisms in their normal context. In contrast, epigenetic defects have been evidenced with a growing frequency (Feinberg et al., 2006). Comparative studies turn out to be quite instructive since there is growing evidence showing that epigenetic defects participate in initial steps and during cancer development (Feinberg et al., 2006).

As an attractive view, Jean-Pierre Issa (2004) had recently proposed what he has called the "CpG-island methylator phenotype" to explain the origins of aberrant DNA methylation and epigenetic silencing of CpG-islands. Briefly, the "methylator phenotype" establishes a central concept in which DNA

methylation patterns should allow certain degree of tumour classification according to expression of genes affected by DNA methylation. In terms of possible mechanisms, it has been suggested that DNA methylation spreading could initiate at “methylation centres” in non-coding regions encompassing large genomic areas including repetitive sequences and in the worst scenario, invading regulatory elements, like tumour suppressor gene promoters (Herman and Baylin, 2003; Issa, 2004). Furthermore, Issa proposed that boundary elements might be present delimiting CpG-islands, protecting them against DNA methylation. In fact, such hypothetical boundary may not be designed exclusively against DNA methylation, but rather could contribute to protect against other abnormal repressive epigenetic modifications like histone deacetylation, or activities displayed by chromatin remodeling complexes or members of the Polycomb family of proteins (Lund and van Lohuizen, 2004). It can be thought that DNA methylation could, not exclusively, stand as one of the latest in the series of events leading to epigenetic silencing (Mutskov et al., 2002). These arguments allow us to insist that DNA methylation is one but not the only source of epigenetic silencing, supporting the multiple levels of epigenetic influence, interdependency and complexity operating during anomalous gene regulation in diseases.

Our findings bear the model in which CTCF can have two main roles in the regulation of tumour suppressor gene promoters (Klenova et al., 2002; Klenova and Ohlsson, 2005). The proposal is a topological function of CTCF on those promoters through their association to the nuclear matrix and a protective role of CpG-islands against DNA methylation and maybe other epigenetic modifications (Recillas-Targa et al., 2006). Then, the possibility of the existence of a novel class of transcription factors is now open, with bivalent or even multivalent properties, with the ability to regulate (in different and versatile ways) epigenetic processes.

## 8. Conclusions and prospects

Comparative studies have provided unexpected amount of biological information at all levels. Here we have presented a comparative analysis of the chicken  $\alpha$ - and  $\beta$ -globin group of genes from the regulatory and chromatin structure perspective. We have also attempted to present in a logical way the relevant findings from our group and other laboratories. We conclude that, even though the highly precise requirements of  $\alpha$ - and  $\beta$ -globin peptide chains for hemoglobin formation, both genomic loci are found in distinct chromosomes and uniquely regulate differential gene expression; accordingly, their domain organization and chromatin structure differs. The  $\beta$ -globin domain is clearly defined and shielded by chromatin boundaries or insulators. In contrast, the  $\alpha$ -globin locus presents a more complex chromatin organization, with no LCR and clear boundaries defined yet, besides intergenic transcripts with no-clear function. All these observations motivated us to raise some questions about the participation of distinct epigenetic processes, like silencing of tumour suppressor gene promoters in human diseases. We propose that a novel kind of transcription factors may contribute

to maintain a CpG-island in a state that avoids abnormal gene expression silencing. The relevance of these phenomena is based on the fact that such scenario turns to be present more commonly in cancer than previously anticipated (Feinberg et al., 2006; Baylin and Ohm, 2006).

New data concerning the eukaryotic epigenome supports the domain hypothesis for genome organization (Razin et al., 2003). Cell nucleus and sub-nuclear compartments, chromatin remodeling processes and genome domain organization are complementary and interdependent phenomena affecting the epigenome leading to a regulated expression of a multifamily gene domain like the globin genes.

Epigenome topology and nuclear dynamics seem to be relevant aspects on epigenetic regulation. This is supported by data showing that inter- and intrachromosomal physical association through loop formation seems to occur with consequences in gene expression coordination (Chakalova et al., 2005; Ling et al., 2006; Kurukuti et al., 2006; Esteller, 2006). This observation supports the concept of three-dimensional chromosomal territories organization of the epigenome in inter-phase nuclei (Cremer and Cremer, 2001; Chakalova et al., 2005) and that chromosomal domains should be coordinated to attract or dissociate genomic domains to active transcription nuclear factories inside the nucleus. One of the most attractive challenges is the need to integrate all these phenomena in a rational way trying to define hierarchies among each process taking into consideration the design of potential therapeutic strategies for treatment of distinct malignancies.

## Acknowledgments

We thank Miriam Rosenberg for critical reading of the manuscript. We would like to acknowledge the two anonymous reviewers for their helpful comments and suggestions. This work was supported by the Dirección General de Asuntos del Personal Académico-UNAM (IN203200, IX230104 and IN209403), the Consejo Nacional de Ciencia y Tecnología, CONACyT (33863-N and 42653-Q), the Third World Academy of Sciences (TWAS, Grant 01-055 RG/BIO/LA), and Fundación Miguel Alemán, A.C. ME-D-A, MF-M, ER, IADLR-V, ES-R, HR-A and CV-Q are fellowship recipients from CONACyT.

## References

- Abruzzo, L.V., Reitman, M., 1994. Enhancer activity of upstream hypersensitive site 2 of the chicken  $\beta$ -globin cluster is mediated by GATA sites. *J. Biol. Chem.* 269, 32563–32571.
- Aguirre, X., Vizmanos, J.L., Calasanz, M.J., Garcia-Delgado, M., Larrayoz, M.J., Novo, F.J., 2003. Methylation of CpG dinucleotides and/or CCWGG motifs at the promoter of TP53 correlates with decreased gene expression in a subset of acute lymphoblastic leukemia patients. *Oncogene* 22, 1070–1072.
- Andrews, N.C., Erdjument-Bromage, H., Davidson, M.B., Tempst, P., Orkin, S.H., 1993. Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. *Nature* 362, 722–728.
- Andersen, A.A., Panning, B., 2003. Epigenetic gene regulation by noncoding RNAs. *Curr. Opin. Cell Biol.* 15, 281–289.
- Anguita, E., Johnson, C.A., Wood, W.G., Turner, B.M., Higgs, D.R., 2001. Identification of a conserved erythroid specific domain of histone acetylation

- across the  $\alpha$ -globin gene cluster. *Proc. Natl. Acad. Sci. U. S. A.* 98, 12114–12119.
- Antequera, F., 2003. Structure, function and evolution of CpG island promoters. *Cell. Mol. Life Sci.* 60, 1647–1658.
- Ashie, H.L., Monks, J., Wijgerde, M., Fraser, P., Proudfoot, N.J., 1997. Intergenic transcription and transinduction of the human  $\beta$ -globin locus. *Genes Dev.* 11, 2494–2509.
- Baylin, S.B., Ohm, J.E., 2006. Epigenetic gene silencing in cancer— a mechanism for early oncogenic pathway addiction? *Nat. Rev., Cancer* 6, 107–116.
- Bernstein, E., Allis, C.D., 2005. RNA meets chromatin. *Genes Dev.* 19, 1635–1655.
- Boyes, J., Byfield, P., Nakatani, Y., Vasily, O., 1998. Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature* 396, 594–598.
- Bulger, M., Groudine, M., 1999. Looping versus linking: toward a model for long-distance gene activation. *Genes Dev.* 13, 2465–2477.
- Burgess-Buesse, B., Farrell, C., Gaszner, M., Litt, M., Mutskov, V., Recillas-Targa, F., Simpson, M., West, A., Felsenfeld, G., 2002. The insulation of genes from external enhancers and silencing chromatin. *Proc. Natl. Acad. Sci. U. S. A.* 99 (suppl. 4), 16433–16437.
- Cantor, A.B., Orkin, S.H., 2005. Coregulation of GATA factors by the friend of GATA (FOG) family of multiple zinc finger proteins. *Semin. Cell Dev. Biol.* 16, 117–128.
- Chakalova, L., Debrand, E., Mitchell, J.A., Osborne, C.S., Fraser, P., 2005. Replication and transcription: shaping the landscape of the genome. *Nat. Rev., Genet.* 6, 669–677.
- Chervenak, A.P., Basu, P., Shin, M., Redmond, L.C., Sheng, G., Lloyd, J.A., 2006. Identification, characterization and expression pattern of the chicken EKLf gene. *Dev. Dyn.* 235, 1933–1940.
- Cremer, T., Cremer, C., 2001. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev., Genet.* 2, 292–301.
- Dean, A., 2006. On a chromosome far, far away: LCRs and gene expression. *Trends Genet.* 22, 38–45.
- Escamilla-Del-Arenal, M., Recillas-Targa, F., in preparation. Characterization of the chromatin structure of the silencer-enhancer elements of the chicken alpha-globin gene domain.
- Esteller, M., 2006. The necessity of a human epigenome project. *Carcinogenesis* 27, 1121–1125.
- Farache, G., Razin, S.V., Recillas Targa, F., Scherrer, K., 1990. Organization of the 3' boundary of the chicken  $\alpha$ -globin gene domain and characterization of a CR1-specific protein binding site. *Nucleic Acids Res.* 18, 401–409.
- Feinberg, A.P., Ohlsson, R., Henikoff, S., 2006. The epigenetic progenitor origin of human cancer. *Nat. Rev., Genet.* 7, 21–33.
- Flint, J., Tufarelli, C., Peden, J., Clark, K., Daniels, R.J., Hardisson, R., Miller, W., Philipson, S., Tan-Un, K.C., McMarrow, T., Frampton, J., Alter, B.P., Frischauf, A.M., Higgs, D.R., 2001. Comparative genome analysis delimits a chromosomal domain and identifies key regulatory elements in the  $\alpha$ -globin cluster. *Hum. Mol. Genet.* 10, 371–382.
- Foley, K.P., Engel, J.D., 1992. Individual stage selector element mutations lead to reciprocal changes in  $\beta$ - vs  $\epsilon$ -globin gene transcription: genetic confirmation of promoter competition during globin gene switching. *Genes Dev.* 6, 730–744.
- Francastel, C., Magis, W., Groudine, M., 2001. Nuclear relocation of a transactivator subunit precedes target gene activation. *Proc. Natl. Acad. Sci. U. S. A.* 98, 12120–12125.
- Gribnau, G., Diderich, K., Pruzina, S., Calzolari, R., Fraser, P., 2000. Intergenic transcription and development remodeling of chromatin subdomains in the human  $\beta$ -globin locus. *Mol. Cell* 5, 377–386.
- Grimaud, C., Bantignies, F., Pal-Bhadra, M., Chana, P., Bhadra, U., Cavalli, G., 2006. RNAi components are required for nuclear clustering of Polycomb group response elements. *Cell* 124, 957–971.
- Grosveld, F., van Assendelf, G.B., Greaves, D.R., Kollias, G., 1987. Position-independent, high-level expression of the human  $\beta$ -globin gene in transgenic mice. *Cell* 51, 975–985.
- Haussecker, D., Proudfoot, N.J., 2005. Dicer-dependent turnover of intergenic transcripts from the human  $\beta$ -globin gene cluster. *Mol. Cell. Biol.* 25, 9724–9733.
- Herman, J.G., Baylin, S.B., 2003. Gene silencing in cancer in association with promoter hypermethylation. *N. Engl. J. Med.* 349, 2042–2054.
- Higgs, D.R., Garrick, D., Anguita, E., DeGobbi, M., Hughes, J., Muers, M., Vernimmen, D., Lower, K., Lae, M., Argentaro, A., Deville, M.A., Gibbons, R., 2005. Understanding  $\alpha$ -globin gene regulation: aiming to improve the management of thalassemia. *Ann. N.Y. Acad. Sci.* 1054, 92–102.
- Hong, W., Nakazawa, M., Chen, Y.Y., Kori, R., Vakoc, C.R., Rakowski, C., Blobel, G.A., 2005. FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. *EMBO J.* 24, 2367–2378.
- Huang, S., Litt, M., Felsenfeld, G., 2005. Methylation of histone H4 by arginine methyltransferase PRMT1 is essential *in vivo* for many subsequent histone modifications. *Genes Dev.* 19, 1885–1893.
- Huo, X., Zhang, J., 2005. Important roles of reversible acetylation in the function of hematopoietic transcription factors. *J. Cell. Mol. Med.* 9, 103–112.
- Issa, J.-P., 2004. CpG island methylator phenotype in cancer. *Nat. Rev., Cancer* 4, 988–993.
- Johnson, K.D., Christensen, H.M., Zhao, B., Bresnick, E.H., 2001. Distinct mechanisms control RNA polymerase II recruitment to a tissue-specific locus control region and a downstream promoter. *Mol. Cell* 8, 465–471.
- Klenova, E.M., Ohlsson, R., 2005. Poly(ADP-ribosylation) and epigenetics. Is CTCF part of the plot? *Cell Cycle* 4, 96–101.
- Klenova, E.M., Morse III, H.C., Ohlsson, R., Lobanekov, V.V., 2002. The novel *BORIS+CTCF* gene family is uniquely involved in the epigenetics of normal biology and cancer. *Semin. Cancer Biol.* 12, 399–414.
- Klochkov, D., Rincón-Arango, H., Ioudinkova, E.S., Valadez-Graham, V., Gavrillov, A., Recillas-Targa, F., Razin, S.V., 2006. A CTCF-dependent silencer located in the differentially methylated area may regulate expression of a housekeeping gene overlapping a tissue-specific gene domain. *Mol. Cell. Biol.* 26, 1589–1597.
- Kosak, S.T., Groudine, M., 2004. Form follows function: the genomic organization of cellular differentiation. *Genes Dev.* 18, 1371–1384.
- Kurukuti, S., Tiwari, V.K., Tavosidama, G., Pugacheva, E., Murrell, A., Zhao, Z., Lobanekov, V., Reik, W., Ohlsson, R., 2006. CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to *Igf2*. *Proc. Natl. Acad. Sci. U. S. A.* 103, 10684–10689.
- Liew, C.W., Rand, K.D., Simpson, R.J., Ung, W.W., Mansfield, R.E., Crossley, M., Proetorius-Ibba, M., Nerlov, C., Poulsen, F.M., Mackay, J.P., 2006. Molecular analysis of the interaction between the hematopoietic master transcription factors GATA-1 and PU.1. *J. Biol. Chem.* 281, 28296–28306.
- Ling, J.Q., Li, T., Hu, J.F., Vu, T.H., Chen, H.L., Qui, X.W., Cherry, A.M., Hoffman, A.R., 2006. CTCF mediates interchromosomal colocalization between *Igf2/H19* and *Wsb1/Nf1*. *Science* 312, 269–272.
- Litt, M.D., Simpson, M., Recillas-Targa, F., Prioleau, M.N., Felsenfeld, G., 2001a. Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *EMBO J.* 20, 2224–2235.
- Litt, M.D., Simpson, M., Gaszner, M., Allis, C.D., Felsenfeld, G., 2001b. Correlation between histone lysine methylation and developmental changes at the chicken  $\beta$ -globin locus. *Science* 293, 2453–2455.
- Lund, A.H., van Lohuizen, M., 2004. Epigenetics and cancer. *Genes Dev.* 18, 2315–2335.
- Mason, M.M., Lee, E., Westphal, H., Reitman, M., 1995. Expression of the chicken  $\beta$ -globin gene cluster in mice: correct developmental expression and distribution control. *Mol. Cell. Biol.* 15, 407–414.
- Mutskov, V.J., Farrell, C.M., Wade, P.A., Wolffe, A.P., Felsenfeld, G., 2002. The barrier function of an insulator couples high histone acetylation levels with specific protection of promoter DNA from methylation. *Genes Dev.* 16, 1540–1554.
- Ohlsson, R., Reinkawitz, R., Lobanekov, V., 2001. CTCF is uniquely versatile transcription regulator linker to epigenetics and diseases. *Trends Genet.* 17, 520–527.
- Ohtani-Fujita, N., Drvja, T.P., Rapaport, J.M., Fujita, T., Matsumura, S., Ozasa, K., Watanabe, Y., Hayashi, K., Maeda, K., et al., 1997. Hypermethylation in the retinoblastoma gene is associated with unilateral, sporadic retinoblastoma. *Cancer Genet. Cytogenet.* 98, 43–49.
- Pikaart, M.J., Recillas-Targa, F., Felsenfeld, G., 1998. Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev.* 12, 2852–2862.
- Ragoczy, T., Bender, M.A., Telling, A., Byron, R., Groudine, M., 2006. The locus control region is required for association of the murine  $\beta$ -globin locus

- with engaged transcription factories during erythroid maturation. *Genes Dev.* 20, 1447–1456.
- Raich, N., Enver, T., Nakamoto, B., Josephson, B., Papayannopoulou, T., Stamatoyannopoulos, G., 1990. Autonomous developmental control of human embryonic globin gene switching and transgenic mice. *Science* 250, 1147–1149.
- Razin, S.V., Farrell, C.M., Recillas-Targa, F., 2003. Genomic domains and regulatory elements operating at the domain level. *Int. Rev. Cyt.* 226, 63–125.
- Razin, S.V., Rynditch, A., Borunova, V., Ioudinkova, E., Smalko, V., Scherrer, K., 2004. The 33 kb transcript of the chicken  $\alpha$ -globin gene domain is part of the nuclear matrix. *J. Cell. Biochem.* 92, 445–457.
- Recillas-Targa, F., 2000. The chicken  $\alpha$ - and  $\beta$ -globin gene domains and their chromatin organization. *Cell. Mol. Biol. Lett.* 5, 451–467.
- Recillas-Targa, F., Razin, S.V., 2001. Chromatin domains and regulation of gene expression: familiar and enigmatic clusters of chicken globin genes. *Crit. Rev. Eukaryot. Gene Expr.* 11, 227–242.
- Recillas Targa, F., De Moura Gallo, C.V., Huesca, M., Scherrer, K., Marcaud, L., 1993. Silencer and enhancer elements located at the 3'-side of the chicken and duck  $\alpha$ -globin-encoding gene domains. *Gene* 129, 229–237.
- Recillas-Targa, F., Pikaart, M.J., Burgess-Beusse, B., Bell, A.C., Litt, M.D., West, A.G., Gaszner, M., Felsenfeld, G., 2002. Position-effect protection and enhancer blocking by the chicken  $\beta$ -globin insulator are separable activities. *Proc. Natl. Acad. Sci. U. S. A.* 99, 6883–6888.
- Recillas-Targa, F., Valadez-Graham, V., Farrell, C.M., 2004. Prospects and implications of using chromatin insulators in gene therapy and transgenesis. *BioEssays* 26, 796–807.
- Recillas-Targa, F., De La Rosa-Velázquez, I.A., Soto-Reyes, E., Benítez-Bribiesca, L., 2006. Epigenetic boundaries of tumour suppressor gene promoters: the CTCF connection. *J. Cell. Mol. Med.* 10, 552–566.
- Reik, W., Murrell, A., Lewis, A., Mitsuya, K., Umlauf, D., Dean, W., Higgins, M., Fiel, R., 2004. Chromosome loops, insulators and histone methylation: new insights into regulation of imprinting in clusters. *Cold Spring Harbor Symp. Quant. Biol.* LXIX, 29–37.
- Reitman, M., Felsenfeld, G., 1988. Mutational analysis of the chicken  $\beta$ -globin enhancer reveals two positive-acting domains. *Proc. Natl. Acad. Sci. U. S. A.* 85, 6267–6271.
- Reitman, M., Lee, E., Westphal, H., Felsenfeld, G., 1990. Site-independent expression of the chicken  $\beta^A$ -globin gene in transgenic mice. *Nature* 348, 749–752.
- Rincón-Arano, H., Valadez-Graham, V., Guerrero, G., Escamilla-Del-Arenal, M., Recillas-Targa, F., 2005. YY1 and GATA-1 interaction modulate the chicken 3'-side  $\alpha$ -globin enhancer activity. *J. Mol. Biol.* 349, 961–975.
- Rincón-Arano, H., Recillas-Targa, F. in preparation. Protection against telomeric position effects by the chicken 5' cHS4 beta-globin insulator.
- Saitoh, N., Bell, A.C., Recillas-Targa, F., West, A.G., Simpson, M., Pikaart, M., Felsenfeld, G., 2000. Structural and functional conservation at the boundaries of the chicken  $\beta$ -globin domain. *EMBO J.* 19, 2315–2322.
- Sanchez-Elsner, T., Gou, D., Kremmer, E., Sauer, F., 2006. Noncoding RNAs of trithorax response elements recruit *Drosophila* Ash1 to ultrathorax. *Science* 311, 1118–1123.
- Scherrer, K., 2003. Historical review: the discovery of 'giant' RNA and RNA processing: 40 years of enigma. *Trends Biochem. Sci.* 28, 566–571.
- Shyu, Y.-C., Lee, T.-L., Ting, C.-Y., Wen, S.-C., Hsieh, L.-J., Li, Y.-C., Hwang, J.-L., Lin, C.-C., Shen, C.-K.J., 2005. Sumoylation of p45NF-E2: nuclear positioning and transcriptional activation of the mammalian  $\beta$ -like globin gene locus. *Mol. Cell. Biol.* 25, 10365–10378.
- Sidhu, S., Martin, R., Gicquel, C., Melki, J., Clark, S.J., Campbell, P., Magarey, C.J., Schulte, K.M., Roher, H.D., Delbridge, L., Robinson, B.G., 2005. Mutation and methylation analysis of TP53 in adrenal carcinogenesis. *Eur. J. Surg. Oncol.* 31, 549–554.
- Singal, R., Wang, S.Z., Sargent, T., Zhu, S.Z., Ginder, G.D., 2002a. Methylation of promoter proximal-transcribed sequences of an embryonic globin gene inhibits transcription in primary erythroid cells and promotes formation of a cell-type-specific methyl cytosine binding complex. *J. Biol. Chem.* 277, 1897–1905.
- Singal, R., vanWert, J.M., Ferdinand Jr., L., 2002b. Methylation of  $\alpha$ -type embryonic globin gene  $\alpha^{\pi}$  represses transcription in primary erythroid cells. *Blood* 100, 4217–4222.
- Stamatoyannopoulos, G., 2005. Control of globin gene expression during development and erythroid differentiation. *Exp. Hematol.* 33, 259–271.
- Stirzaker, C., Millar, D.S., Paul, C.L., Warnecke, P.M., Harrison, J., Vincent, P.C., Frommer, M., Clark, S.J., 1997. Extensive DNA methylation spanning the Rb promoter in retinoblastoma tumors. *Cancer Res.* 57, 2229–2237.
- Stopka, T., Amanatullah, D.F., Papetti, M., Skoultchi, A.I., 2005. PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure. *EMBO J.* 24, 3712–3723.
- Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J., Blobel, G.A., 2005. Proximity among distant regulatory elements at the  $\beta$ -globin locus requires GATA-1 and FOG-1. *Mol. Cell* 17, 453–462.
- Valadez-Graham, V., Razin, S.V., Recillas-Targa, F., 2004. CTCF-dependent enhancer blockers at the upstream region of the chicken  $\alpha$ -globin domain. *Nucleic Acids Res.* 32, 1354–1362.
- Wasserman, W.W., Sanderlin, A., 2004. Applied bioinformatics for the identification of regulatory elements. *Nat. Rev., Genet.* 5, 276–287.
- West, A.G., Fraser, P., 2005. Remote control of gene transcription. *Hum. Mol. Genet.* 14, R101–R111.
- West, A.G., Huang, S., Gaszner, M., Litt, M.D., Felsenfeld, G., 2004. Recruitment of histone modifications by USF proteins at a vertebrate barrier element. *Mol. Cell* 16, 453–463.
- Yi, Z., Cohen-Barak, O., Hagiwara, N., Kingsley, P.D., Fuchs, D.A., Erickson, D.T., Epner, E.M., Palis, J., Brilliant, M.H., 2006. Sox6 directly silences epsilon globin expression in definitive erythropoiesis. *PLoS Genet.* 2, e14.
- Yusufzai, T.M., Felsenfeld, G., 2004. The 5'-HS4 chicken  $\beta$ -globin insulator in a CTCF-dependent nuclear matrix-associated element. *Proc. Natl. Acad. Sci. U. S. A.* 101, 8620–8624.
- Yusufzai, T.M., Tagami, H., Nakatani, Y., Felsenfeld, G., 2004. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol. Cell* 13, 291–298.



---

Flores Herrera O, Riveros Rosas H, Sosa Peinado A, Vázquez Contreras E (eds). **Mensaje Bioquímico, Vol XXVIII**. Depto Bioquímica, Fac Medicina, Universidad Nacional Autónoma de México. Cd Universitaria, México, DF, MÉXICO. (2004).

(<http://bq.unam.mx/mensajebioquimico>)

(ISSN-0188-137X)

---

## PARTICIPACIÓN DE LA ESTRUCTURA DE LA CROMATINA EN LA REGULACIÓN DE LA EXPRESIÓN GÉNICA

Félix Recillas Targa y Martín Escamilla Del Arenal

Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Apartado Postal 70-242, México D.F. 04510.

[frecilla@ifc.unam.mx](mailto:frecilla@ifc.unam.mx)

### ROLE OF CHROMATIN STRUCTURE INTO GENE EXPRESSION

#### Abstract

*Regulation of gene expression in eukaryotes represents a central process in cell differentiation, proliferation and death. Gene regulation includes the binding of sites-specific transcription factors to the DNA and the recruitment of co-factors. In his natural milieu, all this processes occurs in the chromatin context in which histone modifications are one of the primordial targets. More recently the nuclear organization and dynamics with a broad and direct biological implications on gene expression has been the topic of intense research. Thus regulation of gene expression can be visualized in general terms at three different levels all of them interdependently related. The first level involves the genetic information codified by the DNA molecule. The second involves the natural context, in which the DNA is functional. It is based on the genome organization into chromatin and all the associated chromatin remodeling activities also known as epigenetic regulation. And finally, resent set of experimental data supports a more active and dynamic role of the cell nuclei on the regulation of gene expression. In the present chapter, we review the current knowledge on regulatory elements, the general transcription factors involved on transcription initiation, and we integrate all such information taking into consideration the chromatin structure and the contribution of the cell nucleus to the differential gene expression in eukaryotic cells.*

**Keywords:** Transcription, gene regulation, chromatin, nucleus, histone modifications.

#### Introducción

La complejidad adquirida por el genoma eucariota a través de la evolución ha requerido, a la par, el desarrollo de múltiples niveles de regulación, para poder llevar a cabo la expresión

genética de manera ordenada y permitir el desarrollo de un organismo. La regulación transcripcional se ha dividido en 3 niveles (1). El primero lo encontramos a nivel de las secuencias nucleotídicas que se codifican en la molécula de ADN. El segundo nivel involucra la estructura de la cromatina, donde la regulación de la expresión de un gen no solo depende de la información codificada en la secuencia del ADN, sino también de su organización en cromatina y la regulación epigenética asociada a ésta. Considerando a la regulación epigenética como los patrones heredables de la expresión de un gen, que no pueden ser explicados por cambios en la secuencia del ADN (2). El tercer nivel involucra la organización espacial del genoma al interior del núcleo.

En el presente capítulo revisaremos los principales niveles de regulación de la expresión transcripcional de un gen y el papel que desempeña en ella la estructura de la cromatina; todos estos procesos enmarcados en el entorno del núcleo de la célula. También mencionaremos la participación de las diferentes secuencias de control, cuyas funciones se ejecutan a grandes distancias, entre las cuales se encuentran los silenciadores ("silencers"), los potenciadores ("enhancers"), los elementos de control del locus (LCRs) y la formación de dominios transcripcionalmente activos.

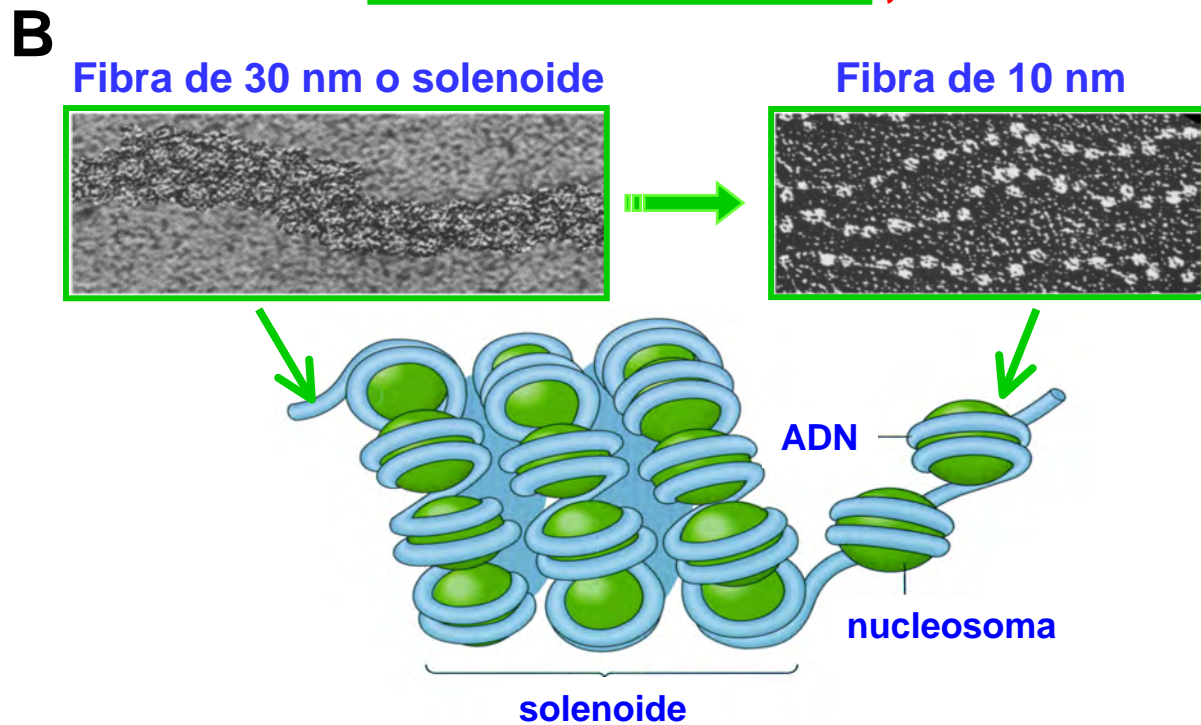
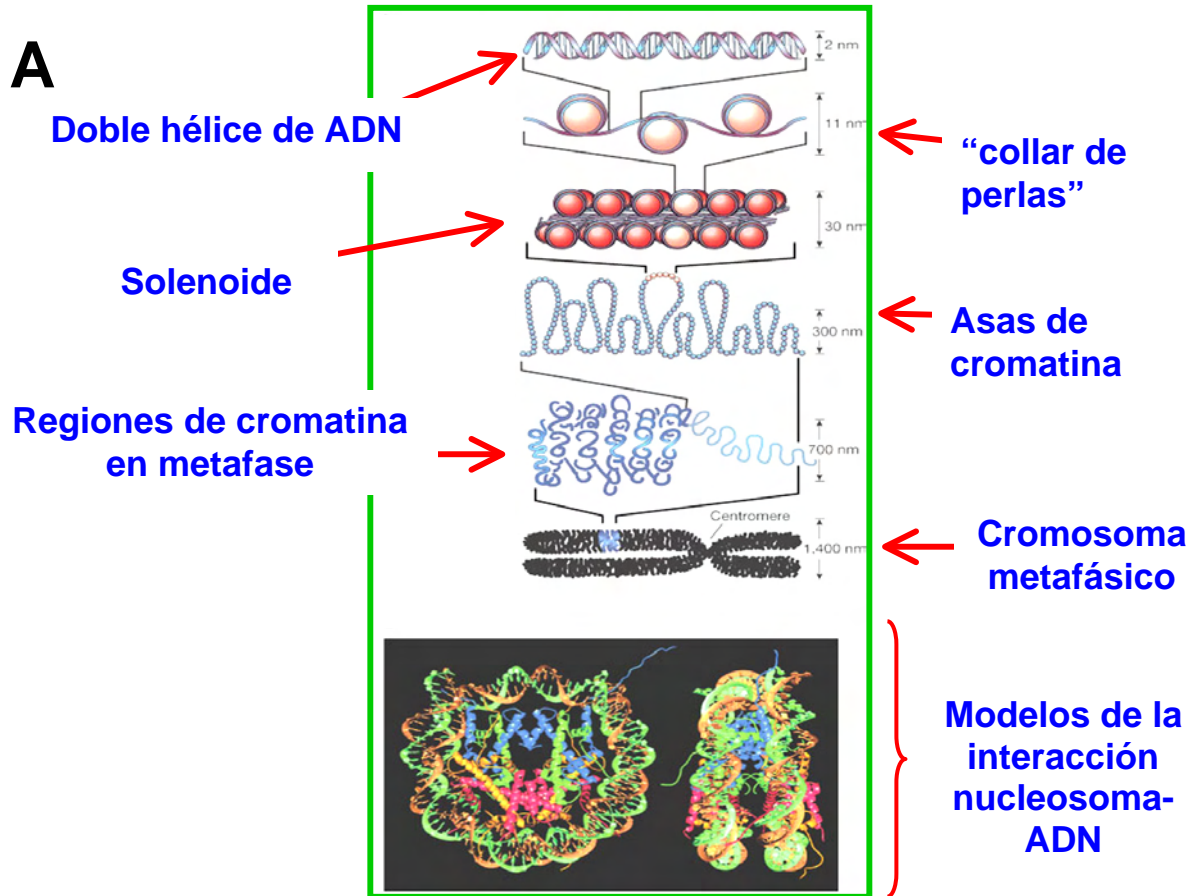
## I- Elementos de Regulación

Las características que definen a una célula están dadas por las proteínas que la constituyen, las cuales son el resultado de patrones específicos de la expresión de sus genes (2). Un gen que codifica para un ARN mensajero está constituido por un segmento determinado de la cadena de ADN, involucrado en la síntesis de una cadena polipeptídica, es decir, de una proteína. Un punto crucial en la determinación de los patrones de expresión, está dado por los factores transcripcionales que reconocen secuencias nucleotídicas en el ADN, principalmente en las regiones promotoras de los genes (2). La interacción de estos factores con sus secuencias reguladoras, desencadena eventos que tendrán como resultado la activación o la represión de un gen ó grupos de genes.

