

4.3 CAPÍTULO 3: Interacción con el ADN de la proteína de la familia PHD finger HAT3.1. Estudio de la especificidad de unión de un homeodominio particular con histidina en la posición 51.

4.3.1 Introducción

La función reguladora de la transcripción llevada a cabo por las proteínas con homeodominio depende de la especificidad de la interacción de este motivo con el ADN. Como se describió en la Introducción general, aunque los homeodominios exhiben cierta variabilidad en la secuencia reconocida, la mayoría de los casos se ajusta a un modelo donde la secuencia unida es del tipo:



Sintéticamente, esta secuencia blanco se define por el contacto del residuo Asn51, casi absolutamente conservado en los homeodominios descritos, con la A de la tercera posición (A₃). Una vez establecido el núcleo en torno a esta posición, puede completarse un esquema en el cual las dos primeras posiciones son contactadas por el brazo N-terminal, mientras que la cuarta es unida por el residuo 47 altamente conservado, isoleucina o valina (Gehring y col., 1994). Las posiciones 5 y 6 son variables (representadas como N₅ y N₆) y son reconocidas por los aminoácidos 50 y 54, menos conservados en los distintos homeodominios (Laughon, 1991; Wilson y col., 1996). En la Figura 4 de la Introducción general se encuentran esquematizadas estas interacciones.

Las proteínas con homeodominio se agrupan en distintas familias. Una de ellas, la familia PHD finger, es exclusiva de plantas y contiene un motivo rico en cisteínas que, si bien se ha encontrado en otros organismos eucariotas (Bienz, 2006), sólo aparece asociado al homeodominio en plantas.

En *Arabidopsis* existen dos proteínas con homeodominio de la familia PHD finger: HAT3.1 (Schindler y col., 1993) y PRHA (Plesch y col., 1997). También se han encontrado miembros de esta familia, como las proteínas ZmHox de maíz (Klinge y col., 1996; Uberlacker y col., 1996) y PRHP de perejil (Korfhage y col., 1994), en otras plantas. En la figura 32 se muestra una comparación de la secuencia de aminoácidos del homeodominio de HAT3.1 con

las de diferentes miembros de la familia PHD finger. Como homeodominio modelo se utilizó el de Antennapedia. Como puede observarse, a diferencia de otras familias, los miembros de este grupo presentan una característica inusual: difieren notoriamente en la secuencia del homeodominio, especialmente en los residuos que normalmente contactan el ADN. Sin embargo, presentan conservadas las posiciones que forman el bolsillo hidrofóbico de manera que puede asumirse que se plegarían adoptando una estructura tridimensional similar a la de otros homeodominios (Figura 32).

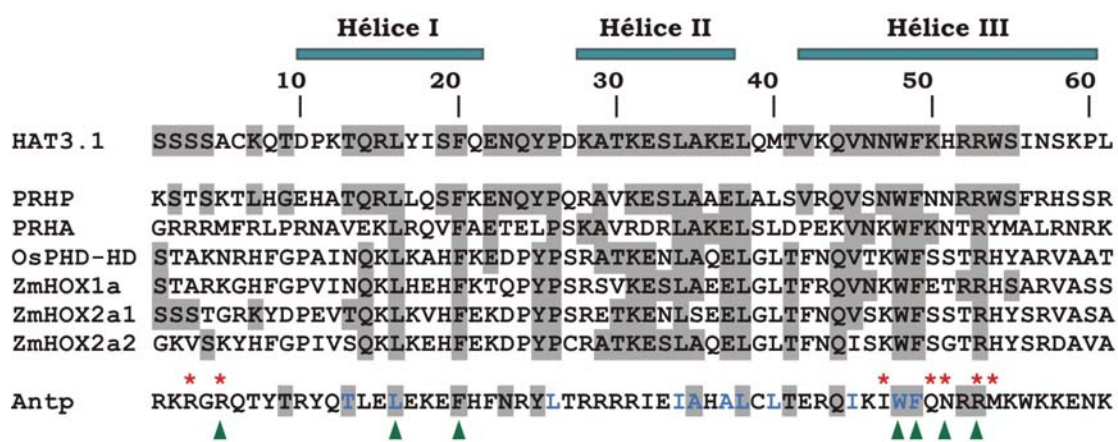


Figura 32. Alineamiento de la secuencia de aminoácidos del homeodominio de HAT3.1 con otros homeodominios. Se realiza una comparación de HAT3.1 con miembros de la familia PHD finger [PRHP (perejil), ID L21975; PRHA (Arabidopsis), ID NM_119140; OsPHD-HD (arroz), ID NM_001063749; ZmHOX1a y ZmHOX2a (maíz), ID X67561 y X89760] y con Antennapedia. Sombreados en gris se muestran los residuos presentes en HAT3.1 y al menos uno de los demás homeodominios. Los números en la parte superior indican la posición dentro del homeodominio y las barras señalan las tres hélices α . Las flechas verdes muestran las posiciones que son invariables en la mayoría de los homeodominios y los asteriscos rojos indican los aminoácidos normalmente involucrados en la interacción con el ADN. En celeste aparecen los residuos que forman el bolsillo hidrofóbico de Antennapedia.

A partir de este análisis surge rápidamente como interrogante si los homeodominios de las proteínas PHD finger son capaces de comportarse como motivos de unión a ADN y, de ser así, si reconocen secuencias específicas. Al menos para ZmHox2a este parece ser el caso (Kirch y col., 1998). Esta proteína contiene dos homeodominios funcionales, HD1 y HD2, que reconocen las secuencias TCCT y GATC, respectivamente, que difieren del núcleo TAAT. Sin embargo, para el homeodominio de HAT3.1 no pudo demostrarse una interacción específica con el ADN (Schindler y col., 1993). Como se describió en la Introducción general, estos autores llevaron a cabo varios experimentos de los cuales concluyeron que esta proteína reconocería fragmentos de ADN mayores a 100 pb, independientemente de la secuencia de

nucleótidos. Asimismo, encontraron que para que la unión ocurra de manera eficiente era necesaria la presencia del motivo PHD finger (Schindler y col., 1993).

Resulta interesante pensar que las características particulares que presentan los homeodominios de las proteínas PHD finger cumplirían un papel esencial en la función de estos factores de transcripción. En consecuencia, nos propusimos caracterizar la interacción con el ADN de HAT3.1 y los resultados obtenidos se presentan en este capítulo.

En primer lugar, estudiamos la capacidad del homeodominio de HAT3.1 de reconocer, tanto *in vitro* como *in vivo*, una secuencia de ADN de manera específica. Por otra parte, a fin de analizar si la unión ocurre a través mecanismos conservados en otros homeodominios, identificamos los aminoácidos que participan de la interacción y los nucleótidos de la secuencia de ADN blanco que son contactados.

Para determinar si HAT3.1 interacciona con secuencias de ADN específicas utilizamos la técnica de selección de oligonucleótidos de secuencia al azar empleada al comienzo de este trabajo de tesis. Además, estudiamos la capacidad de esta proteína de interactuar con su secuencia blanco *in vivo*, en ensayos en *Saccharomyces cerevisiae*. A continuación, caracterizamos la unión al ADN mediante estudios de pérdida de nucleósido por ataque con radicales hidroxilo (interferencia). Finalmente, hemos analizado en detalle la importancia que poseen los residuos de aminoácidos que participan en la unión en otros homeodominios en la determinación de la especificidad de unión al ADN.

4.3.2 Resultados

4.3.2.1 Análisis de la capacidad del homeodominio de HAT3.1 de discriminar específicamente secuencias de ADN a partir de una población de oligonucleótidos de 50 pb de longitud

Para estudiar la capacidad de unión específica al ADN de HAT3.1, decidimos expresar el homeodominio de esta proteína como producto de fusión con GST en *E. coli* DH5 α . Para esto, una región que codifica el homeodominio y secuencias adyacentes (aminoácidos 542 a 660) se amplificó por PCR y se clonó en los sitios *Bam*HI y *Eco*RI del vector pGEX-3X. Como

cebadores se utilizaron los oligonucleótidos HAT31-1 y HAT31-2, detallados en la tabla 3 de la sección Materiales y Métodos.

La proteína recombinante se denominó HAT3.1_{HD} y, una vez purificada, fue utilizada para realizar estudios de selección de secuencias de ADN específicas a partir de una población de 4¹² oligonucleótidos diferentes. Como se describe en Materiales y Métodos (sección 3.21) y siguiendo la estrategia de SELEX empleada para hallar la secuencia reconocida por ATH1, se realizaron ensayos de retardo en gel para separar los complejos HAT3.1_{HD}-ADN de los oligonucleótidos libres. A continuación, las moléculas de ADN de movilidad reducida (unidas) fueron amplificadas por PCR para ser usadas en nuevas rondas de selección.

Como se muestra en la Figura 33 (panel izquierdo), a medida que se fueron realizando los sucesivos pasos de selección, purificación y amplificación se produjo un aumento en la proporción de oligonucleótidos unidos por HAT3.1_{HD}. Esto demuestra que el homeodominio de HAT3.1 es capaz de interactuar específicamente con el ADN. De manera similar a lo descrito para ATH1_{HD}, la especificidad de la unión obtenida en cada ciclo se comprobó a través de ensayos de competencia realizados con un exceso molar de 20 veces de las diferentes mezclas de oligonucleótidos obtenidas sin marcar (no mostrado).

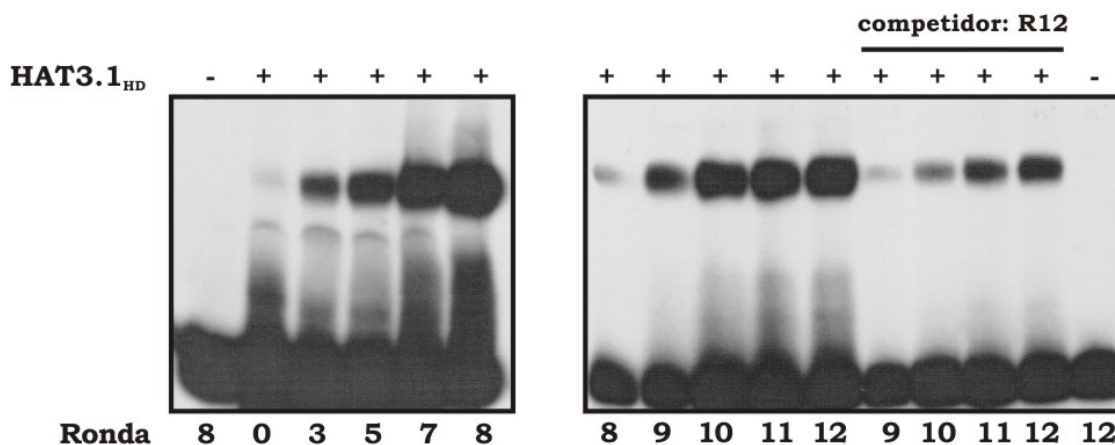


Figura 33. Unión de HAT3.1_{HD} a poblaciones de oligonucleótidos obtenidas en distintos ciclos de selección. Ensayos de retardo en gel realizados con 200 ng (panel izquierdo) y 20 ng (panel derecho) de HAT3.1_{HD}. Los números en cada calle indican el paso de selección al que corresponde la mezcla de oligonucleótidos empleada en la reacción de unión.

A partir del décimo paso de selección no se observaron diferencias en la relación ADN unido/ADN libre con respecto al paso anterior y se pensó que a partir de este ciclo la unión

permanecía constante (Figura 33, panel derecho). Sin embargo, cuando se realizaron los ensayos de competencia con mezclas de oligonucleótidos no marcados de las distintas rondas, pudo deducirse que las moléculas de ADN de los ciclos once y doce fueron unidas con mayor afinidad (Figura 33). Por consiguiente, se decidió clonar la mezcla de oligonucleótidos del decimosegundo paso en el vector pCR 2.1-TOPO (Invitrogen) y se eligieron al azar 28 clones para ser secuenciados.

4.3.2.2 Determinación de la secuencia consenso unida por HAT3.1_{HD}

En la Figura 34 se muestran las secuencias de los 28 clones analizados. La parte variable de cada oligonucleótido aparece en letra mayúscula y en minúscula se muestran los brazos conservados.

Del alineamiento de las regiones variables de 12 pb de cada oligonucleótido se deduce que la secuencia consenso es ACCATCTAAACC (donde las posiciones subrayadas aparecen en al menos 27 de los 28 clones obtenidos). Además, esta secuencia está presente siempre en la misma hebra de ADN respecto de los brazos fijos. Esto sugiere que las regiones conservadas de los oligonucleótidos podrían haber influido en el proceso de selección, probablemente porque formaron parte de la secuencia unida por la proteína. Sin embargo, debe descartarse la posibilidad de que exista una parte importante del sitio de unión en los brazos invariables, dado que la proporción de ADN unido en los primeros pasos de selección fue muy baja (Figura 33).

Cuando el análisis de las secuencias obtenidas se extendió a los nucleótidos que forman los brazos conservados, resultó evidente que las moléculas seleccionadas contenían dos copias en tandem de la secuencia 5'-TRMACCA-3' (R=A o G; M=A o C). En función de esto, puede pensarse que cada una de estas secuencias corresponde al sitio de unión de una molécula de HAT3.1_{HD} y que existiría una clara preferencia por A sobre C en la tercera posición (TRAACCA) cuando este nucleótido es variable (Figura 34).

Debe destacarse que la secuencia de ADN consenso de HAT3.1_{HD}, T(A/G)(A/C)ACCA, no contiene el núcleo TAAT reconocido por muchos homeodominios y difiere de la obtenida para los homeodominios de ZmHox2a (TCCT y GATC) (Kirch y col., 1998). Como hasta el

momento no se han reportado secuencias similares a este consenso como blanco de otros homeodominios, puede decirse que HAT3.1 presenta una especificidad de unión diferente.

C1	gaattctacagtggctgc	ACCAAGTGAACC	attgtccaggaagcttcac
C2	gaattctacagtggctgc	ACCAGGTGAACC	attgtccaggaagcttcac
C3	gaattctacagtggctgc	ACCATTGAACC	attgtccaggaagcttcac
C4	gaattctacagtggctgc	ACCATCTAAACC	attgtccaggaagcttcac
C7	gaattctacagtggctgc	ACCACCTAAACC	attgtccaggaagcttcac
C8	gaattctacagtggctgc	ACCATCTAAACC	attgtccaggaagcttcac
C9	gaattctacagtggctgc	ACCATCTAAACC	attgtccaggaagcttcac
C12	gaattctacagtggctgc	ACCAGCTGAACC	attgtccaggaagcttcac
C13	gaattctacagtggctgc	ACCATCTAAACC	attgtccaggaagcttcac
C14	gaattctacagtggctgc	ACCAAGTGAACC	attgtccaggaagcttcac
C15	gaattctacagtggctgc	ACCACCTACACC	attgtccaggaagcttcac
C16	gaattctacagtggctgc	ACCAGCTAAACC	attgtccaggaagcttcac
C17	gaattctacagtggctgc	ACCGCCTAAACC	attgtccaggaagcttcac
C18	gaattctacagtggctgc	ACCAATTGAACC	attgtccaggaagcttcac
C20	gaattctacagtggctgc	ACCACTTAAACC	attgtccaggaagcttcac
C21	gaattctacagtggctgc	ACCAAGTGAACC	attgtccaggaagcttcac
C22	gaattctacagtggctgc	ACCAAGTGAACC	attgtccaggaagcttcac
C49	gaattctacagtggctgc	ACCATGTGAACC	attgtccaggaagcttcac
C50	gaattctacagtggctgc	ACCATCTAAACC	attgtccaggaagcttcac
C51	gaattctacagtggctgc	ACCAGGTGAACC	attgtccaggaagcttcac
C52	gaattctacagtggctgc	ACCACGTAAACC	attgtccaggaagcttcac
C56	gaattctacagtggctgc	ACCATCTAAACC	attgtccaggaagcttcac
C57	gaattctacagtggctgc	ACCAACTAAACC	attgtccaggaagcttcac
C58	gaattctacagtggctgc	ACCATCTAAACC	attgtccaggaagcttcac
C59	gaattctacagtggctgc	ACCATGTAAACC	attgtccaggaagcttcac
C60	gaattctacagtggctgc	ACCAACTAAACC	attgtccaggaagcttcac
C61	gaattctacagtggctgc	ACCAATTAAACC	attgtccaggaagcttcac
C62	gaattctacagtggctgc	ACCATCTAAACC	attgtccaggaagcttcac

CONSENSO	Sitio de unión 1							Sitio de unión 2								
	T	G	C	A	C	C	A	(T/A)	C	T	A	A	A	C	C	A
A				28	-	-	27	8	-	-	18	27	28	-	-	
G				-	-	-	1	4	9	-	10	-	-	-	-	
C				-	28	28	-	5	15	-	-	1	-	28	28	
T				-	-	-	-	11	4	28	-	-	-	-	-	
%				100	100	100	96	(39/29)	54	100	64	96	100	100	100	

Figura 34. Alineamiento de las secuencias de 28 clones obtenidos luego del clonado de la población de oligonucleótidos seleccionada por HAT3.1_{HD} correspondiente a la decimosegunda ronda. La parte central de 12 pb aparece en letra mayúscula. Los brazos conservados se muestran en minúscula. A la izquierda se indica el número de cada clon (arbitrario). Los nucleótidos que coinciden con el consenso están en negrita. La tabla en la parte inferior muestra la frecuencia de nucleótidos en cada posición y las secuencias consenso resultantes.