Los organismos superiores tales como la levadura y el humano están constituidos por células de origen eucariota; lo anterior significa que tienen su ADN compartimentalizado en el núcleo. Al interior de este núcleo, el ADN no se encuentra disperso, por el contrario, se estructura a través de interacciones proteína-ADN formando el complejo que conocemos como cromatina la cual permite contener al genoma en el interior del núcleo celular (Fig. 1). La unidad

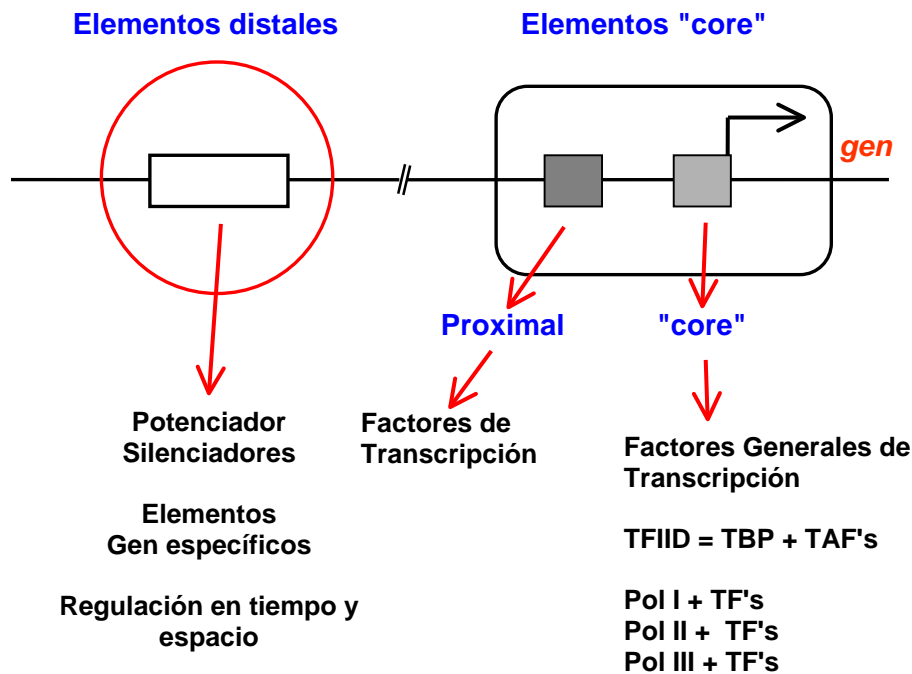
---

**Figura 1.** Distintos niveles de estructuración del genoma en la cromatina. A- El nivel primario de compactación del genoma está dado por el nucleosoma donde se encuentran interaccionando 147 pb de ADN, a lo que se conoce también como la fibra de 10 nm. Con la incorporación de una quinta histona, la histona H1 se pasa al nivel superior de compactación representada por un grupo de seis nucleosomas llamado solenoide o fibra de 30 nm. Posteriormente y en su estructura de solenoide se forman asas de cromatina que abarcan centenas de kilo bases. Poco se sabe sobre los mecanismos de compactación entre el solenoide y el nivel superior de compactación que es el cromosoma metafásico, estructura necesaria para la segregación de los cromosomas durante la mitosis. En la parte inferior se muestra un esquema de la estructura nucleosomal asociada al ADN (2). B- El solenoide constituye el medio ambiente natural donde deben llevarse a cabo todas las eventos de regulación, en particular el remodelaje de su propia estructura. En este esquema se muestran microscopías electrónicas de la fibra de 10 y 30 nm.





fundamental de la cromatina es el nucleosoma, el cual esta constituido por el octámero de histonas (dos de cada una de las proteínas H2A, H2B, H3 y H4) alrededor del cual interaccionan 147 pares de bases (pb) de ADN. Esta estructuración en cromatina, traerá como consecuencia una inaccesibilidad a las proteínas reguladoras que reconocen secuencias blanco en el ADN (2). Existen múltiples niveles de compactación del ADN siendo el cromosoma en metafase el nivel más elevado (Fig. 1A; 2, 3). Cabe mencionar que la gran mayoría de los eventos de remodelaje de la estructura de la cromatina ocurren en la fibra de 30 nm, también conocida como solenoide (Fig. 1B).



**Figura 2.** Esquema de los dos tipos de elementos de regulación que intervienen en el control de la expresión génica. El promotor está compuesto por el elemento "core" que se localiza cercano al sitio de inicio de la transcripción y donde encontramos a la caja TATA, secuencia a la cual se asocian el TBP, el TFIID, la RNAPol-II y el resto del complejo de preiniciación e iniciación de la transcripción. En un intervalo de 10 a 200 pb se encuentra el elemento proximal donde se incorporan factores transcripcionales tanto específicos como generales con la capacidad de formar puentes moleculares tanto con los TAF's del complejo TFIID o con los componentes del Mediador/srb asociado a la CTD de la RNA pol-II. Para completar el escenario, los elementos distales pueden localizarse a gran distancia tanto en el extremo 5' como en el 3' en relación con el gen y tienen la función de aumentar tanto los niveles de expresión de un gen como su especificidad (6).

Evolutivamente la organización del genoma eucariote en la cromatina, permite que éste sea contenido al interior del núcleo. Dado el elevado nivel de compactación del genoma, la

estructura de la cromatina se vuelve un obstáculo para la actividad transcripcional, es decir, para la expresión de los genes. Por lo anterior, es necesario remodelar dicha estructura represora para permitir la expresión regulada de los genes. Por lo tanto, el entender cómo se expresa un gen depende en gran medida de la capacidad de remodelar la estructura de la cromatina tanto al nivel de sus regiones control como de sus secuencias codificantes y no codificantes.

A continuación describiremos a los diferentes elementos de regulación para tratar de entender cómo los complejos de iniciación de la transcripción pueden activar de manera regulada la expresión de un gen.

#### *Regulación al nivel de secuencia.*

El desarrollo, crecimiento y sobrevivencia de los organismos eucariotes, requiere de la regulación y expresión coordinada de cientos de genes. La expresión de éstos es controlada por una gran variedad de mecanismos, donde la cascada de eventos que preceden a la transcripción tiene como blanco final a las secuencias de ADN conocidas como promotores y al el sitio de inicio de la transcripción asociado a éstos (Fig. 2; 4-6).

La regulación al nivel de secuencia nucleotídica es la mejor estudiada. En ella se involucran a los elementos en *cis*, tales como los promotores, los potenciadores y los silenciadores. La función de estos elementos es dependiente de la atracción de las proteínas denominadas como los factores transcripcionales que se unen a las secuencias específicas en el ADN, de las interacciones proteína-proteína entre cofactores y de un conjunto de moléculas cuya finalidad es remodelar la estructura de la cromatina (1, 2).

### **Promotores**

Un promotor se define como la secuencia continua de ADN que es suficiente para dirigir la correcta iniciación de la transcripción de un gen mediada por la RNA polimerasa II (RNA pol-II) (4-9). Estas secuencias suelen encontrarse cercanas al sitio de inicio de la transcripción dentro de un intervalo que va de +20 a -30 en relación con el sitio de inicio de la transcripción (también conocido como el +1) las cuales se denominan elementos "core" del promotor (Fig. 2).

En las secuencias promotoras, existen diferentes motivos consenso de unión al ADN (Fig. 2; 8, 9), entre los más comunes encontramos a la caja TATA, los elementos iniciadores (Inr), los elementos de reconocimiento para el factor TFIIB (BRE) y los elementos promotores río abajo (DPE). Estos motivos de unión al ADN, suelen encontrarse en las secuencias promotoras, ya sea de manera individual o en diferentes combinaciones o acompañadas de otras secuencias que se mencionaran más adelante (5, 8, 9). Distintas combinaciones de estas secuencias constituyen los que se conoce como la "arquitectura" del promotor la cual permite establecer una combinación única y selectiva para el reconocimiento de moléculas reguladoras.

Un segundo grupo de las secuencias de regulación, es conocido como "elementos distales de regulación génica" (3, 6, 10). Éstos participan en la regulación otorgando una expresión en el tiempo y tejido específicos. Entre estas secuencias encontramos a los potenciadores, los silenciadores y los elementos aisladores. Todos estos elementos contienen

motivos de reconocimiento en el ADN, tanto para factores transcripcionales generales como específicos, que participarán de manera precisa y coordinada en la regulación transcripcional.

### **Potenciadores o "Enhancers"**

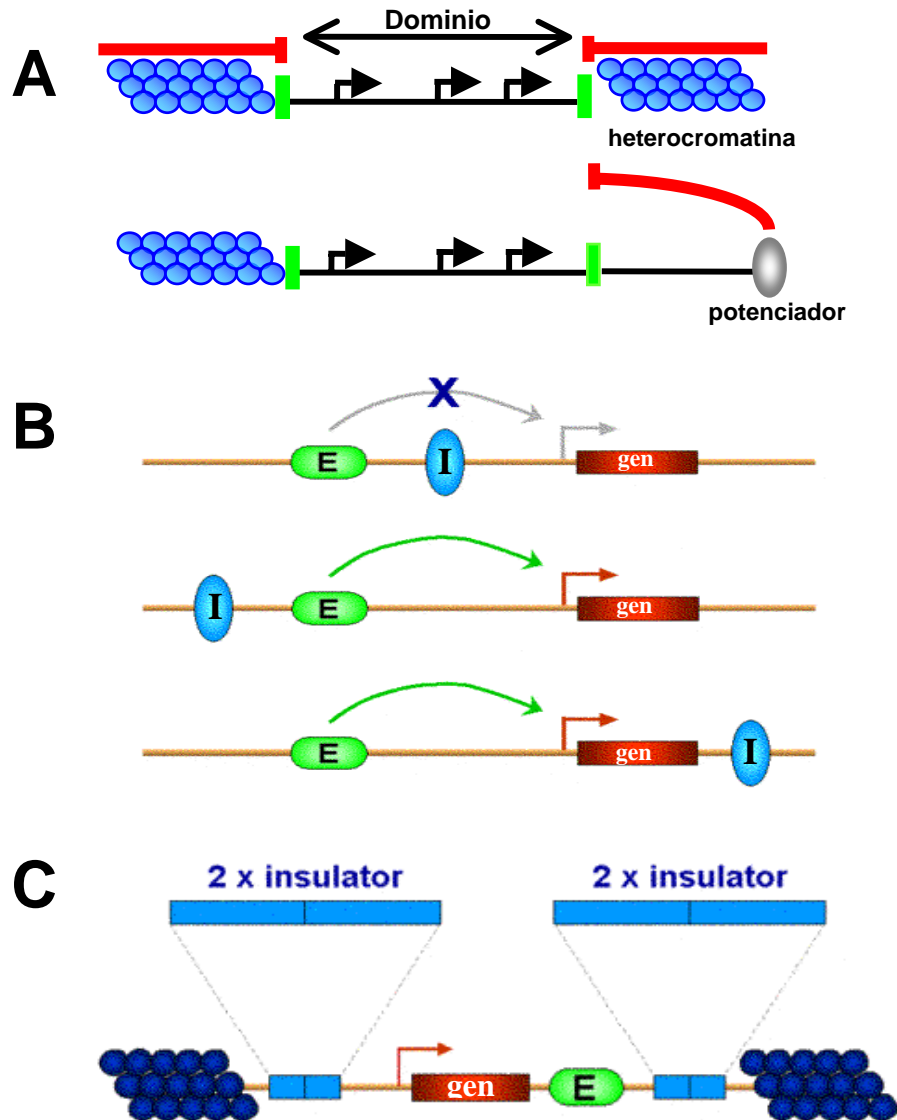
Los potenciadores son elementos de regulación en cuya secuencia interaccionan proteínas que transmitirán a distancia señales de activación específica (8). Los potenciadores fueron originalmente identificados como secuencias en *cis*, que pueden incrementar la transcripción de un gen, independientemente de la orientación y de la distancia a la que se encuentren en relación con el sitio de inicio de la transcripción (8).

Al encontrar elementos de regulación que se podían situar a grandes distancias dentro de un intervalo que puede ir desde 200 pb hasta decenas e incluso centenas de kilo bases (kb) de los genes que están regulando, surgió el interés por comprender sus mecanismos de acción. Dentro de las diferentes propuestas para su funcionamiento en términos del incremento en la tasa transcripcional, el grupo de Kadonaga propuso que estos elementos podrían estar incrementando la probabilidad para que un gen se active en un momento dado de la diferenciación celular (8). Lo anterior significa que el potenciador estaría logrando que un mayor número de células activen transcripcionalmente a un gen en particular, lo que se conoce como el modelo de "encendido y apagado" (8). Un modelo alternativo sugiere que los potenciadores aumentan los niveles transcripcionales de manera homogénea en todas las células (8). Evidencias experimentales apoyan ambos modelos por lo que en la actualidad aún no se descifra el mecanismo *in vivo* de la activación a distancia por parte de un potenciador.

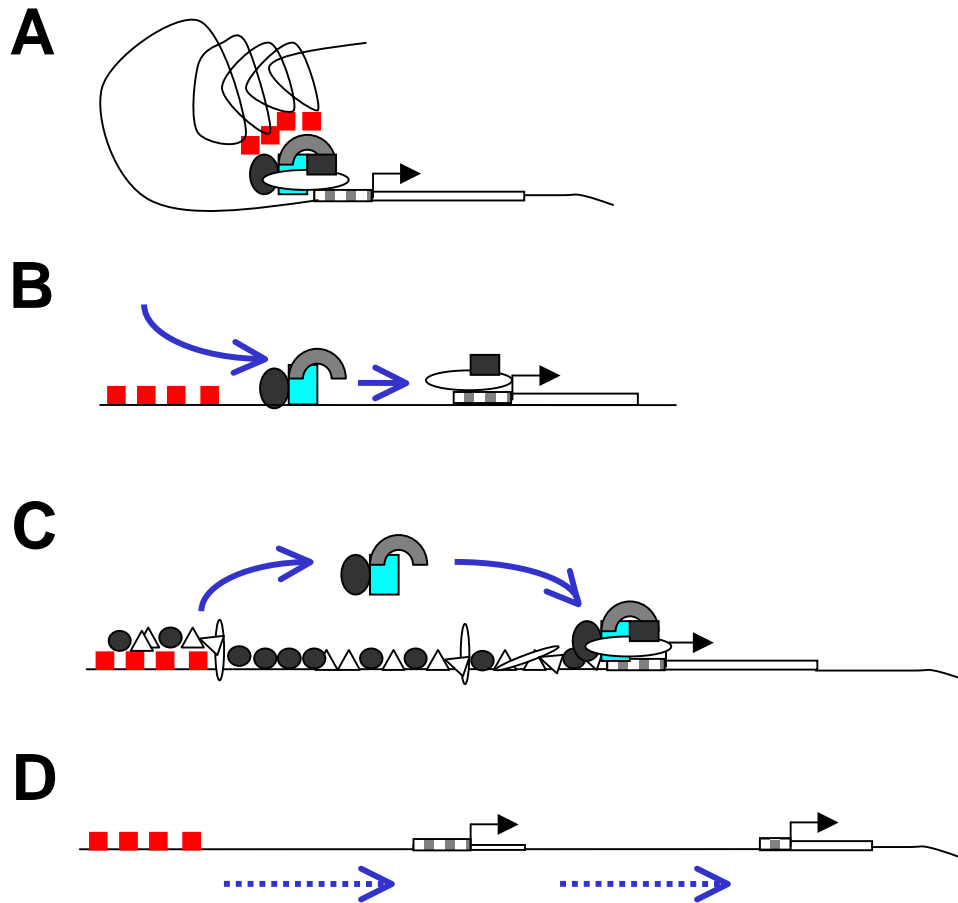
Las evidencias experimentales apoyan la idea de que estos elementos podrían estar interaccionando físicamente con la secuencia promotora del gen, a través de interacciones proteína-proteína, facilitando la formación de un asa. Lo anterior traería como consecuencia la formación de un complejo transcripcional, también conocido como holocomplejo, que facilita los procesos de la iniciación y la elongación en los que participa la RNA pol-II (3, 8). Esta interacción podría generar modificaciones a nivel de la estructura de la cromatina y de una manera regulada contrarrestar el efecto represivo por parte de ésta. Otra alternativa para la transcripción es la de facilitar o atraer diferentes actividades enzimáticas a la región promotora, como por ejemplo, la acetilación de histonas para relajar la estructura de la cromatina y favorecer la formación del complejo de preiniciación de la cromatina (CPT; 2, 3, 8, 9).

Uno de los modelos más novedosos para el funcionamiento de un potenciador, tiene que ver con el contexto tri-dimensional del núcleo. Donde el papel de los potenciadores sería el de relocalizar a un gen de las zonas transcripcionalmente inactivas, a regiones subnucleares donde se fomenta su activación (3, 10). Hay que recordar que los modelos antes descritos no son excluyentes entre ellos y que falta mucho trabajo para lograr comprender de manera detallada sus mecanismos de acción.

**Dominios cromosómicos y sus límites**



**Figura 3.** Organización del genoma eucariota en un dominio. Evidencias tanto bioquímicas como citológicas apoyan el concepto de una organización del genoma en dominios independientes. A- Recientemente se han descubierto un grupo de secuencias llamadas en inglés "insulators" o barrera (cajas verdes en el esquema) cuya función teórica es por un lado favorecer la formación de un dominio, es decir, la de facilitar la apertura de la cromatina al nivel de un dominio compuesto por uno o más genes. También se ha propuesto que sirven para mantener abierta esa región genómica y así evitar que la cromatina, en particular segmentos de heterocromatina, invadan y repriman la expresión de genes contenidos en el dominio. Además evitan la influencia inespecífica de elementos de regulación, como por ejemplo un potenciador, ajenos al dominio (17). B- Como parte de las propiedades funcionales de los "insulators" tenemos la actividad de bloqueo del potenciador, la cual es dependiente de su posición dado que la interferencia en la comunicación entre un enhancer (E) y su promotor ocurre únicamente cuando el "insulator" se encuentra entre ellos (17). C- La segunda propiedad funcional tiene que ver con la capacidad por parte de algunos "insulators" de proteger a un transgen de los efectos de posición causados por los distintos medios ambientes de la cromatina asociados al sitio de integración. Colocando a estas secuencias (2 x "insulators") es posible proteger al transgen del efecto de la cromatina circundante y a su vez mantener niveles de expresión constantes (17).



**Figura 4.** Modelos para el funcionamiento del LCR. A- Modelo de asa. Los factores transcripcionales interaccionan con el LCR, representado por las cajas rojas. El LCR forma un asa a través de múltiples interacciones entre factores transcripcionales con el promotor (cajas con diversas formas). B- Modelo de encarrilamiento ó "tracking" (3). Los factores transcripcionales se unen al LCR, forman un complejo que se encarrila a lo largo del ADN hasta encontrar al promotor que se activará. C- Modelo de relevo ó "linking" (3). Los diferentes factores transcripcionales se unen consecutivamente o forman un relevo molecular para favorecer interacciones a distancia entre el LCR y el promotor. D- Se ha planteado que el LCR podría estar contribuyendo en la apertura de la cromatina, dado que se han identificado transcritos intergénicos (flechas con líneas punteadas) a lo largo del *locus* que además inician en el LCR y son estadio y tejido específicos.

## Silenciadores

Los silenciadores se han definido como elementos a los que se unen factores nucleares que tienen un efecto negativo sobre la transcripción de manera independiente de su posición y orientación en relación con el sitio de inicio de la transcripción (11-13). Estos elementos poseen

una organización modular al igual que los potenciadores, pero con una función negativa sobre la expresión génica.

Además de los silenciadores clásicos se han definido secuencias denominadas elementos reguladores negativos, cuya actividad si es dependiente de su posición. Se ha propuesto que los mecanismos de acción de ambos elementos son diferentes, pues se cree que los silenciadores clásicos interfieren con el ensamblaje de la maquinaria transcripcional (13), mientras que los elementos reguladores negativos reclutan proteínas que interfieren con la fijación de otros factores transcripcionales a su sitio de unión al ADN. Esto ocurre en los elementos de regulación del extremo 5' afectando procesos de la regulación post-transcripcional como la edición de los intrones, la poliadenilación, los sitios de inicio de la traducción, la elongación y la estructura de la cromatina (14).

### **Delimitadores**

En años recientes se ha propuesto que los genes se encuentran organizados en regiones de la cromatina que se encuentran "abiertos" en un tiempo y espacio determinado del desarrollo. A estas regiones se les conoce como dominios cromosómicos transcripcionalmente activos (15). Se postula que éstos tienen límites bien definidos e incluyen a los elementos de regulación necesarios para llevar a cabo la regulación transcripcional de sus genes (Fig. 3A; 16).

Los límites están constituidos por lo que conocemos como los elementos frontera o delimitadores llamados en inglés: "insulators" o "boundary elements". Estos elementos frontera incluyen a las secuencias nucleotídicas que van de 0.5 a 3 kb, a las cuales se unen proteínas nucleares (17). Teóricamente los elementos frontera o "insulator" han sido propuestos como secuencias responsables de la formación y mantenimiento de un dominio o región en el cual un gen o grupo de genes se encuentran activos transcripcionalmente (15, 18). Funcionalmente los elementos frontera se definen por dos propiedades: 1) La capacidad de bloquear la acción positiva de un potenciador sobre su promotor siempre y cuando las secuencias tipo "insulator" se localicen entre el potenciador y el promotor (Fig. 3B) y 2) los elementos barrera o frontera tienen la propiedad de proteger a un transgen, es decir un gen exógeno, de los efectos de posición causados por los distintos ambientes de la cromatina en el entorno del sitio de integración al genoma (Fig. 3C).

En resumen los elementos barrera o "insulator" parecen contribuir a la formación de los dominios transcripcionalmente activos, por lo que se consideran como uno de los primeros niveles de regulación de la expresión del material genético. Lo anterior se basa en el hecho de que una vez definido el dominio, se llevarán a cabo al interior de éste, los eventos finos de regulación. Cabe señalar que no todos los genes o grupos de genes se encuentran de manera natural enmarcados por secuencias tipo "insulator" y que posiblemente será la organización del genoma la que dicte su presencia en regiones donde el ambiente cromatínico pueda influir en la regulación génica (18).

### **Los elementos de control del locus**

Al igual que los potenciadores, las regiones control del locus (LCR del inglés: "Locus Control Region"), contienen múltiples sitios de unión para los activadores transcripcionales

secuencia específica, que en su conjunto forman un sitio abierto en la cromatina que puede ser identificado por la accesibilidad del ADN al corte por la endonucleasa DNasa I. La DNasa I tiene la capacidad de cortar la doble hebra del ADN reconociendo al surco menor de la doble hélice del ADN. Los LCR's están formados generalmente por grupos de sitios de hipersensibilidad en zonas del genoma que abarcan de 200 a 600 pb donde la estructura de la cromatina es más laxa y por lo tanto la unión de los factores transcripcionales se ve favorecida (3, 16). Los LCR's están constituidos por dos funciones principales: una actividad tipo potenciador y otra que fomenta la apertura de la estructura de la cromatina al nivel de un dominio (3, 8, 16, 18,19). Resulta importante señalar que la función de los LCR's es tejida específica. Un aspecto que apoya la relevancia de este tipo de secuencias, es el número creciente de dominios que presentan un LCR entre sus elementos de regulación (16, 19).

Un aspecto práctico de los LCR's es que al colocarlos en presencia de un transgen, la expresión de éste en un contexto integrado al genoma, se ve favorecida y es dependiente del número de copias del transgen e independiente del sitio de integración de una manera tejida específica (3, 8, 16, 19). Al igual que los potenciadores, se ha propuesto que estos podrían estar incrementando la probabilidad de la expresión de un gen o grupo de genes en un mayor número de células (8, 19).

Actualmente y derivado de varios años de trabajo de un número considerable de grupos de investigación, se ha propuesto cuatro modelos que tratan de explicar la función de un LCR:

- 1- El modelo de formación de un asa o "Looping" (Fig. 4A). Este modelo favorece la idea de las interacciones físicas entre el LCR y los promotores individuales de los genes de manera regulada a lo largo del desarrollo. En este modelo el LCR actuaría como un holocomplejo para estimular la transcripción de genes individuales dentro del *locus* y a lo largo del desarrollo acercando al promotor del gen que se va a activar formando un asa que contribuye a aumentar la concentración de la maquinaria transcripcional de manera específica (3).
- 2- Modelo de encarrilamiento o "Tracking" (Fig. 4B). En este modelo se propone que diversos factores transcripcionales específicos y generales, así como cofactores, reconocen las secuencias presentes en el LCR formando un complejo de activación. Dicho complejo migraría o se "encarrilaría" a lo largo de la cromatina hasta topar con la maquinaria basal de la transcripción, que estaría situada en el promotor en turno, cercano al sitio de inicio de la transcripción (3).
- 3- El modelo de relevo o "Linking" (Fig. 4C). Este modelo sugiere una cadena o relevo de interacciones moleculares, que fomentan la formación de un "puente" molecular acercando el LCR al promotor. Esto se logra a través de interacciones entre los factores expresados en un estado de diferenciación específico, facilitando la formación de la estructura de la cromatina de manera específica y regulada para cada gen del dominio (3).
- 4- Los transcritos intergénicos (Fig. 4D). En el dominio  $\beta$ -globina humano, se han identificado transcritos intergénicos, es decir, en regiones no codificantes (que tomaron como molde el ADN situado entre los genes y los elementos de regulación distales, en particular el LCR  $\beta$ -globina). Aunque se desconoce su función, se ha demostrado una correlación entre transcritos no codificantes y la formación de la cromatina abierta. Se cree que estos podrían ser los responsables de contrarrestar el efecto represor de la cromatina compacta (3). El dominio  $\beta$ -globina de pollo, fue uno

de los primeros donde se describió la existencia de transcritos intergénicos en regiones no codificantes (3).

Como podemos observar la expresión de un gen constituye un evento biológico muy preciso y sofisticado donde ha sido necesario crear sistemas cada vez más complejos de control para lograr en tiempo y en espacio la transcripción de un gen. Pero los elementos de la regulación a distancia no son los únicos responsables, el ensamblaje de los complejos de iniciación de la transcripción en las zonas cercanas al inicio de ésta resulta fundamental y añade otro nivel de sofisticación como veremos a continuación.

## **II- La RNA polimerasa y los factores generales de la transcripción**

Una adecuada y eficiente transcripción, requiere de la RNA polimerasa y un conjunto de factores peptídicos auxiliares, mejor conocidos como factores “generales” o “basales” de la transcripción (TF's), entre los cuales se encuentran los complejos: TFIIA, TFIID, TFIIE, TFIIIF, TFIIH y TFIIIB, éste último formado por un solo péptido (5, 7, 20). Estos factores pueden unirse a las secuencias promotoras antes mencionadas y formar lo que conocemos como el complejo de preiniciación de la transcripción (Fig. 5).

Uno de los principales paradigmas en la regulación transcripcional, se encuentra en entender como esta regulación puede llevar a un gen a ser expresado de manera específica en tejidos, niveles de expresión y etapa del desarrollo. El descubrimiento de los diferentes secuencias promotoras y tres formas diferentes de RNA polimerasa (I, II y III), localizadas en distintos compartimentos de la célula, ayudo a comprender que la especificidad necesaria para el control fino de la regulación transcripcional, está dada por combinaciones particulares en la composición de la maquinaria basal de la transcripción y los factores transcripcionales tanto específicos como generales (4-7, 20). Cabe mencionar que recientemente Orphanides y Reinberg han propuesto una teoría unificadora para el control de la expresión génica donde se propone que los procesos que llevan a la activación de un gen son secuenciales e inter-dependientes y que ocurren a múltiples niveles (21).

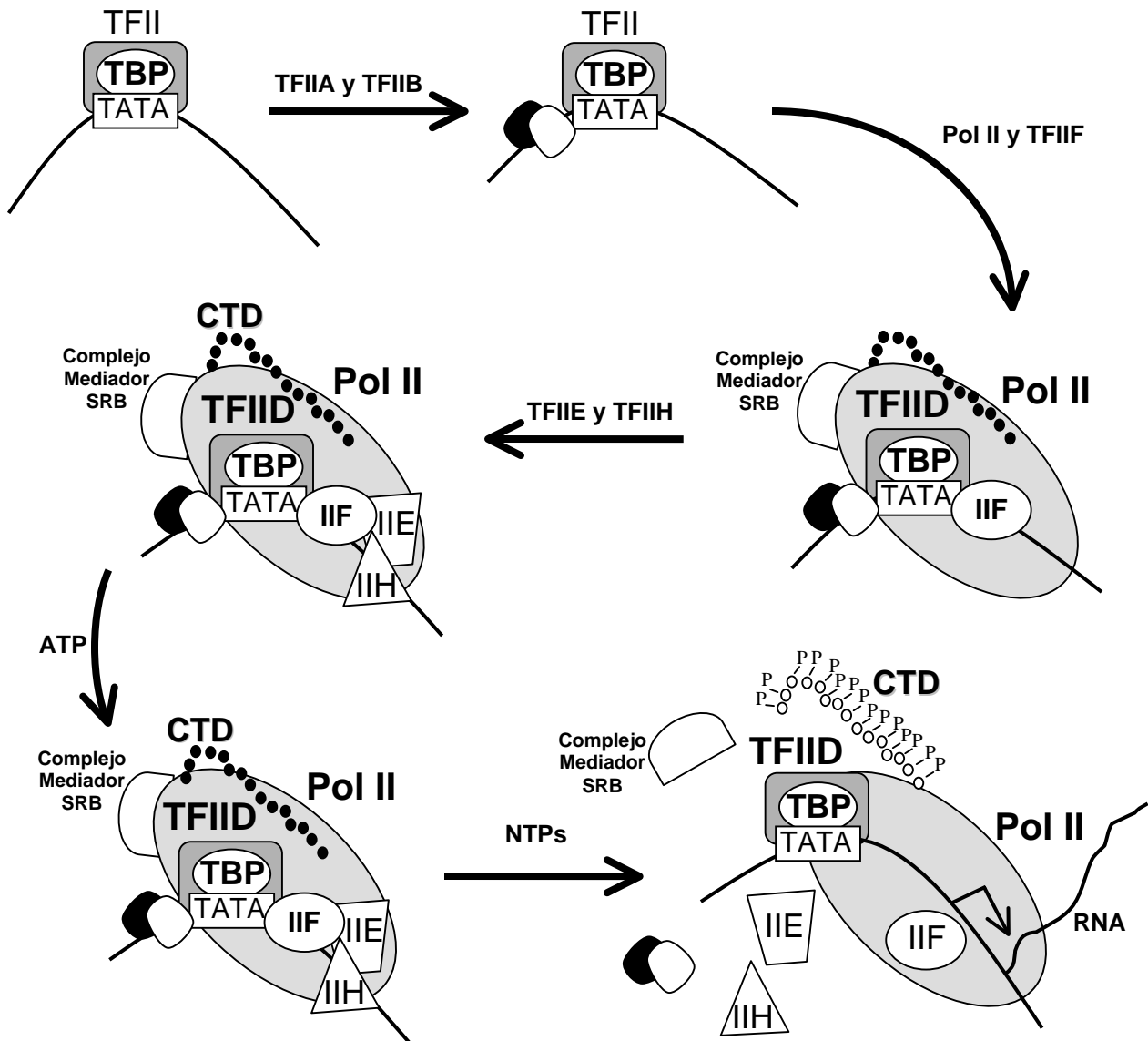
La idea de tres tipos de RNA polimerasas, con funciones distintas, generó una primera clasificación en grupos de genes, que dependen del tipo de RNA polimerasa por el que son transcritos (6, 7):

Clase I, genes que codifican para las subunidades largas de los RNA ribosomales (RNAr). Estos genes son transcritos por la RNA polimerasa I, la cual se encuentra localizada en el nucleolo.

Clase II, genes que codifican para algunos RNA nucleares pequeños, y principalmente para los RNA mensajeros (RNAm) que codifican para proteínas estructurales. Estos son transcritos por la RNA polimerasa II.

Clase III, genes que codifican para RNA estructurales pequeños y componentes de la maquinaria de procesamiento o "splicing" de los RNA's pre-mensajeros. Estos genes son transcritos por la RNA polimerasa III.





**Figura 5.** Modelo secuencial del ensamblaje del complejo de inicio de la transcripción por parte de la RNA polimerasa II. A través de la unión de TBP a la caja TATA se forma el complejo TFIID con la incorporación de un grupo de proteínas conocidas como TAF's, algunas de las cuales pueden ser incluso tejido específicas. Una vez asociado el complejo TFIID se incorporan TFIIA y TFIIB cuya función es la de determinar la distancia y orientación en la cual la RNA pol-II se posicionará en el "core" del promotor. Posteriormente la RNA pol-II, escoltada por el complejo TFIIF, es atraída al promotor. Cabe señalar que el complejo Mediador/srb viene acoplado a la RNA pol-II en su región carboxilo terminal (CTD). Resulta importante comentar que las interacciones moleculares generadas por las TAF's (en el complejo TFIID) y el Mediador/srb representan un "panel de control" donde pueden ocurrir múltiples interacciones moleculares, con un elevado grado de especificidad, entre factores transcripcionales generales, específicos e incluso con cofactores (coactivadores o corepresores). El siguiente paso es la incorporación de los complejos TFIIE y TFIIH, este último de suma importancia, dado que, en presencia de ATP y nucleótidos, TFIIH fosforila a la región CTD de la RNA pol-II permitiendo el proceso de elongación y por lo tanto la síntesis del RNA pre-mensajero (6, 7, 20, 21).

Como se ha mencionado, las polimerasas se asocian con distintos factores generales de transcripción los cuales forman el complejo de preiniciación de la transcripción que está compuesto por más de 100 proteínas (4-7, 20). En particular, para el grupo de genes transcritos por la RNA polimerasa II, al menos se requiere de 6 factores generales de transcripción, que en conjunto formarán lo que conocemos como el CPT (Fig. 5). La asociación estable de la RNA pol-II con la secuencia promotora, requiere de la unión previa de TFIID, en particular TBP y sus factores asociados conocidos como TAF's (del inglés: "TBP-associated factors"), TFIIB y TFIIF. Sin embargo para que la transcripción se pueda llevar a cabo se necesita de los factores transcripcionales TFIIE y TFIIH (20).

El inicio de la transcripción requiere del previo reconocimiento de la caja TATA, por la proteína TBP (del inglés: "TATA-binding protein") y su unión genera una distorsión en la estructura de ADN (Fig. 5). Este complejo formado por la TBP y el ADN ahora distorsionado, será el blanco que reconocerá el TFIIB. Se ha propuesto que TFIIB no reconoce una secuencia en específico, sino la curvatura generada por la interacción de la TBP con la caja TATA. Además se propone que la unión del TFIIB podría estar indicando la direccionalidad del promotor (7, 20).

La asociación de la RNA polimerasa II al complejo ADN-TBP-TFIIB, requiere de la interacción del TFIIF con la polimerasa, su función será la de asistir la unión al promotor. Lo anterior es el resultado de la capacidad del TFIIF para interactuar con el TFIIB, la RNA pol-II y el ADN. Cabe señalar que el TFIIF interactúa río abajo del elemento TATA (Fig. 5) (7, 20).

Una vez que se ha establecido la unión de la RNA pol-II a la secuencia promotora, se necesita de la unión de los factores TFIIE y TFIIH para lograr el proceso conocido como "alarge" del promotor (Fig. 5). Este evento permitirá liberar a la RNA polimerasa II de la secuencia promotora y facilitar el inicio de la transcripción, en un proceso mejor conocido como elongación (Fig. 5). El complejo multiprotéico TFIIH resulta ser imprescindible, dado que es el único factor general de transcripción que presenta una actividad enzimática clave para el inicio de la transcripción, consiste en la fosforilación de la región carboxilo terminal de la RNA polimerasa II, conocida como CTD (7, 20, 22). El complejo TFIIH además de incluir a la cinasa responsable de fosforilar a la región CTD de la RNA pol-II, tiene la capacidad de reparar el ADN mediante el proceso conocido como reparación por escisión de nucleótidos (22). Finalmente debemos recordar que la transcripción se realiza en el contexto natural en la presencia de los nucleosomas (Fig. 1) por lo que se ha propuesto que el proceso de elongación debe estar acoplado y coordinado por las actividades de remodelaje de la estructura de la cromatina (23).

En relación con lo anterior, el segundo nivel de regulación de la actividad transcripcional está dado por la modulación de la estructura de la cromatina. Para que la maquinaria general de la transcripción pueda unirse a la región promotora, primero se necesita contrarrestar el efecto represor de la estructura de la cromatina poniendo al descubierto las secuencias blanco para la unión de múltiples y variados factores transcripcionales. A continuación presentaremos una descripción de los eventos ligados al remodelaje de la cromatina.