4.3.2.3 Estudio de la capacidad de HAT3.1 de unir la secuencia de ADN seleccionada *in vivo* en levaduras

Una vez identificada la secuencia de ADN unida por el homeodominio de HAT3.1 *in vitro* se examinó la capacidad del mismo para reconocer, de manera específica, esta secuencia en el contexto del núcleo de una célula eucariota. Para llevar esto a cabo decidimos emplear el sistema de simple híbrido en levaduras (Y1H) desarrollado inicialmente por Li y col. (1993).

Por un lado, un fragmento de ADN que codifica una región de HAT3.1 que contiene el homeodominio (aminoácidos 344 a 660) se clonó en fase con la secuencia que codifica el dominio de activación de GAL4 (AD) en el vector pGADT7 (sección 3.28.2, Materiales y Métodos). El plásmido resultante se denominó HAT3.1-AD.

Por otra parte, tres copias de un oligonucleótido que contiene la secuencia blanco de HAT3.1 se clonaron en tandem en el vector pHIS3NX (Meijer y col., 1998) de acuerdo a la estrategia descrita en Materiales y Métodos (sección 3.28.1). A continuación, la construcción que contenía este inserto y el gen reportero *HIS3* fusionado a un promotor mínimo (BS_{x3} :*HIS3*) se transfirió al vector pINT1, para permitir su inserción en el genoma de la cepa Y187 de *S. cerevisiae* (Gietz et al., 1992). La cepa resultante se denominó BS_{x3} :*HIS3*-Y187.

Como se ha detallado en Materiales y Métodos, en el sistema de Y1H empleado la interacción entre la proteína de fusión y el ADN produce la activación transcripcional del gen reportero *HIS3*. Este evento permite a la levadura completar la síntesis de histidina y sobrevivir en ausencia de este aminoácido. Por lo tanto, para determinar la expresión del gen *HIS3* en la cepa obtenida, se realizaron diluciones seriales a partir de un cultivo inicial de células que fueron plaqueadas en medio de cultivo mínimo en presencia, o ausencia, de histidina (sección 3.28.3, Materiales y Métodos).

Como se observa en la Figura 35A, la cepa BS_{x3} :*HIS3*-Y187 transformada con HAT3.1-AD fue capaz de crecer en medio de cultivo carente de histidina, pero esto no ocurrió con la cepa transformada con el plásmido pGADT7 vacío. Este resultado indica que HAT3.1 es capaz de unirse al promotor BS_{x3} y, gracias a la presencia del AD de GAL4, activar la expresión del gen reportero.

Para analizar la especificidad de la interacción detectada se utilizaron las cepas *HIS3*-Y187, que no posee la secuencia de unión de HAT3.1, y BS^* :*HIS3*-Y187, que contiene tres

copias de la secuencia blanco de ATH1, no unida *in vitro* por HAT3.1 (no mostrado). Como se muestra en la Figura 35A, estas cepas (transformadas con el plásmido HAT3.1-AD) fueron incapaces de crecer en medio de cultivo sin histidina. Además, no se detectó unión al sitio BS cuando se empleó la mutante HAT3.1_{HD}-K50A fusionada al AD (K50A-AD en la Figura 35) en lugar de HAT3.1-AD. Como se verá más adelante, esta mutante es incapaz de unir ADN de manera eficiente *in vitro*.

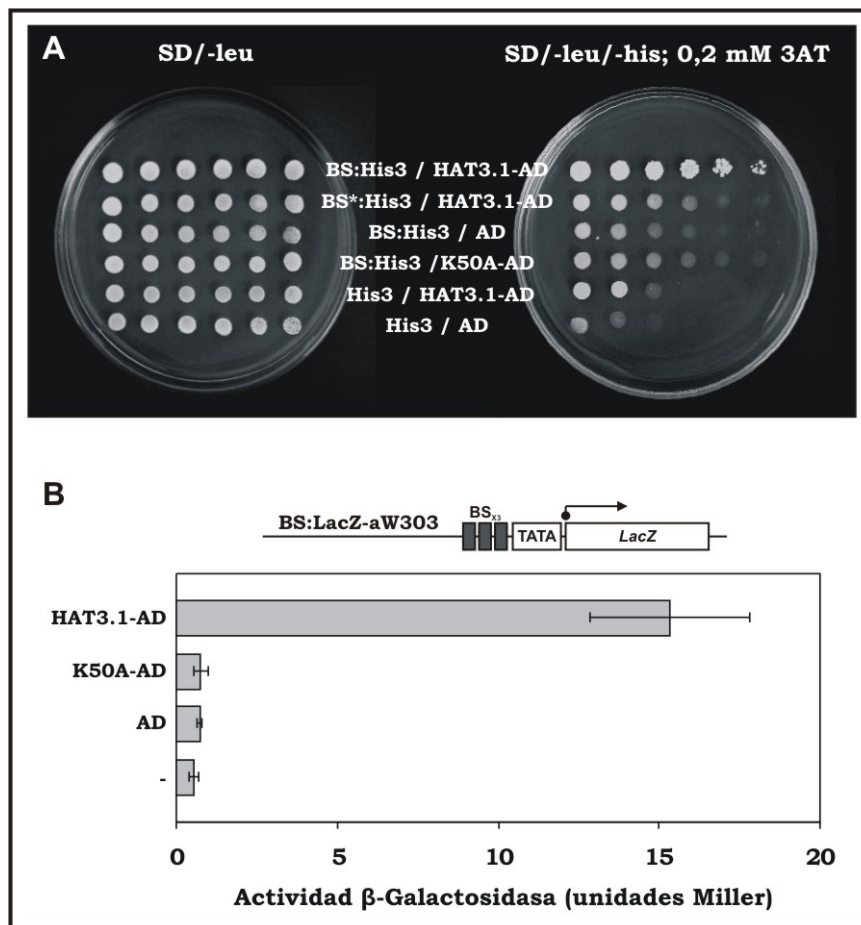


Figura 35. Interacción de HAT3.1 con su secuencia blanco *in vivo* en levadura. **A.** Ensayo de simple híbrido realizado con cepas Y187 de *S. cerevisiae* que contenían las construcciones: promotor mínimo:His3 (HIS3:Y187), promotor mínimo precedido por la secuencia blanco de HAT3.1 (BS:HIS3-Y187), o la secuencia reconocida por ATH1 (BS*:HIS3-Y187). Estas cepas fueron transformadas con plásmidos que permiten expresar un fragmento que contiene el homeodominio de HAT3.1 nativo o mutado en el residuo 50 fusionados al AD de GAL4 (HAT3.1-AD y K50A-AD, respectivamente), o solamente el AD (AD). Diferentes diluciones de cultivos crecidos hasta una DO₆₀₀ de 0,5-1 fueron sembradas en placas con medio mínimo selectivo que contenía, o no, histidina (fotos izquierda y derecha, respectivamente). **B.** Ensayo de actividad β -galactosidasa de la cepa de levadura aW303 que contenía, integrada en el locus *URA*, la secuencia blanco de HAT3.1 fusionada al gen *LacZ* (BS:LacZ-aW303). Estas células se transformaron con los plásmidos descriptos en (A). Para todos los casos se muestra el valor promedio (+/- DS) obtenido de tres ensayos independientes.

A continuación, se realizaron ensayos de Y1H utilizando como reportero el gen *LacZ*. La estrategia desarrollada para la obtención de las construcciones necesarias y la metodología empleada para la medición de la actividad β -galactosidasa se encuentran descritas en Materiales y Métodos (secciones 3.28.1 y 3.28.4, respectivamente). Los resultados obtenidos señalan que las células que expresan HAT3.1 fusionada al AD presentan un incremento de diez veces en la actividad β -galactosidasa respecto de las células que expresan el AD solo, o fusionado a la mutante HAT3.1_{HD}-K50A (Figura 35B).

En conjunto, los resultados obtenidos indican que el homeodominio de HAT3.1 es capaz de interactuar *in vivo* de manera específica con la secuencia consenso obtenida a partir de los ensayos *in vitro*.

4.3.2.4 Análisis de la unión de HAT3.1 a su secuencia de ADN consenso mediante ensayos de interferencia en la unión por ataque con radical hidroxilo

Con el objeto de caracterizar con más detalle la unión del homeodominio de HAT3.1 a la secuencia blanco obtenida en los experimentos de selección, se realizaron ensayos de interferencia en la unión por pre-tratamiento del ADN con radicales hidroxilo. Como se ha descrito, en esta técnica se emplea un oligonucleótido doble cadena marcado en un extremo y tratado con radical hidroxilo a fin de generar una población de moléculas de ADN con pérdidas de nucleósidos en posiciones únicas de cada hebra. De esta manera pueden determinarse cuáles son los nucleótidos necesarios para la unión y establecer las contribuciones individuales de cada nucleótido a la estabilidad global del complejo proteína-ADN (Dixon y col., 1991; Draganescu y col., 1995).

Para realizar este ensayo se empleó la proteína recombinante HAT3.1_{HD} en reacciones de unión con el oligonucleótido BS como se describió en la sección 3.25. El fragmento de ADN empleado deriva del clon c4 y fue preparado según se detalla en la sección 3.19 de Materiales y Métodos.

En la Figura 36 se muestra el resultado obtenido al analizar el patrón de moléculas que fueron unidas por HAT3.1_{HD} respecto de las que permanecieron libres. Cuando se utilizó el oligonucleótido marcado en la hebra que contiene ACCA (denominada hebra superior) se vio que los nucleótidos que conforman ambas secuencias TRMACCA son necesarios para la unión.

Sin embargo, la pérdida del dinucleótido TC localizado entre estas dos secuencias no afectó la interacción de la proteína con el ADN. De manera similar, el dinucleótido GC del brazo conservado (GCACCAAGTGAACCA) tampoco resultó importante para la unión. Al realizar un examen detallado del patrón de interferencia obtenido se observa que las posiciones relativamente más importantes en cada secuencia de unión son TGCACCA en la región 5' y TAAACCA en la región 3'.

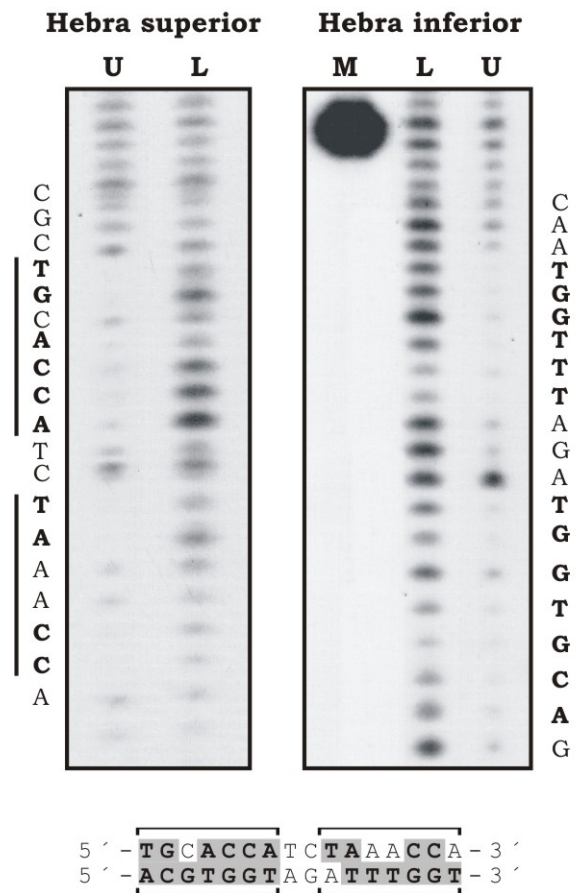


Figura 36. Experimentos de interferencia en la unión de HAT3.1_{HD} al oligonucleótido BS, que contiene la secuencia consenso 5'-TGCACCATCTAAACCA-3', por ataque con radical hidroxilo. Como marcador de posición (M) se utilizó el oligonucleótido marcado digerido con enzimas de restricción (Materiales y Métodos, 3.25). Para cada ensayo se indica la secuencia de ADN de la hebra correspondiente. En la parte inferior se muestra un esquema del sitio de unión de BS1 donde las posiciones contactadas aparecen sombreadas y en negrita. B, ADN unido; F, ADN libre.

En cuanto a la hebra inferior, el patrón de interferencia resultó más uniforme y se observó que las regiones complementarias a las dos secuencias RMACCA son necesarias para la interacción con HAT3.1_{HD} (Figura 36, panel derecho). Además, el nucleótido A

complementario a T en la secuencia de unión de la región 5' de BS (TGCACCA) también interfirió en la formación del complejo proteína-ADN al ser eliminado.

En conclusión, los patrones de interferencia observados con las dos hebras de ADN refuerzan la idea de que el sitio de unión obtenido en los ensayos de selección representa una duplicación en tandem de la secuencia TRMACCA.

4.3.2.5 Caracterización de la unión de HAT3.1 a la secuencia TGCACCATCTAAACCA

La presencia de dos regiones similares en la secuencia consenso obtenida para HAT3.1 podría indicar que ambas secuencias son requeridas para la unión por parte de esta proteína. Para estudiar en más detalle si esto es así o si, por el contrario, una sola de estas regiones es suficiente para la unión, se decidió analizar el efecto de mutaciones en los dinucleótidos CC de cada una de las secuencias de reconocimiento presentes en BS.

En la Figura 37A se muestra un ensayo de retardo en gel en el que puede verse que la mutación de CC a TA en la secuencia 5' (oligonucleótido BS-8) afecta menos significativamente la unión de HAT3.1_{HD} que cuando se muta el par CC del sitio presente hacia el 3' (BS-9). En consecuencia, especulamos que la región 3' sería unida en forma preferencial por la proteína. Este hecho no resulta inesperado si se tiene en cuenta que la secuencia de la región 3' fue seleccionada en forma casi completa, mientras que la repetición localizada hacia el 5' está formada por tres nucleótidos del brazo conservado.

La presencia de dos secuencias de unión podría, además, indicar que HAT3.1_{HD} une ADN como dímero. La posible formación de dímeros fue analizada en un experimento de retardo en gel realizado con una mezcla de proteínas de diferente tamaño. En este ensayo, se realizó una reacción de unión al ADN con la proteína HAT3.1_{HD}-GST y la resultante del tratamiento de esta fusión con Factor Xa (sección 3.16.3 de Materiales y Métodos), a fin de observar bandas de retardo de movilidad intermedia como causa de la formación de dímeros de proteínas de distinto tamaño. Sin embargo, como lo indica la Figura 37B, solamente se observaron dos bandas correspondientes a la unión de cada proteína al ADN. Además, dado que el oligonucleótido BS-8, que contiene solamente un sitio TAAACCA intacto, fue unido eficientemente por HAT3.1_{HD}, concluimos que la interacción de un único homeodominio con el ADN es suficiente para que ocurra una unión específica.

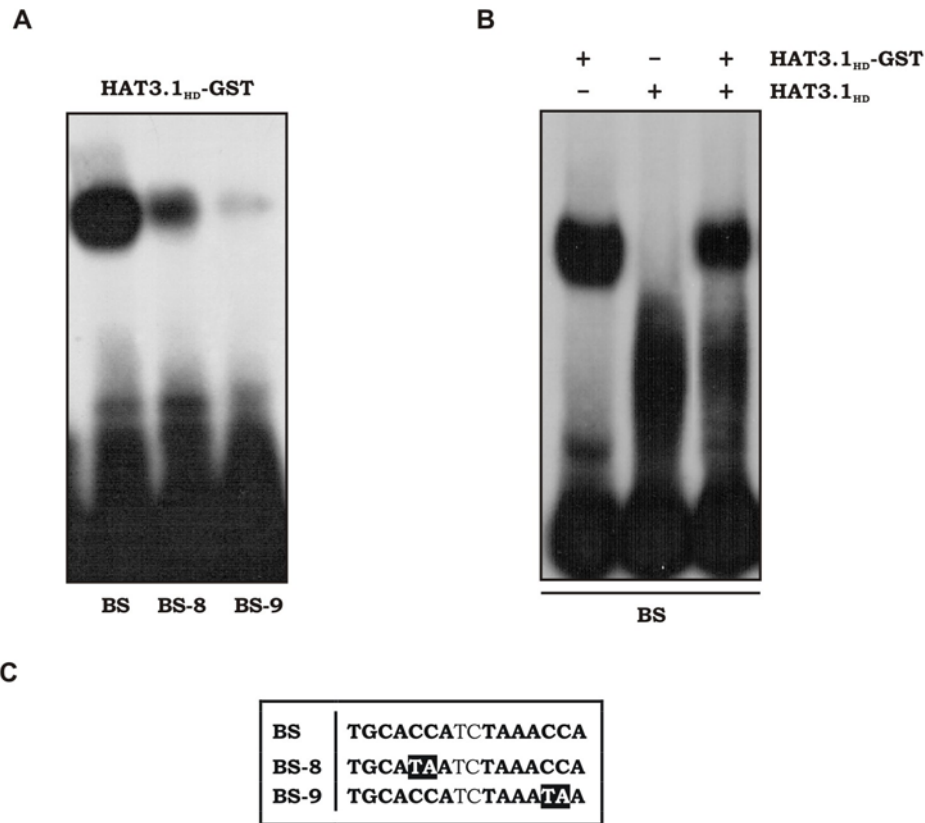


Figura 37. A. Unión de HAT3.1_{HD} a oligonucleótidos que contienen mutaciones en el dinucleótido CC. Ensayo de retardo en gel realizado con oligonucleótidos doble hebra sintéticos con cambios de CC a TA en cada secuencia ACCA. **B. Unión de dos formas de HAT3.1_{HD} al ADN.** En las reacciones de unión se empleó la proteína de fusión recombinante HAT3.1_{HD}-GST o la obtenida luego del clivaje de la misma con Factor Xa (HAT3.1_{HD}). **C. Secuencia de unión de los oligonucleótidos empleados en cada calle.**

4.3.2.6 Análisis de la especificidad de unión de HAT3.1_{HD} a la secuencia TAAACCA

Como el sitio de unión localizado en la región 3' de la secuencia consenso fue contactado de manera preferencial por HAT3.1_{HD}, nos enfocamos en la secuencia TAAACCA para estudiar la importancia de los distintos nucleótidos en la especificidad de unión al ADN de esta proteína.

Se diseñaron oligonucleótidos sintéticos con mutaciones puntuales en posiciones específicas, cuya secuencia de reconocimiento se muestra en la tabla de la Figura 38A (la secuencia completa se detalla en Materiales y Métodos, sección 3.5). Como se observa en el ensayo de retardo en gel de esta figura, el mayor efecto sobre la unión de HAT3.1_{HD} se observó cuando se modificaron cualquiera de las C de la secuencia ACCA (oligonucleótidos BS-5 y BS-

6). Este resultado es similar al obtenido con el oligonucleótido BS-9, que posee mutadas las dos C (CC→TA) del sitio TAAACCA. Los cambios en las posiciones 1 y 4 (oligonucleótidos BS-1 y BS-4, respectivamente) produjeron una disminución algo menor, pero significativa, de la unión. Mientras que la modificación de A a T en la posición 7 (BS-7) casi no afectó la unión al ADN de HAT3.1_{HD} (Figura 38A).

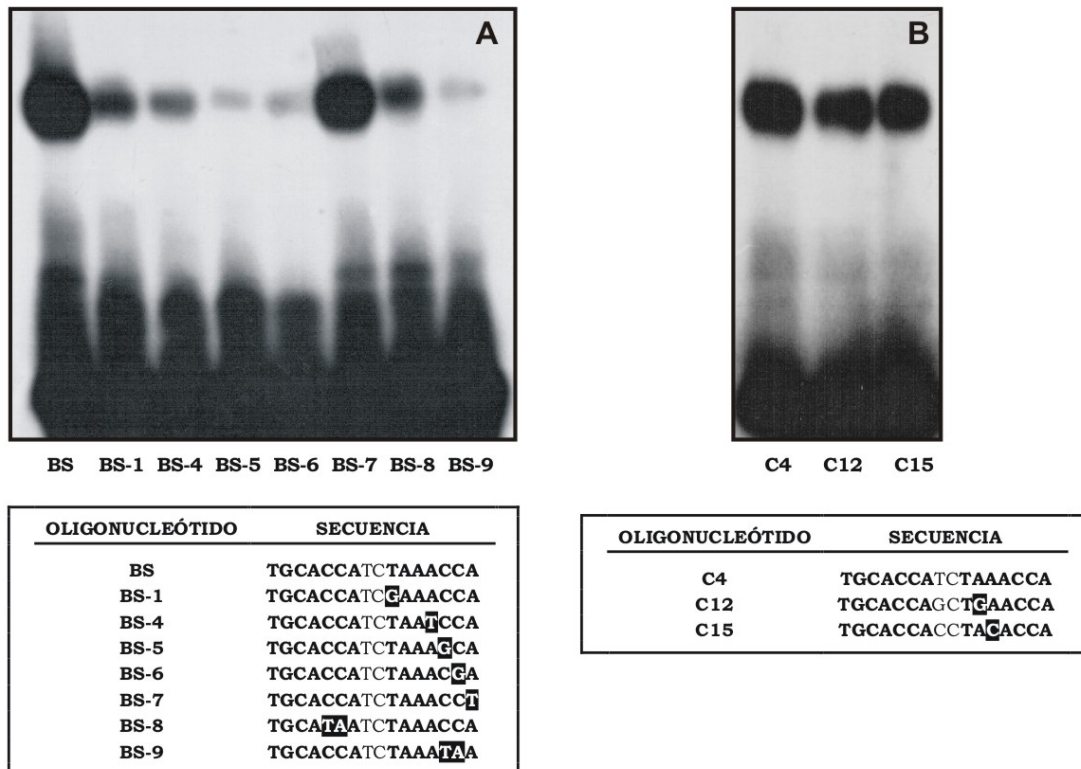


Figura 38. Unión de HAT3.1_{HD} a oligonucleótidos con modificaciones dentro de la secuencia consenso. A. Ensayo de retardo en gel realizado con oligonucleótidos sintéticos doble hebra con cambios puntuales en la secuencia TAAACCA (BS-1 a BS-7) y cambios de CC a TA en cada secuencia ACCA (BS-8 y BS-9). La tabla muestra la secuencia de unión de los oligonucleótidos empleados en cada calle. B. Experimento similar en el que se utilizaron oligonucleótidos doble cadena (44 nt) obtenidos de los clones C4, C12 y C15. La secuencia de estos fragmentos de ADN se indica en la parte inferior de la figura.

Para analizar el efecto de cambios en los nucleótidos de las posiciones 2 y 3 no conservadas en el consenso obtenido (TRMACCA), se emplearon los oligonucleótidos con diferencias en estas posiciones obtenidos en el ensayo de selección del sitio de unión. Como se observa en la Figura 38B, estas secuencias de ADN fueron unidas por la proteína casi tan eficientemente como el oligonucleótido que contiene la secuencia TAAACCA.

4.3.2.7 Identificación de los residuos de aminoácidos de la hélice III que son necesarios para la unión eficiente de HAT3.1 al ADN.

Si se asume que la secuencia TAAACCA constituye el sitio de unión preferido por HAT3.1, resta por establecer cómo el homeodominio de esta proteína interactúa con el ADN. Cuando se comparó la secuencia de aminoácidos del dominio de unión al ADN de HAT3.1 con la del homeodominio de Antennapedia (usado habitualmente como modelo) se vio que solamente 14 posiciones son idénticas (23%) (Figura 32). Entre éstas, se encuentran las 11 posiciones consideradas invariables o altamente conservadas en todos los homeodominios, como por ejemplo los aminoácidos L16, L38, V45, W48 y F49, que conforman el bolsillo hidrofóbico encargado de mantener la estructura terciaria de este motivo proteico. Por su parte, la presencia de una metionina 40 en lugar de leucina puede considerarse como una sustitución conservadora. Entre las posiciones no conservadas se encuentran los residuos ubicados en el brazo N-terminal y la histidina 51, es decir en posiciones que participan directamente de la unión al ADN en otros homeodominios. Como puede observarse en la Figura 32, el extremo N-terminal del homeodominio de HAT3.1 no contiene ninguno de los residuos que, se sabe, interactúan con el ADN en otros homeodominios (como R3, R5, R7 o Y/F8) (Gehring y col., 1994). Por lo tanto, es improbable que HAT3.1 utilice esta porción del homeodominio para contactar el ADN.

En función de los residuos de aminoácidos que posee el homeodominio de HAT3.1, suponemos entonces que éste se plegaría adoptando la estructura de un homeodominio típico, pero que las interacciones proteína-ADN podrían diferir de las establecidas por otros homeodominios.

Para poder estudiar la manera en que el homeodominio de HAT3.1 interactúa con el ADN, decidimos construir mutantes de la proteína HAT3.1_{HD} en residuos de la hélice III que normalmente están involucrados en los contactos específicos con el ADN (posiciones 47, 50, 51 y 54). Empleando la estrategia de mutagénesis sitio-dirigida mencionada en los capítulos anteriores, y utilizando como cebadores de las reacciones de PCR los oligonucleótidos detallados en la tabla 3 de Materiales y Métodos, se construyeron las mutantes: H51N, H51A, K50Q, K50A, N47A, N47V y W54A. Estas proteínas se expresaron en *E. coli* fusionadas a GST (detallado en la sección 3.15.1 de Materiales y Métodos) y, una vez purificadas, se emplearon

en ensayos de retardo en gel con un oligonucleótido sintético de 28 pb que contiene la secuencia TCTAAACCATT.

Como se observa en el ensayo de la Figura 39, la presencia de alanina en las posiciones 50, 51 ó 54 del homeodominio genera proteínas con una muy baja afinidad por la secuencia de unión empleada. Esto sugiere que estos residuos de aminoácidos establecen contactos importantes con el ADN. Por otra parte, la inclusión de Ala47 produce un efecto menor pero significativo sobre la unión de HAT3.1_{HD} al ADN.

Como se mencionó más arriba, los residuos 47, 50, 51 y 54 del homeodominio de HAT3.1 fueron modificados a otros que aparecen con frecuencia en los homeodominios. Con estas proteínas mutantes se observó que las sustituciones a valina 47 o glutamina 50 (N47V y K50Q, respectivamente) eliminan casi por completo la unión al ADN. El efecto producido por la presencia de valina en la posición 47 podría corresponder a un impedimento estérico causado por la introducción de este residuo hidrofóbico y a la pérdida de los contactos establecidos por la asparagina 47 en la proteína nativa. En el caso de Q50, es probable que este reemplazo elimine los contactos realizados por K50 puesto que la unión de K50Q es similar a la observada cuando hay alanina en esta posición (K50A).

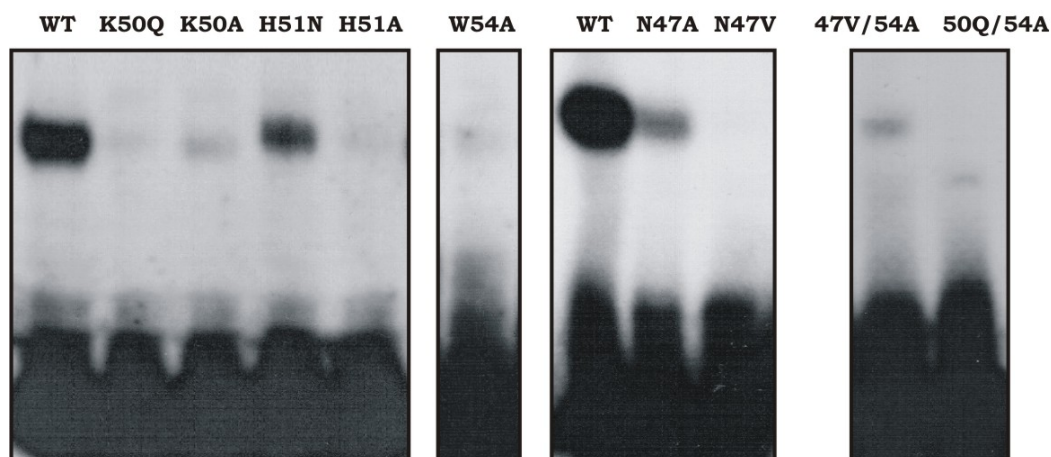


Figura 39. Unión al ADN de mutantes de HAT3.1_{HD} en las posiciones 47, 50, 51 y 54 del homeodominio. Ensayos de retardo en gel realizados con un oligonucleótido que contiene la secuencia TCTAAACCATT. En la parte superior se indican las proteínas analizadas.

La inclusión de asparagina en la posición 51, en cambio, tiene un efecto mínimo sobre la unión al ADN (H51N, Figura 39). A partir de esto, puede especularse que la histidina 51 participa en la interacción de modo similar al observado para N51 en otros homeodominios. Sin

embargo, la presencia de H51 parece ser más apropiada en el contexto del homeodominio de HAT3.1.

Finalmente, se realizaron ensayos de retardo en gel con mutantes dobles (N47V/W54A y K50Q/W54A) y triples N47V/H51N/W54A de HAT3.1_{HD}. Como resultado, se observó que la unión al ADN de estas proteínas recombinantes era casi imperceptible (Figura 39 y ensayos no mostrados, respectivamente).

4.3.2.8 Influencia del estado de protonación de H51 sobre la eficiencia de unión de HAT3.1_{HD} al ADN.

Como ya se sabe, los residuos de histidina son capaces de cambiar su estado de protonación a pH casi neutro (pK_a entre 6.0 y 7.0). En función de esto decidimos estudiar si la unión al ADN de HAT3.1_{HD} era dependiente del pH.

Con esto en mente, se llevaron a cabo reacciones de unión entre HAT3.1_{HD} y el oligonucleótido BS a distintos pH. Como puede observarse en la Figura 40, se produjo un incremento en la unión al ADN a medida que se disminuyó el pH de la reacción (de 8,2 a 4,9), con un punto de inflexión entre pH 6,0 y 6,5. Esto sugiere que la protonación del residuo de histidina mejoraría la interacción de HAT3.1_{HD} con el ADN. En concordancia con esta suposición, cuando se realizaron ensayos similares con la mutante HAT3.1_{HD}-H51N se vio que el cambio a asparagina 51 originaba una proteína incapaz de responder a las variaciones de pH (Figura 40). En función de estos resultados puede decirse que el estado de protonación del residuo de histidina 51 modula la eficiencia de unión de HAT3.1 al ADN.