### **III- La expresión génica en el contexto de la estructura de la cromatina**

Como se mencionó, el genoma eucariota se encuentra organizado en cromatina donde la unidad mínima de estructuración y por lo tanto de compactación está dada por el nucleosoma (2). En la actualidad la regulación de la expresión génica no puede ser entendida sin tomar en consideración a la estructura de la cromatina.

### **La regulación epigenética**

La regulación epigenética gira en torno a la estructuración del genoma en la cromatina y su efecto regulador sobre la expresión de los genes. Se consideran procesos de regulación epigenéticos, todos aquellos que influyen tanto de manera normal como anormal en la expresión heredable de los genes sin que ocurran cambios en la secuencia del ADN. Esto es, la información genética que viene codificada en la molécula de ADN no se ve afectada y se transmite su información sin problema alguno (por ejemplo el marco abierto de lectura que conforma la secuencia de un gen que dará origen a una proteína).

Los elementos que participan en la regulación epigenética responsables de permitir que la expresión génica tenga un vínculo muy estrecho con la estructura de la cromatina son básicamente cuatro: 1) las modificaciones post-traduccionales de las histonas, 2) la metilación del ADN, 3) los complejos de remodelaje ATP-dependientes y 4) los complejos originalmente descritos en *Drosophila melanogaster*, la mosca de la fruta, llamados Polycomb y Trithorax (PcG y trxG). A continuación presentaremos las características de cada uno de estos elementos.

#### **Modificaciones post-traduccionales de las histonas**

Los nucleosomas son uno de los constituyentes asociados a la cromatina más dinámicos. Son blanco de modificaciones post-traduccionales de las histonas (Fig. 1 y 6). Las regiones amino terminales de las histonas sobresalen de la estructura globular del nucleosoma. Mediante modificaciones post-traduccionales se favorecen o se relajan las uniones de estas proteínas básicas con el ADN. Estas regiones de las histonas son ricas en residuos de lisina, blanco de modificaciones covalentes tales como: acetilación, metilación, fosforilación y más recientemente se han descubierto modificaciones tales como la ubiquitinación y sumoilación, que marcan a las histonas como forma para regular su funcionamiento y degradación (Fig. 6) (24).

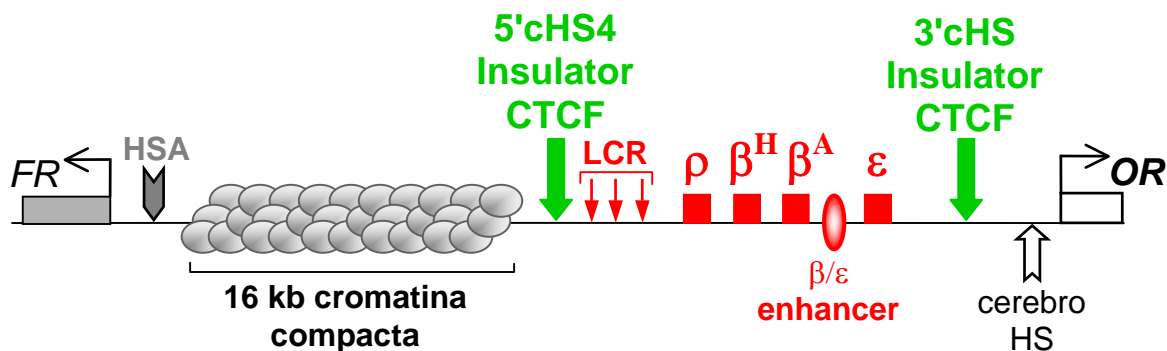
En la actualidad se ha establecido claramente la presencia de distintas combinaciones de éstas modificaciones, las cuales afectan tanto positiva como negativamente los distintos niveles de compactación de la estructura de la cromatina (Fig. 1 y 6). Estas combinaciones han incluso llevado a proponer su contribución en procesos celulares tan variados como la activación regulada de un gen hasta la segregación de los cromosomas ó el mismo control del ciclo celular (25).



**Figura 6.** Modificaciones covalente de las histonas. Las histonas H3, H4, H2A y H2B pueden ser modificadas en sus regiones amino terminales. Las principales modificaciones son: la acetilación y desacetilación de residuos lisina; la metilación (incluso existe la mon-, di- y tri-metilación); la fosforilación; la ubiquitinización y más recientemente la sumoilación (24-26). Todas estas modificaciones pueden darse en distintas combinaciones formando lo que Allis ha propuesto como el código de histonas que lleva a diversa respuestas regulatorias a nivel epigenético (24-26)

Recientemente se ha demostrado que las modificaciones postraduccionales favorecen la inter-comunicación entre nucleosomas afectando y regulando localmente la estructura de la cromatina, pero también al nivel de un dominio e incluso de un cromosoma completo (24). El efecto local se considera el que tiene relación con las regiones cercanas a los genes (Fig. 2), donde se localizan los elementos de control de la transcripción y a los propios genes donde las histonas se encuentran hiperacetiladas, característica consistente con una cromatina "abierta" y por lo tanto apta para ser transcrita. Por el contrario, en condiciones en las cuales la expresión

de un gen debe ser reprimida, sus histonas se encontrarán desacetiladas lo que favorece una cromatina compacta. Otro par de modificaciones involucra la metilación de los residuos de lisina 4 y 9 de la histona H3 cuyos efectos biológicos son antagónicos. La metilación de la lisina 4 de la histona H3 se asocia a la formación de eucromatina facilitando la actividad transcripcional (25, 26). Por el contrario, la metilación de la lisina 9 de la histona H3, por parte de la metil-transferasa de histonas SUV39H1, induce la incorporación de la proteína de heterocromatina 1 logrando como consecuencia un efecto represor en la actividad transcripcional (HP1; 26). Esta manera epigenética de inducir la formación de heterocromatina, es utilizada para silenciar promotores de genes como la ciclina E que es parte dependiente de la acción de E2F y Rb, moléculas reguladores del ciclo celular (27).



**Figura 7.** Esquema del dominio  $\beta$ -globina de pollo. El dominio abarca una región genómica de 55 kb y representa uno de los primeros ejemplos que validan el concepto de organización del genoma eucariota en dominios. El dominio  $\beta$ -globina de pollo está compuesto por un gen que se expresa en la etapa fetal ( $\rho$ ), dos genes adultos ( $\beta^H$  y  $\beta^A$ ) y un gen embrionario ( $\epsilon$ ), con un patrón de expresión diferencial a lo largo del desarrollo (17, 30). Los límites, tanto en el extremo 5' como en el 3', están definidos por dos elementos tipo "insulator" o barrera. De manera interesante, a cada uno de los "insulators" (5'cHS4 y 3'cHS) se le asocia el factor nuclear CTCF, lo cual hace pensar una posible interacción física y la formación de un asa cromatínica entre los extremos que definen al dominio  $\beta$ -globina de pollo (17). Río arriba en relación con el límite 5' del dominio  $\beta$ -globina, se identifican 16 kb de cromatina compacta, cuyo ADN se encuentra metilado y su estructura de la cromatina inaccesible al corte por parte de la endonucleasa DNasa I (18, 30, 31). Más "arriba" se identificó un gen que codifica para uno de los receptores de folato (*FR*) con un patrón de expresión eritroide pero en una etapa de diferenciación previa en la cual los genes  $\beta$ -globina se expresan (31). HSA, representa un sitio de hipersensibilidad con una actividad aparente de potenciador. Río abajo en relación con el límite 3' del dominio  $\beta$ -globina se identificó uno de los genes que codifican para uno de los miembros de la familia de receptores olfatorios (*OR*) con un patrón de expresión no eritroide y que es cerebro específico (32).

A nivel de un dominio, existen ejemplos en la literatura tanto en vertebrados como en hongos, donde las histonas que definen regiones del genoma que contienen a un gen o grupo de genes pueden encontrarse hiperacetiladas (en un intervalo de varias ó incluso decenas de kilo bases; 18, 28, 29). Dichos dominios se encuentran delimitados por las histonas con marcas de cromatina abierta y sus límites definen claramente una transición entre la cromatina abierta y la

cerrada (Fig. 7; 29). Uno de los ejemplos más claros se presenta en el dominio  $\beta$ -globina de pollo el cual constituye una familia multigénica que se expresa de manera diferencial a lo largo del desarrollo (30). En su extremo 5' se ha identificado un elemento tipo "insulator" seguido río arriba de 16 kb de cromatina compacta y un gen que codifica para uno de los receptores de folato que posee una expresión eritroide específica y que se encuentra activo en una etapa previa a la expresión de los genes para la globina (Fig. 7; 31). En el extremo 3', se identificó otro elemento "insulator" que delimita al dominio  $\beta$ -globina de un gen que codifica para uno de los receptores olfatorios con una expresión cerebro específico (Fig. 7; 32). Este ha sido uno de los pocos ejemplos, que muestran la organización del genoma en dominios, cuyos límites se encuentran definidos por secuencias tipo "insulator" y con un patrón bien definido en cuanto al estado de acetilación de las histonas (29, 33).

Finalmente, al nivel de un cromosoma completo el ejemplo clásico es la determinación del sexo mediante la inactivación del cromosoma X (34). Para lograr dicha inactivación convergen múltiples señales de silenciamiento empezando por la presencia de transcritos antisentido que parecen ser la señal para atraer modificaciones de la cromatina tales como la desacetilación de histonas, metilación del ADN y metilación de la lisina 9 de la histona H3 entre otras (34). Así el cromosoma X se verá silenciado en la expresión de sus genes, por medio de la formación de una cromatina compacta, en la mayor parte de su extensión.

En resumen, queda de manifiesto la interdependencia entre las distintas modificaciones covalentes que además constituyen, como lo ha propuesto David Allis, las señales moleculares que pueden propagarse de nucleosoma a nucleosoma construyendo así una respuesta a las señales de regulación que llevan a la modulación de la estructura de la cromatina.

### ***La metilación del ADN***

La metilación del ADN es una de las modificaciones epigenéticas más estudiadas en los últimos años e incorpora un nuevo concepto a la regulación conocida como "memoria epigenética" o "memoria celular". Este nuevo concepto de memoria se suma a la definición de epigenética en la cual los cambios en la expresión de los genes son heredables de una generación celular a otra.

La metilación del ADN desempeña un papel central en la transmisión de la información epigenética a través de la incorporación de un grupo metilo a la posición 5' de la citosina, siempre y cuando se encuentre en el contexto del dinucleótido CpG (Fig. 8; 35). La metilación del ADN es sinónimo de represión transcripcional mediada por la formación de una cromatina compacta. En términos generales la metilación del ADN contribuye a mantener la estabilidad del genoma, asegurando que las secuencias repetidas correspondientes a las secuencias de transposones no "salten", con lo cual se evita la inestabilidad genómica (36).

La metilación del ADN también participa en el silenciamiento de genes de manera permanente e irreversible en distintas etapas del desarrollo. Al momento del nacimiento se lleva a cabo una fuerte oleada de metilación la cual contribuye al "apagado" de genes embrionarios que no serán expresados en la etapa adulta del organismo (37). Otro ámbito en el cual la metilación del ADN contribuye a la memoria epigenética, es en el fenómeno conocido como "impronta genómica" en donde un grupo de dominios involucrados en la expresión de genes en

etapas tempranas del desarrollo poseen una expresión mono-alélica. Es decir, que la expresión de los genes improntados se lleva a cabo en el alelo paterno o en el alelo materno pero no en ambos (38). Recientemente se ha demostrado que para la mayoría de los loci improntados, una región de metilación diferencial contribuye en gran medida a que se realice una expresión mono-alélica (38).

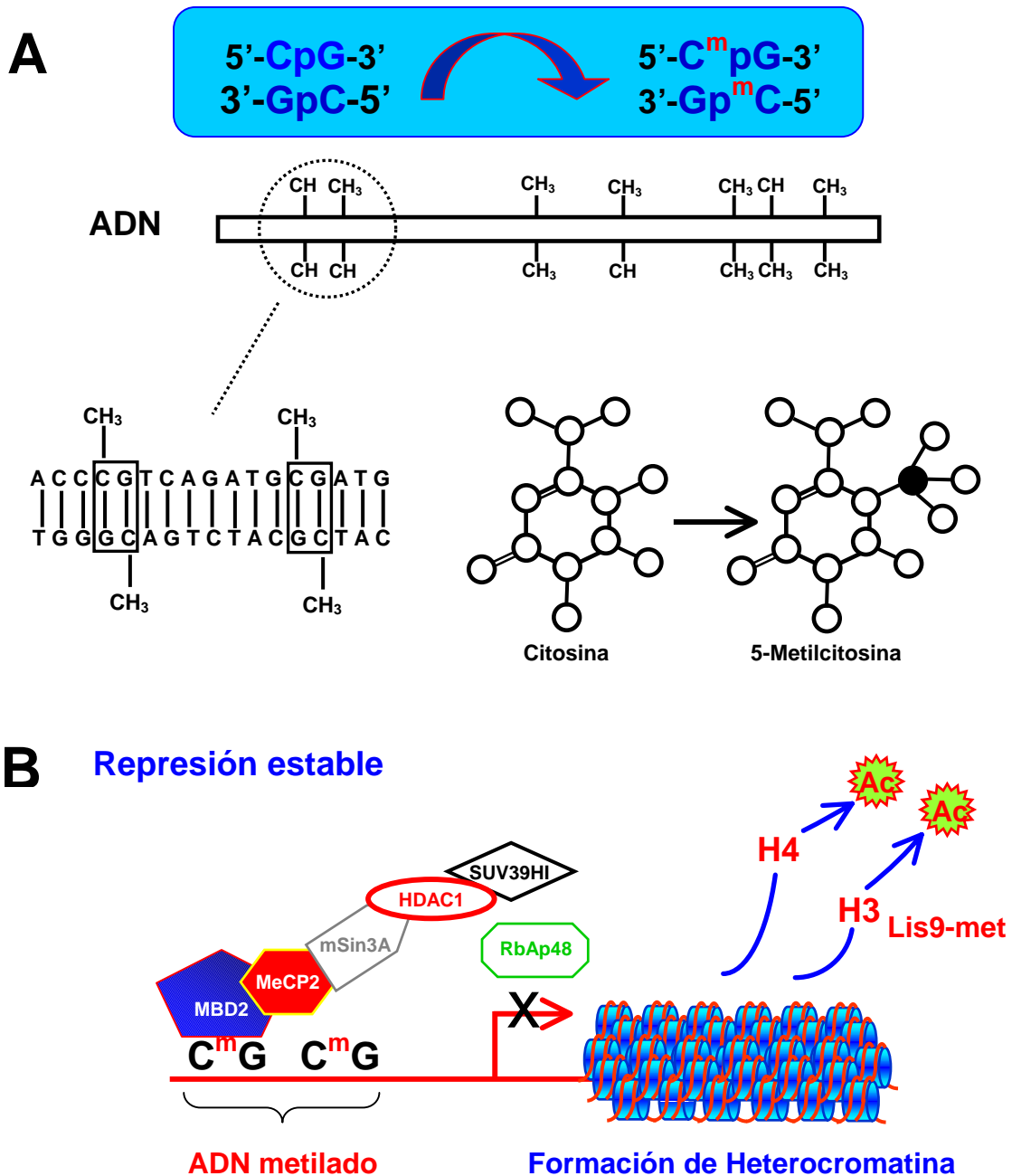
Investigaciones recientes han permitido identificar a una familia de proteínas con la capacidad de unirse al ADN conocidas como MBD's o MeCP's (del inglés: "Methyl-Binding Proteins" o "Methyl-CpG-binding Proteins", respectivamente (39)). Al interactuar con el ADN metilado y con la ayuda de los corepresores, se inician las actividades de remodelaje de la cromatina que inducen una cromatina compacta, en particular, a través de la acción de desacetilasas de las histonas (Fig. 8; 39).

Finalmente se ha demostrado que las modificaciones post-traduccionales de las histonas se encuentran epigenéticamente ligadas a la metilación del ADN. En la planta *Arabidopsis thaliana* y en el hongo *Neurospora crassa* se ha demostrado que la metilación de la lisina 9 de la histona H3 constituye una señal molecular para atraer a las metil-transferasas del ADN y metilarlo localmente (25). Incluso en una publicación reciente se demostró en *Neurospora crassa* que el homólogo de la proteína HP1 es atraída por la metilación de la lisina 9 de la histona H3, evento necesario para la metilación del ADN (40). Lo anterior demuestra no sólo la interdependencia entre las modificaciones covalentes de las histonas sino también con otros procesos epigenéticos como el caso de la metilación del ADN.

### **Los complejos de remodelaje ATP-dependientes**

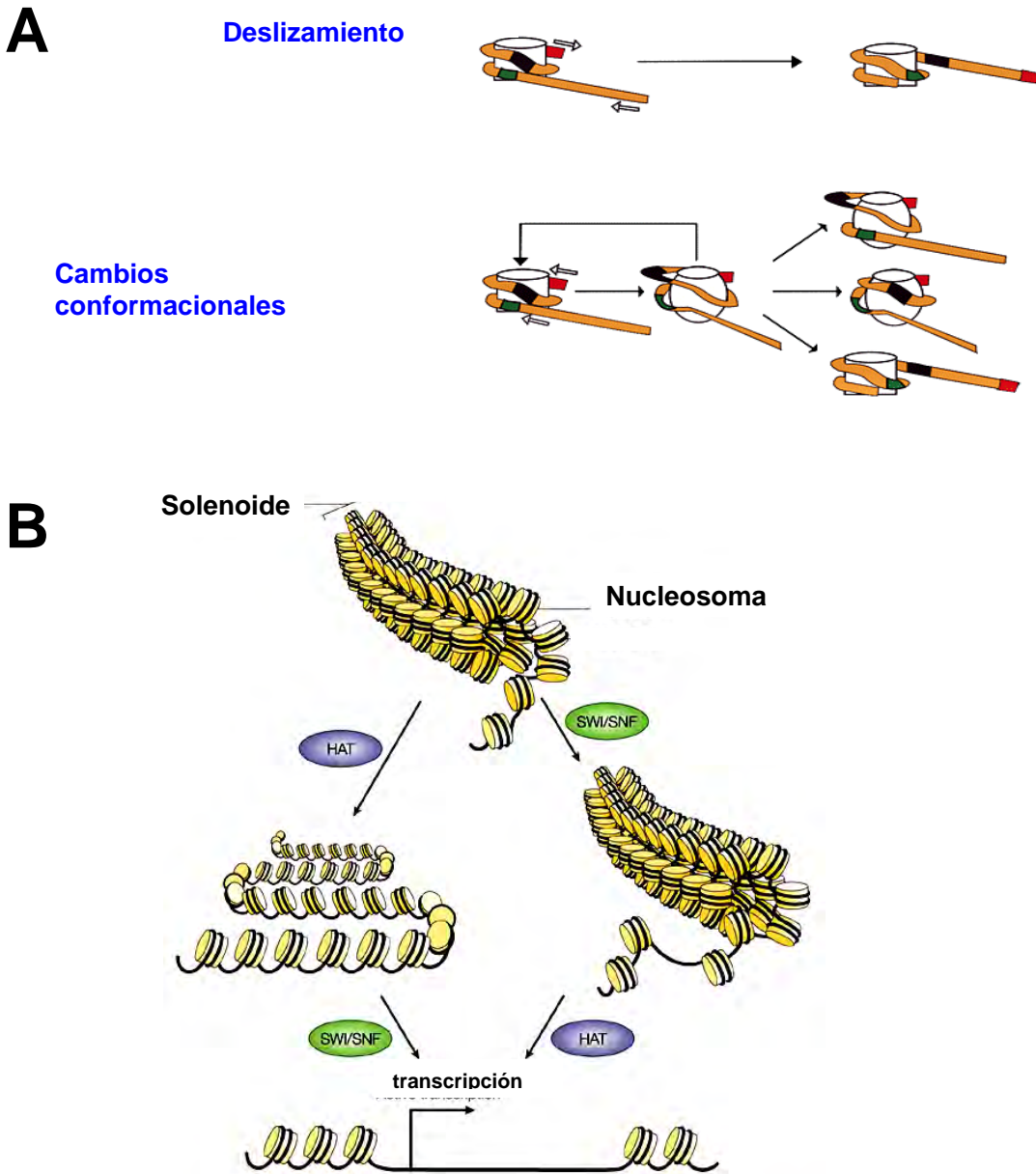
Para activar o reprimir la expresión de un gen es necesario que factores transcripcionales tanto generales como específicos tengan acceso, reconozcan y puedan unirse a sus secuencias blanco en la molécula de ADN. En el contexto de la estructura de la cromatina este proceso se ve obstaculizado por la organización nucleosomal. Recientemente se han descrito una serie de complejos multipéptidicos cuya función, dependiente de ATP, es la de liberar dichas secuencias desplazando o re-posicionando a los nucleosomas (Fig. 9). Estos desplazamientos ocurren a distancias cortas y sin que los nucleosomas se disocien por completo de sus interacciones con el ADN (41). SWI/SNF originalmente estudiado en levadura y *Drosophila*, constituye el complejo pionero seguido por NURF, CHRAC y ACF entre otros (42). El complejo Mi-2/NuRD está compuesto por desacetilasas de las histonas y las proteínas de unión al ADN metilado entre otras, por lo tanto posee tres actividades remodeladoras en un solo complejo (43).

Los modelos propuestos para el funcionamiento de estos complejos son básicamente dos. El primero atribuido a la capacidad de estos complejos para deslizarse a los nucleosomas mediante un proceso giratorio tanto del propio nucleosoma como de la molécula de ADN (Fig. 9A). El segundo modelo tiene que ver con un cambio conformacional del propio nucleosoma lo que facilita la exposición de secuencias en el ADN (Fig. 9A) (41, 42). Aparentemente la finalidad de estos complejos es la de liberar secuencias en el ADN que se encuentran ocultas por la estructuración del genoma en cromatina. Curiosamente lo opuesto también ocurre, es decir, se ha demostrado que estos complejos pueden actuar reposicionando a los nucleosomas de tal



**Figura 8.** La metilación del ADN constituye uno de los elementos centrales de la regulación epigenética. A- La metilación del ADN se refiere a la incorporación de un grupo metilo (CH<sub>3</sub>) en la posición 5 de la citosina siempre y cuando se encuentre formando el dinucleótido CpG. Cabe señalar que hasta ahora, esta reacción se considera como irreversible (35). B- La metilación del ADN debe considerarse como uno de los elementos de regulación epigenética con un efecto silenciador o represor sobre la expresión génica. Recientemente se ha favorecido el modelo en el cual la metilación del ADN se considera como uno de los factores críticos en el silenciamiento génico induciendo la formación de una estructura de la cromatina altamente compacta. Lo anterior se logra por medio de la unión de las proteínas que reconocen al ADN metilado (conocidas como MeCP's o MBD's) cuya función es la de atraer contactos moleculares con co-represores como el Sin3A, los cuales a su vez reclutan desacetilasas de histonas favoreciendo la compactación de la cromatina (35).





**Figura 9.** Los complejos de remodelaje ATP-dependientes en la modulación de la estructura de la cromatina. A- Actualmente se han descubierto un número creciente de complejos de remodelaje con distintas actividades. Hay dos modelos principales (aunque no son los únicos) para explicar la función de dichos complejos. El primero propone el deslizamiento del nucleosoma o "sliding" a lo largo del ADN sin que ocurra una disociación de las histonas del nucleosoma (41). Se ha demostrado que tanto el nucleosoma como la misma doble hebra del ADN giran facilitando dicho movimiento. Recordemos que la finalidad de estos complejos es modular tanto positiva como negativamente a la estructura de la cromatina, permitiendo que secuencias nucleotídicas sean expuestas o en su defecto ocultadas, para su reconocimiento por parte de los factores de transcripción e incluso el propio complejo de iniciación de la transcripción. El segundo modelo sugiere, a diferencia del anterior, un cambio conformacional transitorio que puede darse a múltiples niveles sin llegar al extremo de disociar por completo al nucleosoma. B- Datos experimentales recientes han demostrado una acción orquestada entre las acetilasas de las histonas (HAT's) y los complejos de remodelaje ATP-dependientes (como por ejemplo SWI/SNF) con la finalidad de "abrir" la estructura de la cromatina de manera más eficiente (41, 42, 44).

forma que ahora las secuencias blanco para algún factor transcripcional, se oculten. Lo anterior trae como consecuencia la represión de la actividad de un gen. Por lo tanto los complejos de remodelaje ATP-dependientes con base a interacciones con cofactores son capaces de regular la expresión génica a nivel epigenético, tanto positiva como negativamente. Como si fuera poco se han demostrado procesos cooperativos entre los complejos de remodelaje y las acetilasas de las histonas sin que exista un orden predominante en la acción de estas proteínas sobre la región de control (42, 44). Lo que parece volverse una constante es que estas dos actividades de remodelaje de la cromatina son cooperativas (Fig. 9B) (42, 44).

En resumen, los complejos de remodelaje ATP-dependientes representan un eslabón más dentro de la regulación epigenética. Aún falta mucho por comprender en relación con estos complejos ya que su complejidad los hace un tema difícil pero a la vez fascinante de investigación.

### **Regulación epigenética y cáncer**

Entre un 50 y 60 % de los promotores de genes humanos poseen una alta densidad de dinucleótidos CpG's. Su característica principal es la de corresponder a una región no-metilable. A estas regiones de control se les conoce como "islas CpG's" y regulan a un gran número de genes entre los cuales se encuentran genes constitutivos, oncogenes y genes supresores de tumores. En el resto del genoma humano encontramos en promedio de entre 70 y 80 % de CpG's las cuales en su gran mayoría están metiladas. En este contexto se pretende entender a la metilación del ADN como una modificación epigenética del genoma a partir de la cual puede darse el origen y progresión a procesos tumorales con base a la metilación anormal de las islas CpG's localizadas en promotores críticos de genes involucrados en el control del ciclo celular. Dentro de estos promotores que se metilan de manera anormal se encuentra un grupo nada despreciable, aunque no son los únicos, de promotores de genes supresores de tumores tales como: *Rb* (retinoblastoma); *p16<sup>INK4a</sup>* (un inhibidor de las cinasas dependientes de ciclinas); *BRCA1* (involucrado en la reparación del ADN) entre otros (45, 46). Como se mencionó, los mecanismos que rigen a la regulación epigenética tienen que ver con la metilación del ADN y eventos de remodelamiento de la cromatina. Recientemente se ha demostrado que el uso individual de inhibidores de la metilación del ADN, como la 5-azacitidina e inhibidores de las desacetilasas, como la tricostatina A (TSA), permite la reactivación de la expresión génica. Un punto importante que muestra la convergencia de estos mecanismos ligados a la estructura de la cromatina es que al utilizar simultáneamente ambos inhibidores se logran niveles de reactivación más elevados. Este último punto nos hace pensar seriamente en la aplicación de estos compuestos como agentes terapéuticos. A pesar de conocer poco sobre el espectro de acción de estas drogas y sus posibles efectos colaterales, su uso se encuentra ya en distintas fases de análisis clínico para su futura aplicación en pacientes.

La importancia al nivel celular de las enzimas responsables de la metilación del ADN está claramente demostrada por la mutación de dichas proteínas, lo cual provoca un fenotipo letal en ratón. Basados en estas mutaciones tipo "knock-out" y los resultados experimentales, se ha propuesto a la *Dnmt3a* y la *Dnmt3b* como las metil-transferasas de ADN responsables de la metilación *de novo* del ADN, mientras que la *Dnmt1* se considera como la enzima responsable de mantener a la metilación en cada división celular. Estas actividades concuerdan con la

definición de la epigenética dado que la información molecular derivada de sus actividades es heredable ya que se sugiere que dicha información o "memoria molecular" puede ser transmitida de generación en generación lo cual representa un mecanismo epigenético de regulación (35, 46).

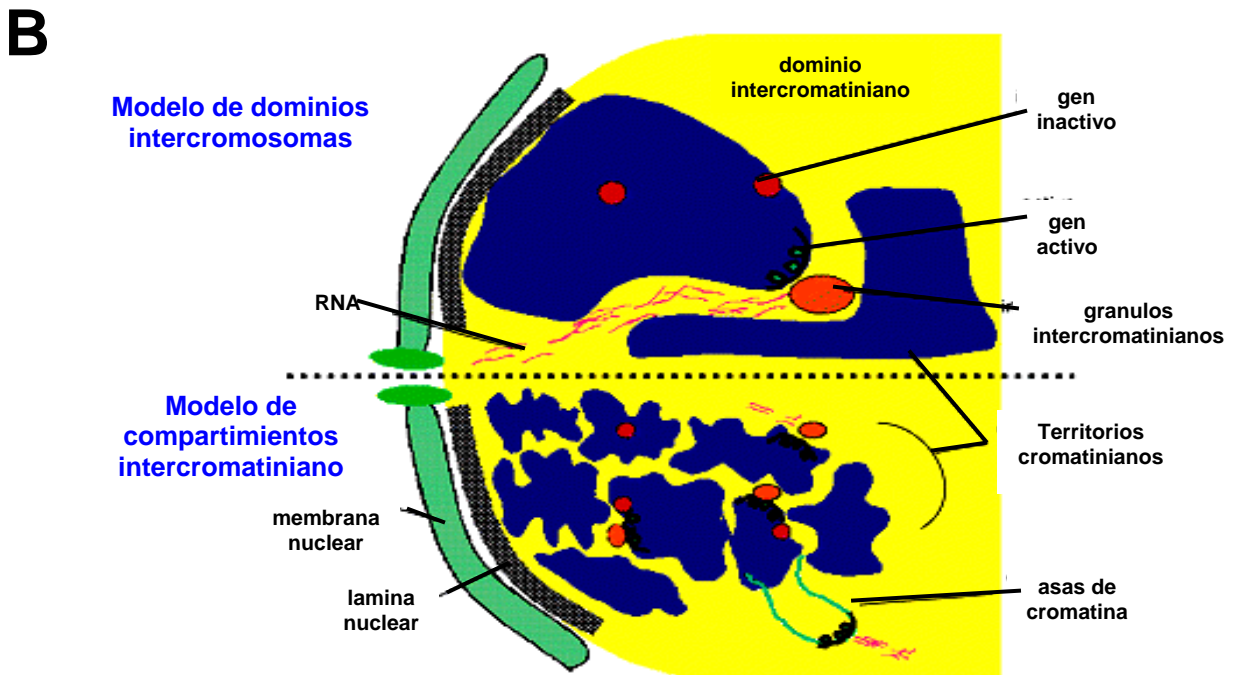
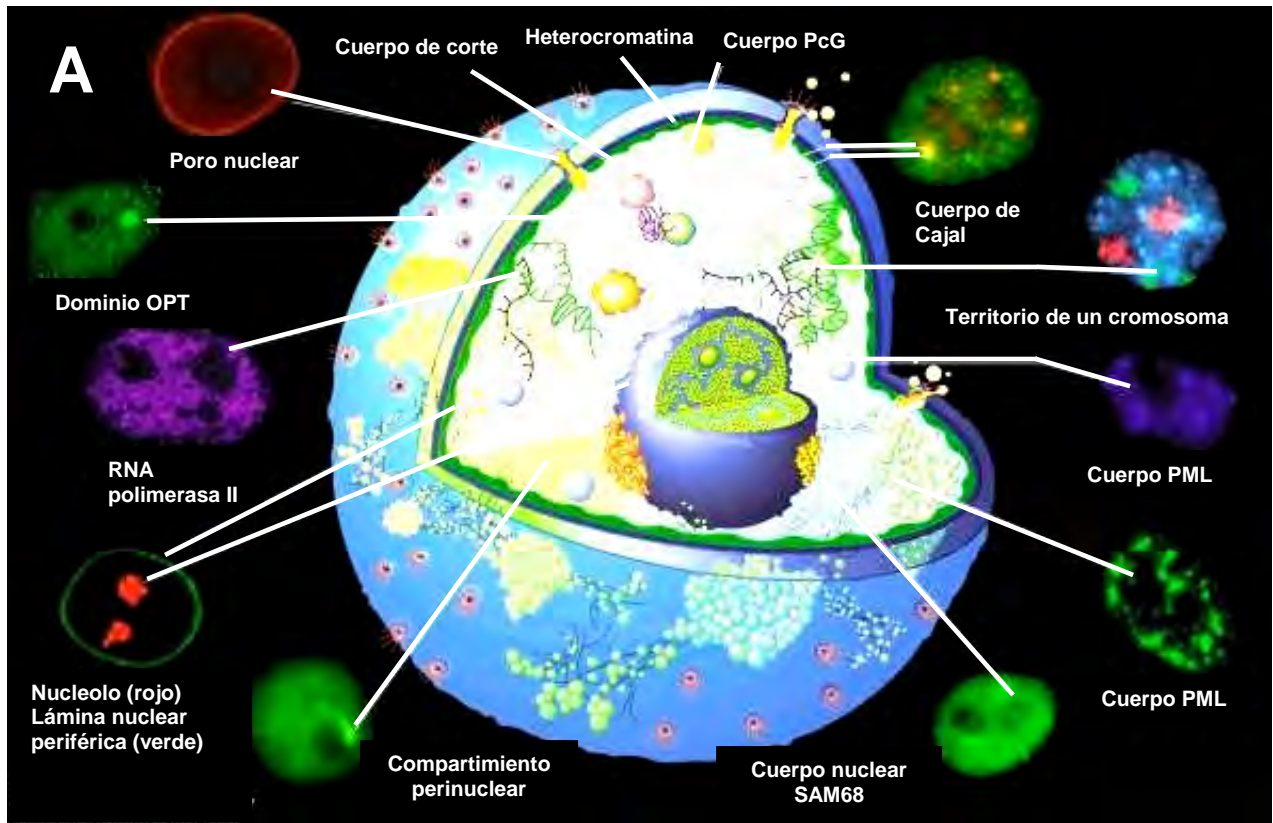
Otro fenómeno biológico donde la regulación epigenética desempeña un papel crítico, es la impronta genómica o "imprinting" que se refiere a la expresión mono-alelica de genes. En particular con los grupos de genes involucrados en el desarrollo embrionario de los mamíferos. Parte de los mecanismos que definen la expresión específica de un alelo y el silenciamiento del otro alelo esta dada por una región conocida como el dominio de metilación diferencial el cual por lo general se encuentra no metilado en el alelo que va a ser expresado mientras que el alelo silenciado se encuentra metilado. Por lo tanto, parte de la regulación de la expresión en los dominios improntados está dada por los mecanismos epigenéticos. La pérdida de esta expresión mono-alelica trae como consecuencia la desregulación de muchos genes incluyendo oncogenes y genes supresores de tumores al igual que la aparición de ciertos síndromes (38, 47).