En otros homeodominios, la asparagina 51 es esencial en la interacción de la hélice III con el ADN. Este residuo establece un puente de hidrógeno bidentado con la adenina de la tercera posición del núcleo reconocido (por ejemplo, TAAT) (Gehring, 1994). Como en los experimentos llevados a cabo con HAT3.1_{HD} se vio que el residuo asparagina 51 puede de alguna manera reemplazar a la histidina 51, puede pensarse que este aminoácido también reconocería una A dentro de la secuencia TAAACCA. En principio, el residuo de histidina tiene la capacidad de establecer puentes de hidrógeno a través del N δ y el N ϵ cuando éstos se encuentran protonados. Así, la interacción con la adenina podría ocurrir a través del N7 de la misma. El N restante de la histidina podría establecer contactos específicos con otro nucleótido

de la secuencia de unión, o bien, interactuar con el esqueleto azúcar fosfato del ADN a través de la carga positiva. Otra posibilidad es que interactúe con otro aminoácido favoreciendo la unión de este al ADN, o estabilizando la estructura del homeodominio.

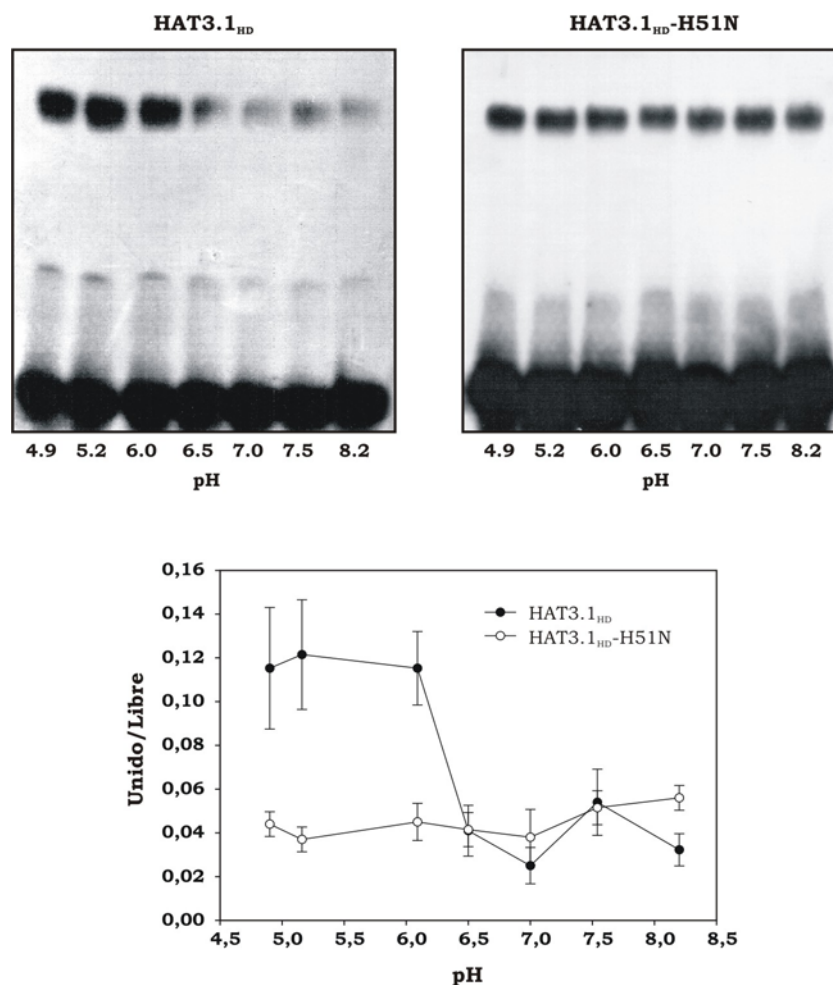


Figura 40. Efecto del pH en la interacción de HAT3.1 con el ADN. Unión de HAT3.1_{HD} nativa o HAT3.1_{HD}-H51N a un oligonucleótido que contiene la secuencia consenso (BS). Las reacciones de unión se realizaron en soluciones que contenían una mezcla de HEPES y MOPS, 0.2 M cada uno, a diferentes pH (indicado en la parte inferior de cada panel). El gráfico en la parte inferior muestra un análisis cuantitativo de la unión a los distintos pH. La determinación de la proporción de ADN unido se realizó por medición en contador de centelleo de la radioactividad presente en las bandas de retardo y del ADN total empleado. Las barras de error representan la desviación estándar de tres experimentos diferentes.

A continuación, se estudió la interacción de HAT3.1_{HD}-H51N con el ADN mediante ensayos de pérdida de nucleósido por ataque con radical hidroxilo (Figura 41A). A diferencia de lo obtenido en la hebra superior con la proteína nativa, se observó que la A₄ de la secuencia TRMACCA muestra una interferencia relativa menor que el dinucleótido CC que le sucede y que la A₃ o C₃ adyacente a A₄ (TRMACCA). Por el contrario, en el experimento de interferencia

realizado con la proteína nativa (Figura 36) se observa que la pérdida del nucleósido A₄ interfiere de modo similar a la eliminación de A₃ en TAAACCA y que la interferencia es mayor que cuando se pierde la C₃ en TGCACCA.

En conjunto, estas observaciones sugieren que al reemplazar la histidina 51 por asparagina disminuye en forma relativa el contacto con A₄. Además, la eliminación del nucleótido A₃ en TAAACCA produce la mayor interferencia en la unión de la mutante H51N al ADN, en relación al resto de las posiciones de la secuencia.

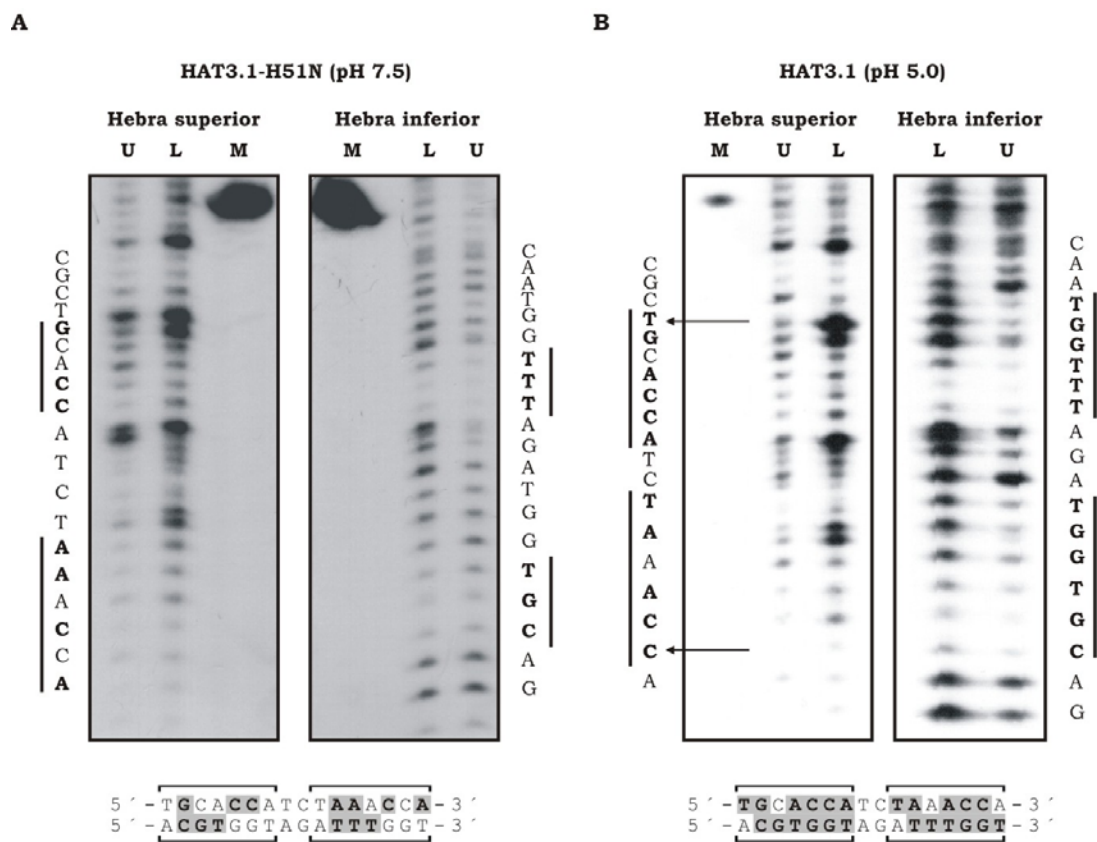


Figura 41. Experimentos de interferencia en la unión de HAT3.1_{HD} al oligonucleótido BS por ataque con radical hidroxilo. El oligonucleótido marcado en un extremo fue sometido al ataque por radicales hidroxilo antes de la unión con HAT3.1_{HD}-H51N a pH 7,5 (**A**) o con HAT3.1_{HD} nativa a pH 5,0 (**B**). Como marcador de posición (M) se utilizó el oligonucleótido digerido con enzimas de restricción (Materiales y Métodos, 3.25). Al costado de cada panel se indica la secuencia de ADN desde el extremo 5'. En la parte inferior se indica el sitio de unión y las posiciones contactadas aparecen sombreadas y en negrita. B, ADN unido; F, ADN libre

Por otra parte, cuando se examinó el patrón de interferencia resultante de la unión de HAT3.1_{HD} nativa al ADN a pH 5.0, se observó un aumento de la importancia relativa de la presencia de A en la cuarta posición para la unión de la proteína en estas condiciones (Figura 41B).

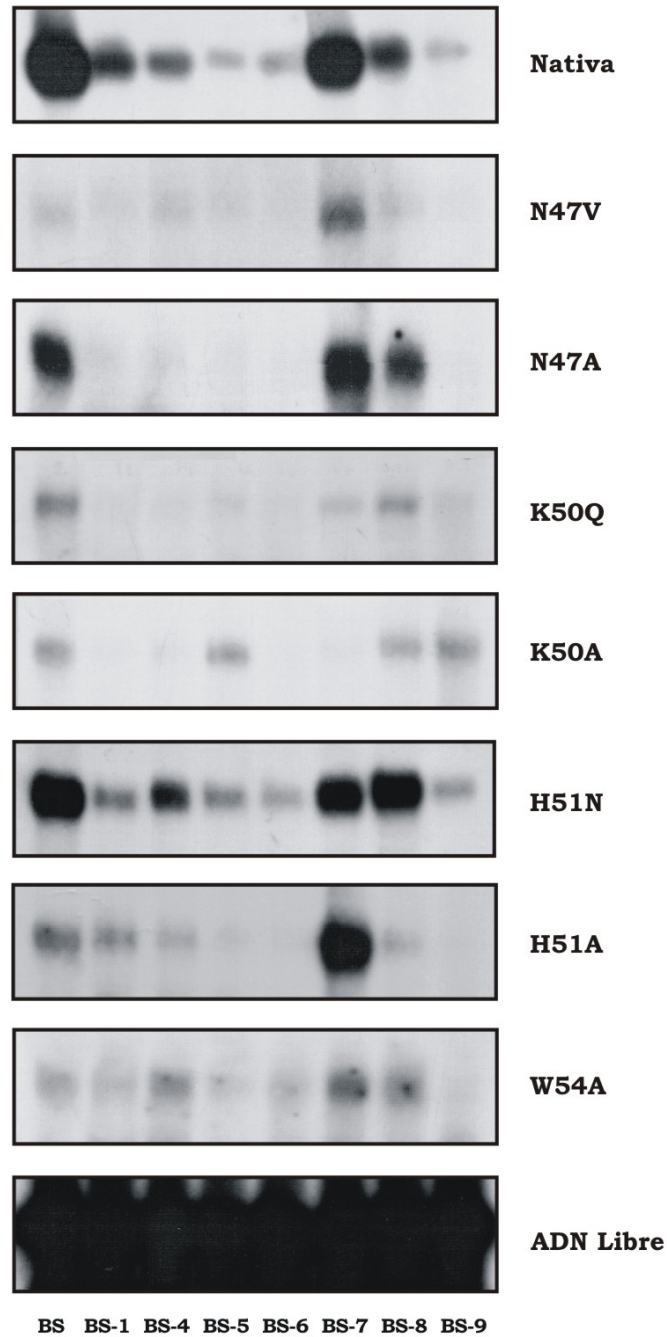
A partir de todo esto postulamos que el residuo de histidina 51 establecería contactos con la A₄ de la secuencia de ADN reconocida y que estos contactos se estabilizarían con la protonación de este aminoácido. Esto explicaría la preferencia por histidina en lugar de asparagina en esta posición en el contexto del homeodominio de HAT3.1.

4.3.2.9 Análisis de la especificidad de unión al ADN de las mutantes de HAT3.1_{HD}

Como se verá a continuación, cuando se estudiaron las características de la unión de las mutantes en las posiciones 47, 50, 51 y 54, se vio que, además de una afinidad reducida, estas proteínas mostraban diferencias respecto de HAT3.1_{HD} nativa en las preferencias de unión por oligonucleótidos con variaciones en la secuencia TAAACCA.

De acuerdo a los geles de retardo que se muestran en la Figura 42, la mutante HAT3.1_{HD}-H51N resultó menos afectada que la proteína nativa por el cambio A→T de la posición 4, que genera el núcleo TAAT reconocido por numerosos homeodominios (oligonucleótido BS-4). Esto apoya la idea planteada anteriormente de que la histidina 51 establecería contactos más fuertes con A₄ que asparagina 51. Además, la unión de la mutante H51N no se ve afectada por el cambio de CC a TA en la región 5' de la secuencia (BS-8), mientras que un cambio similar en la región 3' (BS-9) produce un efecto significativo sobre la interacción. Esto señala que esta mutante presenta una preferencia mayor por el segundo sitio de unión (es decir, por AA en lugar de GC en las posiciones 2 y 3). En consecuencia, puede asumirse que el residuo de asparagina 51 establece contactos preferenciales con una A localizada en las posiciones 2 ó 3, mientras que una histidina en esta posición interaccionaría principalmente con una A en la cuarta posición.

La unión de la mutante K50A a los distintos oligonucleótidos mostró que esta proteína no se ve afectada por el cambio de C a G en la posición 5 (BS-5, Figura 42) o por la modificación CC→TA en alguna de las secuencias presentes en el consenso (BS-8 y BS-9). Esto puede explicarse si se asume que el residuo lisina 50 interacciona con uno o ambos miembros del dinucleótido CC, mientras que una alanina 50 no. La pérdida de esta interacción en la mutante HAT3.1_{HD}-K50A sería la responsable de la significativa disminución en la unión al consenso y el limitado efecto que se observa cuando se cambian las posiciones 5 y 6 de esta secuencia.



OLIGONUCLEÓTIDO	SECUENCIA
BS	TGCACCATCTAAACCA
BS-1	TGCACCATC G AAACCA
BS-4	TGCACCATCTAA A CCA
BS-5	TGCACCATCTAA A CCA
BS-6	TGCACCATCTAA A CCA
BS-7	TGCACCATCTAA A CCA
BS-8	TGC A TAATCTAAACCA
BS-9	TGCACCATCTAA A TA

Figura 42. Ensayos de retardo en gel realizados con HAT3.1^{HD} nativa, N47A, N47V, K50Q, K50A, H51N, H51A, W54A y oligonucleótidos con variaciones puntuales de la secuencia 5'-TAAACCA-3'. A la derecha de cada figura se indican las proteínas analizadas en cada experimento. En el último panel se muestran las bandas correspondientes al ADN libre de uno de los ensayos. En la parte inferior aparecen los oligonucleótidos utilizados en cada calle y la secuencia central de cada uno de ellos se muestra en la tabla. Los recuadros negros indican las mutaciones realizadas.

Para el caso de la mutante HAT3.1_{HD}-W54A no se observaron cambios en la unión al modificarse la A₄ (BS-4, Figura 42). Esto demuestra que W54 interactúa con una A en la posición 4.

El comportamiento de las mutantes HAT3.1_{HD}-H51A y N47V resultó inesperado dado que mostraron una preferencia significativa por el oligonucleótido BS-7. Este aumento en la unión podría responder al establecimiento de nuevos contactos con el nucleótido modificado. Mientras que el residuo de valina 47 podría por sí mismo estar involucrado en estas nuevas interacciones, es improbable que la cadena lateral de alanina 51 contacte específicamente el ADN. Puede pensarse que a partir de la remoción del residuo de histidina 51 la mutante H51A establecería una nueva interacción entre otro aminoácido del homeodominio y la posición 7 modificada.

En relación a las mutantes dobles (N47V/W54A y K50Q/W54A) y triple (N47V/H51N/W54A), no se observó una interacción significativa con ninguna de las secuencias de unión ensayadas (no mostrado). Esto último resulta notable dado que la mutante N47V/H51N/W54A posee en las posiciones 47 a 54 los mismos residuos que el homeodominio de PITX2, una proteína que reconoce la secuencia TAATCC (Chaney y col., 2005).

4.3.3 Discusión

El factor de transcripción con homeodominio HAT3.1 de *Arabidopsis* fue aislado durante una búsqueda en una biblioteca de expresión con un fragmento del promotor del gen *cab-E* (Schindler y col., 1993). Sin embargo, una vez expresada en forma recombinante, esta proteína fue incapaz de unir ADN específicamente, mostrando afinidad por cualquier secuencia mayor a 50 pb (Schindler y col., 1993). Debido a la divergencia de secuencia que posee el homeodominio de HAT3.1, en ese momento quedó abierto el interrogante de si este dominio era capaz de funcionar como una unidad independiente y específica de unión a ADN.

En la presente tesis hemos demostrado que el homeodominio de HAT3.1 es capaz de seleccionar secuencias específicas de ADN a partir de una población de moléculas de secuencia al azar. A partir de este análisis se determinó que HAT3.1 une secuencias del tipo TRMACCA con una preferencia por A en las posiciones dos y tres. Pensamos que esta información, junto con los datos de expresión de plantas con cambios en los niveles de

HAT3.1, pueden ser de utilidad para identificar los posibles genes blanco de este factor de transcripción. En efecto, la región del promotor del gen *cab-E* empleada por Schindler y col. (1993) contiene la secuencia TACACCA. Esto sugiere que probablemente esta porción haya sido la responsable de la identificación del ADNc de HAT3.1.

A pesar de que el consenso obtenido contiene dos copias en tandem de la secuencia de unión, hemos demostrado que HAT3.1 puede interactuar con un oligonucleótido que posee un único sitio TAAACCA. Esto indica que una sola molécula de homeodominio es capaz de unir ADN de manera eficiente. Aunque no debe descartarse la posibilidad de que esta proteína interactúe con el ADN como dímero, no hemos encontrado evidencias de dimerización del homeodominio de HAT3.1. En consecuencia, la preferencia por moléculas con dos secuencias en tandem durante el proceso de selección indica que un aumento en el número de sitios de unión específicos por molécula conduce a un incremento de la eficiencia de unión.

La mayor parte de los residuos conservados dentro del homeodominio de la proteína estudiada corresponden a aquellos que forman el núcleo hidrofóbico que mantiene la estructura terciaria de las tres hélices α . De acuerdo a esto, los mecanismos de reconocimiento proteína-ADN deberían corresponderse con los establecidos por otros homeodominios. Sin embargo, HAT3.1 presenta variaciones en los residuos conservados que generalmente participan en la unión al ADN. De acuerdo a nuestros resultados, los aminoácidos presentes en las posiciones 47, 50, 51 y 54, involucrados en el contacto con el surco mayor del ADN en otros homeodominios, también son importantes para HAT3.1.

A partir de los estudios de interacción proteína-ADN de HAT3.1_{HD} y de las respectivas mutantes con la secuencia de ADN consenso y variaciones de la misma concluimos que los residuos K50, H51 y W54 establecen contactos con nucleótidos específicos. Esto se ilustra en el esquema de la Figura 43A. El eje central del modelo planteado está dado por la interacción de lisina 50 con las posiciones 5 y 6 de la secuencia blanco, puesto que este residuo está presente en la posición 50 en otros homeodominios. Por ejemplo, en el homeodominio de Bicoid (que posee K50) este aminoácido es responsable de la especificidad de unión por la secuencia TAATCC, en vez de TAATTG, mediante el contacto con el dinucleótido GG complementario a CC (Baird-Titus y col., 2006; Figura 43B). A partir de nuestros experimentos

proponemos entonces que el residuo **K50** del homeodominio de HAT3.1 también interacciona con las posiciones 5 y 6 (CC) de la secuencia reconocida. Esta afirmación se origina en la comparación directa de la secuencia TAATCC con el sitio blanco de HAT3.1, TAAACC (Figura 43).

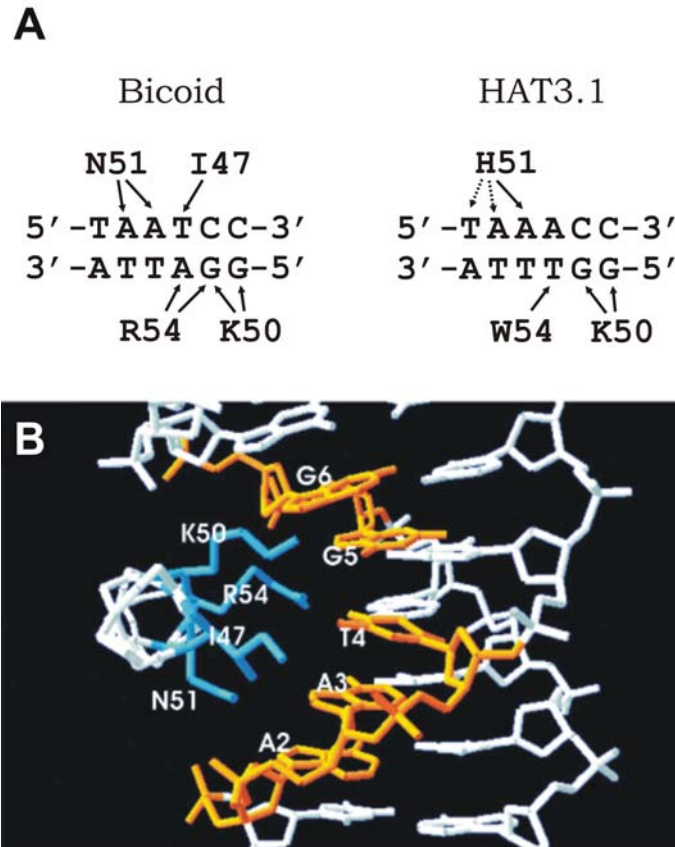


Figura 43. Modelo de la interacción del homeodominio de HAT3.1 con el ADN. A. Contactos establecidos entre los residuos de la hélice III del homeodominio de Bicoid y los nucleótidos de su secuencia de ADN blanco, realizado a partir de la estructura del complejo proteína-ADN en solución (Baird-Titus y col., 2006). A la derecha se esquematizan los contactos putativos del homeodominio de HAT3.1 con la secuencia TAAACC, deducidos de los experimentos descritos en este capítulo. Las flechas completas señalan los contactos con bases específicas del ADN y las flechas incompletas señalan posibles interacciones efectuadas por H51. **B.** Modelo de la interacción de la hélice III del homeodominio de Bicoid con el ADN. Las cadenas laterales de los residuos 47, 50, 51 y 54 se muestran en azul. Los nucleótidos específicos de la secuencia de unión TAATCC se muestran en naranja. G5 y G6 son complementarios al dinucleótido CC.

Como se comentó en la Introducción general, en Bicoid el residuo N51 establece contactos con el dinucleótido AA en TAATCC, mientras que R54 une los nucleótidos complementarios a T₄ y C₅ (Baird-Titus y col., 2006). Esto concuerda con los resultados que hemos obtenido con las mutantes de HAT3.1_{HD} que muestran que N51 es importante para el

reconocimiento de las posiciones dos y tres, mientras que el residuo **W54** estaría involucrado en la unión a la cuarta posición. Además, puede deducirse que la presencia de triptofano en la posición 54 es necesaria para la unión eficiente de HAT3.1 al ADN, dado que la mutación a alanina 54 elimina completamente la capacidad de esta proteína para unir ADN, aún cuando A54 se encuentra en otros homeodominios (Kissinger y col., 1990). Por ejemplo, el homeodominio de PITX2 contiene los residuos V47, K50, N51 y A54 y reconoce la secuencia TAATCC. En este caso, V47 establece contactos hidrofóbicos con la T en la cuarta posición (Chaney y col., 2005). Una interacción similar se ha observado para Bicoid y muchos otros homeodominios (Kissinger y col., 1990; Gehring y col., 1994; Baird-Titus y col., 2006). Aunque la mutación de N47 a alanina en HAT3.1_{HD} produce una significativa disminución en la unión, **N47** no parece establecer contactos similares a los reportados para valina en esta posición. En efecto, la introducción de V47 en HAT3.1_{HD} origina una proteína con muy baja afinidad de unión tanto para TAAACC como para TAATCC (oligonucleótidos BS y BS-4, respectivamente). En consecuencia, puede decirse que el aminoácido W54 sería uno de los determinantes de la unión a la secuencia TAAA, en lugar de TAAT como en la mayoría de los homeodominios, y que además impide la unión a TAAT aún en presencia V47.

La unión a las hemi-secuencias TAAATG en lugar de TAATTG ha sido adjudicada a interacciones combinatorias entre los residuos de las posiciones 47 y 54 (fenilalanina y treonina, respectivamente) en proteínas de la familia HD-Zip IV (Tron y col., 2001). Para HAT3.1 la situación parece ser más compleja ya que las mutantes N47V/W54A y N47V/H51N/W54A fueron incapaces de unir oligonucleótidos que contenían las secuencias TAAA o TAAT (resultados no mostrados).

El homeodominio de HAT3.1 representa un caso extraordinario de divergencia dentro de la estructura de estos motivos. La introducción de W54 y H51 constituye una innovación puesto que, hasta el momento, estos residuos no se han encontrado en estas posiciones en ningún otro homeodominio. Mientras que la presencia de triptofano en la posición 54, la más variable entre los homeodominios, originaría una nueva especificidad de unión (en efecto, las posiciones 50 y 54 se han reportado como las causantes principales de los cambios en las preferencias de unión; Laughon, 1991), la presencia de H51 resulta llamativa, ya que la asparagina es casi invariable en esta posición (Gehring y col., 1994).

Una pregunta interesante es: ¿por qué la histidina 51 no está presente en otros homeodominios si puede reemplazar de manera eficiente a la asparagina 51? Pensamos que la respuesta radica en que la presencia de histidina en esta posición requeriría de cambios compensatorios en otras partes de la molécula, o bien, que representa una ventaja sólo en el contexto de HAT3.1. En este sentido, por ejemplo, H51 podría estar interaccionando con la cadena lateral de otro/s aminoácido/s estabilizando la estructura del homeodominio. Este efecto combinatorio explicaría también por qué la introducción de residuos usualmente presentes en otros homeodominios (V47, Q50, A54), en distintas combinaciones, provoca una disminución significativa de la eficiencia de unión de HAT3.1 al ADN.

En conclusión, el homeodominio de HAT3.1 constituye un ejemplo del considerable grado de plasticidad que posee este motivo y, por lo tanto, de su importancia desde el punto de vista evolutivo para generar nuevas especificidades de interacción proteína-ADN.

CONCLUSIONES

5. CONCLUSIONES

A partir de los resultados experimentales obtenidos en el presente trabajo de Tesis, podemos enunciar las siguientes conclusiones:

- La secuencia consenso de ADN reconocida por el homeodominio de la proteína ATH1 de la familia BEL es 5´ TGACAGGT- 3´, y contiene el núcleo TGAC reconocido por las proteínas KNOX.

- Los homeodominios de ATH1 y de la proteína KNOX de clase I STM reconocen secuencias de ADN similares, pero existen diferencias en el modo en el que éstos interaccionan con el ADN.

- Modificaciones en las posiciones 2 y 3 del núcleo TGAC impiden por completo la unión de los homeodominios de ambas proteínas, mientras que la misma no se ve afectada por cambios en la posición 1 (TGAC).

- En cuanto a la especificidad por la cuarta posición, ninguna de las dos proteínas reconoce la secuencia TGAG. Sin embargo, a diferencia de STM, ATH1 es capaz de unir la secuencia TGAA.

- Modificaciones en las posiciones 5 y 6 (TGACAG) producen un efecto menor sobre la unión de STM que sobre la de ATH1.

- La presencia de una G en la séptima posición de la secuencia unida (TGACAGGT) es importante para la unión al ADN de ambas proteínas.

- El cambio de T a A en la posición 8 no afecta la unión de ATH1, pero reduce la formación del complejo ADN-proteína por parte de STM.

- ATH1 contacta los nucleótidos TGACAGGT en la hebra de ADN que contiene la secuencia TGAC y una región de 7 pb, que incluye el nucleótido complementario a T en la primera posición, en la otra hebra.

- STM contacta una región de ocho nucleótidos en la hebra que contiene el núcleo TGAC, TGACAGGTG, extendiéndose una posición más hacia el extremo 3´ de la secuencia consenso, y ocho nucleótidos complementarios a la secuencia GTGACAGG en la hebra inferior del ADN.

- El residuo 54 de los homeodominios de ATH1 y STM es el principal responsable de las diferencias observadas en la afinidad de unión y la especificidad en las posiciones 4, 5 y 6 de la secuencia de ADN reconocida por estas proteínas.

- El residuo valina 54 de ATH1 establecería contactos principalmente con las posiciones 5 y 6 (TGACAG), y no (o en forma menos importante) con la posición 4. Esto conduciría a una especificidad menos estricta por el nucleótido presente en esta posición.

- El aminoácido lisina 54 presente en STM establecería uniones puente de hidrógeno con la G complementaria a C en TGAC, mientras que el aminoácido valina 54 establecería uniones más débiles como las de van der Waals o mediadas por moléculas de agua. Esta sería la causa de la menor afinidad de unión mostrada por las proteínas con valina 54 respecto de las que poseen lisina en esta posición.

- Otros residuos de aminoácidos, o combinaciones de ellos, participarían del contacto con la octava posición de la secuencia TGACAGGT.

- *Los resultados obtenidos pueden ser extendidos a otros miembros de las familias BEL y KNOX y muestran que ambos tipos de proteínas interactúan en forma similar, pero no idéntica, con el ADN. Estas diferencias pueden atribuirse principalmente al aminoácido de la posición 54 del homeodominio y, probablemente, son esenciales para la función de estas proteínas in vivo si se tiene en cuenta la conservación de este residuo en las proteínas de cada familia.*

- STM (completa) y la proteína de la familia BEL BLH3 son capaces de interactuar *in vitro* con la secuencia TGACAGGT y con un oligonucleótido que contiene dos sitios TGAC en tandem separados por un nucleótido.

- La formación del complejo BLH3-STM aumenta la eficiencia de unión a moléculas de ADN que contienen tanto uno como dos sitios TGAC.

- Las regiones responsables de la heterodimerización entre las proteínas BEL y KNOX son necesarias para producir el aumento en la afinidad de unión del complejo BLH3-STM a la secuencia TGACAGGT.
- Al aumentar la concentración de BLH3 la afinidad de unión de esta proteína se acerca a la del complejo BLH3-STM. Sin embargo esto no ocurre con la interacción de STM con el ADN.
- BLH3 establece contactos con los nucleótidos TGACAGGT en la hebra de ADN que contiene la secuencia TGAC y a lo largo de los nucleótidos TCACCTGTC, complementarios a la secuencia GACAGGTGA.
- Los dos sitios TGAC presentes en la secuencia TGACTTGACAGGT son contactados por BLH3 con la misma eficiencia y mediante interacciones similares.
- Con la formación del complejo no se establecen interacciones con nucleótidos adicionales de las secuencias de unión.
- Dentro del complejo cualquiera de los homeodominios puede interactuar con el ADN, pero existe una preferencia del homeodominio de STM para contactar la secuencia con dos núcleos TGAC en tandem.
- Los ensayos *in vitro* señalan que la formación del complejo BLH3-STM genera un aumento en la afinidad de unión al ADN que no requiere la presencia de dos motivos TGAC. Pero esto no ocurre en ensayos *in vivo* en levaduras:
 - No existen diferencias significativas en la unión *in vivo* a la secuencia TGACAGGT de las proteínas individuales y el complejo BLH3-STM.
 - STM y BLH3 no interactúan, o lo hacen muy pobremente, con los sitios en tandem presentes en TGACTTGACAGGT, pero ocurre un incremento notable en la eficiencia de unión a esta secuencia cuando se co-expresan las dos proteínas.