### **Regulación epigenética a través de los complejos Polycomb y Trithorax**

Los grupos de genes Polycomb y Trithorax fueron originalmente descritos en *Drosophila* como proteínas que regulan la expresión de los genes homeóticos. El grupo de genes Polycomb (PcG) codifica para un conjunto de proteínas que regulan negativamente a los genes homeóticos mientras que Trithorax (trxG) lo hace positivamente (48-50). Se ha postulado que PcG y trxG son parte de los mecanismos que contribuyen a la memoria epigenética. En particular, las proteínas del grupo PcG son responsables de silenciar no sólo a los genes homeóticos fomentando una

---

**Figura 10.** El núcleo participa activamente en la regulación de la expresión génica. A- El núcleo constituye hoy en día un organelo subcelular mucho más complejo de lo imaginado pero sobre todo muy dinámico y que influye en la expresión de los genes. En este esquema se resaltan varios ejemplos de compartimientos y estructuras subnucleares, en particular la organización del genoma eucariota en dominios. Cabe hacer notar que el avance en las técnicas de marcaje fluorescente y de la microscopia de fluorescencia han permitido incluso proponer una organización de los cromosomas en territorios en un núcleo en interfase (53-55). B- La organización del genoma en territorios ha llevado a proponer los modelos de dominios inter cromosómicos o de compartimientos inter cromosómicos. Ambos modelos proponen que el relajamiento de la estructura de la cromatina y la transcripción ocurre en la periferia de los territorios en contacto con los canales intercromatinianos donde se localiza una mayor densidad de los factores tanto de transcripción como de procesamiento e incluso organelos sub-nucleares como los cuerpos de Cajal y los "speckles" entre otros. Debemos recordar que este modelo contempla el facilitar el procesamiento de los transcritos nacientes y su localización en los canales intercromatinianos facilita su transporte hacia los poros nucleares y la subsecuente migración hacia el citoplasma. En el modelo de dominios inter cromosómicos los territorios abarcan grandes áreas y en sus regiones centrales los genes se encuentra silenciados e inmersos en una estructura de la cromatina compacta. Consecuentemente la activación de un gen requerirá de una re-localización de dicha secuencias, de zonas centrales a la periferia de un territorio. Por el contrario el modelo de compartimientos inter cromosómicos prevé que los territorios estén formados por subterritorios y el acceso a la periferia de estos se vería facilitado (53, 55).



estructura de la cromatina represora evitando así la incorporación de factores de transcripción y por lo tanto la transcripción. Uno de los modelos propuestos contempla la formación de complejos multipéptídicos que reconocen a las secuencias llamadas PRE (del inglés: "Polycomb Response Elements"). Proteínas homólogas al grupo PcG y trxG han sido identificadas en diversos organismos incluyendo al humano (42, 49). En particular, los miembros del grupo PcG han sido asociados a la regulación de genes ligados al control del ciclo celular vía la proteína retinoblastoma (Rb; 51). Por otra parte, la sobre expresión anormal de los miembros de PcG se asocia directamente a la proliferación celular relacionado con la progresión del cáncer de próstata (52). De igual manera mutaciones en los miembros de trxG han mostrado tener relación con la aparición de tumores pediátricos, en particular a través del complejo hBrm que corresponde a un elemento de remodelaje ATP-dependiente asociado al grupo trxG (42).

En resumen, poco se sabe a cerca de estos complejos en cuanto a su función en la regulación epigenética pero resulta indudable que son parte de ella y que su acción es una vez más a través de modificar y modular diversas formas a la estructura de la cromatina.

#### **IV- La estructura de la cromatina y la organización nuclear**

Durante años el núcleo celular fue considerado como un organelo subcelular con la única función de contener en su interior a la información genética de la célula.

Apoyados en los avances de la microscopia electrónica y recientemente de la microscopia de fluorescencia, ha sido posible determinar que el núcleo celular es una entidad dinámica y que en definitiva participa en la regulación de la expresión génica (Fig. 10A; 53, 54). El núcleo de una célula eucariota se encuentra compartimentalizado en zonas de heterocromatina y eucromatina pero también por estructuras subnucleares tales como los cuerpos de Cajal, los cuerpos promielocíticos (PML), los "speckles" y el nucleolo. Cada uno de estos compartimientos poseen funciones bien definidas que van desde almacenar a los factores de transcripción, los factores ligados al procesamiento de los RNA pre-mensajeros hasta los genes ribosomales en el caso del nucleolo entre otros. De manera general la relación que existe entre estos compartimientos es prácticamente una, el llevar a cabo de manera regulada, la expresión génica y el procesamiento de los RNA's mensajeros.

Uno de los aspectos más novedosos al respecto, es la organización del genoma en territorios al interior del núcleo. Estudios realizados por microscopia de fluorescencia, con base a la posibilidad de marcar de manera fluorescente cromosomas completos han demostrado su localización en territorios bien definidos (Fig. 10B; 53). En un mismo territorio el genoma varía en cuanto a su organización de la cromatina. Se ha propuesto que en las regiones más centrales al interior de un territorio, la estructura de la cromatina se encuentra en un alto grado de compactación. Por otra parte al interior de un territorio se presentan zonas de heterocromatina mayoritariamente asociadas a la periferia nuclear y perinucleolar (Fig. 10B). A partir del surgimiento del concepto de organización del genoma en territorios al interior del núcleo se planteó un nuevo modelo en el cual las regiones entre territorios, conocidas como regiones intercromatinianas participan activamente en la expresión génica (55). El modelo propone que las regiones del genoma que serán transcritas se relocalizan (por mecanismos aún desconocidos) en la periferia de los territorios y en contacto con los compartimientos

intercromatinianos donde además se encuentran posas de los factores transcripcionales, las RNA polimerasas y los factores de procesamiento (Fig. 10B; 55). Experimentalmente se ha demostrado que las histonas asociadas al genoma en las zonas internas en un territorio se encuentra mayoritariamente desacetiladas mientras que las zonas más expuestas y en contacto con las regiones intercromatinianas se encuentran hiperacetiladas (55, 56). Este modelo resulta ser muy atractivo dado que al citolocalizar tanto al ADN a ser transcrito, los elementos de regulación y de procesamiento, al generar un transcrito maduro podría ser a su vez transportado por estos "canales" pericromatinianos dirigiendo y facilitando su transporte a través de los poros nucleares hacia el citoplasma para que dicho mensaje sea traducido en una proteína funcional (Fig. 10B; 55, 56).

Un aspecto más reciente que confirma la participación de la dinámica nuclear en la regulación de la expresión génica, son los procesos de relocalización (57). La relocalización se refiere al hecho de asociar, tanto regiones de cromatina como factores transcripcionales a las zonas de heterocromatina o eucromatina, ocasionando la represión o activación de la expresión génica, respectivamente. Actualmente son tres los ejemplos que apoyan a la relocalización al interior del núcleo como un nivel novedoso de regulación:

1- Represión por relocalización mediada por la proteína Ikaros. Ikaros es una proteína específica del linaje celular linfocítico y cuya función es reprimir la expresión de los genes relacionados con el sistema inmune (58, 59). Ikaros actúa reclutando genes que no deben ser transcritos, asociando dichos genes a regiones centroméricas y por lo tanto heterocromatinizándolas, causando su represión a nivel transcripcional (58-60).

2- Utilizando al modelo de los genes  $\beta$ -globina se ha demostrado que los elementos de la regulación a distancia, como lo es un potenciador o "enhancer", en células no eritroides o en caso de las mutaciones en el potenciador, éste es inactivo con base a su asociación con las regiones de heterocromatina centromérica (61).

3- En las células eritroides se ha demostrado que el factor transcripcional NF-E2p45, el cual se dimeriza con el factor NF-E2p18, se encuentra en compartimientos distintos al interior del núcleo (62). En particular el factor NF-E2p45 se encuentra mayoritariamente asociado a zonas de heterocromatina al interior del núcleo cuando los genes globina no se encuentran activos. Al diferenciarse las células eritroides y los genes globina volverse activos transcripcionalmente, NF-E2p45 se relocaliza en las regiones transcripcionalmente activas asociándose con el factor NF-E2p18 y uniéndose a sus secuencias blanco para llevar a cabo sus funciones de regulación (62).

En resumen, resulta indudable que la visión actual del núcleo es mucho más dinámica de lo imaginado y con una mayor ingerencia en la regulación de la expresión de los genes.

## ***V- Conclusiones y perspectivas***

La regulación de la expresión de los genes influye en un gran número de los procesos, tanto normales como anormales, a lo largo de la vida de una célula. En sus inicios el estudio de la expresión génica parecía considerar únicamente la necesidad de atraer a las RNA polimerasas a regiones cercanas al sitio de inicio de la transcripción y llevar a cabo la expresión de un gen. Actualmente nos encontramos en un momento histórico donde las investigaciones de los últimos 10 años nos muestran un escenario insospechado, mucho más complejo de lo anticipado.

La activación coordinada en el tiempo y el espacio de un gen requiere de múltiples niveles y de interacciones secuenciales e interdependientes entre un gran número de moléculas. Además son necesarios los elementos de regulación, tanto proximales como distales, con una arquitectura particular elevando los niveles de especificidad. Por si fuera poco, todo esto ocurre en el contexto fisiológico representado por la estructura de la cromatina. Además, el núcleo celular también participa en la expresión específica de los genes de una manera más activa de lo antes previsto.

El ambiente en el cual deben ocurrir la gran mayoría de los eventos de la regulación, de control de la progresión del ciclo celular, de la duplicación y la reparación del ADN, los eventos de recombinación, de impronta genómica y muchos otros tiene lugar en el contexto de la estructura de la cromatina. Los descubrimientos relacionados con la modulación de la estructura de la cromatina han mostrado su papel crítico en la regulación epigenética. Por lo tanto y de manera integral la transmisión de la información genética, principalmente codificada en la molécula del ADN, no puede entenderse sin una estricta dependencia con procesos de la regulación epigenética y viceversa.

De lo anterior surgen evidencias que demuestran que la desregulación al nivel de la estructuración del genoma en la cromatina representa hoy en día una de las causas del origen y progresión de procesos tumorigénicos. Durante muchos años los esfuerzos fueron canalizados en los defectos genéticos, minimizando la contribución de los procesos epigenéticos en el desarrollo de diversas enfermedades principalmente en el cáncer. Con base a lo anterior se pone de manifiesto la necesidad de invertir tiempo y esfuerzo en conocer mejor los mecanismos de regulación epigenética, para en conjunción con la información genética se logren desarrollar estrategias predictivas y terapéuticas en beneficio del sector salud.

### **Agradecimientos**

Agradecemos a Irene Patricia del Arenal y Héctor Rincón Arano por la lectura crítica del manuscrito. Las investigaciones realizadas en nuestro laboratorio son apoyadas económicamente por: Dirección General de Asuntos del Personal Académico-UNAM (IN203200 y IN209403), Consejo Nacional de Ciencia y Tecnología, CONACyT (33863-N y 42653-Q), Third World Academy of Sciences (TWAS, Grant 01-055 RG/BIO/LA) y Fundación Miguel Alemán. MEDA es apoyado por parte del CONACYT (166354) con una beca de doctorado.

### **Referencias**

- 1- van Driel R, Fransz PF, Verschurre PJ (2003) The eukaryotic genome: a system regulated at different hierarchical levels. *J Cell Sci* 116: 4067-4075.
- 2- Felsenfeld G, Groudine M (2003) Controlling the double helix. *Nature* 421: 448-453.
- 3- Bulger M, Groudine M. Looping versus linking: toward a model for long-distance gene activation. *Genes Dev* 13:2465-2477.
- 4- Butler JEF, Kadonaga JT (2002) The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev* 16:2583-2592.
- 5- Kadonaga JT (2004) Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. *Cell* 116:247-257.
- 6- Roeder RG (1998) Role of general and gene-specific cofactors in the regulation of eucariotic transcription. In *Mechanisms of Transcription*. Cold Spring Harbor Lab Press, Vol. LXIII pp. 201-218.
- 7- Reinberg D, Orphanides G, Ebright R, Akoulitchev S, Carcamo J, Cho H, Cortes P, Drapkin R, Flores O, Ha I, Inostroza JA, Kim S, Kumar P, Lagrange T, Leroy G, Lu H, Ma D-M, Maldonado E, Merino A,

- Mermelstein F, Olave I, Sheldon M, Shierhattar R, Stone N, Sun X, Weis L, Yeung K, Zavel L. (1998), The RNA polimerasa II general Transcription factors: Past, Present and Future. In Mechanisms of Transcription. Cold Spring Harbor Lab Press, Vol. LXIII pp. 83-103.
- 8-Blackwood EM, James T Kadonaga JT (1998), Going the distance: A current view of enhancer action. *Science* 281:60-63.
  - 9- Smale ST (2001) Core promoters: active contributors to combinatorial gene regulation. *Genes Dev* 15: 2503-1508.
  - 10- Carter D, Chakalova L, Osborne CS, Dai Y-F, Fracer P (2002) Long-range chromatin regulatory interactions *in vivo*. *Nature Genet* 32:623-626.
  - 11- Brand AH, Breeden L, Abraham J, Sternglanz R, Nasmyth K. (1985) Characterization of a "silencer" in yeast: a DNA sequence with properties opposite to those of a transcriptional enhancer. *Cell* 41:41-48.
  - 12- Kamakaka R (1997) Silencers and locus control regions: opposite sides of the same coin. *Trends Biochem Sci* 22: 124-128.
  - 13- Ogbourne S, Antalis TM. (1998) Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem J* 331: 1-14.
  - 14- Clark AR, Docherty K. (1993) Negative regulation of transcription in eukaryotes. *Biochem J* 15: 521-541.
  - 15- Recillas-Targa F, Razin SV. (2001) Chromatin domains and regulation of gene expression: familiar and enigmatic clusters of chicken globin genes. *Crit Rev Eukaryot Gene Expr* 11:227-242.
  - 16- Valadez Graham V, Recillas Targa F (2002) Características y formación de un dominio transcripcionalmente activo en eucariotes. *REB* 21: 21-40.
  - 17- West AG, Gaszner M, Felsenfeld G (2002) Insulators: many functions, many mechanisms. *Genes Dev* 16: 271-288.
  - 18- Razin SV, Farrell CM, Recillas-Targa F (2003) Genomic domains and regulatory elements operating at the domain level. *Int Rev Cytol* 226: 63-125.
  - 19- Li Q, Peterson KR, Fang X, Stamatoyannopoulos G (2002) Locus control regions. *Blood* 100: 3077-3086.
  - 20- Lemon B, Tjian R (2000) Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev* 14: 2551-2569.
  - 21- Orphanides G, Reinberg D (2002) A unified theory of gene expression. *Cell* 108: 439-451.
  - 22- Zurita M, Merino C (2003) The transcriptional complexity of the TFIIH complex. *TIGS* 19: 578-584.
  - 23- Carrozza MJ, Kusch T, Workman JL (2003) Repairing nucleosomes during transcription. *Nat Struct Biol* 10: 879-880.
  - 24- Fischle W, Wang Y, Allis CD (2003) Histone and chromatin cross-talk. *Curr Opin Cell Biol* 15: 172-183.
  - 25- Hake SB, Xiao A, Allis CD (2004) Linking the epigenetic "language" of covalent histone modifications to cancer. *Br J Cancer* 90: 761-769.
  - 26- Lachner M, Jenuwein T (2002) The many faces of histone lysine methylation. *Curr Opin Cell Biol* 14: 286-298.
  - 27- Nielsen SJ, Schneider R, Bauer UM, Bannister AJ, Morrison A, O'Carroll D, Firestein R, CLEarly M, Jenuwein T, Herrera RE, Kouzarides T (2002) Rb targets histone H3 methylation and HP1 to promoters. *Nature* 412: 561-565.
  - 28- Noma K, Allis CA, Grewal SI (2001) Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science* 293: 1150-1155.
  - 29- Litt MD, Simpson M, Recillas-Targa F, Prioleau M-N, Felsenfeld G (2001) Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *EMBO J* 20: 2224-2235.
  - 30- Recillas-Targa F (2000) The chicken  $\alpha$ - and  $\beta$ -globin gene domains and their chromatin organization. *Cell Mol Biol Lett* 5: 451-467.
  - 31- Prioleau M-N, Nony P, Simpson M, Felsenfeld G (1999) An insulator element and condensed chromatin region separate the chicken  $\beta$ -globin locus from an independently regulated erythroid-specific folate receptor gene. *EMBO J* 18: 4035-4048.
  - 32- Saitoh N, Bell AC, Recillas-Targa F, West AG, Simpson M, Pikaart M, Felsenfeld G (2000) Structural and functional conservation at the boundaries of the chicken  $\alpha$ -globin domain. *EMBO J* 19: 2315-2322.
  - 33- Litt MD, Gaszner M, Simpson M, Allis CD, Felsenfeld G (2001) Correlation between histone lysine methylation and developmental changes at the chicken  $\beta$ -globin locus. *Science* 293: 2453-2455.
  - 34- Lee JT (2003) Molecular links between X-inactivation and autosomal imprinting: X-inactivation as a driving force for the evolution of imprinting ? *Curr Biol* 13: R242-R254.
  - 35- Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes Dev* 16: 6-21.
  - 36- Wolffe AP, Matzke MA (1999) Epigenetics: regulation through repression. *Science* 286: 481-486.
  - 37- Li E (2002) Chromatin modification and epigenetic reprogramming in mammalian development. *Nat Rev Genet* 3: 662-673.
  - 38- Recillas-Targa F (2002) DNA methylation, chromatin boundaries and mechanisms of genomic imprinting. *Arch Med Res* 33: 428-438.



- 39- Wade PA (2001) Methyl CpG-binding proteins and transcriptional repression. *BioEssays* 23: 1131-1137.
- 40- Frreitag M, Hickey PC, Khalafallah TK, Read ND, Selker EU (2004) HP1 is essential for DNA methylation in *Neurospora*. *Mol Cell* 13: 427-434.
- 41- Narlikar GJ, Fan HY, Kingston RE (2002) Cooperation between complexes that regulate chromatin structure and transcription. *Cell* 108: 475-487.
- 42- Roberts CWM, Orkin SH (2004) The SWI/SNF complex-chromatin and cancer. *Nat Rev Cancer* 4: 133-142.
- 43- Bowen NJ, Fujita N, Kajita M, Wade PA (2004) Mi-2/NuRD: multiple complexes for many purposes. *Biochim Biophys Acta* 1677: 52-57.
- 44- Fry CJ, Peterson CL (2002) Unlocking the gates to gene expression. *Science* 295: 1847-1848.
- 45- Baylin SB, Herman JG (2000) DNA hypermethylation in tumorigenesis: Epigenetics joins genetics. *TIG* 16: 168-173.
- 46- Jones PA, Baylin SB (2002) The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 3: 415-428.
- 47- Tilghman SM (1999) The sins of the fathers and mothers: genomic imprinting in mammalian development. *Cell* 96: 185-193.
- 48- Pirrotta V (1998) Polycomb the genome: PcG, TrxG and chromatin silencing. *Cell* 93: 333-336.
- 49- Francis NJ, Kingston RE (2001) Mechanisms of transcriptional memory. *Nat. Rev Mol Cell Biol* 2: 409-421.
- 50- Orlando V (2003) Polycomb, epigenomes and control of cell identity. *Cell*, 112, 599-606.
- 51- Dahiya A, Wong S, Gonzalo S, Gavin M, Dean DC (2001) Linking the Rb and Polycomb pathways. *Mol Cell* 8:557-568.
- 52- Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG, Ghosh D, Pienka KJ, Sewalt RGAB, Otte AP, Rubin MA, Chinnaiyan AM (2002) The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* 419: 624-629.
- 53- Cremer T, Cremer C (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* 2: 292-301.
- 54- O'Brien TP, Bult CJ, Cremer C, Grunze M, Knowles BB, Langowski J, McNally J, Pederson T, Politz JC, Pombo A, Schmahl G, Spatz JP, van Driel R (2003) Genome function and nuclear architecture: from gene expression to nanoscience. *Genome Res* 13: 1029-1041.
- 55- Kosak ST, Groudine M (2002) The undiscovered country: chromosome territories and the organization of transcription. *Dev Cell* 2: 690-692.
- 56- Parada LA, Misteli T (2002) Chromosome positioning in the interphase nucleus. *Trends Cell Biol* 12: 425-432.
- 57- Francastel C, Schubeler D, Martin DI, Groudine M (2000) Nuclear compartmentalization and gene activity. *Nat Rev Mol Cell Biol* 1: 137-143.
- 58- Hahmk K, Cobb BS, McCarty AS, Brown KE, Klug CA, Lee R, Akashik A, Weissman IL, Fisher AG, Smale ST (1998) Helios, a T cell-restricted Ikaros family member that quantitatively associates with Ikaros at centromeric heterochromatin. *Genes Dev* 12: 782-796.
- 59- Georgopoulos K (2002) Haematopoietic cell-fate decisions, chromatin regulation and Ikaros. *Nat Rev Immunol* 2: 162-174.
- 60- Brown KE, Guest SS, Smale ST, Hahn K, Markenschlager M, Fisher AG (1997) Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* 91: 845-854.
- 61- Schubeler D, Francastel C, Cimbora DM, Reik A, Martin DR, Groudine M (2000) Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human  $\beta$ -globin locus. *Genes Dev* 14: 940-950.
- 62- Francastel C, Magis W, Groudine M (2001) Nuclear relocation of a transactivator unit precedes target gene activation. *Proc Natl Acad Sci USA* 98: 12120-12125.

## PARTICIPACIÓN DE LA ESTRUCTURA DE LA CROMATINA EN LA REGULACIÓN DE LA EXPRESIÓN GÉNICA

### Resumen.

La regulación de la expresión genética en los eucariontes representa un proceso central en la diferenciación, la proliferación y la muerte celular. La regulación genética incluye la unión de los factores de transcripción a sitios específicos del ADN y el reclutamiento de diferentes cofactores. En su medio ambiente natural, todo este proceso ocurre en el contexto de la cromatina en el cual, las modificaciones de las histonas son uno de los principales objetivos. Actualmente, la organización y dinámica del núcleo, así como sus implicaciones biológicas sobre la expresión genética, son tópicos de estudio por parte de diferentes grupos científicos. La regulación de la expresión genética puede ser visualizada, en términos generales, en tres diferentes niveles, todos ellos interrelacionados. El primero de ellos involucra la información genética codificada por la molécula del ADN. El segundo involucra el contexto natural en el cual el ADN es funcional. Esta es la base de la organización del genoma en la estructura de la cromatina y todas las actividades asociadas al remodelaje de ésta, conocido también como la regulación epigenética. Finalmente, un importante número de datos experimentales apoyan la idea de un papel más activo y dinámico del núcleo celular en la regulación de la expresión de los genes. En el presente capítulo, revisamos el conocimiento actual sobre los diferentes elementos relacionados con la regulación de la expresión genética, los factores generales involucrados en el inicio de la transcripción, e integramos esta información tomando en consideración la estructura de la cromatina y la contribución del núcleo celular en la expresión genética diferencial en una célula eucarionte.

**Palabras clave:** Transcripción, regulación genética, cromatina, núcleo, modificación de histonas.

### Semblanza del Dr. Félix Recillas Targa.



Nació el 11 de julio de 1961 en la Ciudad de México. Es biólogo de la Facultad de Ciencias de la Universidad Nacional Autónoma de México (1986) y doctor en Ciencias con especialidad en Bioquímica, por la Universidad de París 7, Francia (1993). Tesis realizada en el laboratorio del Dr. Klaus Scherrer en el Instituto Jacques Monod del CNRS. Realizó una primera estancia post-doctoral en el Instituto de Biotecnología de la UNAM de 1993 a 1996. Realizó una segunda estancia post-doctoral en el laboratorio del Dr. Gary Felsenfeld en los Institutos Nacionales de Salud de los Estados Unidos (NIH) de 1996 a 1999 en Bethesda, Maryland, EU. Actualmente es investigador titular del Departamento de Genética Molecular del Instituto de Fisiología Celular de la UNAM desde 1999. Es miembro del Sistema Nacional de Investigadores Nivel 2. Su desarrollo científico versa sobre el estudio del efecto de la estructura de la cromatina y la regulación génica a distintos niveles. En particular, emplea los dominios  $\alpha$ - y  $\beta$ - de las globinas como sistema experimental. Su trabajo experimental se ha publicado en 40 artículos internacionales y 10 nacionales, que han recogido más de 600 citas internacionales. Ha dirigido 2 tesis de licenciatura, 2 tesis de maestría y actualmente 5 tesis de doctorado. Es miembro de diferentes cuerpos académicos, entre los que se encuentran la Sociedad Mexicana de Bioquímica, la Sociedad Mexicana de Biología del Desarrollo y la Academia Mexicana de Ciencias; asimismo, es Miembro del Consorcio promotor del Instituto Mexicana de Medicina Genómica, Miembro del Comité de Difusión de Información sobre el Genoma Humano de la Comisión Nacional para el Genoma Humano y Socio Fundador y Tesorero de la Sociedad Mexicana de Medicina Genómica (SOMEGEN).

## Transcriptional regulation: regulatory elements and their way of action

Martín Escamilla-Del-Arenal and Félix Recillas-Targa\*

Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Apartado Postal 70-242, México D.F. 04510.

\* Corresponding autor: Félix Recillas-Targa  
Instituto de Fisiología Celular  
Departamento de Genética Molecular  
Universidad Nacional Autónoma de México  
Apartado Postal 70-242  
México D.F. 04510  
México  
Phone: (52 55) 56 22 56 74  
FAX: (52 55) 56 22 56 30  
Email: [frecilla@ifc.unam.mx](mailto:frecilla@ifc.unam.mx)

## **Abstract**

**Transcriptional regulation is a highly sophisticated and complex process that requires the coordinated action of multiple and inter-dependent components to reach tissue and developmental gene expression in eukaryotic cells. Such regulation is associated to multiple levels where genetic and epigenetic processes are needed. Here we reviewed how promoter architecture is relevant for short and/or long distance interactions and the different models associated to enhancer's way of action. At the present time the main challenge is to integrate all the novel data generated in particular through novel methodologies allowing the demonstration of the association distal genomic sequences and the fluorescent nuclear studies. These results confirm the relevance of the organization and topology of the nucleus and its constituents and demonstrate a new dimension of the regulation of gene expression that is intimately associated to epigenetic processes.**

## Introduction

Whole genome sequences are now available from different organisms facilitating their genetic material to the scientific and biomedical community. Comparative genome analyses revealed a surprising consistency in their genetic content (Levine and Tjian, 2003). Indicating that the increasing gene number, do not necessary account for increase in morphological and behavioral complexity. For example, nematode worms like *Caenorhabditis elegans* posses 20,000 genes, but lack the full range of cell-type and tissue-specific genes found in the fruit fly *Drosophila*, which contains less than 14,000. Thus, equilibrated and coordinated patterns of gene expression are required for proper development of an organism. The complexity is growing when genetic information is addressed on its natural context, that is, the chromatin organization of genome. Furthermore, the understanding of sophisticated patterns of gene expression is challenged when we consider that the great majority of some genomes, like the human genome, are composed by non-coding sequences that in addition are frequently transcribed (Yazzan and Krebs, 2007).

Transcriptional regulation is generally affected by *cis*-acting DNA sequence that directs the assembly of the protein machinery responsible for the differential control of transcription. Alterations in these *cis*-acting regulatory sequences are thought to dictate morphological diversification and evolution of regulatory mechanisms (1, 19). Four general types of *cis*-acting elements control the activity of the RNA Polymerase II (RNA Pol II): a) basal promoters corresponding to DNA sequences located near the initiation site (Fig. 1), b) enhancers and silencers that contain combinatorial DNA binding sites associated to transcriptional activators and/or repressors, c) Locus control regions, that contributes to transcriptional activation in conjunction to chromatin remodeling and topological organization of chromatin fibers, and d) boundary elements or insulators with multiple functions, in particular, the one associated to chromatin domains definition (West and Fraser, 2005). All these regulatory elements exemplified the emerging evidence that suggest that organism complexity arise from progressively and more elaborated regulatory networks of gene expression (Arnosti 2003; Levine and Tjian 2003).

*Cis*-regulatory sequences represent sophisticated and complex genetic information of the eukaryotic genome and are responsible for the control of chromosome replication, condensation, pairing, segregation and regulation of gene expression among many others (Levine and Tjian 2003; Hahn 2004). As an example, a typically genetic locus in *Drosophila* contains several enhancer (along with others *cis*-regulatory sequences as insulators) scattered over an average distance of 10 kb of genomic DNA with transcribed DNA comprising only 2 to 3 kb (Levine and Tjian 2003; Hahn 2004). By contrast, the regulatory sequences of eukaryotes such as yeast are usually composed in a few hundred base pairs or even less, located immediately adjacent to the core promoter and the transcription start site (28). Additionally the yeast genome encodes around ~300 transcriptional factors, in contrast the genome of *C. elegans* and *Drosophila* reveals the need of more than 1,000 transcriptional factors (Levine and Tjian 2003; Hahn 2004). In human cells, around 3,000 transcription factors have been suggested to be required for normal gene expression. Yeast contains an average of one transcriptional factor per 20 genes while human appear to contain one factor for every 10 genes (Levine and Tjian 2003; Hahn 2004). Such diversity is probably responsible for the lack of general algorithms for the regulation of gene expression.

An unresolved question in eukaryotic transcriptional regulation is how transcription units are organized in the chromosome so that promoters are regulated by their cognate enhancers in the context of multiple and complex DNA-protein interactions, in the three dimensional universe of the cell nucleus. Although how such task occurs is poorly understood; *cis*-regulatory elements

such as enhancers, silencers and insulators are certainly contributing to the now so-called epigenome (Kato and Sasaki, 2005).

In the present review we would like to explore systematically the features and integrate the functional aspects of diver's regulatory elements with particular emphasis in their mechanisms of action in the context of their chromatin structure and nuclear organization and dynamics.

### **Basal promoter elements**

A basal or core promoter can be defined, as a ~100-200 bp DNA sequence surrounding the transcriptional initiation site responsible for the general transcriptional machinery assembly (Arnosti 2003). The core promoter encompasses the RNA synthesis start site, understood as the minimal stretch of contiguous DNA sequence sufficient to direct transcription initiation (but not in a regulated way) by the basal RNA Pol II transcriptional machinery (Fig. 1A; Blackwood and Kadonaga 1998; Jennifer et al. 2002; Lim et al. 2004).

Recently a renovate interest on the elements of the basal promoter promoter have emerged (Smale and Kadonaga, 2003). The basal promoters provide nucleation sites for binding by basal transcriptional machinery. There are several core promoter elements, including the TATA box, TFIIB recognition element (BRE), Initiator (Inr), and downstream promoter element (DPE) and more recently the motive ten element (MTE) (Fig. 1B). The TATA box was the first eukaryotic core promoter motif to be identified (12, 42, 43). The TATA box is an A+T-rich sequence motif that is located about 30 nucleotides (nt) upstream of the transcription start site (consensus TATAAA centered at -30) and is recognized by the TATA-binding protein (TBP) subunit of the TFIID general transcription factor (70). The BRE is located immediately upstream of the TATA box of some TATA-containing promoters, and it increase the affinity of TFIIB for the core promoter (73). The Inr is a conserved sequence, encompassing the transcription start site (at +1, consensus TCAGT) that acts to direct accurate transcription initiation either alone or in conjunction with a TATA or DPE motif (7, 8, 44). The initiator regions contact some TAF's (TBP-associated factors) sub-components of the TFIID and the RNAPolII (71). The DPE is located about 30 nucleotides downstream of the start site (at +30 consensus A/GGA/TC/TGT) which is also recognized by distinct sets of TAFs and has been suggested to also interact with TFIID (1, 8). Motif 10 was initially identified as a downstream motif in a computational analysis of *Drosophila* core promoters (6, 45). The MTE motif ten elements is a distinct and novel downstream core promoter motif that functions in conjunction with the Inr element to promote transcription and acts independently of the DPE and TATA motifs, but can also work synergistically with the DPE or the TATA-box. In addition, the MTE appears to be conserved from *Drosophila* to human (6). Not all basal promoters contain all these elements but combinations of them seem to be relevant for the so-called promoter architecture that contributes to the pre-initiation complex formation but most importantly, to guide specific long range interactions with distal regulatory elements.

In *Drosophila*, core promoters fall into two distinct classes (7). Roughly, half contain a TATA-box located 25 to 30 bp upstream of the transcription start site with the presence of an initiator (Inr) element overlapping the +1 start site. For instance, it was estimated that approximately 43% of a total of 205 core promoters analyzed in *Drosophila* contain TATA-box (12). The other half contain an Inr element combined with a downstream promoter element (DPE), which is located approximately 30 pb downstream of the start site (7). In mammals, core

promoter structure appears to be even more diverse. Precise calculations remain difficult since transcription start site have been determined accurately for only a relative small fraction of genes. Nevertheless, the available data suggest that a smaller percentage of mammalian promoters like *Drosophila* promoters contain-TATA boxes (7). Indeed, in human it was found that about 32% of 1031 analyzed promoters regions contain a putative TATA-box motif (12, 50). In addition, TATA-boxes are paired with Inr elements in a smaller percentage of mammalian promoters (7). Furthermore, DPE elements exist in mammalian promoters, but have been difficult to identify, and (4) many promoters, including a number of promoters within CpG island, appears to lack all three of these core elements (7).

### ***CpG-islands and core promoter***

Genes can be divided by their expression patterns in two groups, tissue-specific and ubiquitous. Promoters with CpG-island have been associated with housekeeping genes, tumor suppressor genes and more recently microRNA promoters (14; Lujambo and Estelleer, 2007). CpG-islands are relatively GC-rich and overrepresented in CpG dinucleotides that are mostly unmethylated (12; Recillas-Targa et al. 2006). CpG-islands generally range in size from 0.5 to 2 kb they typically lack TATA or DPE core promoter elements, but contain multiple GC dinucleotides that are frequently bound by Sp1 and related transcription factors (12). A paradox emerged when we consider that a CpG-island, enriched in CpG-dinucleotides, their sequences should never be methylated (Recillas-Targa et al. 2006; De La Rosa-Velázquez et al. 2007). Our group and others have shown that abnormal DNA methylation of those CpG-islands may cause epigenetic silencing (De La Rosa-Velázquez et al. 2007; Filippova 2007). We believe that CpG-island may have mechanisms to protect them against DNA methylation. We have proposed that CTCF nuclear factor can contribute to such protection creating a methylation boundary against abnormal DNA methylation (ref). The analysis of 1031 human genes revealed that about half of the potential promoter regions are located in CpG islands (12, 50). Thus, this is a clear example in which genetic information is intimately associated to epigenetic events that are part of the multiple levels involved in the regulated gene activation or repression.

### **Promoter proximal element attributes**

Core promoter components are apparently not sufficient for regulated transcription activation. Immediately upstream of the core promoter from about -50 to -200 bp relative to the transcription start site (known as the proximal promoter element), multiple recognition sites for a subgroup of sequences-specific and ubiquitous DNA-binding transcription factors can be found. Collectively we will refer to the core promoter and the promoter proximal sequences, as the promoter (5). As mentioned the promoter proximal sequences starts to confer a molecular platform with specific properties to the core promoter, such as interactions with enhancer and/or long distance regulatory elements (6). The molecular bridge between the proximal and the core promoter elements seem to be mainly due by physical contacts between general and specific factor. Those interactions occur mainly with components of the TFIID (in particular TAF's) and the mediator components. In some cases basal promoter region are bound by sequence-specific transcriptional factors that play important roles in the regulation of individual genes (Hahn 2004; Baek et al. 2006). That gives a mechanism for cell-type-specific transcription regulation of gene expression. (Fig. 2).

Activator proteins are presumed to target one or more components of the transcription machinery, and two main targets such as: the TAF hypothesis was challenged by genetics studies in yeast, showing that TAF mutations do not affect transcriptional activation of a particular subset of yeast genes (25). The mediator hypothesis, on the other hand, has been supported by genetic studies; 13 of the 20 subunits of isolated mediator are products of genes identified in screens for mutations affecting transcriptional regulation in yeast (25).

The mediator associates exclusively with transcriptionally active and not inactive GAL1 genes. This association maps to the upstream activating sequence, rather than the core promoter, and is independent of RNA pol II general transcriptional factors, and promoter sequences. These findings support the idea of the mediator acting as a primary conduit of regulatory information from enhancers to promoters in eukaryotic cell (25). The mediator interacts with GAL1 promoter more rapidly than the poll, TBP, TFIIE or TFIIH following the induction of transcription (25).