- Las proteínas BEL de Arabidopsis ATH1 y BEL1 reconocen la secuencia TGACAGGT *in vivo* en levaduras y forman heterodímeros con STM que presentan propiedades de unión análogas a las mostradas por el complejo BLH3-STM.
- *Existen diferencias específicas de secuencia en las propiedades de unión al ADN de las proteínas STM y BLH3 individuales y del complejo formado por las mismas. La formación de un complejo entre la proteína BEL BLH3 y la proteína KNOX STM produce un aumento en la afinidad de unión al ADN cuya magnitud depende de la secuencia del sitio de unión y, frente a la secuencia con un único núcleo TGAC, de las concentraciones relativas de cada una de estas proteínas. La existencia de un comportamiento similar dentro de la planta otorgaría una mayor plasticidad a las redes de regulación de las que forman parte estas proteínas.*
- El homeodominio de HAT3.1, una proteína de la familia PHD finger, es capaz de seleccionar secuencias específicas de ADN a partir de una población de moléculas de secuencia al azar.
- La secuencia consenso reconocida por HAT3.1_{HD} es T(A/G)(A/C)ACCA, con una preferencia por A sobre C en la tercera posición. Esta secuencia no contiene el núcleo TAAT unido por muchos homeodominios.
- Los nucleótidos de las posiciones 1, 4, 5 y 6 de la secuencia TAAACCA son importantes para la unión de HAT3.1.
- Los residuos de la hélice III ubicados en posiciones usualmente involucradas en la interacción con el surco mayor del ADN en otros homeodominios (N47, K50, H51 y W54) son importantes para la unión de HAT3.1.
- La protonación del residuo de histidina 51 mejora la eficiencia de unión al ADN de HAT3.1.

- El residuo K50 del homeodominio de HAT3.1, al igual que en otros homeodominios que contienen lisina en esta posición, interacciona con las posiciones 5 y 6 (TAAACCA) de la secuencia reconocida.
- Mientras que la histidina 51 reconocería los nucleótidos 2 a 4, el residuo W54 estaría involucrado en la interacción con la cuarta posición y sería uno de los determinantes de la unión a la secuencia TAAA, en lugar de TAAT como en la mayoría de los homeodominios.
- *La secuencia de ADN obtenida para HAT3.1 difiere del núcleo TAAT reconocido por otros homeodominios. Esto se debe, en parte, por la presencia de los residuos H51 y W54, no encontrados en ningún otro homeodominio hasta el momento. Los resultados obtenidos señalan que HAT3.1 interacciona específicamente con el ADN empleando mecanismos conservados de unión, constituyendo un claro ejemplo de la plasticidad de este dominio como motivo de unión a ADN y, por lo tanto, de su importancia en términos evolutivos para generar nuevas especificidades de interacción proteína-ADN.*

RESUMEN

6 RESUMEN

Las proteínas que contienen homeodominio constituyen un grupo de factores de transcripción eucariotas involucrados en la regulación de los procesos de desarrollo.

El homeodominio es un dominio proteico de 60 aminoácidos muy conservado que interacciona de manera específica con el ADN. Este motivo adopta en solución una estructura tridimensional compacta compuesta por tres hélices α y un brazo amino terminal flexible. Las hélices I y II se disponen en forma antiparalela separadas mediante un giro abierto, mientras que la hélice III (o de reconocimiento) se ubica de manera perpendicular respecto de las otras dos. Las hélices II y III, unidas por una vuelta apretada, forman un motivo hélice-giro-hélice similar al descrito en proteínas reguladoras de la transcripción en procariontes. Los contactos entre el homeodominio y el ADN son establecidos por la hélice III, que se inserta en el surco mayor de la doble hélice, y por el brazo amino terminal flexible, el cual interacciona con bases presentes en el surco menor del ADN.

A pesar de la diversidad funcional de las proteínas que contienen homeodominio, una gran parte de ellas reconoce secuencias de ADN que contienen el núcleo TAAT, pero difieren en las preferencias por los nucleótidos que suceden en dirección 3' a esta secuencia. La unión a un núcleo conservado está de acuerdo con el alto grado de conservación que presentan los residuos involucrados en el contacto con el ADN. Los aminoácidos variables del homeodominio serían los responsables de interactuar con los nucleótidos adyacentes a TAAT. Así, las diferencias en la especificidad de unión se deben, en parte, a la interacción del aminoácido de la posición 50 del homeodominio, que contacta los nucleótidos 5 y 6 de la secuencia de ADN reconocida. Otro residuo que participa en la determinación de la especificidad de unión es el presente en la posición 54, que en la mayoría de los homeodominios estudiados contacta los nucleótidos 4, 5 y/o 6 del sitio de unión.

Las proteínas que contienen homeodominio se clasifican en diferentes familias y subfamilias de acuerdo a la conservación de aminoácidos de este dominio y a la presencia de otros motivos adyacentes al mismo.

Un grupo particular de proteínas con homeodominio constituye la superclase TALE (*Three Amino acid Loop Extension*) cuyo rasgo más característico es la presencia de tres

aminoácidos extra altamente conservados (prolina-tirosina-prolina) en el rulo que conecta las hélices I y II del homeodominio. En plantas existen dos familias de genes que codifican proteínas TALE: *KNOX* y *BEL*.

Los homeodominios de las proteínas TALE presentan características particulares y generalmente interaccionan con secuencias de ADN que contienen el núcleo TGT. Sin embargo, existen entre ellas diferencias importantes en algunos aminoácidos del homeodominio y en el tipo de secuencia reconocida por fuera de este núcleo. Esto hace que los conocimientos obtenidos en levaduras y animales no se puedan extrapolar a las proteínas de plantas *KNOX* y *BEL*.

Las proteínas con homeodominio y PHD finger contienen un motivo rico en cisteínas que, si bien se ha encontrado en otros organismos eucariotas, sólo aparece asociado a un homeodominio en plantas. A diferencia de otras familias, los miembros de este grupo difieren notoriamente entre sí en la secuencia del homeodominio, especialmente en los residuos que normalmente contactan el ADN. Sin embargo, como presentan conservadas las posiciones que forman el bolsillo hidrofóbico, se plegarían adoptando una estructura tridimensional similar a la de otros homeodominios. Probablemente, las características particulares que presentan los homeodominios de las proteínas PHD finger poseen un papel esencial en la función de estos factores de transcripción.

Teniendo en cuenta que la regulación de la transcripción llevada a cabo por las proteínas con homeodominio depende de la especificidad de sus interacciones con el ADN, nos planteamos como objetivo general de este trabajo de Tesis profundizar el conocimiento sobre los mecanismos de interacción con el ADN de proteínas que poseen aminoácidos particulares en sus respectivas hélices de reconocimiento. Elegimos entonces para nuestro estudio proteínas pertenecientes a las familias *BEL*, *KNOX* y PHD finger de *Arabidopsis thaliana*.

En relación con la especificidad de unión al ADN de las proteínas TALE de plantas, en los últimos años se han realizado estudios de selección del sitio de unión y mutagénesis que indican que las proteínas de la familia *KNOX* unen secuencias de ADN que contienen el núcleo TGAC. Sin embargo, no se han llevado a cabo estudios similares con las proteínas de la familia

BEL. Hasta el momento sólo se ha reportado que la proteína StBel5 de papa interacciona con una secuencia que contiene el núcleo TGAC presente en el promotor del gen *ga20ox1* de papa.

Es por esto que nos propusimos en primer lugar determinar la secuencia de ADN reconocida por ATH1, un factor de transcripción perteneciente a la familia BEL. Para ello se aplicó el método de selección de oligonucleótidos de secuencia al azar (SELEX). A través del apilamiento de las secuencias de 30 de los clones obtenidos luego de 7 rondas de selección se determinó que la secuencia de ADN preferida por el homeodominio de ATH1 (ATH1_{HD}) es 5'-TGACAGGT-3'. Ésta es idéntica a la informada para KN1, una proteína KNOX de clase I de maíz, y contiene el núcleo TGAC reconocido por otras proteínas de las familias KNOX y BEL. En consecuencia, decidimos estudiar comparativamente las propiedades de unión al ADN de los homeodominios de ATH1 y STM, una proteína KNOX de clase I de *Arabidopsis thaliana*.

A partir de estudios de mutagénesis dirigida y de protección (e interferencia) frente al ataque con radicales hidroxilo se determinó que ATH1_{HD} establece contactos directos en ambas hebras del ADN con las posiciones 2, 3 y 5 (TGACAGGT), y efectúa interacciones adicionales con los nucleótidos que se encuentran hacia el extremo 3' del núcleo TGACA. Sin embargo, tolera cambios de C a A en la posición 4. Por su parte, el homeodominio de STM (STM_{HD}) interacciona con las posiciones 2, 3, y 4 del núcleo TGAC y no tolera cambios en la cuarta posición. Cuando se introdujeron modificaciones en las posiciones 5 y 6 (TGACAG) de la secuencia consenso se observó que producían menos efecto sobre la unión de STM_{HD} que sobre la de ATH1_{HD}. Aunque la presencia de una G en la séptima posición de la secuencia unida (TGACAGGT) es importante para la unión al ADN de ambas proteínas, el cambio de T a A en la posición 8 sólo afecta la unión de STM_{HD}.

Los experimentos de protección, y de interferencia en la unión, mostraron que ATH1_{HD} interacciona con una región de siete nucleótidos que contiene el núcleo GAC (TGACAGGT) y con siete nucleótidos en la otra hebra que incluyen el nucleótido complementario a T en la primera posición. Sin embargo, STM_{HD} establece contactos con una región de ocho nucleótidos, extendiéndose una posición más hacia el extremo 3' de la secuencia (TGACAGGTG) respecto de ATH1_{HD}. En la hebra inferior interacciona con ocho nucleótidos complementarios a la secuencia GTGACAGG.

Al comparar la naturaleza de los residuos de la hélice III de ATH1 y STM observamos una diferencia en la posición 54 y decidimos analizar la importancia que posee el residuo ubicado en esta posición en las propiedades de unión a ADN de ambas proteínas. Para ello, se construyeron mutantes en las que el homeodominio de STM fue modificado de lisina 54 a valina 54 y el de ATH1 fue modificado de valina 54 a lisina 54. En función de los resultados obtenidos en ensayos de retardo en gel con oligonucleótidos con mutaciones puntuales en la secuencia de unión, puede decirse que la menor afinidad por el ADN que presenta ATH1_{HD} se debe a la presencia de valina 54 en lugar de lisina 54. Además, esta diferencia es la responsable de la menor selectividad por el nucleótido de la posición 4 y de la mayor especificidad por la posición 5 de la secuencia de ADN reconocida por esta proteína. En cambio, otros residuos, o combinaciones de ellos, participarían del contacto con la octava posición de la secuencia TGACAGGT. Nosotros planteamos que la presencia de valina 54 en ATH1 establecería contactos con las posiciones 5 y 6, más que con la posición cuatro, generando una especificidad más relajada por este nucleótido. La capacidad de la lisina 54 para establecer enlaces puente de hidrógeno con la G complementaria a C (TGAC) explicaría la mayor afinidad de STM por el ADN. De modo diferente, el aminoácido valina presente en ATH1 sólo establecería uniones más débiles como las de van der Waals o mediadas por moléculas de agua.

Los resultados obtenidos, que pueden ser extendidos a otros miembros de las familias BEL y KNOX, indican que ambos tipos de proteínas interactúan en forma similar, pero no idéntica, con el ADN. Estas diferencias pueden atribuirse principalmente al aminoácido de la posición 54 del homeodominio y, probablemente, son esenciales para la función de estas proteínas *in vivo* si se tiene en cuenta la conservación de este residuo en las proteínas de cada familia.

Las proteínas de la familia KNOX interactúan con las de la familia BEL a través de dominios proteicos presentes en porciones amino terminales de las mismas (dominio MEINOX en las proteínas KNOX y MID en las BEL). Si bien diversos complejos BEL-KNOX han sido reportados en distintas especies, son escasos los estudios realizados en relación a la interacción de estos complejos con el ADN. El complejo formado por las proteínas de papa POTH1 (KNOX) y StBEL5 (de la familia BEL) interactúa con la secuencia TGACTTGAC, que

contiene dos sitios TGAC en tandem, presente en el promotor del gen *ga20ox1*. Por otra parte, la formación del complejo KN1-KIP produce un aumento en la afinidad de unión a un fragmento de ADN con la secuencia TGACAGGT en ensayos *in vitro*.

Teniendo en cuenta los resultados obtenidos sobre las características de la interacción de ATH1 y STM con el ADN, decidimos estudiar la interacción del complejo formado por las proteínas STM y BLH3 de *Arabidopsis* (pertenecientes a las familias KNOX y BEL, respectivamente) con secuencias de ADN que contenían uno o dos núcleos TGAC (TGACAGGT y TGACTTGACAGGT, respectivamente). Los ensayos de retardo en gel realizados mostraron que tanto BLH3 como STM son capaces de interactuar con ambos oligonucleótidos y que la formación del complejo BLH3-STM aumenta la eficiencia de unión a moléculas de ADN que contienen tanto uno como dos sitios TGAC. Además, observamos que las regiones responsables de la heterodimerización entre las proteínas BEL y KNOX son necesarias para producir el aumento en la afinidad de unión del complejo BLH3-STM a la secuencia TGACAGGT.

Cuando se analizó la unión con el ADN de las proteínas individuales se vio que en bajas concentraciones BLH3 y STM unen de manera similar y con muy poca eficiencia. La utilización de mayores concentraciones de proteína en las reacciones de unión genera un aumento pronunciado de la proporción de ADN unido por BLH3, mientras que esto no ocurre con la unión de STM. La inclusión de una cantidad constante de la otra proteína produce un aumento en la proporción de ADN unido en todas las concentraciones ensayadas.

Los experimentos de interferencia por ataque con radical hidroxilo llevados a cabo sugieren que con la formación del complejo no se establecen interacciones con nucleótidos adicionales de la secuencia de unión. Además, para el caso del oligonucleótido que contiene los dos sitios TGAC, se vio que ambos sitios son contactados con la misma eficiencia y mediante interacciones similares.

La construcción de proteínas mutadas en el residuo 55 del homeodominio (de arginina a lisina), incapaces de unir ADN, permitió demostrar que dentro del complejo cualquiera de los homeodominios puede participar del reconocimiento del ADN, pero existiría una clara preferencia del homeodominio de STM para contactar el oligonucleótido con dos núcleos TGAC en tandem.

Los resultados obtenidos en los ensayos *in vitro* señalan que la formación del complejo BLH3-STM genera un aumento en la afinidad de unión al ADN que no requiere la presencia de dos motivos TGAC. Sin embargo, ensayos de unión a la secuencia TGACAGGT realizados *in vivo* en levaduras mostraron que no existen diferencias significativas en los niveles de expresión del gen reportero producidos por las proteínas individuales y el complejo BLH3-STM. Además, las propiedades de unión a la secuencia TGACAGGT difieren en función de las concentraciones de proteína. Si la co-expresión de BLH3 y STM en la levadura es dirigida por un promotor fuerte, se observa una competencia de las proteínas individuales por la unión al ADN. Mientras que bajos niveles de expresión de BLH3 conducen a un aumento en la activación del reportero en presencia de STM.

Las proteínas BEL de *Arabidopsis* ATH1 y BEL1 también reconocen la secuencia TGACAGGT *in vivo* en levaduras y son capaces de formar heterodímeros con STM que presentan propiedades de unión análogas a las mostradas por el complejo BLH3-STM.