The core promoter is the ultimate target of the vast network of regulatory factors that contribute to the initiation of transcription by RNA pol II (6, 12). The functional role of the basal promoter appears to be context dependent (1). Meanwhile in yeast a typical transcriptional unit, usually depends on DNA sequences located in proximity and immediately 5' of the transcription start site. Only a few genes in yeast genome, such as HO, contain distal regulatory sequences (3, 46). In contrast, animal cell-type-specific transcription of genes is often regulated by enhancer (or locus control regions (LCR)) sequences that act over large distances or even in *trans* upon promoters on a different chromosome (2, 39). Thus, the complexity increases proportional to the specificity of expression of the gene or group of genes to be regulated.

## Promoter specificity

Why does core promoter diversity exist? One possible explanation is that during the evolution of complex genomes, an interaction site for at least one component of the general transcriptional machinery may have been required for each promoter, but the recognition site(s) that evolved in a given promoter may have had no special significance other than to facilitate the binding and proper positioning of the general machinery (12). Early evidence supports the hypothesis that core promoter elements can be functionally interchangeable, for example some TATA elements can be substituted by DPE (7, 12). A second hypothesis that attempt to explain the existence of diverse core promoters is that they may be recognized by different protein complexes, each of them contributing to the transcription of a relatively small set of genes (7). A third scenario is now emerging with the participation of nuclear organization and the formation of transcriptional foci contributing to the co-expression of many genes (see below).

There are three major strategies for regulating the binding and function of the RNA Pol II complex at the core promoter that are not mutually exclusive (3):

1. *Diversification of core promoter transcription complexes.* It is now apparent that metazoans have evolved multiple related TFIID complexes that can function at distinct promoters through the use of tissue-specific TAF's. For example, in yeast there is only one TBP gene and no TRF's, in contrast in *Drosophila* there have been described to date four TRF's (TRF, 72) in addition to TBP and tissue specific TAFs (40; Chen et al. 2005).
2. *Transcription co-factor complexes.* The yeast mediator is a multi-subunit co-activator complex that is thought to facilitate the binding and/or function of RNA Pol II at the core promoter (47, 48). While yeast has one such co-factor complex, metazoans contains



several related complexes: TRAP, CRSP, ARC/DRIP, SMCC and hMcd (49). In the case of human mediator components such as ARC and CRSP, the smaller CRSP complex can augment transcription *in vitro*, while the larger ARC complex might be involved in repression (53). Thus, inducing an elevated degree of plasticity in terms of possible regulatory responses.

3. *Chromatin remodeling and modifying complexes.* The chromatin structure is at the centre of transcriptional regulation and needs to be remodeled in a highly coordinated way. The remodeler complexes that either displace or induce nucleosomes conformational changes (for example, SWI/SNF, Bcf/Brm, ACT and NURF) and covalent histone acetylation (for example, the histone acetyltransferases (CBP/p300/pCAF, GCN5) represent another potential source of regulatory diversification during metazoan evolution (61).

Thus, core promoter appears not only to mediate the initiation of the transcriptional machinery. It also mediates core promoter specificity for gene expression via diversification of core promoter transcription factors, transcription co-factor complexes and chromatin remodeling activities responsible for enhancer promoter-specific interactions. The most intriguing hypothesis to explain the variable distribution of sequence promoter element is that diversity makes an important contribution to combinatorial gene regulation. Selective communication between transcriptional factors bound to distal sites and core promoters could be of considerable benefit to such combinatorial strategies (Smale, 2001; Smale and Kadonaga, 2003; Lim et al. 2004).

### ***Short range enhancer promoter communication***

One additional implication is the participation of distant control elements on gene expression that does not exist solely for the purpose of altering the chromatin structure. Rather, the core promoter selectively and specific can be guided by enhancer-bound factors that must communicate directly or indirectly with transcriptional factors bound to the core promoter in particular combinatorial conformations (8). A clear example is provided by the analysis of TAF recruitment by upstream activating sequences (UAS in *S.cerevisiae*) has revealed a mechanism for enhancer-promoter specificity based on the differential interactions between TAF's and enhancer components (12, 51).

Alternatively, in *S.cerevisiae* it has been demonstrated a differential usage of two TATA-motifs that has also been observed with mammalian enhancers and activators (12, 51; Jennifer et al. 2002; Li et al. 2002). For instance, the analysis of the *trans*-activation of the human *hsp70* promoter by E1A, showed that substitution of the *hsp70* TATA (TATAA) with the SV40 early promoter TATA (TATTTAT) resulted in a loss of E1A inducibility (12; Jennifer et al. 2002). Additional diversity in gene regulation is achieved by the use of multiple promoters for a single gene, for example, the differential regulation of the *Hunchback* gene (52; Schroder et al. 1988).

It is easy to imagine but difficult to support (or explain) a promoter-proximal “code”, where one combination of regulatory factors recruits some enhancers, while another combination recruits some others enhancers (29; Keller et al. 1990). We cannot discard that the regulatory DNA's can be recruited closer to the “wrong” promoter, than the proper one (29). But, there are at least four three mechanisms for ensuring that the right enhancer interacts with the right promoter: insulator elements with their enhancer blocking property, gene promoter competition, promoter tethering elements for recruiting distal enhancers and relocation inside of the nucleus to active and specific transcriptional factories. Thus, a more recent model implicates re-localization

processes inside of the nucleus as a mechanism of action for enhancer activity. Some evidences point out to the responsibility of an enhancer to separate a genomic domain from an heterochromatin milieu, for example, away from Centromeric heterochromatin or the nuclear periphery, rendering the gene or genes to be regulated transcriptionally active (Schübeler et al. 2000; Taddei et al. 2006).

As previously mentioned it is generally accepted that basal promoters provide nucleation sites for the transcriptional machinery (6; Hahn 2004; Lim et al. 2004). But is there enough promoter sequence diversity to explain cell-type-specific gene transcription of higher eukaryotes? On one side, it is likely that there are more core promoter elements waiting to be discovered and on the other, the answer to this question is no, since several other regulatory elements are needed, like enhancers and LCRs. Thus, the functional role of the basal promoters appears to be context dependent.

Intriguingly, with some frequency, the proximal and core promoter sequences are not sufficient to reach coordinated gene expression. Additional elements, frequently acting over long distances (from the linear DNA point of view) to complement the specificity needed to reach time- and tissue-specific gene expression regulation. One of those elements, not yet clearly understood in terms of their mechanisms of action is the enhancers.

### **Promoter-enhancer selectivity**

One of the recurrent questions is how promoter-enhancer contacts are made in an specific manner? Core promoters not only direct the initiation of transcription, they also participate in the specificity recognition and association with enhancer to transduce specific regulatory signals (12). But, which is the basis of the promoter-enhancer specificity if some TATA elements can substitute DPE. Only a small number of enhancers can discriminate between identical basal promoters containing either TATA or DPE (1, 12; Arnosti 2003; Jennifer et al. 2002). In *Drosophila* the AE1 and IAB5 enhancers can function with a TATA-less core promoter, but exhibit a strong preference for a TATA-containing promoter relative to TATA-less promoter on a competitive situation (Fig. !; 12, 47; Jennifer et al. 2002; Ohtsuki et al. 1998). The specificity of an enhancer may depend on specific topological constrains. This is in part the reason why enhancers are located hundred of kilo bases away from its target promoter, because a particular tri-dimensional structure must be needed to reach specific conformation required for specific gene expression. Furthermore, the auto-regulatory element 1 (AE1) enhancer in *Drosophila* provides another example of preferential interactions between an enhancer and a core promoter. In its natural context, this enhancer is equidistant from both the *sexcombs reduced* (*Scr*) and *Fushi tarazu* (*ftz*) promoters, but preferentially activates *ftz* gene expression. The *Scr* and *ftz* genes differ in their core promoters elements. The *ftz* contains a TATA box, whereas the *Scr* promoter lacks a TATA box, but instead it contains an initiator and DPE sequence elements. In synthetic test constructions, the AE1 enhancer can activate transcription from a TATA-less promoter in the absence of a competing TATA-promoter. Thus, in the presence of both TATA-containing and TATA-less promoters, the AE1 enhancer preferentially activates transcription from the TATA-containing promoter (5). This example clearly illustrates the relevance of core promoter architecture and its composition to render the enhancer action highly selective.

The hypothesis rose by Butler and Kadonaga (2001) stated that core promoter diversity contributes to the selective communication between promoters and enhancers (8). They

determined the enhancer promoter specificity in *Drosophila* and found that 14 out of 18 enhancers tested did not show enhancer specificity for DPE or TATA motif. Thus, core promoter-specificity appears to be a particular property of a selected group of transcriptional enhancers (8). One potential obstacle for their hypothesis is that only a small subset of enhancers, exhibit core promoter selectivity. An alternative explanation for the 14 promoters that did not show enhancer specificity is because their endogenous context must be driving regulatory signals, to complement the enhancer-promoter specificity needed for proper gene regulation.

### **Promoter-Enhancers connection**

Enhancers were originally discovered as regulatory elements on viral genomes. Today enhancers are defined as DNA sequences that increased expression of a linked gene in a manner that is independent of their orientation and distance relative to the transcription start site (1, 5, 13). In addition, they present a modular organization in terms of their combinatorial DNA-binding motifs for transcription factors (39). A typical enhancer varies from 50 bp to 1.5 kb in size (39). Tissue-specific enhancers are active over long distances going from a few base pairs to 100 kb or more (3, 9, 23). Enhancers usually activate nearby target genes, although there are examples where they ignore the most proximal promoters and interact with distantly linked genes, like the 3' enhancers of the *dpp* gene and the *T1* enhancer of *Scr* (36).

The relative independence of enhancer action regarding distance and orientation to the promoter presents a particular problem for enhancer regulatory specificity (21, 48). Enhancer specificity may be determined by competition among multiple promoters through which the most preferred promoter(s) preclude others (21, 22, 49). This is well illustrated by the developmental switch of the vertebrate and in particular the chicken  $\beta$ -globin genes in which two *cis*-linked  $\epsilon$  and  $\beta^A$  globin promoters compete for a shared  $\beta^A/\epsilon$  enhancer (ref). During the embryonic stage, the  $\epsilon$  and  $\beta^A$  gene promoters compete for the  $\beta^A/\epsilon$  enhancer reaching a preferential activation of the embryonic gene. In adult stages, binding of an adult-specific transcription factors near the basal promoter of the  $\beta^A$  gene augments the enhancer-promoter interaction, causes activation of the adult  $\beta^A$  gene and the concomitant shut-off of the embryonic  $\epsilon$  gene (21, 26). These results demonstrate a competition way of action for an enhancer in which tissue- and stage-specific factors interacting to individual promoters define the enhancer task.

Interestingly, this competition model does not explain all the regulatory possibilities. In fact, an alternative model has been proposed based on the autonomous silencing, in particular, in the case of multi-family group of genes like the globin genes (ref). Recent evidences demonstrated that Sox6 is required to silence the mouse epsilon globin gene in definitive erythropoiesis allowing the expression of adult genes (Yi et al. 2006). Results from our research groups shows that the chicken  $\alpha$ -globin 3' enhancer has the same probability to activate any of the three globin genes (the embryonic  $\pi$  gene and the adult  $\alpha^D$  and  $\alpha^A$ ) and that the epigenetic silencing of the embryonic gene allows the regulated action of the 3' enhancer over the adult genes (Guerrero et al. 2007 and unpublished observations). Thus, as in the previous example, an epigenetic based autonomous silencing may contribute to the enhancer decision to which promoter it should act.

Enhancer's can promote the assembly of productive and stable transcription complexes in a repressive chromatin environment (5). The result of these activities has direct consequences for a gene from an inactive to an active state (on/off effect) or to modulate transcriptional levels (rheostat effect) (1). The study of transcriptional activation by enhancers on a single-cell basis, has led to the proposal that enhancers can increase the probability that a gene will be transcribed

in any particular cell while not affecting the level of transcription in the cell in which the gene is active. This “on or off” behavior, contrasts with the uniform and progressive enhanced gene expression model. Such model suggests that enhancers might also quantitatively regulate transcription rates through a continuous spectrum, depending on the amount and nature of interacting factors with an intrinsic capacity to counteract a repressive chromatin structure environment (5, 8, 33).

### **Alternative models for enhancer activity**

One of the decisive problems for understanding gene regulation is to explain how specific sets of genes are selected for expression during cell growth differentiation, or in response to environmental cues (34). All the information required for the regulatory functions mediated by enhancers and promoters is encoded in the DNA sequences, but is also subjected to their epigenetic environment.

#### ***The enhanceosome model***

This model, proposes that the DNA sequences of the enhancer operate as a scaffold to form a unified nucleoprotein complex. The enhanceosome feature take into consideration a high degree of cooperation between enhancer-bound proteins, such that alterations in individual binding sites can have a drastic effect on enhancer output (31, 32, 33, 34). Thus specificity in gene transcription via combinatorial theory of expression would arise from assembly of a unique enhancer-transcription factor complex, termed “enhanceosome” (Merika and Thanos 2001). The virus-inducible enhancer of the interferon- $\beta$  (IFN- $\beta$ ) gene provides one of the best examples of how combinatorial interactions between transcription factors can lead to a highly specific gene expression program. The gene is normally silent but it is induced at very high levels upon viral infection (34; Merika and Thanos 2001). The enhancer contains three overlapping positive regulatory domains recognized by NF- $\kappa$ B, members of the IRF family, and the ATF-2/c-Jun heterodimer nuclear factors. None of these regulatory domains acts on their own but two or more copies of any one of them can contribute to the virus-inducible enhancer activity (34). Single point mutations that modify or remove the binding sites of individual proteins disable this enhancer, suggesting that an overall protein-DNA their sequential and ordered assembly leading to specific superstructure is critical (55). Enhanceosomes assemble requires precise spacing between binding sites ensuring that each of the enhanceosome constituents assemble properly, allowing specific DNA recognition. Significantly, the precise positioning of the enhanceosome as a unit with respect to the core promoter seems not to be critical, as the enhancer can work from various positions, even when placed thousand of bases away from the promoter (34).

#### ***The billboard model***

In the second model, the enhancer acts as a “billboard”, in which the exact position of the binding sites is not strictly necessary. In this model the interactions of the transcriptional factors will be consecutive (33). The billboard model of enhancer suggests that there are multiple configurations in which enhancer-bound factors can interact with the basal transcriptional machinery (33). In contrast the billboard model suggests that binding sites are flexibly disposed, because the proteins bound to the enhancer do not operate as a single unit, with a strictly defined overall architecture, but rather as a set of separately acting factors or small groups of factors that independently interact with their target, as for example the basal transcriptional machinery (33). In agreement with this model, the chicken  $\alpha$ -globin enhancer presents a 120 bp DNA fragment that may be an

integral component of the enhancer, not presenting by itself any enhancer activity, but in the full enhancer context it is able to modulate the enhancer's activity, depending on the erythroid stage, probably based on its combinatorial of transcription factor interactions and chromatin conformation (Rincón-Arano et al. 2005; Escamilla-Del-Arenal and Recillas-Targa, 2007).

Supporting the billboard model, the *in vivo* footprinting demonstrated extensive TRC  $\alpha$  enhancer ( $E\alpha$ ) occupancy, not only in T cells representing a developmental stage where  $E\alpha$  enhancer is known to be active, but surprisingly, also in cells at an early developmental stage where  $E\alpha$  is not active (17). Latter, alterations in the structure of the pre-assembled nucleoprotein complex correlate with the onset of  $E\alpha$  activity, exemplifying one mechanism by which enhancers can rapidly response to incoming stimuli (17).

Considering the “looping” and “tracking” models there are recent evidences that connect both models in the so-called “facilitated tracking” model in which the combinatorial association of transcription factors and co-factors track along chromatin until the target promoter is encountered facilitating at the same time the long distance interactions needed for chromosomal loop formation (Dean 2006; Wang et al. 2005; Demers et al. 2007).

### **Enhancers and chromatin remodeling**

At this point it is important to recall that the enhanceosome should work in the presence of a chromatinized templates that for some instances are situated at the level of the solenoid. Such environment needs to be, in a controlled way, remodeled to allow recognition of target sequences. The enhanceosome exposes a unique activating surface, which can initiate with remarkable precision and efficiency a program of sequential recruitment events that counteracts the repressive chromatin environment, leading to efficient downloading of the basal transcriptional machinery (34). The complementary or maybe primary molecular function of the enhanceosome is to modify nucleosomal organization in a way that allows recruitment of the basal transcriptional apparatus. Following its assembly, the IFN- $\beta$  enhanceosome displays a second layer of stereospecificity in which the activation domains create a surface that initiates a program of recruitment events aimed at nucleosome modifications (56). First, the GCN5 complex is recruited acetylating the nucleosomes and HMG-I(Y), then HMG I(Y) acetylation further stabilizes the enhanceosome, thereby promoting the ability of the enhanceosome to recruit the SWI/SNF chromatin-remodeling complex (56).

Studies of enhancer's way of action demonstrated that chromatin remodeling evaluated by nuclease hypersensitivity formation is primarily determined by tissue-specific nuclear factors and their associated co-factors (responsible to recruit chromatin remodeling activities). Such activity should facilitate the subsequent incorporation of complementary specific and general transcription factors associated to the enhancer activity (Boyes and Felsenfeld, 1996; Escamilla-Del-Arenal and Recillas-Targa, 2007). Consequently, the chromatin structure associated to the enhancer represents an additional component of the multiple layers that conforms an enhancer activity.

In summary, a prevailing view of enhancer action is that they ensure cooperatively and synergism between *trans*-activators by organizing them into a tightly clustered nucleoprotein complex, know as enhanceosomes (17). The billboard enhancer model appears to describe more accurately many developmentally active enhancers, whose specific target binding sites, are subjected to exact position of binding to rapid change in evolutionary time, even as the overall output remains constant (33). The enhanceosome, is proposed to give a rapidly, inducible transcriptional response, such as inflammation or infection response may through a pre-

assembled composition what needs to be recalled is that all those processes occur in a tri-dimensional context of the cell nucleus and novel observations supports a dynamic contribution of the nucleus to transcriptional regulation. On the other side, the billboard model renders the enhancer activation more flexible.

### **Three-dimensional model for long distance regulation**

The enhanceosome and the billboard models allow to understand the specificity of the interactions needed for gene activation but does not explain how long distance contacts occur. Novel experimental data supports the physical contact model between distal regulatory elements (9). The contact model suggests that initial communication between an enhancer and its regulatory targets leads to formation of a stable DNA-protein complex in the vicinity and /or over a promoter element based on chromatin loop formation (Fraser and Bickmore, 2007; Kim et al. 2007). This association may facilitate subsequent rounds of transcription serving as a “memory” of the initial enhancer-promoter interaction (16). Alternatively, the average distance between promoter and enhancer could be considerably decreased if the intervening DNA is super-coiled or bent in the context of the higher-order chromatin structure (74). Studitsky and collaborators demonstrated that DNA super-coiling greatly facilitates enhancer-promoter communication, probably by bringing enhancer and promoter into close proximity. In contrast, when the enhancer-promoter spacing is shorter (110 bp), the enhancer works equally well on relaxed and supercoiled DNA. Thus, DNA supercoiling seems to be the critical factor for enhancer action over a large distance that is frequently under estimated (16).

### ***Classic promoter-enhancer contact models***

The looping model, which is based on extensive studies of well-defined systems, assumes direct physical contacts, based on their interacting domains and relative affinities between protein-protein and protein-DNA complexes formed on distant enhancers and proximal promoters (40, 57). This contact may be generated by random collisions that loop out the intervening DNA and are stabilized by productive protein-protein interactions (40, 58; Recillas-Targa et al. 2006; Filippova 2007). The random collisions model is not necessary in agreement with the loop model when the chromosomal loop is formed, in a regulated manner, through the action of MAR sequences or insulator elements (West et al. 2004; Capelson and Corces, 2004).

The scanning or tracking model, assumes that regulatory proteins (transcription factors, co-activators, or RNA Pol II) are initially recruited to the enhancer and then slide along the DNA until they encounter their cognate promoter. Such a model requires the association of chromatin remodeling activities to allow transcription progression through the action of ATP-dependent remodeling complexes like FACT and the posterior proper reassembly and positioning of nucleosomes after RNA Pol II passage (Reinberg and Sims 2006). As previously mentioned, it was recently proposed the facilitated tracking hypothesis (39). This model assumes that the enhancer-bound complex tracks via small steps, probably through a sort of processing chromatin remodeling along the chromatin until it encounters the cognate promoter, at which a stable looped structure is formed. Important components of the tracking complex are histone acetyltransferases CBP and P/CAF. These proteins may modify the chromatin as they move along the DNA (40).

An alternative model, based on genetic evidences is the linking model in which enhancer-promoter communication involves facilitator proteins, which are recruited by enhancer binding factors and then propagated along the DNA or not, thus organizing the chromatin structure of the intervening sequences in a way that allows the transmission of the enhancer signals to the promoter (40, 59). The major concern regarding both the scanning and the linking models is that

they assume no physical interactions between the enhancer and the promoter, a feature that has been supported by a number genetic and more recently, physical contacts (60; Ling and Hoffman 2007; Fraser and Bickmore 2007). But there also exist the possibility that one or more models could co-exist.

These models seem to be supported by the presence of an additional component: the intergenic non-coding transcripts (Miles et al. 2007; see below). Interestingly, a large portion of the vertebrate non-coding sequences are transcribed generating a large family of diverse RNAs, some of them involved in distinct regulatory functions such as X-chromosome inactivation, genomic imprinting, heterochromatin formation, RNA interference and microRNAs among others (Willingham and Gingeras, 2006; Wiemer 2007). In particular, long non-coding RNAs have been implicated in the positive and negative gene expression regulation on a growing list of genomic loci. *Drosophila* bithorax complex and the human  $\beta$ -globin loci have been studied in more detail (Schmitt et al. 2005; Gribnau et al. 2000; Miles et al. 2007). In the bithorax domain it has been postulated that intergenic transcripts are responsible for the activation and cellular memory of the locus (Schmitt et al. 2005). In the human  $\beta$ -globin domain intergenic transcripts define, in a regulated manner, sub-domains associated to the stage of globin gene expression and active chromatin modifications (Miles et al. 2007). Thus, intergenic transcription can be viewed as a more frequent event than previously anticipated that needs to be considered as one of the mechanisms, in close association with control elements, involved in the regulation of gene expression and necessary to maintain an open chromatin conformation in eukaryotes.

### **Enhancers and nuclear relocalization**

Alteration of the position of a gene relative to heterochromatin can affect its expression status (35). Recent studies have revealed that gene silencing by heterochromatin association can also involve the movement of genomic loci within the nucleus to a repressive nuclear compartment close to pericentromeric heterochromatin (Schübeler et al. 2000). In yeast, Sir-dependent telomeric and mating-type locus silencing are also associated with compartmentalization of silenced genes in the nuclear periphery, where Sir proteins are concentrated (35; Akhtar and Gasser, 2007).

Functional compartmentalization of interphase nuclei supports the association of transcription active genes to the limits of chromosomal territories, in close contact with transcription “factories” (Misteli 2007; Osborne et al. 2004; Ling and Hoffman 2007; Fraser and Bickmore 2007). In contrast, inactive genes are located in proximity or associated to heterochromatin domains like nuclear periphery or pericentromeric and centromeric regions (Fisher and Merckenschlager 2002). At the present time there are still some discussions concerning the long range directional movements inside the nucleus and relative to a chromosome territory (Misteli, 2007). What seems to be clear is that the mechanisms associated to those re-localization processes remain poorly understood. On the other side, recent evidences confirms interphase chromosome movements and that those displacements are not only dependent on protein-protein interactions as assumed until recently, but that they depend directly or indirectly on actin/myosin chains (Chuang et al. 2006). This discovery opens new questions and incorporates an additional component to the nuclear dynamics like the polarity and directionality of chromosome movements.

At the present time, it is generally accepted that in interphase nuclei, chromosomes are decondensed and occupy distinct and discrete chromosome territories (Cremer and Cremer, 2001; Lanctot et al. 2007). Moreover, it has become evident that nuclear architecture and

dynamics are important components associated to the regulation of gene expression. With the development and improvement of two complementary experimental methodologies, the fluorescence *in situ* hybridization and the chromosome conformation capture (3C) technique we are on a better position to confirm that nuclear organization and long-range interactions are part of the regulation of gene expression in eukaryotes (Dekker et al. 2002; Fraser and Bickmore 2007). RNA- and DNA-FISH and immunofluorescence assays allowed the *in situ* determination of chromosome territories separated by interchromosomal compartments and associate such a genome distribution inside of the nucleus with transcription, processing and localization of regulatory and processing factors. Co-localization experiments have been very instructive to determine for example genomic domains associated to sites enriched on heterochromatin, like some portion of the nuclear envelope, or in contrast the one associated to transcription factories where active co-regulated transcription and replication occurs (Chakalova et al. 2005).

The development of the 3C technology allowed the determination of the physical interaction between two or more distal genomic loci supporting the interchromosomal loops formation between regulatory elements (Carter et al. 2002; Tolhuis et al. 2002; Murrell et al. 2004). Two of the most outstanding examples are the demonstration of such interactions in the  $\beta$ -globin and the *Igf2/H19* loci. In particular, in the murine  $\beta$ -globin locus it has been demonstrated that the  $\beta$ -globin LCR, the far upstream DNase I hypersensitive sites (-62.5 and -60.7 kb HSS) and the 3'HS are together forming physical contacts in proximity to the transcriptional active genes *in vivo* through the formation of intrachromosomal loops (Tolhuis et al. 2002). In other words, the majority, if not all, the hypersensitive sites of the loci, spread over more than 100 kilo bases physically interact to form a cluster in a determined nuclear space, called the active chromatin hub (deLaat and Grosveld 2003). Furthermore, the proper spatial organization allowing loops formation requires the presence of protein factors with the appropriate timing of incorporation and affinities for each developmental and erythroid differentiation stages (Palstra et al. 2003). Not worthy, part of the mechanism of loop formation has been shown to be dependent on the 11-zinc finger CCCTCC-binding factor (CTCF) who has been involved in boundary functions, enhancer blocking activity and genomic topology (Splinter et al. 2006; Recillas-Targa et al. 2006).

Concerning the imprinted *Igf2/H19* domain, Reik and collaborators demonstrated that the mouse paternal allele, the *Igf2* differentially methylated region-2 (DMR-2), to which CTCF binds, loops out to interact with distant differentially methylated domain (DMD) of the imprinted *Igf2/H19* domain, thereby favoring the long distance intrachromosomal approximation between the *Igf2* promoter and the downstream *H19* enhancer and consequently, paternal allele expression of *Igf2* gene (Murrell et al. 2004). In contrast, in the maternal allele, an exclusion loop is formed between the DMR-1 and the DMD avoiding *Igf2* expression and thus, favoring maternal expression of *H19*.

Interestingly, in addition to intrachromosomal interactions there are now several evidences demonstrating that interchromosomal contacts are possible (Ling et al. 2006; Spilianakis et al. 2005; Simonis et al. 2006; Zhao et al. 2006). One more time CTCF is involved since its association to the differential methylation region of the *Igf2* and *H19* on mouse chromosome 7 and the paternal DMR on chromosome 11 between *Wsb1* and *Nf1* denoted a regulated interchromosomal contacts between two distal and separated chromosome imprinted loci (Ling et al. 2006). Another illustrative example is the interchromosomal interactions between the promoter element of the murine IFN- $\gamma$  gene on chromosome 10 and the developmentally regulated T<sub>H</sub>2 cytokine locus control region on chromosome 11 (Spilianakis et



al. 2005). More recently, the development of the 4C protocol allowing large-scale genome analysis of inter- and intrachromosomal associations suggest that this type of long-distance interactions may be a general phenomenon occurring at multiple genomic domains *in vivo*, inside the interphase nucleus, with a certain frequency on gene loci that require a differential and coordinated pattern of gene expression (Simonis et al. 2006; Zhao et al. 2006).

In conclusion, the predicted model of chromosome loop formation seems to be validated and partially confirms the enhancer-promoter communication through physical contacts. The 3C and 4C experimental approaches add sophistication to the long-distance regulation, incorporating inter- and intra-chromosomal association in conjunction with the chromosomal territories, interchromatinian compartments and transcriptional active foci as critical structures of the genetic and epigenetic regulation and nuclear dynamics.

### **Boundary elements or insulators**

Cell-type-specific transcription of genes is often regulated by enhancer sequences that act over large distances or even in *trans* upon promoters on different chromosome (5, 10). This long-range regulatory potential of enhancers raises the question of how their influence is restricted to their cognate promoter and whether or how their range of action can be defined hence their regulatory activity must be restricted in order to prevent unspecific action over non-target genes (11). This is achieved by specialized DNA sequences, termed chromatin boundary or insulators. Chromatin boundary elements or insulators are needed in complex genetic loci by organizing specialized chromatin structures (West et al. 2004; Gaszner and Felsenfeld 2006; Valenzuela and Kamakaka, 2006). Insulators are functionally defined by their ability to block the activity of an enhancer in a position-dependent-manner (2, 61) and/or by the capacity to protect a transgene against position effect (42, 62), and more recently experimental observations support their role in the topological organization of the genome (Dorman et al. 2007).

They are also considered as a novel transcriptional regulator (75, 76). Boundary elements or insulators are regulatory sequences that frequently contain cluster binding sites for large zinc fingers proteins, such as Su(Hw) in *Drosophila* and CTCF in vertebrates (3), establishing a barrier between regulatory elements on one side and basal promoters on the other (1). How can this occur without inactivating the enhancer or promoter ? (24). Boundary elements are thought to function as transcriptionally neutral elements that block, or insulate, the regulatory effects of positive (such as enhancers) or negative elements (such as silencers) (5). Boundaries also can be the result of a balance between countervailing chromatin opening and condensation processes that are nucleated at neighboring gene loci probably blocking the spreading of heterochromatin through an active recruitment of open chromatin marks (Gaszner and Felsenfeld 2007; 9). It has been hypothesized that insulators may interact with each other to form chromatin loop domains, restricting interactions among neighboring regulatory elements (24) (Gaszner et al. 2003).

### **Models of enhancer-blocking activity**

In the context of the present review insulator way of action could have direct consequences on enhancer activity. In particular, the enhancer-blocking property can contribute to the enhancer selectivity, by blocking undesired, enhancer-promoter communication. There is an example in which such regulation can be done in hormone-dependent manner (77). Even though, the mechanism of insulator action remains unknown; two models have been proposed to explain the enhancer-blocking activity of insulators. First, the “promoter decoy” model suggests that insulators neutralize enhancer action by capturing or titrating functional active components of the enhancer (10). Second, the “chromatin boundary” model proposes that insulators foment higher

order chromatin structure, by specific chromatin loops formation (65). Furthermore, and associated to the chromatin boundary model, Studitsky and collaborators have build an *in vitro* insulator system of study who clearly favor a model of DNA loop structure inducing the isolation of an enhancer (10). This proposal was later supported *in vivo* (11), by conditionally looping out an enhancer from his target promoter in an inducible manner, resulting in a DNA structure-dependent gene regulation (11). These results are fully consistent with a recently described chromatin-loop model for the imprinted expression of *Igf2* (66). In addition, the insulator proteins Zw5 and BEAF, which bind to the insulator element *scs* and *scs'* in *Drosophila*, interact with each other *in vitro*, and more recently, it has been demonstrated that Zw5 and BEAF are located in close proximity to each other *in vivo* favoring the formation of a well defined looped domain (11; Blanton et al. 2003).

### **Relationship between enhancers and insulators**

We just argued that enhancer-blocking can be an strategy to discriminate and regulate enhancer action. Opposing to such a model we propose that insulators, through the multimerization of nuclear factors like CTCF, can topologically bring together, to certain relative proximity, groups of genes to a particular transcription foci or to take advantage of one or several enhancers with the purpose of gene co-regulation (West and Fraser 2005; Chakalova et al. 2005). Such view is supported by inter- and intra-chromosomal interaction rendering much more efficient the transcriptional regulation of genes that are distal or even in distinct chromosomes (Fraser and Bickomre 2007, Ling and Hoffman 2007). Tethering of an enhancer blocker to foci could interfere with an enhancer's access to a gene promoter in two ways. First, an enhancer may be "miss-directed". Secondly, enhancers may be "obstructed" from making interactions with promoters beyond their own loop by tensional constraints or steric interference at the base of the loop, like in the imprinted *Igf2/H19* loci (Murrell et al. 2004) (9). However, in a particular case, enhancer-blocking activity is copy number and position sensitive. When two copies of the Su(Hw) insulator element, instead of a single one, are inserted between enhancer and promoter, insulator activity is neutralized and the enhancer-promoter interaction may instead be facilitated. This paradoxical phenomenon could be explained by interactions between proteins complexes bound at the insulator (27).

Some studies have shown that distinct *cis* elements in the core promoter such as a TATA, Inr and DPE, and their associated *trans*-factors determine the ability of the promoter to interact and compete for regulatory enhancers (47). Promoter competition, which depends on the core promoter sequences, is also important for enhancer-promoter specification, in other words, insulator function could be affected by the balance of promoter competition among neighboring genes (21). For example, in transgenic flies the *gypsy* insulator mediates blockage of the AE1 enhancer, but depending on its context, as shown by the complete blockage of AE1 from white, but not from eve or ftz promoters when opposed by *Scr* or the complete blockage of AE1 from eve, but not from ftz when opposed by white (21) This evidence provided a mechanism through which the effectiveness of an insulator is determined by regulatory composition of the neighboring genome (21). Therefore it is now our understanding that insulators, including those identified as boundary elements and those identified as enhancers blockers, are not impenetrable walls for transcriptional interactions, even more, they participate actively on transcription regulation (21; Recillas-Targa et al. 2006; Filippova 2007) .

## Insulator and epigenetics: CTCF connection

Enhancer blocking activities can now be considered as part of the transcriptional regulatory mechanisms, even more when such activity seems to be regulated (Lutz et al. 2003). One of the clearest examples in the binding of the vertebrate enhancer-blocking protein CTCF can be abrogated by CpG methylation (Recillas-Targa 2000; Recillas-Targa et al. 2006; Filippova 2007). More recently, CTCF-dependent enhancer-blocking has also been shown to be modulated by a number of proteins co-factors, in particular, through the action of the CHD8, a chromodomain helicase family member (Ishihara et al. 2006). Insulators or chromatin boundaries are frequently found in multiple copies, flanking enhancers or genetic loci they regulate, such as the *scs* and *scs'* elements, the *Msp-1* and *Fab* boundaries and the chicken  $\beta$ -globin 5' and 3' boundaries (24; Saitoh et al. 2000; Burgess-Beusse et al. 2002). This distribution incorporates a novel concept associated to insulators that is, their topological organization of the epigenome (Dorman et al. 2007; Fraser and Bickmore 2007). Supporting this vision, CTCF has the capacity to multimerize and contact the nuclear matrix associated protein nucleophosmin contributing to the insulator-CTCF mediated topological organization of the globin domain in the nucleolus periphery (Yusufzai et al. 2004). In summary, novel results support a model in which insulators might organize high order chromatin structures, participate in DNA loops formation, in addition to delimiting functional units of independent gene expression. Alternatively, it looks like insulators block the basal transcriptional possibility that a promiscuous promoter-enhancer interaction could happen. Then, we can build a model in which insulator elements appear to control the balance between competing promoters, helping to the establishment of stable transcriptional (or pre-transcriptional) domains that makes more efficient the gene regulation and expression processes.

CTCF turns also to be an important component in the determination of epigenetic control of diverse imprinted loci, partly in a DNA methylation-dependent manner, by its enhancer blocking and through long distance chromatin interactions (Murrell et al. 2004; Recillas-Targa et al. 2006; Ling and Hoffman 2007; Filippova 2007). Moreover, CTCF had been involved on gene activation and repression through its action over promoter regions (Klenova and Ohlsson 2005; Recillas-Targa et al. 2006; Filippova 2007). From our point of view, one of the most relevant aspects of CTCF is its participation, in a sort of methylation boundary promoter elements (frequently corresponding to CpG-islands) of the human tumour suppressor genes like *p19<sup>ARF</sup>*, *BRCA1* and *Rb* (Butcher et al. 2004; Filippova et al. 2002; De La Rosa-Velázquez et al. 2007). Interestingly, for human *BRCA2* gene promoter, CTCF defines a DNA methylation transition zone (Butcher et al. 2004). When sporadic breast tumours samples were analyzed, abnormal cytoplasmic localization of CTCF correlates with DNA hypermethylation of the *BRCA1* gene promoter and overexpression of the DNMT3b DNA methyltransferase (Butcher and Rodenhiser, 2007).

Results from our laboratory involve CTCF in the epigenetic regulation of the human retinoblastoma gene promoter. We have found that CTCF has a protective role against epigenetic silencing of the human *Rb* gene promoter. When such a promoter is DNA hypermethylated it rapidly acquires a repressive chromatin conformation, CTCF is then dissociated from its binding site and replaced by the methyl-CpG-binding protein Kaiso and its associated repressive complex (Fig. 3; De La Rosa-Velázquez et al. 2007). These results suggest that CTCF is inducing a local chromatin structure that contributes to the proper responsiveness of the promoter. Based on the *BRCA1* and *Rb* gene promoters we propose a model in which CTCF might have the capacity to prevent DNA methylation and/or other epigenetic repressive marks, and we speculate that this could be part of the mechanisms that maintain any CpG-island unmethylated (Fig. 3; Herman and

Baylin 2003; Recillas-Targa et al. 2006; Filippova 2007). This model does not seem to be so unlikely in regards to large scale survey of CTCF distribution along the human genome and based on CTCF protective role of the escape genes in the epigenetic silenced environment of the X inactive chromosome (Filippova et al. 2005; Mukhopadhyay et al. 2004; Barski et al. 2007).

A more recent high-resolution profiling of CTCF in the human genome identified in a primary human resting CD4<sup>+</sup> T cells, 20,262 CTCF *in vivo* target sites, out of which 8,308 are located in intergenic regions, 6,305 in transcribed regions and 5,642 within 2 kb around transcription start sites (Barski et al. 2007). These observations suggest that around one fourth of the CTCF-binding sites are found on promoter regions. Based on such results it turns to be very attractive to determine if those sites coincide with transition zones of CpG-island promoters and if the majority of them correspond to tumour suppressor gene promoters and possess a protective role against DNA methylation (Fig. 3; De La Rosa-Velázquez et al. 2007). If our model is correct, all these evidences incorporate a novel level of transcriptional regulation in which, nuclear factors like CTCF, are in charge of protecting CpG-rich promoters against aberrant DNA methylation and epigenetic silencing.

### **Insulators, enhancers, nuclear dynamics and non-coding RNA's**

Corces and collaborators using FISH with DNA probes located between two endogenous Su(Hw) binding sites, revealed that the intervening chromatin is arranged in a chromatin loop, with the two insulators located at the loop base (67). *In situ* insertion of an additional insulator in the centre of the loop results in the formation of two smaller loops, and mutations of the insulator-binding proteins lead to the disruption of the loop (67). Thus, there are convincing experimental evidences, in particular in *Drosophila*, supporting the topological role of insulators. Genome organization mediated by *gypsy* insulator comprise other components including the insulator protein CP190, SUMO and more recently the RNAi machinery, this latter proposed as modulator of nuclear architecture capable to induce changes in gene expression (Capelson and Corces 2006; Lei and Corces 2006). Complementing this vision, Cavalli and collaborators showed in *Drosophila* that the RNAi machinery contributes to the nuclear organization and long distance interactions between chromosomes through the action of the repressive Polycomb group (PcG) of proteins (Grimaud et al. 2006). Interestingly, after chromosome pairing is established, double-stranded RNA is generated and the small interfering RNA machinery is attracted inducing a highly stable chromosomal pairing structure and the maintenance of a silencing state of two distal located loci (Grimaud et al. 2006). In concordance with these observations, an outstanding investigation revealed the long distance silencing and chromosomal domain demarcation by non-coding RNA's (Rinn et al. 2007). A systematic survey of *HOX* gene clusters transcription in human fibroblasts from different body locations revealed a significant proportion of intergenic transcripts. Based on this strategy, the *HOX Antisense Intergenic RNA* (HOTAIR) transcript was defined (Rinn et al. 2007). Interestingly, HOTAIR, a long non-coding RNA is able to act in *trans* to silence Hox chromatin domains located on distinct chromosomes. HOTAIR guides the Polycomb PRC2-repressive complex as a mechanism of epigenetic silencing to regulate gene expression in *trans*.

Altogether, insulator and non-coding transcripts constitute complementary components associated to the enhancer, LCR and promoter regulatory elements. All these evidences are guiding us to the need to reach an integral view of all these processes to understand transcriptional regulation in terms of the hierarchy and interdependence of the components that now includes nuclear dynamics, RNAi machinery and intergenic transcripts.

## Conclusions and perspectives

The development of an organism from a fertilized single cell requires sophisticated orchestration of multiple networks of gene expression at the right time, place, and levels (14). *Cis*-acting DNA sequences have evolved to give access to more elaborated gene regulation control pathways and in this way play a driving role in morphological diversification. *Cis*-regulatory sequences produce complexity and diversification at the genetic and epigenetic levels. As an example, recent evidences indicates that stem cell chromatin shows a bivalent configuration allowing the acquisition of specific decisions contributing to activated or repressive programs of gene expression involved in proliferation and differentiation. In other words, a bivalent chromatin conformation on ES cells may allow rapid responses to genetic and epigenetic signs for early cell proliferation and differentiation (Bernstein et al. 2007).

In the analysis of gene expression, the core promoter should be considered critical component but not sufficient for complete regulation of gene expression. The requirement for time- and tissue-specific gene expression could be part of the explanation for core promoter architecture diversity and frequently the need of additional regulatory elements. In fact, core promoters are the ultimate target of the vast network of regulatory components (12). It remains unclear how an enhancer interacts with those factors. At the present time TAF's and Mediator components are considered a key elements participating in this highly specific and complex recognition. Therefore, TAF and Mediator components are considered as control panel dictating highly specific responses regulating promoter responsiveness specificity in concurrence with distal regulatory elements.

Transcriptional control regions often contain multiple, autonomous enhancer elements. Each of these enhancers appears to be designed to perform specific functions, such as activation of its cognate gene in a specific cell-type or at a particular developmental stage, in a manner that is independent of their orientation and distance relative to the transcription start site (5). Such regulatory strategies increment the complexity but at the same time the specificity for gene activation or repression (Orphanides and Reinberg, 2002). The main questions are at the level of how the activity of these regulatory elements and their associated multi-peptide complex elements are modulated, and how they work. The way in which they are organize (enhanceosome or billboard models) and the way in which they transmit their effect (contact –looping, linking, tracking or facilitated tracking- or no contact models) to the target promoter are still a subject of debate but apparently there is no unique or predominant model. Alternative models to explain the enhancer action such as promoter re-localization or DNA super-coiling has also been proposed. But what seems to be the main question is how promoter-enhancer specific interactions are achieved in the correct developmental time and cell-type. Because promoter-enhancer interactions could be under certain circumstances promiscuous, complementary components like insulators, promoter competence or even their own activity modulation may be required to restrict it sphere of actions.

This prompted us to propose a model in which the stronger enhancer-promoter or boundaries interaction must structure the local epigenome context to make possible regulated gene expression. In this model an enhancer or a promoter could be captured or looped out insulator related loop. Otherwise, an insulator or silencer could be captured by a stronger enhancer-promoter loop. At the present time non-coding RNAs, including RNAi machinery and intergenic transcripts need to be taken into consideration as regulatory components capable to act

in *cis* but also in *trans* (Rinn et al. 2007; Zhu et al. 2007). Based on more recent results we cannot discard the regulated participation of inter- and intrachromosomal interactions (Ling and Hoffman 2007; Fraser and Bickmore, 2007).

All these evidence confirms the relevance of the long distance interaction, re-localization of genomic sequences, inter- and intra-chromosomal contacts that directs us to the constant appreciation of the complexity of the regulation of differential gene expression. The interconnection between multiple epigenetic processes need also to be at the front of our considerations but today, we need to accept that the chromatin contribution to the epigenome organization and the dynamic three-dimensional organization of the nucleus contribute in *cis* and *trans* to the functional responsibilities of the cell with direct regulatory consequences. Thus, the biggest challenge in this area, is the integration of the different levels at which gene regulation is studied, with the future directions are in our understanding that the processes are taking place in the context of a nucleus, in which the chromatin architecture imposes a complex conformation in which gene expression through the development is regulated.

### **Acknowledgements**

We thank Mayra Furlan-Magaril and Paul Delgado-Olguín for helpful comments on the manuscript. This work was supported by grants from the Dirección General de Asuntos del Personal Académico-UNAM (IN203200, IX230104, IN209403 and IN214407), Consejo Nacional de Ciencia y Tecnología, CONACyT (33863-N, 42653-Q and 58767), and the Third World Academy of Sciences (TWAS, Grant 01-055 RG/BIO/LA). ME-D-A was a fellowship recipient from CONACyT.

### **References**

1. Arnosti DN. (2003) Analysis and function of transcriptional regulatory elements: insights from *Drosophila*. *Annu Rev Entomol.* 48: 579-602. Review.
2. Ameres SL, Druempel L, Pfeleiderer K, Schmidt A, Hillen W, Berens C. (2005) Inducible DNA-loop formation blocks transcriptional activation by an SV40 enhancer. *EMBO J.* 24(2):358-67.
3. Levine M, Tjian R. (2003) Transcription regulation and animal diversity. *Nature.* 424(6945): 147-51. Review.
5. Blackwood EM, Kadonaga JT. (1998) Going the distance: a current view of enhancer action. *Science.* 281(5373): 60-3. Review.
6. Lim CY, Santoso B, Boulay T, Dong E, Ohler U, Kadonaga JT. (2004) The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes Dev.* 18(13): 1606-17.
7. Smale ST. (2001) Core promoters: active contributors to combinatorial gene regulation. *Genes Dev.* 15(19): 2503-8. Review.

8. Butler JE, Kadonaga JT. (2001) Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs. *Genes Dev.* 15(19): 2515-9.
9. West AG, Fraser P (2005) Remote control of gene transcription. *Hum Mol Genet.* 14 Spec No 1: R101-11.
10. Bondarenko VA, Jiang YI, Studitsky VM. (2003) Rationally designed insulator-like elements can block enhancer action in vitro. *EMBO J.* 22(18): 4728-37.
11. Ameres SL, Druempel L, Pfliegerer K, Schmidt A, Hillen W, Berens C. (2005) Inducible DNA-loop formation blocks transcriptional activation by an SV40 enhancer. *EMBO J.* 24(2): 358-67.
12. Jennifer EF Butler and James T Kadonaga, (2002) The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes and Development.* 16: 2583-92.
13. Morris JR, Petrov DA, Lee AM, Wu CT. (2004) Enhancer choice in cis and in trans in *Drosophila melanogaster*: role of the promoter. *Genetics.* 167(4): 1739-47.
14. Schug J, Schuller WP, Kappen C, Salbaum JM, Bucan M, Stoeckert CJ Jr. (2005) Promoter features related to tissue specificity as measured by Shannon entropy. *Genome Biol.* 6(4): R33.
15. White RJ, Jackson SP (1992) The TATA-binding protein: a central role in transcription by RNA polymerases I, II and III. *Trends Genet.* 8(8): 284-8. Review.
16. Liu Y, Bondarenko V, Ninfa A, Studitsky VM. (2001) DNA supercoiling allows enhancer action over a large distance. *Proc Natl Acad Sci U S A.* 98(26): 14883-8.
17. Spicuglia S, Payet D, Tripathi RK, Rameil P, Verthuy C, Imbert J, Ferrier P, Hempel WM. (2000) TCRalpha enhancer activation occurs via a conformational change of a pre-assembled nucleo-protein complex. *EMBO J.* 19(9):2034-45.
18. Kato Y, Sasaki H. (2005) Imprinting and looping: epigenetic marks control interactions between regulatory elements. *Bioessays.* 27(1): 1-4. Review
19. Tautz D. (2000) Evolution of transcriptional regulation. *Curr Opin Genet Dev.* 10(5):575-9. Review.
20. Conte C, Dastugue B, Vaury C. (2002) Promoter competition as a mechanism of transcriptional interference mediated by retrotransposons. *EMBO J.* 21(14): 3908-16.
21. Cai HN, Zhang Z, Adams JR, Shen P. (2001) Genomic context modulates insulator activity through promoter competition. *Development.* 128(21): 4339-47.
22. Sharpe J, Nonchev S, Gould A, Whiting J, Krumlauf R. (1998) Selectivity, sharing and competitive interactions in the regulation of Hoxb genes. *EMBO J.* 17(6): 1788-98.

23. Krebs JE, Dunaway M. (1998) The scs and scs' insulator elements impart a cis requirement on enhancer-promoter interactions. *Mol Cell*, 1(2): 301-8.
24. Cai HN, Shen P. Effects of cis arrangement of chromatin insulators on enhancer-blocking activity. *Science*. 291(5503): 493-5.
25. Kuras L, Borggreffe T, Kornberg RD. (2003) Association of the Mediator complex with enhancers of active genes. *Proc Natl Acad Sci U S A*. 100(24): 13887-91.
26. Herendeen DR, Kassavetis GA, Geiduschek EP. (1992) A transcriptional enhancer whose function imposes a requirement that proteins track along DNA. *Science*. 256(5061): 1298-303.
27. Muravyova E, Golovnin A, Gracheva E, Parshikov A, Belenkaya T, Pirrotta V, Georgiev P. Loss of insulator activity by paired Su(Hw) chromatin insulators. *Science*. 291(5503): 495-8.
28. Mueller-Sturm HP, Sogo JM, Schaffner W. (1989) An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge. *Cell*. 58(4): 767-77. Erratum in: *Cell* 1989 Oct 20;59(2):405.
29. Keller RJ III, Flanagan PM and Kornberg RD. (1990) .*Cell*. 61, 1209-1215.
30. Henikoff S. (2000) Heterochromatin function in complex genomes. *Biochim. Biophys. Acta* 1470: 1-8
31. Rincon-Arano H, Valadez-Graham V, Guerrero G, Escamilla-Del-Arenal M, Recillas-Targa F. (2005) YY1 and GATA-1 Interaction Modulate the Chicken 3'-Side alpha-Globin Enhancer Activity. *J Mol Biol*. 349: 961-75.
32. Struhl K. (2001). Gene regulation. A paradigm for precision. *Science* 252: 138-87.
33. Arnosti DN, Kulkarni MM. (2005) Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? *J Cell Biochem*. 94(5): 890-8.
34. Merika M, Thanos D. (2001) Enhanceosomes. *Curr Opin Genet Dev*. 11(2):205-8. Review.
35. Francastel C, Walters MC, Groudine M, Martin DI. (1999) A functional enhancer suppresses silencing of a transgene and prevents its localization close to centromeric heterochromatin. *Cell*. 99(3): 259-69.
36. Calhoun VC, Levine M. (2003) Long-range enhancer-promoter interactions in the Scr-Antp interval of the *Drosophila* Antennapedia complex. *Proc Natl Acad Sci U S A*. 100(17): 9878-83.
37. Foley KP, Engel JD. (1992) Individual stage selector element mutations lead to reciprocal changes in beta- vs. epsilon-globin gene transcription: genetic confirmation of promoter competition during globin gene switching. *Genes Dev*. 6(5): 730-44.



38. Chen JL, Huisinga KL, Viering MM, Ou SA, Wu CT, Geyer PK. (2002) Enhancer action in trans is permitted throughout the *Drosophila* genome. *Proc Natl Acad Sci U S A*. 99(6): 3723-8.
39. Blackwood EM, Kadonaga JT. (1998) Going the distance: a current view of enhancer action. *Science*. 281(5373): 60-3. Review.
40. Hatzis P, Talianidis I. (2002) Dynamics of enhancer-promoter communication during differentiation-induced gene activation. *Mol Cell*. 10(6):1467-77.
41. Wedel A, Weiss DS, Popham D, Droge P, Kustu S. (1990) A bacterial enhancer functions to tether a transcriptional activator near a promoter. *Science*. 248(4954):486-90.
42. Recillas-Targa F, Valadez-Graham V, Farrell CM. Prospects and implications of using chromatin insulators in gene therapy and transgenesis. *Bioessays*. 2004 Jul;26(7):796-807. Review.
43. Breathnach R, Chambon P. (1981) Organization and expression of eucaryotic split genes coding for proteins. *Annu Rev Biochem*. 50:349-83. Review.
44. Smale ST, Baltimore D. (1989) The "initiator" as a transcription control element. *Cell*. 57(1):103-13.
45. Ohler U, Liao GC, Niemann H, Rubin GM. (2002) Computational analysis of core promoters in the *Drosophila* genome. *Genome Biol*. 3(12): RESEARCH0087.
47. Ohtsuki S, Levine M, Cai HN. (1998) Different core promoters possess distinct regulatory activities in the *Drosophila* embryo. *Genes Dev*. 12(4): 547-56.
48. Gindhart JG Jr, King AN, Kaufman TC. (1995) Characterization of the cis-regulatory region of the *Drosophila* homeotic gene *Sex combs reduced*. *Genetics*. 139(2): 781-95.
49. Walker SS, Shen WC, Reese JC, Apone LM, Green MR. Related (1997) Yeast TAF(II)145 required for transcription of G1/S cyclin genes and regulated by the cellular growth state. *Cell*. 90(4): 607-14.
50. Suzuki Y, Tsunoda T, Sese J, Taira H, Mizushima-Sugano J, Hata H, Ota T, Isogai T, Tanaka T, Nakamura Y, Suyama A, Sakaki Y, Morishita S, Okubo K, Sugano S. (2001) Identification and characterization of the potential promoter regions of 1031 kinds of human genes. *Genome Res*. 11(5): 677-84.
51. Li XY, Bhaumik SR, Zhu X, Li L, Shen WC, Dixit BL, Green MR. (2002) Selective recruitment of TAFs by yeast upstream activating sequences. Implications for eukaryotic promoter structure. *Curr Biol*. 12(14): 1240-4.

52. Schroder C, Tautz D, Seifert E, Jackle H. (1988) Differential regulation of the two transcripts from the *Drosophila* gap segmentation gene hunchback. *EMBO J.* 1988 Sep;7(9): 2881-7.
53. Freidman R, Hotaling E, Borack L, Sofer W. (1996) Interactions between the regulatory regions of two *Adh* alleles. *Genetica.* 97(1): 1-14.
54. Milot E, Strouboulis J, Trimborn T, Wijgerde M, de Boer E, Langeveld A, Tan-Un K, Vergeer W, Yannoutsos N, Grosveld F, Fraser P. (1996) Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell.* 87(1): 105-14.
55. Thanos D, Maniatis T. (1995) Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. *Cell.* 83(7): 1091-100.
56. Agalioti T, Lomvardas S, Parekh B, Yie J, Maniatis T, Thanos D. (2000) Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. *Cell.* 103(4): 667-78.
57. Ptashne M. (1986) Gene regulation by proteins acting nearby and at a distance. *Nature.* 322(6081): 697-701. Review.
58. Ringrose L, Chabanis S, Angrand PO, Woodroffe C, Stewart AF. (1999) Quantitative comparison of DNA looping in vitro and in vivo: chromatin increases effective DNA flexibility at short distances. *EMBO J.* 18(23): 6630-41.
59. Bulger M, Groudine M. (1999) Looping versus linking: toward a model for long-distance gene activation. *Genes Dev.* 13(19): 2465-77. Review.
60. Carter D, Chakalova L, Osborne CS, Dai YF, Fraser P. (2002) Long-range chromatin regulatory interactions in vivo. *Nat Genet.*; 32(4): 623-6.
61. Gerasimova TI, Corces VG. (2001) Chromatin insulators and boundaries: effects on transcription and nuclear organization. *Annu Rev Genet.* 35: 193-208. Review.
62. Recillas-Targa F, Bell AC, Felsenfeld G. (1999) Positional enhancer-blocking activity of the chicken beta-globin insulator in transiently transfected cells. *Proc Natl Acad Sci U S A.* 96(25): 14354-9.
63. Bell AC, West AG, Felsenfeld G. (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell.* 98(3):387-96.
64. West AG, Huang S, Gaszner M, Litt MD, Felsenfeld G. (2004) Recruitment of histone modifications by USF proteins at a vertebrate barrier element. *Mol Cell.* 16(3):453-63.
65. Capelson M, Corces VG. (2004) Boundary elements and nuclear organization. *Biol Cell.* 96(8):617-29. Review.

66. Murrell A, Heeson S, Reik W. (2004) Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat Genet.* 36(8): 889-93.
67. Byrd K, Corces VG. (2003) Visualization of chromatin domains created by the gypsy insulator of *Drosophila*. *J Cell Biol.* 162(4):565-74.
68. Valadez-Graham V, Razin SV, Recillas-Targa F. CTCF-dependent enhancer blockers at the upstream region of the chicken alpha-globin gene domain. *Nucleic Acids Res.* 2004 Feb 23;32(4):1354-62. Print 2004.
69. Yusufzai TM, Felsenfeld G. The 5'-HS4 chicken beta-globin insulator is a CTCF-dependent nuclear matrix-associated element. *Proc Natl Acad Sci U S A.* 2004 Jun 8;101(23):8620-4. Epub 2004 May 28.
70. Aoyagi N, Wassarman DA. 2000. Genes encoding *Drosophila melanogaster* RNA polymerase II general transcription factors: diversity in TFIIA and TFIID components contributes to gene-specific transcriptional regulation. *J. Cell Biol.* 150: F45– 50.
71. Chalkley GE, Verrijzer CP. 1999. DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. *EMBO J.* 18: 4835– 45.
72. Arnosti DN. 2002. Design and function of transcriptional switches in *Drosophila*. *Insect Biochem. Mol. Biol.* 32(10): 1257– 73.
73. Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D., and Ebright, R.H. 1998. New core promoter element in RNA polymerase II-dependent transcription: Sequence-specific DNA binding by transcription factor IIB. *Genes & Dev.* **12**: 34-44.
74. Rippe, K, von Hippel, PH and Langowski, J (1995) Action at a distance: DNA-looping and initiation of transcription. *Trends Biochem. Sci.* 20, 500-506.
75. West AG, Gaszner M, Felsenfeld G. Insulators: many functions, many mechanisms. *Genes Dev.* 2002 Feb 1;16(3):271-88.
76. Capelson M, Corces VG. The ubiquitin ligase dTopors directs the nuclear organization of a chromatin insulator. *Mol Cell.* 2005 Oct 7;20(1):105-16.
77. Lutz M, Burke LJ, LeFevre P, Myers FA, Thorne AW, Crane-Robinson C, Bonifer C, Filippova GN, Lobanenkova V, Renkawitz R. Thyroid hormone-regulated enhancer blocking: cooperation of CTCF and thyroid hormone receptor. *EMBO J.* 2003 Apr 1;22(7):1579-87.
78. Freidman R, Hotelling E, Borack L, Sofer W. Interactions between the regulatory regions of two *Adh* alleles. *Genetica.* 1996 Jan;97(1):1-14.

79. Milot E, Strouboulis J, Trimborn T, Wijgerde M, de Boer E, Langeveld A, Tan-Un K, Vergeer W, Yannoutsos N, Grosveld F, Fraser P. Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell*. 1996 Oct 4;87(1):105-14

- Filippova GN, Cheng MK, Moore JM, Truang J-P, Hu YT, Nguyen BK, Tsuchiya KD and Disteche CM (2005) Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development. *Develop. Cell*, 8, 31-42.

- Mukhopadhyay R, Yu W, Whitehead J, Xu J, Lezcano M, Pack S, Kanduri C, Kanduri M, Ginja V, Vostrov A, Quitschke W, Chernukhin I, Klenova E, Lobanenko V and Ohlsson R (2004) The binding site for chromatin insulator protein CTCF map to DNA methylation-free domain genome-wide. *Genome Res.*, 14, 1594-1602.

- Capelson M and Corces VG (2006) SUMO conjugation attenuates the activity of the gypsy chromatin insulator. *EMBO J.*, 25, 1906-1914.

- Lei EP and Corces VG (2006) RNA interference machinery influences the nuclear organization of a chromatin insulator. *Nat. Genet.*, 38, 936-941.

- Grimaud C, Bantignies F, Pal-Bhadra M, Ghana P, Bhadra U and Cavalli G (2006) RNAi components are required for nuclear clustering of Polycomb group response elements. *Cell*, 124, 957-971.

- Yusufzai TM, Tagami H, Nakatani Y and Felsenfeld G (2004) CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol. Cell*, 13, 291-298.

- Blanton J, Gaszner M and Schedl P (2003) Protein:protein interactions and the pairing of boundary elements in vivo. *Genes Dev.*, 17, 664-675.

- Ling JQ, Li T, Hu JF, Vu TH, Chen HL, Qiu XW, Cherry AM and Hoffman AR (2006) CTCF mediates interchromosomal colocalization between *Igf2/H19* and *Wsb1/Nf1*. *Science*, 312, 269-272.

- Ishihara K, Oshimura M and Nakao M (2006) CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Mol. Cell*, 23, 733-742.

- Butcher DT, Mancini-DiNardo DN, Archer TK and Rodenhiser DI (2004) DNA binding sites for putative methylation boundaries in the unmethylated region of the BRCA1 promoter. *Int. J. Cancer*, 111, 669-678.

- Butcher DT and Rodenhiser DI (2007) Epigenetic inactivation of BRCA1 is associated with aberrant expression of CTCF and DNA methyltransferase (DNMT3B) in some sporadic breast tumours. *Eur. J. Cancer*, 43, 210-219.

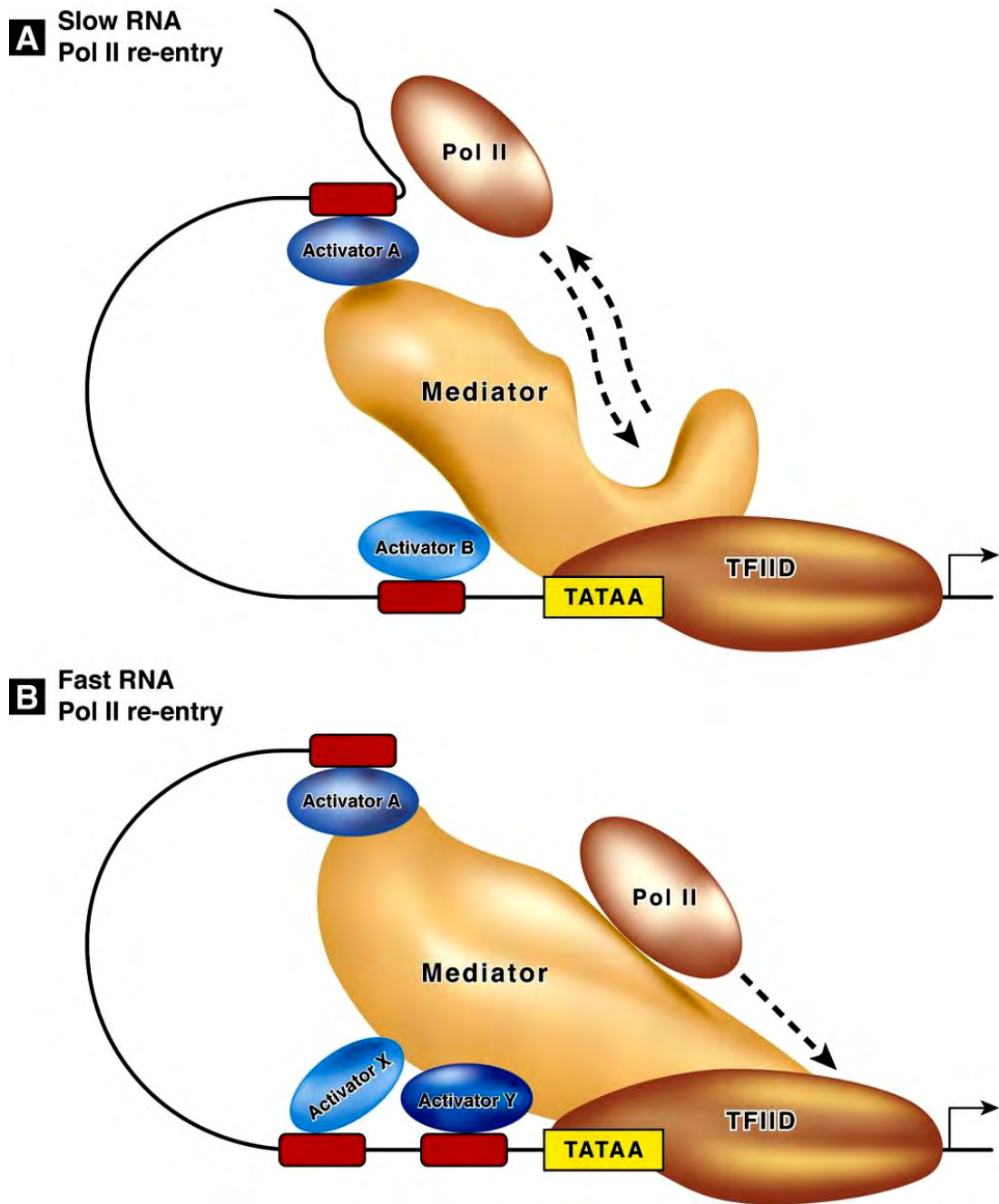
- Recillas-Targa F, De La Rosa-Velázquez IA, Soto-Reyes E and Benítez-Bribiesca L (2006) Epigenetic boundaries of tumour supresor gene promoters: the CTCF connection and its role in carcinogenesis. *J. Cell. Mol. Med.*, 10, 552-566.
- De La Rosa-Velázquez IA, Rincón-Arano H, Benítez-Bribiesca L and Recillas-Targa F (2007) Epigenetic regulation of the human retinoblastoma tumor supresor gene promoter by CTCF. *Cancer Res.*, 67, 2577-2585.
- Herman JG and Baylin SB (2003) Gene silencing in cancer in association with promoter hypermethylation. *N. Engl. J. Med.*, 349, 2042-2054.
- Barshi A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, Wei G, Chepelev I and Zhao K (2007) High-resolution profiling of histone methylation in the human genome. *Cell*, 129, 823-837.
- Misteli T (2007) Beyond the sequence: cellular organization of genome function. *Cell*, 128, 787-800.
- Orphanides G and Reinberg D (2002) A unified theory of gene expression. *Cell*, 198, 439-451.
- West AG, Gaszner M and Felsenfeld G (2002) Insulators: many functions, many mechanisms. *Genes Dev.*, 16, 271-288.
- Gaszner M and Felsenfeld G (2006) Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat. Rev. Genet.*, 7, 703-713.
- Valenzuela L and Kamakaka R (2006) Chromatin insulators. *Ann. Rev. Genet.*, 40, 107-138.
- Capelson M and Corces VG (2004) Boudnary elements and nuclear organization. *Biol. Cll*, 96, 617-629.
- Chakalova L, Debrand E, Mitchell JA, Osborne CS and Fraser P (2005) Replication and transcription: shaping the landscape of the genome. *Nat. Rev. Genet.*, 6, 669-677.
- West AG and Fraser P (2005) Remote control of gene transcription. *Hum. Mol. Genet.*, 14, R101-R111.
- Murrell A, Heeson S and Reik W (2004) Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat. Genet.*, 36, 889-893.
- Ling JQ and Hoffman AR (2007) Epigenetics of long-range chromatin interactions. *Pediatr. Res.*, 61, 11R-16R.
- Fraser P and Bickmore W (2007) Nuclear organization of the genome and the potential for gene regulation. *Nature*, 447, 413-417.

- Yi Z, Cohen-Barak O, Haquiwara N, Kingsley PD, Fuchs DA, Erickson DT, Epner EM, Palis J and Brilliant MH (2006) Sox6 directly silences epsilon globin expression in definitive erythropoiesis. *PLoS Genet.*, 2, e14.
- Schübeler D, Francastel C, Cimbora DM, Reik A, Martin DI and Groudine M (2000) Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human  $\beta$ -globin locus. *Genes Dev.*, 14, 940-950.
- Taddei A, Van Houwe G, Hediger F, Kalck V, Cubizolles F, Schober H and Gasser SM (2006) Nuclear pore association confers optimal expression levels for an inducible yeast gene. *Nature*, 441, 774-778.
- Boyes J and Felsenfeld G (1996) Tissue-specific factors additively increase the probability of the all-or-none formation of a hypersensitive site. *EMBO J.*, 15, 2496-2507.
- Chuang C-H, Carpenter AR, Fuchsova B, Johnson T, de Lanerolle P and Belmont AS (2006) Long-range directional movement of an interphase chromosome site. *Curr. Biol.*, 16, 825-831.
- Osborne CS, Chakalova L, Brown KE, Carter D, Horton A, Debrand E, Goyenechea B, Mitchell JA, Lopes S, Reik W and Fraser P (2004) Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat. Genet.*, 36, 1065-1071.
- Fisher AG and Merckenschlager M (2002) Gene silencing cell fate and nuclear organization. *Curr. Opin. Genet. Dev.*, 12, 193-197.
- Yazgan O and Krebs JE (2007) Noncoding but nonexpendable: transcriptional regulation by large noncoding RNA in eukaryotes. *Biochem. Cell Biol.*, 85, 484-496.
- Hahn S (2004) Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* 11, 394-402.
- Smale ST and Kadonaga JY (2003) The RNA polymerase II core promoter. *Ann. Rev. Biochem.*, 72, 449-479.
- Lujambo A and Esteller M (2007) CpG island hypermethylation of tumor suppressor microRNAs in human cancer. *Cell Cycle*, 6, 1455-1459.
- Baek HJ, Kang YK and Roeder RG (2006) Human mediator enhances basal transcription by facilitating recruitment of transcription factor IIB during preinitiation complex assembly. *J. Biol. Chem.*, 281, 15172-15181.
- Chen X, Hiller M, Saucak Y and Fuller MT (2005) Tissue-specific TAFs counteract Polycomb to form on terminal differentiation. *Science*, 310, 869-872.
- Smale ST (2001) Core promoters: active contributors to combinatorial gene regulation. *Genes Dev.*, 15, 2503-2508.