Por otra parte, STM y BLH3 no interactúan *in vivo*, o lo hacen muy pobremente, con los sitios en tandem presentes en TGACTTGACAGGT, pero ocurre un incremento muy notable en la eficiencia de unión a esta secuencia cuando se co-expresan las dos proteínas.

En conclusión, existen diferencias específicas de secuencia en las propiedades de unión al ADN de las proteínas STM y BLH3 individuales y del complejo formado por las mismas. De esta manera, distintas especies interactuarían con grupos de genes diferentes en función de las secuencias presentes en sus promotores. En el caso de secuencias que presentan núcleos TGAC aislados, y no en tandem, probablemente las características de estas interacciones estarán dadas también por los niveles existentes de cada una de las proteínas.

Finalmente, a través de una búsqueda del genoma de *Arabidopsis* se identificaron los genes funcionalmente relacionados con las proteínas KNOX y BEL que contienen en sus regiones promotoras la secuencia TGACAG o la secuencia TGACNTGAC, que posee dos sitios TGAC en tandem. Se encontró que los promotores de los genes *KNAT2*, *KNAT4*, *BLH2*, y *BLH5* contienen la secuencia TGACAG, mientras que el promotor de *BLH3* posee el motivo TGACTTGAC. De ser estos sitios funcionales, esto indicaría la existencia de una compleja red de interacciones proteína-proteína y proteína-ADN que participaría en la regulación de la acción de las proteínas KNOX y BEL.

HAT3.1 es una proteína de la familia PHD finger cuyo homeodominio presenta conservados los residuos que forman el bolsillo hidrofóbico pero difiere en posiciones conservadas en otros homeodominios, incluso en las que componen la hélice de reconocimiento. En la presente tesis decidimos analizar la capacidad del homeodominio de HAT3.1 (HAT3.1_{HD}) para interactuar de manera específica con el ADN.

A través de la aplicación de la técnica de SELEX determinamos que la secuencia consenso de HAT3.1_{HD} es T(A/G)(A/C)ACCA, con una preferencia por A sobre C en la tercera posición. Esta secuencia no contiene el núcleo TAAT reconocido por muchos homeodominios. Como hasta el momento no se han reportado secuencias blanco similares a la obtenida deducimos que esta proteína posee una especificidad de unión diferente.

Los ensayos de interferencia en la unión por ataque con radical hidroxilo mostraron que los nucleótidos que conforman la secuencia T(A/G)(A/C)ACCA son necesarios para la unión. Asimismo, se vio que modificaciones en las posiciones 1 y 4 afectan la unión al ADN, mientras que cambios en los nucleótidos 5 y 6 de esta secuencia disminuyen severamente la unión al ADN.

Por otra parte, ensayos de simple híbrido en levadura permitieron demostrar que HAT3.1 es capaz de interactuar *in vivo* de manera específica con la secuencia consenso obtenida en los ensayos *in vitro*.

Con el fin de identificar los residuos de aminoácidos de la hélice III que son necesarios para la unión eficiente de HAT3.1 al ADN se construyeron proteínas mutadas en las posiciones que están involucradas en los contactos específicos con el ADN en la mayoría de los homeodominios. Cuando se compararon las propiedades de unión de estas proteínas con las de HAT3.1_{HD} pudimos deducir que los aminoácidos en las posiciones 47, 50, 51 y 54 también son importantes para la unión de HAT3.1. Esta afirmación surge de la reducida afinidad por la secuencia seleccionada que mostraron las proteínas mutadas.

Asimismo, observamos diferencias en las preferencias de unión de las proteínas mutadas por oligonucleótidos con variaciones puntuales en la secuencia TAAACCA. En conjunto, los resultados obtenidos nos permitieron elaborar un modelo de la unión de HAT3.1 al ADN. Concretamente, proponemos que el residuo K50 del homeodominio de HAT3.1, al igual que en otros homeodominios que contienen lisina en esta posición, interactúa con las

posiciones 5 y 6 (TAAACCA) de la secuencia reconocida, mientras que la histidina 51 reconocería los nucleótidos 2 a 4. El residuo W54 estaría involucrado en la interacción con la cuarta posición y sería uno de los determinantes de la unión a la secuencia TAAA, en lugar de TAAT como en la mayoría de los homeodominios. Además, observamos que este residuo es importante para la unión eficiente de HAT3.1_{HD} al ADN.

La presencia de W54 y H51 en HAT3.1_{HD} constituye una innovación puesto que, hasta el momento, estos residuos no se han encontrado en estas posiciones en ningún otro homeodominio. Mientras que la presencia de triptofano en la posición 54, la más variable entre los homeodominios, originaría una nueva especificidad de unión, la presencia de H51 resulta llamativa, ya que la asparagina 51 es casi invariable en los homeodominios. En este sentido, encontramos que el estado de protonación del residuo de histidina 51 modula la eficiencia de unión de HAT3.1 al ADN. Pensamos que el residuo de histidina protonado podría establecer contactos adicionales con otros nucleótidos de la secuencia reconocida, o podría interactuar con el esqueleto azúcar fosfato del ADN a través de la carga positiva. También podría interactuar con otro aminoácido favoreciendo la unión de éste al ADN o estabilizando la estructura del homeodominio. Esto explicaría la preferencia por histidina en lugar de asparagina en esta posición en el contexto del homeodominio de HAT3.1.

Finalmente, la capacidad del homeodominio de HAT3.1 de interactuar específicamente con el ADN constituye un ejemplo del considerable grado de plasticidad que posee este motivo y, por lo tanto, de su importancia desde el punto de vista evolutivo para generar nuevas especificidades de interacción proteína-ADN

ABSTRACT

7 ABSTRACT

Homeobox genes encode a group of eukaryotic transcription factors involved in the regulation of developmental processes. These genes contain a region coding for the homeodomain, a 60-amino-acid protein motif that interacts specifically with DNA. The homeodomain folds into a characteristic three helix structure. Helices I and II are connected by a loop, while helices II and III are separated by a turn, resembling prokaryotic helix-turn-helix transcription factors. The helix III (the recognition helix) is aligned perpendicularly to the first two helices and fits into the major groove of DNA. The N-terminal portion forms a disordered arm that protrudes from helix I and orients towards the DNA, making several specific contacts within the minor groove.

In spite of the different biological functions that are carried out by homeodomain proteins, it has been demonstrated that the recognition sites for most of them have a common TAAT core. This is in agreement with the high conservation of residues involved in homeodomain-DNA interaction. The first two positions of the sequence are contacted by the N-terminal arm, while the third and fourth positions are recognized by the almost invariant N51 and highly conserved isoleucine or valine at position 47, respectively. The next two positions are more variable and confer differential binding specificity. These positions are contacted by residues 50 and 54, which are considerably less conserved among different homeodomains.

Plant homeodomains can be divided into different families according to sequence conservation and the presence of additional motifs besides the homeodomain. A particular group of homeodomain proteins is called the three amino acid loop extension (TALE) superclass because of the three conserved residues proline-tyrosine-proline located between the first and second helix of the homeodomain. There are two groups of TALE homeodomain proteins in plants: the KNOX and BEL families.

Another group of homeobox genes is the PHD finger family, which encodes proteins with a cysteine-rich domain, named the PHD finger domain (for plant homeodomain finger). Although it was discovered in plants, the PHD finger has also been found in other eukaryotic organisms, but not associated with a homeodomain. It is noteworthy that the PHD finger homeodomain proteins differ markedly in sequence (most notably in residues that putatively

interact with DNA). Most conserved residues correspond to the hydrophobic core that maintains the tertiary structure of the three-helix module.

Since the regulation of transcription performed by homeodomain proteins depends on the specificity of the interaction with DNA, we decided to study the DNA binding mechanisms of homeodomain proteins that contain particular amino acids in the third helix. In fact, we chose proteins from the BEL, KNOX and PHD finger families of *Arabidopsis thaliana*.

Regarding interaction with DNA of plant TALE proteins, binding-site selection and mutagenesis studies have indicated that KNOX proteins bind sequences containing a TGAC core. Similar studies with BEL proteins have not been reported, although a recent study demonstrated that the potato BEL protein StBEL5 binds to a TGAC-containing motif in the *ga20ox1* promoter. In the present work, we investigated the interaction of the homeodomain of the *Arabidopsis* BEL protein ATH1 with DNA by a combination of binding-site selection and footprinting studies. Analysis of oligonucleotides selected by the ATH1 homeodomain from a random mixture suggests that ATH1 preferentially binds the sequence TGACAGGT. This sequence contains the TGAC motif recognized by KNOX proteins. In consequence, we performed a comparative analysis of the binding properties of ATH1 and STM, a KNOX protein.

Hydroxyl radical footprinting and missing nucleoside experiments showed that ATH1_{HD} interacts with the positions 2, 3 and 5 on both strands of the consensus TGACAGGT, and has additional interactions with nucleotides located towards the 3' end of the TGACA core. Nevertheless, changes at position 4 are more tolerated. The STM homeodomain interacts with positions 2, 3 and 4 of the TGAC core. Modifications in the fourth position are not tolerated by STM. Changes at positions 5 and 6 (TGACAG) produce a much smaller effect on STM than on ATH1 binding. G at position 7 is particularly important for binding by both proteins, while the change of T for A at position 8, which does not affect ATH1 binding, reduces complex formation by STM.

Hydroxyl radical footprinting and missing nucleoside experiments showed that ATH1_{HD} interacts with a 7-bp region of the strand carrying the GAC core. On the other strand, protection was observed over a 7-bp region, including the nucleotide complementary to T in position 1. The interference pattern of STM covers an additional nucleotide towards the 3' end

(TGACAGGTG) compared with ATH1. In the bottom strand, interference is observed along an 8-bp stretch that covers the nucleotides complementary to GTGACAGG.

In agreement with the fact that both homeodomains recognise similar target sequences, we observed that most putative important positions are conserved. An exception to this is residue 54, which is Lys in STM and Val in ATH1. The role of residue 54 in determining the respective binding properties was then investigated by producing mutant homeodomains with exchanged amino acids at this position. The results obtained indicated that the decreased affinity displayed by the ATH1 homeodomain correlates with the presence of valine (instead of lysine as in STM) at position 54. This difference also explains the decreased and increased selectivities, respectively, at positions 4 and 5, while other residues, or a combination of them, may be involved in the contact of the eighth position.

Changes at positions 5 and 6 of the TGACAGGT sequence significantly affect the interaction when position 54 is Val, while position 7 is very important for all proteins. Our results point to an essential role of residue 54 in determining the different binding properties of BEL and KNOX homeodomains. Based on our analysis we propose that Val54 of ATH1 may establish contacts with positions 5 and 6 (TGACAG), rather than with position 4, thus producing a relaxed specificity at this nucleotide position relative to homeodomains that contain Lys54. Lys54 may be able to establish hydrogen bonds with a G complementary to C (TGAC) in STM. The establishment of these hydrogen bonds may also explain the increased affinity for DNA displayed by proteins with Lys54. The Val side chain present in ATH would only make less energetic van der Waals or water-mediated contacts with DNA.

Our results demonstrated that, although the DNA binding properties of ATH1, and probably all BEL proteins, show similarities with those presented by KNOX proteins, some differences in affinity and selectivity are also evident. These different properties, mainly attributable to the amino acid present at homeodomain position 54, are probably essential for the respective functions of these proteins *in vivo*, considering the conservation observed at this position within each protein family.

Several reports have shown that KNOX and BEL proteins are able to interact in a DNA-independent manner through conserved domains present in the N-terminal portion of the proteins (MEINOX and MID domains, respectively). Although many reports have shown that

KNOX and BEL proteins are able to interact, the impact of this interaction on the DNA binding properties of the complex has been addressed only in two systems involving proteins from maize and potato. *In vitro* studies have shown that the interaction of KN1 and KIP (KNOX and BEL proteins from maize, respectively) produces an increase in affinity for a DNA fragment containing a single TGAC binding site. In the potato system, however, a DNA sequence with two tandem binding sites present in the *ga20ox1* promoter is required for the interaction with a complex formed by the KNOX and BEL proteins POTH1 and StBEL5. In the present study, we have analyzed the interaction of the complex formed by the *Arabidopsis* proteins STM and BLH3 with DNA sequences containing one or two TGAC cores. The results obtained from electrophoretic mobility shift assays indicate that complex formation produces an increase in binding affinity for both types of sequences and suggest that sequences involved in protein-protein interaction are necessary, and sufficient, to increase binding to DNA. However, complex formation is not correlated with the establishment of new detectable contacts as deduced from missing nucleoside experiments. In addition, in the oligonucleotide with two binding sites, both TGAC sites are contacted by the protein(s) with similar efficiency and establishing similar contacts.

We found that, at low protein concentration, binding of BLH3 and STM alone was similar and very inefficient. However, DNA binding by BLH3 increased more pronouncedly with protein concentration, while binding by STM remained low, showing only a modest increase. Inclusion of the second protein produced a net increase in DNA binding.

Proteins with mutations in residue 55 of the homeodomain were unable to bind DNA but increased the affinity of their wild-type partners, suggesting that only one of the homeodomains makes contacts with DNA in the complex. In addition, STM has a more relevant role in the interaction of the complex with the double sequence.

The results obtained in *in vitro* assays indicate that the presence of a second binding site does not seem to be required for efficient interaction of the STM-BLH3 complex with DNA. *In vivo* DNA binding assays using the yeast one-hybrid system showed that binding of the complex to the single TGAC site was not significantly different from binding of BLH3 or STM alone. Additionally, this behaviour depends on the protein expression level. At high level of BLH3 and STM expression, a competition for binding to DNA between both proteins was

observed, whereas low levels of BLH3 led to an increase of reporter activity in the presence of high levels of STM. The BEL proteins ATH1 and BEL1 were also able to interact *in vivo* with the sequence TGACAGGT and to produce complexes with STM that have similar binding properties as those observed with BLH3.

On the contrary, single proteins were unable to bind, or did so only very weakly, the tandem binding sites present in TGACTTGACAGGT *in vivo*, and a pronounced increase in binding efficiency to DNA was obtained when the two proteins were co-expressed.

We conclude that complex formation between the BEL protein BLH3 and the KNOX protein STM increases binding affinity for DNA in a manner that depends on the sequence of the DNA binding site. This behavior may reflect the existence of different regulatory modes by KNOX and BEL proteins, either alone or in combination, depending on specific sequences present in their target genes. For single TGAC cores this effect will also depend on the relative concentrations of the two proteins within the cell.

Finally, a genome survey allowed the identification of genes related with KNOX and BEL protein functions that contain the single or double target sites in their promoters. We found that the *KNAT2*, *KNAT4*, *BLH2*, and *BLH5* genes and several genes implicated in the control of KNOX function contain the TGACAG sequence, whereas the *BLH3* promoter contains the TGACTTGAC motif. These observations suggest that a complex network of protein-protein and protein-DNA crossed interactions may participate in the regulation of TALE homeodomain protein action.

HAT3.1 is a member of the PHD finger homeodomain protein family. The HAT3.1 homeodomain presents conserved hydrophobic residues involved in maintaining the tertiary structure, but is highly divergent in sequence even at positions that are almost invariable among other homeodomains. In the present Thesis, we have applied the random oligonucleotide selection technique to investigate if the HAT3.1 homeodomain is able to recognize specific DNA sequences. This analysis suggests that HAT3.1 binds the sequence T(A/G)(A/C)ACCA, with a preference for A at the second and third positions. This sequence is different from those bound by other homeodomains.

Missing nucleoside experiments indicate that nucleosides comprising the T(A/G)(A/C)ACCA sequences are required for binding. Modifications at positions 1 and 4 affect DNA recognition, whereas changes in nucleotides 5 and 6 severely decrease the interaction of HAT3.1 with DNA.

The study of the interaction of mutants at positions that usually participate in DNA binding in most homeodomains indicates that residues at positions 47, 50, 51 and 54 are also important for HAT3.1-DNA binding. The interaction of mutants at these positions with DNA carrying nucleotide changes at specific sites suggested that H51 and K50 most likely interact with nucleotides 2 to 4 and 5 to 6, respectively, while W54 would establish contacts with position 4. The presence of H51 and W54 represents an innovation among homeodomain structures. It can be proposed that W54 may be one of the determinants of binding to TAAA instead of TAAT, as most homeodomains do.