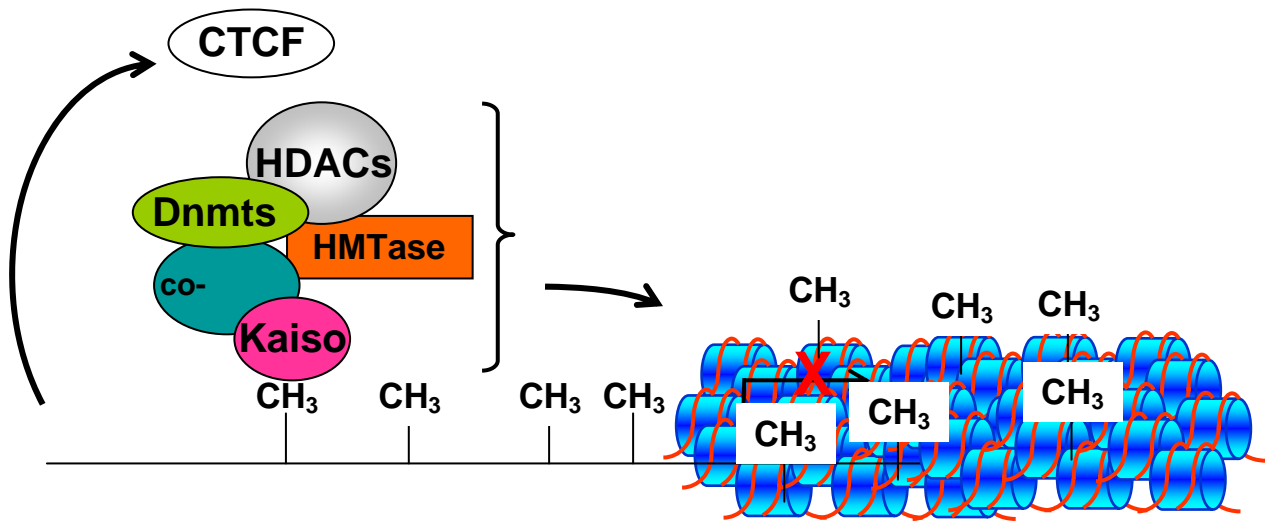
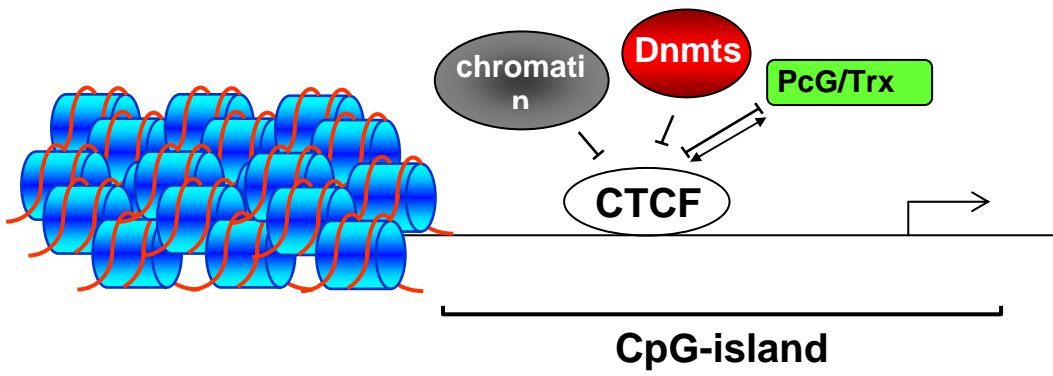
- Lim CY, Santoso B, Boulay T, Dong E, Ohler U and Kadonaga JT (2004) The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes Dev.*, 18, 1606-1617.
- Kim SI, Bultman SJ, Jing H, Blobel GA and Bresnick EH (2007) Dissecting molecular steps in chromatin domain activation during hematopoietic differentiation. *Mol. Cell. Biol.*, 27, 4551-4565.
- Reinberg D and Sims RJ3rd (2006) de FACTo nucleosome dynamics. *J. Biol. Chem.*, 281, 23397-23301.
- Dorman ER, Busher AM and Corces VG (2007) The role of insulator elements in large-scale chromatin structure in interphase. *Semin. Cell Dev. Biol.*, 18, 682-690.
- Lutz M, Burke LJ, LeFevre P, Myers FA, Thorne AW, Crane-Robinson C, Bonifer C, Filippova GN, Lobanenko and Renkawitz R (2003). Thyroid hormone-regulated enhancer blocking cooperation of CTCF and thyroid hormone receptor. *EMBO J.*, 22, 1579-1587.
- Saitoh N, Bell AC, Recillas-Targa F, West AG, Simpson M, Pikaart M and Felsenfeld G (2000) Structural and functional conservation at the boundaries of the chicken  $\beta$ -globin domain. *EMBO J.*, 19, 2315-2322.
- Burgess-Beusse B, Farrell C, Gaszner M, Litt M, Mutskov V, Recillas-Targa F, Simpson M, West A and Felsenfeld G (2002) The insulation of genes from external enhancers and silencing chromatin. *Proc. Natl. Acad. Sci. USA*, 99 suppl.4, 16433-16437.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL and Lander ES (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, 125, 315-326.







**Figure 2**



**Figure 3**

## Bibliografía

1. Stamatoyannopoulos JA, Goodwin A, Joyce T, Lowrey CH. (1995) *NF-E2 and GATA binding motifs are required for the formation of DNase I hypersensitive site 4 of the human beta-globin locus control region*. EMBO J. 14(1):106-16.
2. Boyes J, Felsenfeld G. (1996) *Tissue-specific factors additively increase the probability of the all-or-none formation of a hypersensitive site*. EMBO J. 15(10):2496-507.
3. Boyes J, Omichinski J, Clark D, Pikaart M, Felsenfeld G. (1998) *Perturbation of nucleosome structure by the erythroid transcription factor GATA-1*. J Mol Biol 279(3):529-44.
4. Felsenfeld G, Groudine M (2003) *Controlling the double helix*. Nature 421: 448-453.
5. Bulger M, Groudine M. *Looping versus linking: toward a model for long-distance gene activation*. Genes Dev 13:2465-2477.
6. Darnell J, Berk A, Matsudaira P, Kaiser CA, Krieger M, Scott MP. (2004) Molecular Cell Biology, 5ta Edición, W.H. Freeman and Company, NY, pp 424-433.
7. van Driel R, Fransz PF, Verschurre PJ (2003) *The eukaryotic genome: a system regulated at different hierarchical levels*. J.Cell Sci 116: 4067-4075.
8. Butler JEF, Kadonaga JT (2002) *The RNA polymerase II core promoter: a key component in the regulation of gene expression*. Genes Dev 16:2583-2592.
9. Kadonaga JT (2004) *Regulation of RNA polymerase II transcription by sequence-specific ADN binding factors*. Cell 116:247-257.
10. Roeder RG (1998) *Role of general and gene-specific cofactors in de regulation of eucariotic transcription*. In Mechanisms of Transcription. Cold Spring Harbor Lab Press, Vol. LXIII pp. 201-218.
11. Li B, Carey M, Workman JL. (2007) *The role of chromatin during transcription*. Cell. 128(4):707-19. Review.
12. Hahn S. (2004) *Structure and mechanism of the RNA polymerase II transcription machinery*. Nat Struct Mol Biol (5):394-403. Review.
13. Felsenfeld G, Clark DU, Studitsky Vasily. (2000) *Transcription through nucleosomes*. Biophysical Chemistry 86, 231-237.
14. Allshire, RC, Javerzat JP, Redhead NJ, and Cranston G. (1994) *Position effect variegation at fission yeast centromeres*. Cell 76:157-69.
15. Levine M, Tjian R. (2003) *Transcription regulation and animal diversity*. Nature. 424(6945): 147-51. Review.
16. Reinberg D, Orphanides G, Ebricht R, Akoulitchev S, Carcamo J, Cho H, Cortes P, Drapkin R, Flores O, Ha I, Inostroza JA, Kim S, Kumar P, Lagrange T, Leroy G, Lu H, Ma D-M, Maldonado E, Merino A, Mermelstein F, Olave I, Sheldon M, Shierhattar R, Stone N, Sun X, Weis L, Yeung K, Zawel L. (1998), *The RNA polimerasa II general Transcription factors: Past, Present and Future*. In *Mechanisms of Transcription*. Cold Spring Harbor Lab Press, Vol. LXIII pp. 83-103.
17. Blackwood EM, James T Kadonaga JT (1998), *Going the distance: A current view of enhancer action*. Science 281:60-63.
18. Smale ST (2001) *Core promoters: active contributors to combinatorial gene regulation*. Genes Dev 15: 2503-1508.
19. Weintraub H. (1988) *Formation of stable transcription complexes as assayed by analysis of individual templates*. Proc Natl Acad Sci U S A. 85(16):5819-23.
20. Lim CY, Santoso B, Boulay T, Dong E, Ohler U, Kadonaga JT. (2004) *The MTE, a new core promoter element for transcription by RNA polymerase II*. Genes Dev. 18(13): 1606-17.
21. Jennifer EF Butler and James T Kadonaga, (2002) *The RNA polymerase II core promoter: a key component in the regulation of gene expression*. Genes and Development. 16: 2583-92.
22. Sandelin A, Carninci P, Lenhard B, Ponjavic J, Hayashizaki Y, Hume DA. (2007) *Mammalian RNA polymerase II core promoters: insights from genome-wide studies*. Nat Rev Genet 8(6):424-36.
23. Shiraki T, Kondo S, Katayama S, Waki K, Kasukawa T, Kawaji H, Kodzius R, Watahiki A, Nakamura M, Arakawa T, Fukuda S, Sasaki D, Podhajaska A, Harbers M, Kawai J, Carninci P, Hayashizaki Y. (2003) *Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage*. Proc Natl Acad Sci U S A. 100(26):15776-81.

24. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C, Kodzius R, Shimokawa K, Bajic VB, Brenner SE, Batalov S, Forrest AR, Zavolan M, Davis MJ, Wilming LG, Aidinis V, Allen JE, Ambesi-Impiombato A, Apweiler R, Aturaliya RN, Bailey TL, Bansal M, Baxter L, Beisel KW, Bersano T, Bono H, Chalk AM, Chiu KP, Choudhary V, Christoffels A, Clutterbuck DR, Crowe ML, Dalla E, Dalrymple BP, de Bono B, Della Gatta G, di Bernardo D, Down T, Engstrom P, Fagiolini M, Faulkner G, Fletcher CF, Fukushima T, Furuno M, Futaki S, Gariboldi M, Georgii-Hemming P, Gingeras TR, Gojobori T, Green RE, Gustincich S, Harbers M, Hayashi Y, Hensch TK, Hirokawa N, Hill D, Huminiecki L, Iacono M, Ikeo K, Iwama A, Ishikawa T, Jakt M, Kanapin A, Katoh M, Kawasawa Y, Kelso J, Kitamura H, Kitano H, Kollias G, Krishnan SP, Kruger A, Kummerfeld SK, Kurochkin IV, Lareau LF, Lazarevic D, Lipovich L, Liu J, Liuni S, McWilliam S, Madan Babu M, Madera M, Marchionni L, Matsuda H, Matsuzawa S, Miki H, Mignone F, Miyake S, Morris K, Mottagui-Tabar S, Mulder N, Nakano N, Nakauchi H, Ng P, Nilsson R, Nishiguchi S, Nishikawa S, Nori F, Ohara O, Okazaki Y, Orlando V, Pang KC, Pavan WJ, Pavesi G, Pesole G, Petrovsky N, Piazza S, Reed J, Reid JF, Ring BZ, Ringwald M, Rost B, Ruan Y, Salzberg SL, Sandelin A, Schneider C, Schonbach C, Sekiguchi K, Semple CA, Seno S, Sessa L, Sheng Y, Shibata Y, Shimada H, Shimada K, Silva D, Sinclair B, Sperling S, Stupka E, Sugiura K, Sultana R, Takenaka Y, Taki K, Tammoja K, Tan SL, Tang S, Taylor MS, Tegner J, Teichmann SA, Ueda HR, van Nimwegen E, Verardo R, Wei CL, Yagi K, Yamanishi H, Zabarovsky E, Zhu S, Zimmer A, Hide W, Bult C, Grimmond SM, Teasdale RD, Liu ET, Brusic V, Quackenbush J, Wahlestedt C, Mattick JS, Hume DA, Kai C, Sasaki D, Tomaru Y, Fukuda S, Kanamori-Katayama M, Suzuki M, Aoki J, Arakawa T, Iida J, Imamura K, Itoh M, Kato T, Kawaji H, Kawagashira N, Kawashima T, Kojima M, Kondo S, Konno H, Nakano K, Ninomiya N, Nishio T, Okada M, Plessy C, Shibata K, Shiraki T, Suzuki S, Tagami M, Waki K, Watahiki A, Okamura-Oho Y, Suzuki H, Kawai J, Hayashizaki Y; FANTOM Consortium; RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group). (2005) *The transcriptional landscape of the mammalian genome*. *Science*. 309(5740):1559-63.
25. Schug J, Schuller WP, Kappen C, Salbaum JM, Bucan M, Stoeckert CJ Jr. (2005) *Promoter features related to tissue specificity as measured by Shannon entropy*. *Genome Biol*. 6(4): 47.
26. Suzuki Y, Tsunoda T, Sese J, Taira H, Mizushima-Sugano J, Hata H, Ota T, Isogai T, Tanaka T, Nakamura Y, Suyama A, Sakaki Y, Morishita S, Okubo K, Sugano S. (2001) *Identification and characterization of the potential promoter regions of 1031 kinds of human genes*. *Genome Res*. 11(5): 677-84.
27. Arnosti DN. (2003) *Analysis and function of transcriptional regulatory elements: insights from Drosophila*. *Annu Rev Entomol*. 48: 579-602. Review.
28. Ameres SL, Drueppel L, Pfliederer K, Schmidt A, Hillen W, Berens C. (2005) *Inducible ADN-loop formation blocks transcriptional activation by an SV40 enhancer*. *EMBO J*. 24(2):358-67.
29. Bernstein BE, Liu CL, Humphrey EL, Perlstein EO, Schreiber SL. (2004) *Global nucleosome occupancy in yeast*. *Genome Biol*. 5
30. Ohler U, Liao GC, Niemann H, Rubin GM. (2002) *Computational analysis of core promoters in the Drosophila genome*. *Genome Biol*. 3(12): 0087.
31. Morris JR, Petrov DA, Lee AM, Wu CT. (2004) *Enhancer choice in cis and in trans in Drosophila melanogaster: role of the promoter*. *Genetics*. 167(4): 1739-47.
32. Conte C, Dastugue B, Vaury C. (2002) *Promoter competition as a mechanism of transcriptional interference mediated by retrotransposons*. *EMBO J*. 21(14): 3908-16.
33. Sharpe J, Nonchev S, Gould A, Whiting J, Krumlauf R. (1998) *Selectivity, sharing and competitive interactions in the regulation of Hoxb genes*. *EMBO J*. 17(6): 1788-98.
34. Cai HN, Shen P. (2001) *Effects of cis arrangement of chromatin insulators on enhancer-blocking activity*. *Science*. 291(5503): 493-5.
35. Walker SS, Shen WC, Reese JC, Apone LM, Green MR. (1997) *Yeast TAF(II)145 required for transcription of G1/S cyclin genes and regulated by the cellular growth state*. *Cell*. 90(4): 607-14.
36. Cai HN, Zhang Z, Adams JR, Shen P. (2001) *Genomic context modulates insulator activity through promoter competition*. *Development*. 128(21): 4339-47.
37. Herendeen DR, Kassavetis GA, Geiduschek EP. (1992) *A transcriptional enhancer whose function imposes a requirement that proteins track along ADN*. *Science*. 256(5061): 1298-303.

38. Kuras L, Borggreffe T, Kornberg RD. (2003) *Association of the Mediator complex with enhancers of active genes*. Proc Natl Acad Sci U S A. 100(24): 13887-91.
39. Chiang DY, Nix DA, Shultzaberger RK, Gasch AP, Eisen MB. (2006) *Flexible promoter architecture requirements for coactivator recruitment*. Mol Biol. 28;7:16.
40. Keller RJ III, Flanagan PM and Kornberg RD. (1990) *A novel mediator between activator proteins and the RNA polymerase II transcription apparatus*. Cell. 61, 1209-1215.
41. Freidman R, Hotaling E, Borack L, Sofer W. (1996) *Interactions between the regulatory regions of two Adh alleles*. Genetica. 97(1): 1-14.
42. Lewis EB. (1998). *The bithorax complex: the first fifty years*. Int. J. Dev. Biol. 42: 403–15
43. Lemon B, Tjian R (2000) *Orchestrated response: a symphony of transcription factors for gene control*. Genes Dev 14: 2551-2569.
44. Orphanides G, Reinberg D (2002) *A unified theory of gene expression*. Cell 108: 439-451.
45. Zurita M, Merino C (2003) *The transcriptional complexity of the TFIID complex*. TIGS 19: 578-584
46. West AG, Fraser P (2005) *Remote control of gene transcription*. Hum Mol Genet. No 1: 261-11.
47. Calhoun VC, Levine M. (2003) *Long-range enhancer-promoter interactions in the Scr-Antp interval of the Drosophila Antennapedia complex*. Proc Natl Acad Sci U S A. 100(17): 9878-83.
48. Felsenfeld, G., J. Boyes, J. Chung, D. Clark, and V. Studitsky. (1996). *Chromatin structure and gene expression*. Proceedings of the National Academy of Sciences of the United States of America 93:9384-8.
49. Walker SS, Shen WC, Reese JC, Apone LM, Green MR. Related (1997) *Yeast TAF(II)145 required for transcription of G1/S cyclin genes and regulated by the cellular growth state*. Cell. 90(4): 607-14.
50. Rincon-Arano H, Valadez-Graham V, Guerrero G, Escamilla-Del-Arenal M, Recillas-Targa F. (2005) *YY1 and GATA-1 Interaction Modulate the Chicken 3'-Side alpha-Globin Enhancer Activity*. J Mol Biol. 349: 961-75.
51. Struhl K. (2001). *Gene regulation. A paradigm for precision*. Science 252: 138-87.
52. Arnosti DN, Kulkarni MM. (2005) *Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards?* J Cell Biochem. 94(5): 890-8.
53. Spicuglia S, Payet D, Tripathi RK, Rameil P, Verthuy C, Imbert J, Ferrier P, Hempel WM. (2000) *TCRalpha enhancer activation occurs via a conformational change of a pre-assembled nucleoprotein complex*. EMBO J. 19(9):2034-45.
54. Hatzis P, Talianidis I. (2002) *Dynamics of enhancer-promoter communication during differentiation-induced gene activation*. Mol Cell. 10(6):1467-77.
55. Walters MC, Fiering S, Eidemiller J, Magis W, Groudine M, Martin DI. (1995) *Enhancers increase the probability but not the level of gene expression*. Proc Natl Acad Sci U S A. 92(15):7125-9.
56. Liu Y, Bondarenko V, Ninfa A, Studitsky VM. (2001) *ADN supercoiling allows enhancer action over a large distance*. Proc Natl Acad Sci U S A. 98(26): 14883-8.
57. Ptashne M. (1986) *Gene regulation by proteins acting nearby and at a distance*. Nature. 322 (6081): 697-701. Review.
58. Ringrose L, Chabanis S, Angrand PO, Woodrooffe C, Stewart AF. (1999) *Quantitative comparison of ADN looping in vitro and in vivo: chromatin increases effective ADN flexibility at short distances*. EMBO J. 18(23): 6630-41.
59. Bulger M, Groudine M. (1999) *Looping versus linking: toward a model for long-distance gene activation*. Genes Dev. 13(19): 2465-77. Review.
60. Fischle W, Wang Y, Allis CD (2003) *Histone and chromatin cross-talk*. Curr Opin Cell Biol 15: 172-183.
61. Dekker, J. (2003). *A closer look at long-range chromosomal interactions*. Trends in Biochemical Sciences 28:277-80.
62. Murrell A, Heeson S, Reik W. (2004) *Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops*. Nat Genet. 36(8): 889-93.
63. Kato Y, Sasaki H. (2005) *Imprinting and looping: epigenetic marks control interactions between regulatory elements*. Bioessays. 27(1): 1-4. Review

64. Francastel C, Walters MC, Groudine M, Martin DI. (1999) *A functional enhancer suppresses silencing of a transgene and prevents its localization close to centromeric heterochromatin*. *Cell* 99(3): 259-69.
65. Li, Q., S. Harju, and K. R. Peterson. (1999). *Locus control regions: coming of age at a decade plus*. *Trends in Genetics* 15:403-8.
66. Valadez Graham V, Recillas Targa F (2002) *Características y formación de un dominio transcripcionalmente activo en eucariotes*. *REB* 21: 21-40.
67. Li Q, Peterson KR, Fang X, Stamatoyannopoulos G (2002) *Locus control regions*. *Blood* 100: 3077-3086.
68. Dean, A. (2006). *On a chromosome far, far away: LCRs and gene expression*. *Trends in Genetics* 22:38-45.
69. Bulger, M., and M. Groudine. (1999). *Looping versus linking: toward a model for long-distance gene activation*. *Genes & Development* 13:2465-77.
70. Bulger, M., and M. Groudine. (2002). *TRAPping enhancer function*. *Nature Genetics* 32:555-6.
71. Habermann FA, Biet C, Fries R. (2001) *Chromosomal localization of the genes encoding SCNN1A, BTG1, IFNG and MAOA on chicken chromosome 1 by fluorescence in-situ hybridization*. *Chromosome Res.*;9(6):515-8.
72. Brand AH, Breeden L, Abraham J, Sternglanz R, Nasmyth K. (1985) *Characterization of a "silencer" in yeast: a ADN sequence with properties opposite to those of a transcriptional enhancer*. *Cell* 41:41-48.
73. Ogbourne S, Antalis TM. (1998) *Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes*. *Biochem J* 331: 1-14.
74. Clark AR, Docherty K. (1993) *Negative regulation of transcription in eukaryotes*. *Biochem J* 15: 521-541.
75. Recillas-Targa F, Razin SV. (2001) *Chromatin domains and regulation of gene expression: familiar and enigmatic clusters of chicken globin genes*. *Crit Rev Eukaryot Gene Expr* 11:227-242.
76. Osborne CS, Chakalova L, Brown KE, Carter D, Horton A, Debrand E, Goyenechea B, Mitchell JA, Lopes S, Reik W, Fraser P. (2004) *Active genes dynamically colocalize to shared sites of ongoing transcription*. *Nat Genet.* 36(10):1065-71.
77. Labrador, M., and V. G. Corces. (2002). *Setting the boundaries of chromatin domains and nuclear organization*. *Cell* 111:151-4.
78. West AG, Gaszner M, Felsenfeld G (2002) *Insulators: many functions, many mechanisms*. *Genes Dev* 16: 271-288.
79. Gerasimova, T. I., and V. G. Corces. (2001). *Chromatin insulators and boundaries: effects on transcription and nuclear organization*. *Annual Review of Genetics* 35:193-208.
80. Razin SV, Farrell CM, Recillas-Targa F (2003) *Genomic domains and regulatory elements operating at the domain level*. *Int Rev Cytol* 226: 63-125.
81. Peng JC, Karpen GH. (2007) *H3K9 methylation and RNA interference regulate nucleolar organization and repeated ADN stability*. *Nat Cell Biol.* 9(1):25-35.
82. Recillas\_Targa F., Pikaart MJ, Burgess\_Beusse B, Bell BA, Litt MD, West AG, Gaszner M., and Felsenfeld G.. (2002). *Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities*. *Proceedings of the National Academy of Sciences of the United States of America* 99:6883-8.
83. Espada J, Esteller M. (2007) *Epigenetic control of nuclear architecture*. *Cell Mol Life Sci.* 64(4):449-57. Review.
84. Lutz M, Burke LJ, LeFevre P, Myers FA, Thorne AW, Crane-Robinson C, Bonifer C, Filippova GN, Lobanenko V, Renkawitz R. (2003) *Thyroid hormone-regulated enhancer blocking: cooperation of CTCF and thyroid hormone receptor*. *EMBO J.* 22(7):1579-87
85. Bondarenko VA, Jiang YI, Studitsky VM. (2003) *Rationally designed insulator-like elements can block enhancer action in vitro*. *EMBO J.* 22(18): 4728-37.
86. Capelson M, Corces VG. (2004) *Boundary elements and nuclear organization*. *Biol Cell.* 96(8):617-29. Review.
87. Razin SV, Iarovaia OV, Sjakste N, Sjakste T, Bagdoniene L, Rynditch AV, Eivazova ER, Lipinski M, Vassetzky YS. (2007) *Chromatin domains and regulation of transcription*. *Mol Biol.* 369(3):597-607. Review.

88. Farrell, C. M., A. G. West, and G. Felsenfeld. (2002). *Conserved CTCF insulator elements flank the mouse and human beta-globin loci*. *Molecular and Cellular Biology* 22:3820-31.
89. Recillas-Targa, F., A. C. Bell, and G. Felsenfeld. (1999). *Positional enhancer-blocking activity of the chicken beta-globin insulator in transiently transfected cells*. *Proceedings of the National Academy of Sciences of the United States of America* 96:14354-9.
90. Zhao, H., A. Kim, S. H. Song, and A. Dean. (2006). *Enhancer blocking by chicken beta -globin 5' HS4: Role of enhancer strength and insulator nucleosome depletion*. *Journal of Biological Chemistry*. In Press, Corrected Proof.
91. Ling, J. Q., T. Li, J. F. Hu, T. H. Vu, H. L. Chen, X. W. Qiu, A. M. Cherry, and A. R. Hoffman. (2006). *CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1*. *Science* 312:269-72.
92. Lomvardas, S., G. Barnea, D. J. Pisapia, M. Mendelsohn, J. Kirkland, and R. Axel. (2006). *Interchromosomal interactions and olfactory receptor choice*. *Cell* 126:403-13.
93. Vakoc, C. R., D. L. Letting, N. Gheldof, T. Sawado, M. A. Bender, M. Groudine, M. J. Weiss, J. Dekker, and G. A. Blobel. (2005). *Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1*. *Molecular Cell* 17:453-62.
94. Neff, T., F. Shotkoski, and G. Stamatoyannopoulos. (1997). *Stem cell gene therapy, position effects and chromatin insulators*. *Stem Cells* 15 Suppl 1:265-71.
95. Mutskov, V. J., C. M. Farrell, P. A. Wade, A. P. Wolffe, and G. Felsenfeld. (2002). *The barrier function of an insulator couples high histone acetylation levels with specific protection of promoter ADN from methylation*. *Genes & Development* 16:1540-54.
96. West, A. G., S. Huang, M. Gaszner, M. D. Litt, and G. Felsenfeld. (2004). *Recruitment of histone modifications by USF proteins at a vertebrate barrier element*. *Molecular Cell* 16:453-63.
97. Carrozza MJ, Kusch T, Workman JL (2003) *Repairing nucleosomes during transcription*. *Nat Struct Biol* 10: 879-880.
98. Branco, M. R. & Pombo, A. (2006) *Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations*. *PLoS Biol* 4(5):e138.
99. Turek-Plewa, J., and P. P. Jagodzinski. (2005). *The role of mammalian ADN methyltransferases in the regulation of gene expression*. *Cellular and Molecular Biology Letters* 10:631-47.
100. Recillas-Targa F (2000) *The chicken  $\alpha$ - and  $\beta$ -globin gene domains and their chromatin organization*. *Cell Mol Biol Lett* 5: 451-467.
101. Prioleau M-N, Nony P, Simpson M, Felsenfeld G (1999) *An insulator element and condensed chromatin region separate the chicken  $\beta$ -globin locus from an independently regulated erythroid-specific folate receptor gene*. *EMBO J* 18: 4035-4048.
102. Saitoh N, Bell AC, Recillas-Targa F, West AG, Simpson M, Pikaart M, Felsenfeld G (2000) *Structural and functional conservation at the boundaries of the chicken  $\alpha$ -globin domain*. *EMBO J* 19: 2315-2322.
103. Litt MD, Simpson M, Recillas-Targa F, Prioleau M-N, Felsenfeld G (2001) *Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci*. *EMBO J* 20: 2224-2235.
104. Litt MD, Gaszner M, Simpson M, Allis CD, Felsenfeld G (2001) *Correlation between histone lysine methylation and developmental changes at the chicken  $\beta$ -globin locus*. *Science* 293: 2453-2455.
105. Lee JT (2003) *Molecular links between X-inactivation and autosomal imprinting: X-inactivation as a driving force for the evolution of imprinting ?* *Curr Biol* 13: 392-404.
106. Wolffe AP, Matzke MA (1999) *Epigenetics: regulation through repression*. *Science* 286: 481-486.
107. Adrian Bird A (2002) *ADN methylation patterns and epigenetic memory*. *Genes Dev* 16: 6-21.
108. Dean, W., D. Lucifero, and F. Santos. (2005). *ADN methylation in mammalian development and disease*. *Birth Defects Res C Embryo Today* 75:98-111.
109. Klose, R. J., and A. P. Bird. (2006). *Genomic ADN methylation: the mark and its mediators*. *Trends in Biochemical Sciences* 31:89-97.
110. Doerfler, W. (2005). *On the biological significance of ADN methylation*. *Biochemistry* 70:505-24.
111. D'Alessio AC, Szyf M. (2006) *Epigenetic tete-a-tete: the bilateral relationship between chromatin modifications and ADN methylation*. *Biochem Cell Biol*. 84(4):463-76. Review.

112. Szyf M, Weaver IC, Champagne FA, Diorio J, Meaney MJ. (2005) *Maternal programming of steroid receptor expression and phenotype through ADN methylation in the rat*. *Front Neuroendocrinol.* 26(3-4):139-62. Review.
113. Wade PA (2001) *Methyl CpG-binding proteins and transcriptional repression*. *BioEssays* 23: 1131-1137.
114. Bulger, M., and M. Groudine. (1999). *Looping versus linking: toward a model for long-distance gene activation*. *Genes & Development* 13:2465-77.
115. Frreitag M, Hickey PC, Khalafallah TK, Read ND, Selker EU (2004) *HP1 is essential for ADN methylation in Neurospora*. *Mol Cell* 13: 427-434.
116. Narlikar, G. J., H. Y. Fan, and R. E. Kingston. (2002). *Cooperation between complexes that regulate chromatin structure and transcription*. *Cell* 108:475-87.
117. Narlikar GJ, Fan HY, Kingston RE (2002) *Cooperation between complexes that regulate chromatin structure and transcription*. *Cell* 108: 475-487.
118. Roberts CWM, Orkin SH (2004) *The SWI/SNF complex-chromatin and cancer*. *Nat Rev Cancer* 4: 133-142.
119. Bowen NJ, Fujita N, Kajita M, Wade PA (2004) *Mi-2/NuRD: multiple complexes for many purposes*. *Biochim Biophys Acta* 1677: 52-57.
120. de la Serna, I. L., Y. Ohkawa, and A. N. Imbalzano. (2006). *Chromatin remodelling in mammalian differentiation: lessons from ATP-dependent remodelers*. *Nature Review in Genetics* 7:461-73.
121. Varga-Weisz, P. D., and P. B. Becker. (2006). *Regulation of higher-order chromatin structures by nucleosome-remodelling factors*. *Current Opinion in Genetics & Development Chromosomes and expression mechanisms* 16:151-156.
122. Fry CJ, Peterson CL (2002) *Unlocking the gates to gene expression*. *Science* 295: 1847-1848.
123. Pirrotta V (1998) *Polycomb the genome: PcG, TrxG and chromatin silencing*. *Cell* 93: 333-336.
124. Francis NJ, Kingston RE (2001) *Mechanisms of transcriptional memory*. *Nat. Rev Mol Cell Biol* 2: 409-421.
125. Orlando V (2003) *Polycomb, epigenomes and control of cell identity*. *Cell*, 112, 599-606.
126. Morey, C, Da Silva, Perry NR, & Bickmore WA. (2007) *Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation*. *Development* 134, 909-919.
127. Kosak, S. T., and M. Groudine. (2002) *The undiscovered country: chromosome territories and the organization of transcription*. *Developmental Cell* 2:690-2.
128. Cremer T, Cremer C (2001) *Chromosome territories, nuclear architecture and gene regulation in mammalian cells*. *Nat Rev Genet* 2: 292-301.
129. O'Brien TP, Bult CJ, Cremer C, Grunze M, Knowles BB, Langowski J, McNally J, Pederson T, Politz JC, Pombo A, Schmahl G, Spatz JP, van Driel R (2003) *Genome function and nuclear architecture: from gene expression to nanoscience*. *Genome Res* 13: 1029-1041.
130. Francastel C, Schubeler D, Martin DI, Groudine M (2000) *Nuclear compartmentalization and gene activity*. *Nat Rev Mol Cell Biol* 1: 137-143.
131. Volpi EV, Chevret E, Jones T, Vatcheva R, Williamson J, Beck S, Campbell RD, Goldsworthy M, Powis SH, Ragoussis J, Trowsdale J, Sheer D. (2000) *Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei*. *J. Cell Sci.* 113, 1565-1576.
132. Parada LA, Misteli T (2002) *Chromosome positioning in the interphase nucleus*. *Trends Cell Biol* 12: 425-432.
133. Francastel C, Magis W, Groudine M (2001) *Nuclear relocation of a transactivator unit precedes target gene activation*. *Proc Natl Acad Sci USA* 98: 12120-12125.
134. Palis, J., and G. B. Segel. (1998). *Developmental biology of erythropoiesis*. *Blood Reviews* 12:106-14.
135. Felsenfeld, G. (1993). *Chromatin structure and the expression of globin-encoding genes*. *Gene* 135:119-24.
136. Baur, J. A., Y. Zou, J. W. Shay, and W. E. Wright. (2001). *Telomere Position Effect in Human Cells*. *Science* 292:2075-77.



137. Agalioti, T., S. Lomvardas, B. Parekh, J. Yie, T. Maniatis, and D. Thanos. (2000). *Ordered Recruitment of Chromatin Modifying and General Transcription Factors to the IFN- $\beta$  Promoter*. Cell 103:667-78.
138. Brown JM, Leach J, Reittie JE, Atzberger A, Lee-Prudhoe J, Wood WG, Higgs DR, Iborra FJ, Buckle VJ. (2006) *Coregulated human globin genes are frequently in spatial proximity when active*. J. Cell Biol. 172, 177–187.
139. Tang XB, Liu DP, Liang CC. (2001) *Regulation of the transcription factor GATA-1 at the gene and protein level*. Cell Mol Life Sci. 58(14):2008-17. Review.
140. Michael E. Layon, Catherine J. Ackley, Rachel J. West and Christopher H. Lowrey. (2007) *Expression of GATA-1 in a Non-hematopoietic Cell Line Induces  $\beta$ -Globin Locus Control Region Chromatin Structure Remodeling and an Erythroid Pattern of Gene Expression* J. Mol. Biol. 366, 737–744.
141. Evans T, Felsenfeld G. (1989) *The erythroid-specific transcription factor Eryf1: a new finger protein*. Cell. 58(5):877-85.
142. Mahy, N. L., Perry, P. E. & Bickmore, W. A. (2002) *Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH*. J. Cell Biol. 159, 753–763.
143. Zon LI, Yamaguchi Y, Yee K, Albee EA, Kimura A, Bennett JC, Orkin SH, Ackerman SJ. (1993) *Expression of mRNA for the GATA-binding proteins in human eosinophils and basophils: potential role in gene transcription*. Blood. 81(12):3234-41.
144. Martin DI, Zon LI, Mutter G, Orkin SH. (1990) *Expression of an erythroid transcription factor in megakaryocytic and mast cell lineages*. Nature. 344(6265):444-7.
145. Leonard M, Brice M, Engel JD, Papayannopoulou T. (1993) *Dynamics of GATA transcription factor expression during erythroid differentiation*. Blood. 82(4):1071-9.
146. Ferreira R, Wai A, Shimizu R, Gillemans N, Rottier R, von Lindern M, Ohneda K, Grosveld F, Yamamoto M, Philipsen S. (2007) *Dynamic regulation of Gata factor levels is more important than their identity*. Blood. 109(12):5481-90.
147. George KM, Leonard MW, Roth ME, Lieu KH, Kioussis D, Grosveld F, Engel JD. (1994) *Embryonic expression and cloning of the murine GATA-3 gene*. Development. 120(9):2673-86.
148. Baron MH. (1997) *Transcriptional control of globin gene switching during vertebrate development*. Biochim Biophys Acta. 1351(1-2):51-72. Review.
149. Pevny L, M.C. Simon, E. Robertson, W.H. Klein, S.F. Tsai and V. D'Agati, (1991) *Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1*, Nature 349, pp. 257–260.
150. Weiss M.J. and S.H. Orkin, (1995) *Transcription factor GATA-1 permits survival and maturation of erythroid precursors by preventing apoptosis*, Proc. Natl Acad. Sci. USA 92, pp. 9623–9627.
151. Nichols K.E., J.D. Crispino, M. Poncz, J.G. White, S.H. Orkin, J.M. Maris and M.J. Weiss, (2000) *Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1*, Nature Genet. 24, pp. 266–270.
152. Mehaffey MG, A.L. Newton, M.J. Gandhi, M. Crossley and J.G. Drachman, (2001) *X-linked thrombocytopenia caused by a novel mutation of GATA-1*, Blood 98, pp. 2681–2688.
153. Freson K., G. Matthijs, C. Thys, P. Marien, M.F. Hoylaerts, J. Vermynen and C. Van Geet, (2002) *Different substitutions at residue D218 of the X-linked transcription factor GATA1 lead to altered clinical severity of macrothrombocytopenia and anemia and are associated with variable skewed X inactivation*, Hum. Mol. Genet. 11, pp. 147–152.
154. Letting DL, Chen YY, Rakowski C, Reedy S, Blobel GA. (2004) *Context-dependent regulation of GATA-1 by friend of GATA-1*. Proc Natl Acad Sci U S A. Jan 13;101(2):476-81.
155. Cantor A.B. and S.H. Orkin, (2002) *Transcriptional regulation of erythropoiesis: an affair involving multiple partners*, Oncogene 21, pp. 3368–3376.
156. Hannon R, Evans T, Felsenfeld G, Gould H. (1991) *Structure and promoter activity of the gene for the erythroid transcription factor GATA-1*. Proc Natl Acad Sci U S A. 88 (8):3004-8.
157. Li, J., Ishii, T., Feinstein, P. & Mombaerts, P. (2004) *Odorant receptor gene choice is reset by nuclear transfer from mouse olfactory sensory neurons*. Nature 428, 393–399.

158. Hung HL, Lau J, Kim AY, Weiss MJ, Blobel GA. (1999) *CREB-Binding protein acetylates hematopoietic transcription factor GATA-1 at functionally important sites.* Mol Cell Biol. 19(5):3496-505.
159. Cremer T, Cremer M, Dietzel S, Müller S, Solovei I, Fakan S (2006) *Chromosome territories: a functional nuclear landscape.* Curr. Opin. Cell Biol. 18, 307–316.
160. Crossley M, Merika M, Orkin SH. (1995) *Self-association of the erythroid transcription factor GATA-1 mediated by its zinc finger domains.* Mol Cell Biol. 15(5):2448-56.
161. Mackay JP, Kowalski K, Fox AH, Czolij R, King GF, Crossley M. (1998) *Involvement of the N-finger in the self-association of GATA-1.* J Biol Chem. 273(46):30560-7.
162. Letting DL, Rakowski C, Weiss MJ, Blobel GA. (2003) *Formation of a tissue-specific histone acetylation pattern by the hematopoietic transcription factor GATA-1.* Mol Cell Biol. 23(4):1334-40.
163. John A.Stamatoyannopoulos, Andrew Goodwin, Terrence Joyce and Christopher H.Lowrey (1995) *NF-E2 and GATA binding motifs are required for the formation of ADNse I hypersensitive site 4 of the human 5-globin locus control region.* The EMBO Journal vol. 14 pp. 106-116,
164. Pomerantz O, Goodwin AJ, Joyce T, Lowrey CH. (1998) *Conserved elements containing NF-E2 and tandem GATA binding sites are required for erythroid-specific chromatin structure reorganization within the human beta-globin locus control region.* Nucleic Acids Res. 26(24):5684-91.
165. Li G, Lim KC, Engel JD, Bungert J. (1998) *Individual LCR hypersensitive sites cooperate to generate an open chromatin domain spanning the human beta-globin locus.* Genes Cells., 3(7):415-29.
166. Hui Zhao, AeRiKim, Sang-Hyun Song and Ann Dean (2006) *Enhancer loking by the chicken-globin 5-HS2.* The Journal of Biological Chemistry vol 281 (41) pp 30573.
167. Eggan K, Baldwin K, Tackett M, Osborne J, Gogos J, Chess A, Axel R, Jaenisch R. (2004) *Mice cloned from olfactory sensory neurons.* Nature 428, 44–49.
168. Buckle R, Balmer M, Yenidunya A, Allan J. (1991) *The promoter and enhancer of the inactive chicken beta-globin gene contains precisely positioned nucleosomes.* Nucleic Acids Res., 19(6):1219-26.
169. Fraser P, Bickmore W. (2007) *Nuclear organization of the genome and the potential for gene regulation.* Nature 447(7143):413-7. Review.
170. Lowrey CH, Bodine DM, Nienhuis AW. (1992) *Mechanism of ADNse I hypersensitive site formation within the human globin locus control region.* Proc Natl Acad Sci U S A. 89(3):1143-7.
171. Steger DJ, Workman JL. (1996) *Remodeling chromatin structures for transcription: what happens to the histones?* Bioessays. 18(11):875-84. Review.
172. Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. *Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4.* Mol Cell. 9(2):279-89.
173. Lomvardas S, Thanos D. (2002) *Opening chromatin.* Mol Cell. 9(2):209-11. Review.
174. Shang Y., Myers,M. and Brown,M. (2002) *Formation of the androgen receptor transcription complex.* Mol. Cell, 9: 601–610.
175. Spicuglia S., Kumar,S., Yeh,J.H., Vachez,E., Chasson,L., Gorbach,S., Cautres,J. and Ferrier,P. (2002) *Promoter activation by enhancer-dependent and -independent loading of activator and coactivator complexes.* Mol. Cell, 10:, 1479–1487.
176. Louie M.C., Yang,H.Q., Ma,A.H., Xu,W., Zou,J.X., Kung,H.J. and Chen,H.W. (2003) *Androgen-induced recruitment of RNA polymerase II to a nuclear receptor-p160 coactivator complex.* Proc. Natl Acad. Sci. USA, 100:, 2226–2230.
177. Forrester WC, Perens E, Zallen JA, Garriga G. (1998) *Identification of Caenorhabditis elegans genes required for neuronal differentiation and migration.* Genetics. 148(1):151-65.
178. Simon I, Tenzen T, Mostoslavsky R, Fibach E, Lande L, Milot E, Gribnau J, Grosveld F, Fraser P, Cedar H. (2001) *Developmental regulation of ADN replication timing at the human beta globin locus.* EMBO J. 20(21):6150-7.
179. Feng YQ, Warin R, Li T, Olivier E, Besse A, Lobell A, Fu H, Lin CM, Aladjem MI, Bouhassira EE. (2005) *The human beta-globin locus control region can silence as well as activate gene expression.* Mol Cell Biol. 25(10):3864-74.

180. Szutorisz H, Dillon N, Tora L. (2005) *The role of enhancers as centres for general transcription factor recruitment*. Trends Biochem Sci. 30(11):593-9. Review.
181. Kirby D, Johnson A, Jeffrey A, Grass M, Meghan E, Boyer C, Carol M, Kiekhäfer G, Gerd A, Blobel M, Mitchell J, Weiss H, and Emery H. (2002) *Bresnick Cooperative activities of hematopoietic regulators recruit RNA polymerase II to a tissue-specific chromatin domain* PNAS, vol. 99, no. 18, 11760–11765.
182. Fan X, Chou DM, Struhl K. (2006) *Activator-specific recruitment of Mediator in vivo*. Nat Struct Mol Biol;13(2):117-20.
183. Thompson, C.M. & Young, R.A. (1995) *General requirement for RNA polymerase II holoenzymes in vivo*. Proc. Natl. Acad. Sci. USA 92, 4587–4590.
184. Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA. (1998) *Dissecting the regulatory circuitry of a eukaryotic genome*. Cell 95, 717–728.
185. Boube, M. , Joulia, L. , Cribbs, D.L. & Bourbon, H.M. (2002) *Evidence for a mediator of RNA polymerase II transcriptional regulation conserved from yeast to man*. Cell 110, 143–151.
186. Kornberg, R.D. (2005) *Mediator and the mechanism of transcriptional activation*. Trends Biochem. Sci. 30, 235–239.
187. Bjorklund, S. & Gustafsson, C.M. (2005) *The yeast Mediator complex and its regulation*. Trends Biochem. Sci. 30, 240–244.
188. Kim, Y.J. & Lis, J.T. (2005) *Interactions between subunits of Drosophila Mediator and activator proteins*. Trends Biochem. Sci. 30, 245–249.
189. Conaway, R.C. , Sato, S. , Tomomori-Sato, C. , Yao, T. & Conaway, J.W. (2005) *The mammalian Mediator complex and its role in transcriptional regulation*. Trends Biochem. Sci. 30, 250–255.
190. Malik, S. & Roeder, R.G. (2005) *Dynamic regulation of Pol II transcription by the mammalian Mediator complex*. Trends Biochem. Sci. 30, 256–263.
191. Chadick, J.Z. & Asturias, F.J. *Structure of eukaryotic Mediator complexes*. Trends Biochem. Sci. 30, 264–271
192. Dean A. (2004) *Chromatin remodelling and the interaction between enhancers and promoters in the beta-globin locus*. Brief Funct Genomic Proteomic. 2(4):344-54. Review.
193. Segal E, Fondufe-Mittendorf Y, Chen L, Thastrom A, Field Y, Moore IK, Wang JP, Widom J. (2006) *A genomic code for nucleosome positioning*. Nature. 442(7104):772-8.
194. Fox AH, Kowalski K, King GF, Mackay JP, Crossley M. (1998) *Key residues characteristic of GATA N-fingers are recognized by FOG*. J Biol Chem. 273(50):33595-603.
195. Perry C, Soreq H. (2002) *Transcriptional regulation of erythropoiesis. Fine tuning of combinatorial multi-domain elements*. Eur J Biochem. 269(15):3607-18. Review.
196. Wang X, Crispino JD, Letting DL, Nakazawa M, Poncz M, Blobel GA. (2002) *Control of megakaryocyte-specific gene expression by GATA-1 and FOG-1: role of Ets transcription factors*. EMBO J. 21(19):5225-34.
197. Zhang L, Schroeder S, Fong N, Bentley DL. (2005) *Altered nucleosome occupancy and histone H3K4 methylation in response to 'transcriptional stress'*. EMBO J. 24(13):2379-90.
198. Crispino, J. D., M. B. Lodish, J. P. MacKay, and S. H. Orkin. (1999). *Use of altered specificity mutants to probe a specific protein-protein interaction in differentiation: the GATA-1:FOG complex*. Mol. Cell 3:219-228.
199. Johnson KD, Boyer ME, Kang JA, Wickrema A, Cantor AB, Bresnick EH. (2007) *Friend of GATA-1-independent transcriptional repression: a novel mode of GATA-1 function*. Blood. 109(12):5230-3.
200. Johnson KD, Kim SI, Bresnick EH. (2006) *Differential sensitivities of transcription factor target genes underlie cell type-specific gene expression profiles*. Proc Natl Acad Sci U S A. 103(43):15939-44.
201. Kawabata H, Germain RS, Ikezoe T, Tong X, Green EM, Gombart AF, Koeffler HP. (2001) *Regulation of expression of murine transferrin receptor 2*. Blood. 98(6):1949-54.
202. Raisner RM, Hartley PD, Meneghini MD, Bao MZ, Liu CL, Schreiber SL, Rando OJ, Madhani HD. (2005) *Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin*. Cell.123(2):233-48.

203. Anguita E, Hughes J, Heyworth C, Blobel GA, Wood WG, Higgs DR. (2004) *Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2*. EMBO J. 23(14):2841-52.
204. Liu CL, Kaplan T, Kim M, Buratowski S, Schreiber SL, Friedman N, Rando OJ. (2005) *Single-nucleosome mapping of histone modifications in S. cerevisiae*. PLoS Biol. 3(10).
205. Blank V, Andrews NC. (1997) *The Maf transcription factors: regulators of differentiation*. Trends Biochem Sci. 22(11):437-41. Review.
206. Forsberg EC, Downs KM, Bresnick EH. (2000) *Direct interaction of NF-E2 with hypersensitive site 2 of the beta-globin locus control region in living cells*. Blood. 96(1):334-9.
207. Kiekhaefer CM, Grass JA, Johnson KD, Boyer ME, Bresnick EH. (2002) *Hematopoietic-specific activators establish an overlapping pattern of histone acetylation and methylation within a mammalian chromatin domain*. Proc Natl Acad Sci U S A. 99(22):14309-14.
208. Harju S, McQueen KJ, Peterson KR. (2002) *Chromatin structure and control of beta-like globin gene switching*. Exp Biol Med (Maywood). 227(9):683-700. Review.
209. Gong QH, McDowell JC, Dean A. (1996) *Essential role of NF-E2 in remodeling of chromatin structure and transcriptional activation of the epsilon-globin gene in vivo by 5' hypersensitive site 2 of the beta-globin locus control region*. Mol Cell Biol. 16(11):6055-64.
210. McDowell JC, Dean A. (1999) *Structural and functional cross-talk between a distant enhancer and the epsilon-globin gene promoter shows interdependence of the two elements in chromatin*. Mol Cell Biol. 19(11):7600-9.
211. Yuan GC, Liu YJ, Dion MF, Slack MD, Wu LF, Altschuler SJ, Rando OJ. (2005) *Genome-scale identification of nucleosome positions in S. cerevisiae*. Science. 309(5734):626-30.
212. Anderson KP, Kern CB, Crable SC, Lingrel JB. (1995) *Isolation of a gene encoding a functional zinc finger protein homologous to erythroid Kruppel-like factor: identification of a new multigene family*. Mol Cell Biol. 15(11):5957-65.
213. Turner J, Crossley M. (1999) *Mammalian Kruppel-like transcription factors: more than just a pretty finger*. Trends Biochem Sci. 24(6):236-40. Review.
214. Feng WC, Southwood CM, Bieker JJ. (1994) *Analyses of beta-thalassemia mutant ADN interactions with erythroid Kruppel-like factor (EKLF), an erythroid cell-specific transcription factor*. J Biol Chem. 269(2):1493-500.
215. Peterson KR. (2003) *Hemoglobin switching: new insights*. Curr Opin Hematol. 10(2):123-9. Review.
216. Brown RC, Pattison S, van Ree J, Coghill E, Perkins A, Jane SM, Cunningham JM. (2002) *Distinct domains of erythroid Kruppel-like factor modulate chromatin remodeling and transactivation at the endogenous beta-globin gene promoter*. Mol Cell Biol. 22(1):161-70.
217. Armstrong JA, Bieker JJ, Emerson BM. (1998) *A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF in vitro*. Cell. Oct 2;95(1):93-104.
218. Lee CH, Murphy MR, Lee JS, Chung JH. (1999) *Targeting a SWI/SNF-related chromatin remodeling complex to the beta-globin promoter in erythroid cells*. Proc Natl Acad Sci U S A. 96(22):12311-5.
219. Turner BM. (1999) *Chromatin a target for intracellular signalling pathways*. Semin Cell Dev Biol. 10(2):165-7.
220. Gordon S, Akopyan G, Garban H, Bonavida B. (2006) *Transcription factor YY1: structure, function, and therapeutic implications in cancer biology*. Oncogene. 25(8):1125-42. Review.
221. Wieschaus E, Nusslein-Volhard C, Kluding H. (1984) *Krüppel, a gene whose activity is required early in the zygotic genome for normal embryonic segmentation*. Dev Biol. 104(1):172-86.
222. Gordon S, Akopyan G, Garban H, Bonavida B. (2006) *Transcription factor YY1: structure, function, and therapeutic implications in cancer biology*. Oncogene; 25(8):1125-42. Review.
223. Thomas MJ, Seto E. (1999) *Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key?* Gene. 236(2):197-208. Review.
224. Yao YL, Yang WM, Seto E. (2001) *Regulation of transcription factor YY1 by acetylation and deacetylation*. Mol Cell Biol 21(17):5979-91.
225. Shi Y, Seto E, Chang LS, Shenk T. (1991) *Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A protein*. Cell. 67(2):377-88.

226. Shi Y, Lee JS, Galvin KM. (1997) *Everything you have ever wanted to know about Yin Yang 1.....* Biochim Biophys Acta. 1332(2):F49-66. Review.
227. Kaufman RM, Pham CT, Ley TJ. (1999) *Transgenic analysis of a 100-kb human beta-globin cluster-containing ADN fragment propagated as a bacterial artificial chromosome.* Blood. 94(9):3178-84.
228. Black AR, Black JD, Azizkhan-Clifford J. (2001) *Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer.* J Cell Physiol. 188(2):143-60. Review.
229. Marin M, Karis A, Visser P, Grosveld F, Philipsen S. (1997) *Transcription factor Sp1 is essential for early embryonic development but dispensable for cell growth and differentiation.* Cell. 89(4):619-28.
230. Gribnau J, Diderich K, Pruzina S, Calzolari R, Fraser P. (2000) *Intergenic transcription and developmental remodeling of chromatin subdomains in the human beta-globin locus.* Mol Cell. 5(2):377-86.
231. Recilla-Targa, F. (2000) *The chicken  $\alpha$ -& $\beta$  gene domains and their chromatin organization.* Cell & Moll. Biol. Lett.. 5:451-467.
232. Recillas-Targa F, V. De Moura GC, Huesca M, Scherrer K and Maycaud L. (1993) *Silencer and enhancer elements located at the 3'-side of the chicken and duck  $\alpha$ -globin-encoding gene domains.* Gene., 129: 229-237.
233. Razin SV, Petrov P, Hancock R. (1991) *Precise localization of the alpha-globin gene cluster within one of the 20- to 300-kilobase ADN fragments released by cleavage of chicken chromosomal ADN at topoisomerase II sites in vivo: evidence that the fragments are ADN loops or domains.* Proc Natl Acad Sci U S A. 88 (19):8515-9.
234. Bulger M, Schubeler D, Bender MA, Hamilton J, Farrell CM, Hardison RC, Groudine M. (2003) *A complex chromatin landscape revealed by patterns of nuclease sensitivity and histone modification within the mouse beta-globin locus.* Mol Cell Biol. 23(15):5234-44.
235. Rooks H, Bergounioux J, Game L, Close JP, Osborne C, Best S, Senior T, Height S, Thompson R, Hadzic N, Fraser P, Bolton-Maggs P, Thein SL. (2005) *Heterogeneity of the epsilon gamma delta beta-thalassaemias: characterization of three novel English deletions.* Br J Haematol. 128(5):722-9.
236. Wai AW, Gillemans N, Raguz-Bolognesi S, Pruzina S, Zafarana G, Meijer D, Philipsen S, Grosveld F. (2003) *HS5 of the human beta-globin locus control region: a developmental stage-specific border in erythroid cells.* EMBO J; 22(17):4489-500.
237. Choi OR, Engel JD. (1986) *A 3' enhancer is required for temporal and tissue-specific transcriptional activation of the chicken adult beta-globin gene.* Nature. 323(6090):731-4.
238. Li Q, Barkess G, Qian H. (2006) *Chromatin looping and the probability of transcription.* Trends Genet. 22(4):197-202.
239. Hebbes TR, Clayton AL, Thorne AW, Crane-Robinson C. (1994) *Core histone hyperacetylation co-maps with generalized ADNse I sensitivity in the chicken beta-globin chromosomal domain.* EMBO J. 13(8):1823-30.
240. Lee CK, Shibata Y, Rao B, Strahl BD, Lieb JD. (2004) *Evidence for nucleosome depletion at active regulatory regions genome-wide.* Nat Genet. 36(8):900-5.
241. Jenuwein T, Allis CD. (2001) *Translating the histone code.* Science. 293(5532):1074-80. Review.
242. Sekinger EA, Moqtaderi Z, Struhl K. (2005) *Intrinsic histone-ADN interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast.* Mol Cell. 18(6):735-48.
243. Palstra RJ, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F, de Laat W. (2003) *The beta-globin nuclear compartment in development and erythroid differentiation.* Nat Genet. 35(2):190-4.
244. Guan F, Uboh CE, Soma LR, Luo Y, Jahr JS, Driessen B. (2004) *Confirmation and quantification of hemoglobin-based oxygen carriers in equine and human plasma by hyphenated liquid chromatography tandem mass spectrometry.* Anal Chem. 76(17):5127-35.
245. Clayton AL, Hazzalin CA, Mahadevan LC. (2006) *Enhanced histone acetylation and transcription: a dynamic perspective.* Mol Cell. 23(3):289-96. Review.
246. Knezetic JA, Felsenfeld G. (1989) *Identification and characterization of a chicken alpha-globin enhancer.* Mol Cell Biol. Mar;9(3):893-901.
247. Recillas Targa, F., Razin, S.V., De Moura Gallo, C.V., Marcaud, L. and Scherrer, K. (1995) *Silencer and enhancer elements and the framing structures of the chicken  $\alpha$ -globin gene domain.* In:

Stamatoyannopoulos G., Editor. Ninth Conference on Hemoglobin Switching. Orcas Islands, W.A: Intercept, Andover 203-215.

248. Annunziato AT, Seale RL. (1984) *Presence of nucleosomes within irregularly cleaved fragments of newly replicated chromatin*. Nucleic Acids Res. 12(15):6179-96.
249. Workman JL. (2006) *Nucleosome displacement in transcription*. Genes Dev.;20(15):2009-17. Review.
250. Orkin SH. (1995) *Regulation of globin gene expression in erythroid cells*. Eur J Biochem. 231(2):271-81. Review.
251. Porcu, S., M. Kitamura, E. Witkowska, Z. Zhang, A. Mutero, C. Lin, J. Chang, and K. M. L. Gaensler. (1997). *The human beta globin locus introduced by YAC transfer exhibits a specific and reproducible pattern of developmental regulation in transgenic mice*. Blood 90:4602-4609.
252. Carter D, Chakalova L, Osborne CS, Dai YF, Fraser P. (2002) *Long-range chromatin regulatory interactions in vivo*. Nat Genet. 32(4):623-6.
253. Ling J, Baibakov B, Pi W, Emerson BM, Tuan D. (2005) *The HS2 enhancer of the beta-globin locus control region initiates synthesis of non-coding, polyadenylated RNAs independent of a cis-linked globin promoter*. J Mol Biol. 350(5):883-96.
254. K. Pawlik and T. Townes, (1995) *Autonomous, erythroid-specific ADNse I hypersensitive site formed by human beta-globin locus control region (LCR) 5' HS2 in transgenic mice*, Dev. Biol. 169, pp. 728–732.
255. Pawlik KM, Townes TM. (1995) *Autonomous, erythroid-specific ADNse I hypersensitive site formed by human beta-globin locus control region (LCR) 5' HS 2 in transgenic mice*. Dev Biol. 169(2):728-32.
256. Ling, L. Ainol, L. Zhang, X. Yu, W. Pi and D. Tuan (2004), *HS2 enhancer function is blocked by a transcriptional terminator inserted between the enhancer and the promoter*, J. Biol. Chem. 279, pp. 51704–51713.
257. Kim A, Zhao H, Ifrim I, Dean A. (2007) *Beta-globin intergenic transcription and histone acetylation dependent on an enhancer*. Mol Cell Biol. 27(8):2980-6.
258. Filippova GN, Cheng MK, Moore JM, Truong JP, Hu YJ, Nguyen DK, Tsuchiya KD, Disteche CM. (2005) *Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development*. Dev Cell. 8(1):31-42.
259. Kimura A, Horikoshi M. (2004) *Partition of distinct chromosomal regions: negotiable border and fixed border*. Genes Cells. 9(6):499-508. Review.
260. Cai H, Levine M. (1995) *Modulation of enhancer-promoter interactions by insulators in the Drosophila embryo*. Nature. 376(6540):533-6.
261. Scott K.S. and Geyer,P.K. (1995) *Effects of the su(Hw) insulator protein on the expression of the divergently transcribed Drosophila yolk protein genes*. EMBO J., 14:, 6258–6267.
262. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. (1997) *Crystal structure of the nucleosome core particle at 2.8 Å resolution*. Nature. 389(6648):251-60.
263. Bernstein E, Hake SB. (2006) *The nucleosome: a little variation goes a long way*. Biochem Cell Biol. Aug;84(4):505-17. Review.
264. Arney KL and Fisher AG, (2004) *Epigenetic aspects of differentiation*, J Cell Sci 117, pp. 4355–4363.
265. Kosak ST and Groudine M, (2004) *Form follows function: the genomic organization of cellular differentiation*, Genes Dev 18, pp. 1371–1384.
266. Bannister AJ, Schneider R and Kouzarides T, (2002) *Histone methylation: dynamic or static?*, Cell 109, pp. 801–806.
267. Lachner M, O'Carroll D, Rea S, Mechtler K and Jenuwein T, (2001) *Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins*, Nature 410, pp. 116–120.
268. Berger SL. (2007) *The complex language of chromatin regulation during transcription*. Nature. May 24;447(7143):407-12. Review.
269. Kouzarides T. (2007) *Chromatin modifications and their function*. Cell. 128(4):693-705. Review.
270. Chandy M, Gutierrez JL, Prochasson P, Workman JL. (2006) *SWI/SNF displaces SAGA-acetylated nucleosomes*. Eukaryot Cell. 5(10):1738-47.

271. Hassan AH, Awad S, Prochasson P. (2006) *The Swi2/Snf2 bromodomain is required for the displacement of SAGA and the octamer transfer of SAGA-acetylated nucleosomes.* J Biol Chem. 281(26):18126-34.
272. Ito T, Ikehara T, Nakagawa T, Kraus WL, Muramatsu M. (2000) *p300-mediated acetylation facilitates the transfer of histone H2A-H2B dimers from nucleosomes to a histone chaperone.* Genes Dev. 14 (15):1899-907.
273. Reinke H, Horz W. (2003) *Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter.* Mol Cell. 11(6):1599-607.
274. Zhao J, Herrera-Diaz J, Gross DS. (2005) *Domain-wide displacement of histones by activated heat shock factor occurs independently of Swi/Snf and is not correlated with RNA polymerase II density.* Mol Cell Biol. 25(20):8985-99.
275. Vakoc, C. R., Mandat, S. A., Olenchok, B. A. & Blobel, G. A. (2005) *Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin.* Mol Cell 19, 381–391.
276. Sullivan, B. A. & Karpen, G. H. (2004) *Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin.* Nature Struct. Mol. Biol. 11, 1076–1083.
277. Nowak SJ, Corces VG. (2004) *Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation.* Trends Genet. 20(4):214-20. Review.
278. Strahl BD, Allis CD. (2000) *The language of covalent histone modifications.* Nature. 403(6765):41-5.
279. Turner BM. (2000) *Histone acetylation and an epigenetic code.* Bioessays. 22(9):836-45. Review.
280. Heard E, Disteché CM. (2006) *Dosage compensation in mammals: fine-tuning the expression of the X chromosome.* Genes Dev. 20(14):1848-67. Review.
281. Sterner DE, Berger SL. (2000) *Acetylation of histones and transcription-related factors.* Microbiol Mol Biol Rev. 64(2):435-59. Review.
282. Zhang and Reinberg, (2006) *Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails,* Genes Dev. 15 pp. 2343–2360.
283. Shilatifard A, (2006) *Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression,* Annu. Rev. Biochem. 75, pp. 243–269.
284. Nathan D, Ingvarsdottir K, D.E. Sterner, G.R. Bylebyl, M. Dokmanovic, J.A. Dorsey, K.A. Whelan, M. Krsmanovic, W.S. Lane and P.B. Meluh (2006) *Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications,* Genes Dev. 20, pp. 966–976.
285. Hassa PO, S.S. Haenni, M. Elser and M.O. Hottiger, (2006) *Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going?,* Microbiol. Mol. Biol. Rev. 70, pp. 789–829.
286. Nelson NJ, H. Santos-Rosa and T. Kouzarides, (2006) *Proline isomerization of histone H3 regulates lysine methylation and gene expression,* Cell 126, pp. 905–916.
287. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y. (2004) *Histone demethylation mediated by the nuclear amine oxidase homolog LSD1.* Cell. 119(7):941-53.
288. Lee MG, Wynder C, Cooch N, Shiekhhattar R. (2005) *An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation.* Nature. 437(7057):432-5.
289. Christensen J, Agger K, Cloos PA, Pasini D, Rose S, Sennels L, Rappsilber J, Hansen KH, Salcini AE, Helin K. (2007) *RBP2 belongs to a family of demethylases, specific for tri- and dimethylated lysine 4 on histone 3.* Cell. 128(6):1063-76.
290. Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A, Fundele, R., (2000). *Active demethylation of the paternal genome in the mouse zygote.* Curr. Biol. 10: 475–478.
291. Kirillov, A., Kistler, B., Mostoslavsky, R., Cedar, H., Wirth, T., and Bergman, Y. (1996). *A role for nuclear NF-kappaB in B-cell-specific demethylation of the Igekappa locus.* Nat. Genet. 13: 435–441
292. Paroush, Z., Keshet, I., Yisraeli, J., and Cedar, H. (1990). *Dynamics of demethylation and activation of the alpha-actin gene in myoblasts.* Cell, 63: 1229–1237.
293. Lucarelli, M., Fuso, A., Strom, R., and Scarpa, S. (2001). *The dynamics of myogenin site-specific demethylation is strongly correlated with its expression and with muscle differentiation.* J. Biol. Chem. 276: 7500–7506.

294. Bruniquel, D., and Schwartz, R.H. (2003). *Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process*. Nat. Immunol. 4: 235–240.
295. Kersh, E.N., Fitzpatrick, D.R., Murali-Krishna, K., Shires, J., Speck, S.H., Boss, J.M., and Ahmed, R. (2006). *Rapid demethylation of the IFN- $\gamma$  gene occurs in memory but not naïve CD8 T Cells*. J. Immunol. 176: 4083–4093.
296. Bhattacharya, S.K., Ramchandani, S., Cervoni, N., and Szyf, M. (1999). *A mammalian protein with specific demethylase activity form CpG ADN*. Nature, 397: 579–583.
297. Cervoni, N., and Szyf, M. (2001). *Demethylase activity is directed by histone acetylation*. J. Biol. Chem. 276: 40 778 – 40 787.
298. Detich, N., Theberge, J., and Szyf, M. (2002). *Promoter-specific activation and demethylation by MBD2/demethylase*. J. Biol. Chem. 277: 35 791 – 35 794.
299. Jin J, Cai Y, Li B, Conaway RC, Workman JL, Conaway JW, Kusch T. (2005) *In and out: histone variant exchange in chromatin*. Trends Biochem Sci. 30(12):680-7. Review.
300. Akey CW and K. Luger, (2003) *Histone chaperones and nucleosome assembly*, Curr. Opin. Struct. Biol. 13, pp. 6–14
301. Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y. (2004) *Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of ADN synthesis*. Cell. Jan 9;116(1):51-61.
302. Bruno M, Flaus A, Stockdale C, Rencurel C, Ferreira H, Owen-Hughes T. (2003) *Histone H2A/H2B dimer exchange by ATP-dependent chromatin remodeling activities*. Mol Cell 12(6):1599-606.
303. Malik HS, Henikoff S. (2003) *Phylogenomics of the nucleosome*. Nat Struct Biol. 10(11):882-91. Review.
304. Celeste A, Fernandez-Capetillo O, Kruhlak MJ, Pilch DR, Staudt DW, Lee A, Bonner RF, Bonner WM, Nussenzweig A. (2003) *Histone H2AX phosphorylation is dispensable for the initial recognition of ADN breaks*. Nat Cell Biol. 5(7):675-9.
305. Hake SB, Allis CD. (2006) *Histone H3 variants and their potential role in indexing mammalian genomes: the "H3 barcode hypothesis"*. Proc Natl Acad Sci U S A. 103 (17):6428-35. Review.
306. Leffak M. (1988) *Nonrandom assembly of chromatin during hydroxyurea inhibition of ADN synthesis*. Biochemistry. 27(2):686-91.
307. Henikoff S, Furuyama T, Ahmad K. (2004) *Histone variants, nucleosome assembly and epigenetic inheritance*. Trends Genet. 20(7):320-6. Review.
308. Muller J, Kassis JA. (2006) *Polycomb response elements and targeting of Polycomb group proteins in Drosophila*. Curr Opin Genet Dev. 16(5):476-84. Review.
309. Raich N, Clegg CH, Grofti J, Romeo PH, Stamatoyannopoulos G (1995). *GATA1 and YY1 are developmental repressors of the human epsilon-globin gene*. EMBO J. 14(4):801-9.
310. Beug, H., A. von Kirchbach, G. Doderlein, J. F. Conscience, and T. Graf. (1979). *Chicken hematopoietic cells transformed by seven strains of defective avian leukemia viruses display three distinct phenotypes of differentiation*. Cell 18:375-390.
311. Beug, H., S. Palmieri, C. Freudenstein, H. Zentgraf, and T. Graf. (1982). *Hormone-dependent terminal differentiation in vitro of chicken erythroleukemia cells transformed by ts mutants of avian erythroblastosis virus*. Cell 28:907-919.
312. Feng YQ, Seibler J, Alami R, Eisen A, Westerman KA, Leboulch P, Fiering S, Bouhassira EE. (1999) *Site-specific chromosomal integration in mammalian cells: highly efficient CRE recombinase-mediated cassette exchange*. J Mol Biol. 292(4):779-85.
313. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES. (2006) *A bivalent chromatin structure marks key developmental genes in embryonic stem cells*. Cell. 125(2):315-26.
314. Jáuregui R, Abreu-Goodger C, Moreno-Hagelsieb G, Collado-Vides J, Merino E. (2003) *Conservation of ADN curvature signals in regulatory regions of prokaryotic genes*. Nucleic Acids Res. 31(23):6770-7.
315. Kozobay-Avraham L, Hosid S, Bolshoy A. (2006) *Involvement of ADN curvature in intergenic regions of prokaryotes*. Nucleic Acids Res. 34(8):2316-27.
316. Singal R, vanWert JM, Ferdinand L Jr. (2002) *Methylation of alpha-type embryonic globin gene alpha pi represses transcription in primary erythroid cells*. Blood. 100(12):4217-22.



317. Minie ME, Kimura T, Felsenfeld G. (1992) *The developmental switch in embryonic rho-globin expression is correlated with erythroid lineage-specific differences in transcription factor levels.* *Development.* 115(4):1149-64.
318. Owen-Hughes T, Workman JL. (1996) *Remodeling the chromatin structure of a nucleosome array by transcription factor-targeted trans-displacement of histones.* *EMBO J.* 15(17):4702-12.
319. Dingwall AK, Beek SJ, McCallum CM, Tamkun JW, Kalpana GV, Goff SP, Scott MP. (1995) *The Drosophila snr1 and brm proteins are related to yeast SWI/SNF proteins and are components of a large protein complex.* *Mol Biol Cell.* 6(7):777-91.