The introduction of W54 and H51 represents an innovation since, to our knowledge; these residues are not present in other homeodomains at these positions. While the presence of tryptophan at the more variable position 54 likely originated a new binding specificity (in fact, positions 50 and 54 are usually regarded as the main source of change in binding site preferences), the presence of H51 is more intriguing, since asparagine is almost invariably present at this position. In this sense, we found that the protonation state of H51 influences the efficiency of HAT3.1 binding to DNA.

The fact that the HAT3.1 homeodomain is able to interact with specific DNA sequences is evidence of the inherent plasticity of the homeodomain as a DNA binding unit and indicates that it represents a useful source of evolutionary changes in protein-DNA interaction mechanisms.

BIBLIOGRAFÍA

8 BIBLIOGRAFÍA

Acámpora D., D' Esposito M., Faiella A., Pannese M., Migliaccio E., Morelli F., Stornaiuolo A., Nigro V., Simeone A., y Boncinelli E. (1989) The human HOX gene family. *Nucleic Acids Res.* 17, 10385-10402.

Ades S. y Sauer R. (1994) Differential DNA binding specificity of the Engrailed homeodomain: the role of residue 50. *Biochemistry* 33, 9187-9194.

Affolter M., Percival-Smith A., Müller M., Leupin W., y Gehring W. (1990) DNA binding properties of the purified Antennapedia homeodomain. *Proc. Natl. Acad. Sci. USA* 87, 4093-4097.

Aida M., Ishida T., y Tasaka M. (1999) Shoot apical meristem and cotyledon formation during Arabidopsis embryogenesis: interaction among the *CUP-SHAPED COTYLEDON* and *SHOOT MERISTEMLESS* genes. *Development* 126, 1563-1570.

Ariel F.D., Manavella P.A., Dezar C.A., y Chan R.L. (2007) The true story of the HD-Zip family. *Trends Plant Sci.* 12, 419-26.

Ausubel F., Brent R., Kingston R., Moore D., Seidman J., Smith J., y Struhl K.(1987) *Current Protocols in Molecular Biology.* John Wiley & Sons, N.Y.

Baird-Titus J.M., Clark-Baldwin K., Dave V., Caperelli C.A., Ma J., y Rance M. (2006) The solution structure of the native K50 bicoid homeodomain bound to the consensus TAATCC DNA-binding site. *J. Mol. Biol.* 356, 1137-1151.35.

Banerjee A.K., Chatterjee M., Yu Y., Suh S.G., Miller W.A., y Hannapel D.J. (2006) Dynamics of a mobile RNA of potato involved in a long-distance signaling pathway. *Plant Cell* 18, 3443-57.

Banerjee-Basu S., Morely T., Hsu B., Trout K., y Baxevanis A. (2003) The homeodomain resource: 2003 update. *Nuc. Acids Res.* 31, 304-6.

Bao X., Franks R.G., Levine J.Z., y Liu Z. (2004) Repression of AGAMOUS by BELLRINGER in floral and inflorescence meristems. *Plant Cell* 16, 1478–1489.

Bartel P.L., Chien C.T., Sternglanz R. y Fields S. (1993) Using the two-hybrid system to detect protein-protein interactions. In *Cellular Interactions in Development: A Practical Approach.*, ed. Hartley, D.A. (Oxford University Press, Oxford), 153–179.

Beeman R., Stuart J., Haas M., y Denell R. (1989) Genetic analysis of the homeotic gene complex (HOM-C) in the beetle *Tribolium castaneum*. *Dev Biol.* 133,196-209.

Bellaoui M., Pidkowich M.S., Samach A., Kushalappa K., Kohalmi S.E., Modrusan Z., Crosby W.L. y Haughn G.W. (2001) The Arabidopsis BELL1 and KNOX TALE homeodomain proteins interact through a domain conserved between plants and animals. *Plant Cell* 13, 2455–2470.

Belles-Boix E., Hamant O., Witiak S.M., Morin H., Traas J., y Pautot V. (2006) KNAT6: an Arabidopsis homeobox gene involved in meristem activity and organ separation. *Plant Cell* 18, 1900-7.

Berthelsen J., Zappavigna V., Ferretti E., Mavilio F., y Blasi F. (1998) The novel homeoprotein Prep1 modulates Pbx-Hox protein cooperativity. *EMBO J* 17:1434-1445.

Bhatt A.M., Etchells J.P., Canales C., Lagodienko A., y Dickson H. (2004) VAAMANA a BEL1-like homeodomain protein interacts with KNOX proteins BP and STM and regulates inflorescence stem growth in Arabidopsis. *Gene* 328, 103–111.

Bienz M. (2006) The PHD finger, a nuclear protein-interaction domain. *Trends Biochem. Sci.* 31, 35-40.

Birnboim H. y Dolly J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nuc. Acids Res.* 7, 1513-1515.

Blackwell T.K., y Weintraub H. (1990) Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection, *Science* 250, 1104-1110.

Brambilla V., Battaglia R., Colombo M., Masiero S., Bencivenga S., Kater M., y Colombo L. (2007) Genetic and Molecular Interactions between BELL1 and MADS Box Factors Support Ovule Development in Arabidopsis. *Plant Cell* 19, 2544-56.

Bürglin T. (1994) A comprehensive classification of homeobox genes. In Duboule, D. (ed), *Guidebook to the Homeobox Genes.* Oxford University Press, Oxford, UK, 27-269

Bürglin T. (1997) Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, IROQUOIS, TGIF) reveals a novel domain conserved between plants and animals. *Nucl. Acids Res.* 25, 4173-4180.

Byrne M.E., Groover A.T., Fontana J.R., y Martienssen R.A. (2003) Phyllotactic pattern and stem cell fate are determined by the Arabidopsis homeobox gene BELLRINGER. *Development* 130, 3941–3950.

- Byrne M.E., Barley R., Curtis M., Arroyo J.M., Dunham M., Hudson A. y Martienssen R.A.** (2000) Asymmetric leaves1 mediates leaf patterning and stem cell function in Arabidopsis. *Nature* 408, 967–971.
- Byrne M.E., Simorowski J., y Martienssen R.A.** (2002) ASYMMETRIC LEAVES1 reveals knox gene redundancy in Arabidopsis. *Development* 129, 1957–1965.
- Chan R., Gago G., Palena C., y González D.** (1998) Homeoboxes in plant development. *Biochim. Biophys. Acta* 93134, 1-19.
- Chaney B.A., Clark-Baldwin K., Dave V., Ma J., y Rance M.** (2005) Solution structure of the K50 class homeodomain PITX2 bound to DNA and implications for mutations that cause Rieger syndrome. *Biochemistry* 44, 7497-7511.
- Chang C.P., Jacobs Y., Nakamura T., Jenkins N.A., Copeland N.G., y Cleary M.L.** (1997) Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins. *Mol Cell Biol* 17, 5679-5687.
- Chariot A., van Lint C., Chapelier M., Gielen J., Merville M-P, y Bours V.** (1999a) CBP and histone deacetylase inhibition enhance the transactivation potential of the HOXB7 homeodomain-containing protein. *Oncogene* 18, 4007-4014.
- Chatterjee M, Banerjee A, y Hannapel D.** (2007) A BELL1-like gene of potato is light-activated and wound-inducible. *Plant Physiol.* 145, 1435-43.
- Chen H., Banerjee A.K., y Hannapel D.J.** (2004) The tandem complex of BEL and KNOX partners is required for transcriptional repression of *ga20ox1*. *Plant Journal* 38, 276-284.
- Chen H., Rosin F.M., Prat S., y Hannapel D.J.** (2003) Interacting transcription factors from the TALE superclass regulate tuber formation. *Plant Physiol.* 132, 1391-1404.
- Cole M., Nolte C., y Werr W.** (2006) Nuclear import of the transcription factor SHOOT MERISTEMLESS depends on heterodimerization with BLH proteins expressed in discrete sub-domains of the shoot apical meristem of Arabidopsis thaliana. *Nucl. Acids Res.* 34, 1281-1292.
- Damante G., Fabbro D., Pellizzari L., Civitareale D., Guazzi S., Polycarpou-Schwartz M., Cauci S., Quadrifoglio F., Formisano S., y Di Lauro R.** (1994) Sequence-specific DNA recognition by the thyroid transcription factor-1 homeodomain. *Nuc. Acids Res.* 22, 3075-3083.
- Damante G., Pellizzari L., Esposito G., Fogolari F., Viglino P., Fabbro D., Tell G., Formisano S., y Di Lauro R.** (1996) A molecular code dictates sequence-specific DNA recognition by homeodomains. *EMBO J.* 15, 4992-5000.

- Dave V., Zhao C., Yang F., Tung C., y Ma J.** (2000) Reprogrammable recognition codes in bicoid homeodomain-DNA interaction. *Mol. Cell Biol.* 20, 7673-7684.
- Dean G.S., Casson S., y Lindsey K.** (2004) KNAT6 gene of Arabidopsis is expressed in roots and is required for correct lateral root formation. *Plant. Mol. Biol.* 54, 71–84.
- Dezar C., Fedrigo G., y Chan R.** (2005) The promoter of the sunflower HD-Zip protein gene Hahb4 directs tissue-specific expression and is inducible by water stress, high salt concentrations and ABA. *Plant Science* 169, 447-456.
- Dixon W., Hayes J., Levin J., Weidner M., Dombrosky B. y Tullius T.** (1991) Hydroxyl radical footprinting. *Methods Enzimol* 208, 380-413.
- Dockx J., Quaedvlieg N., Keultjes G., Kock P., Weisbeek P., y Smeekens S.** (1995) The homeobox gene Atk1 of Arabidopsis-thaliana is expressed in the shoot apex of the seedling and in flowers and inflorescence stems of mature plants. *Plant Mol. Biol.* 28, 723-737.
- Dong Y.H., Yao J.L., Atkinson R.G., Putterill J.J., Morris B.A., y Gardner R.C.** (2000) MDH1: an apple homeobox gene belonging to the BEL1 family. *Plant Mol Biol* 42, 623-633.
- Douglas S.J., Chuck G., Dengler R.E., Pelecya L., y Riggs C.D.** (2002) KNAT1 and ERECTA regulate inflorescence architecture in Arabidopsis. *Plant Cell* 14, 547–558.
- Draganescu A., y Tullius T.** (1998) The DNA binding specificity of engrailed homeodomain. *J. Mol. Biol.* 276, 529-536.
- Draganescu A., Levin J., y Tullius T.** (1995) Homeodomain proteins: what governs their ability to recognize specific DNA sequences? *J. Mol. Biol.* 250, 595-608.
- Duboule D.** (1994) *Guidebook to the Homeobox Genes.* Oxford University Press, Oxford.
- Feinberg A. y Vogelstein B.** (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* 132, 6-13.
- Garber R., Kuroiwa A., y Gehring W.** (1983) Genomic and cDNA clones of the homeotic locus Antennapedia in Drosophila. *EMBO J.* 2, 2027-36.
- Gehring W.** (1987) Homeoboxes in the study of development. *Science* 236, 1245-1252.
- Gietz D., St. Jean A., Woods R.A. y Schiestl R. H.** (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.* 20, 1425.
- Graham A., Papalopulu N., y Krumlauf R.** (1989) The murine and Drosophila homeobox gene complexes have common features of organization and expression. *Cell* 57, 367-378.

- Hackbusch J., Richter K., Müller J., Salamini F., y Uhrig J.F.** (2005) A central role of *Arabidopsis thaliana* ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins. *Proc Natl Acad Sci USA*. 102, 4908–4912.
- Hake S., Smith H.M.S., Holtan H., Magnini E., Mele G., y Ramirez J.** (2004) The role of *knox* genes in plant development. *Annu Rev Cell Dev Biol* 20, 125–151.
- Halbach T., Scheer N., y Werr W.** (2000) Transcriptional activation by the PHD finger is inhibited through an adjacent leucine zipper that binds 14-3-3 proteins. *Nucleic Acids Res.* 28, 3542-50.
- Hamant O., Nogue F., Belles-Boix E., Jublot D., Gryjean O., Traas J. y Pautot V.** (2002) The KNAT2 homeodomain protein interacts with ethylene and cytokinin signaling. *Plant Physiol.* 130, 657–665.
- Hanahan D.** (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166, 557-580.
- Hanes S. y Brent R.** (1989) DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9. *Cell* 57, 1275-1283.
- Hanes S. y Brent R.** (1991) A genetic model for interaction of the homeodomain recognition helix with DNA. *Science* 251, 426-430.
- Hareven D., Gutfinger T., Parnis A., Eshed Y. y Lifschitz E.** (1996) The making of a compound leaf: genetic manipulation of leaf architecture in tomato. *Cell* 84, 735–744.
- Harper J.W., Adami G.R., Wei N., Keyomarsi K., y Elledge S.J.** (1993) The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* 75, 805-16.
- Hay A., Craft J., y Tsiantis M.** (2004) Plant hormones and homeoboxes: bridging the gap? *Bioessays* 26, 395–404.
- Hay A., Kaur H., Phillips A., Hedden P., Hake S., y Tsiantis M.** (2002) The gibberellin pathway mediates KNOTTED1-type homeobox function in plants with different body plans. *Curr. Biol.* 12, 1557–1565.
- Hayes J. y Tullius, T.** (1989) The missing nucleoside experiment: a new technique to study recognition of DNA by protein. *Biochemistry* 28, 9521-9527.
- Hoey T. y Levine M.** (1988) Divergent homeobox proteins recognize similar DNA sequences in *Drosophila*. *Nature* 332, 858-861.

Hoffman C.S. y Winston F. (1987) A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene*. 57, 267-272.

Hovde S., Abate-Shen C., y Geiger J. (2001) Crystal structure of the Msx-1 homeodomain/DNA complex. *Biochemistry* 40, 12013-12021.

Jaw T.J., You L.R., Knoepfler P.S., Yao L.C., Pai C.Y., Tang C.Y., Chang L.P., Berthelsen J., Blasi F., Kamps M.P., y Sun Y.H. (2000) Direct interaction of two homeoproteins, Homothorax and Extradenticle, is essential for EXD nuclear localization and function. *Mech Dev* 91, 279- 291.

Kaiser P. y Auer B. (1993) Rapid shuttle plasmid preparation from yeast cells by transfer to *E. coli*. *BioTechniques* 14, 552.

Kanrar S., Onguka O., y Smith H.M. (2006) Arabidopsis inflorescence architecture requires the activities of KNOX-BELL homeodomain heterodimers. *Planta* 224, 1163-1173.

Kaufman T., Levis R., y Wakimoto B. (1980) Cytogenic analysis of chromosome 3 in *Drosophila melanogaster*: the homeotic gene complex in polytene chromosome interval 84-A-B. *Genetics* 94, 115-133.

Ke A. y Wolberger C. (2003) Insights into binding cooperativity of Mata1/MATalpha2 from the crystal structure of a MATa1 homeodomain-maltose binding protein chimera. *Protein Sci.* 12, 306-312.

Kirch T., Bitter S., Kisters-Woike B., y Werr W. (1998) The two homeodomains of the *ZmHox2a* gene from maize originated as an internal gene duplication and have evolved different target site specificities. *Nucleic Acids Res.* 26, 4714-4720.

Kissinger C., Liu B., Martin-Blanco E., Kornberg T., y Pabo C. (1990) Crystal structure of an engrailed homeodomain-DNA complex at 2,8 Å resolution: a framework for understanding homeodomain-DNA interactions. *Cell* 63, 579-590.

Klemm J., Rould M., Aurora R., Herr W., y Pabo C. (1994) Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell* 77, 21-32.

Klinge B., Uberlacker B., Korfhage C., y Werr W. (1996) *ZmHox*: a novel class of maize homeobox genes, *Plant Mol. Biol.* 30, 439-453.

Kneale G. (1994) *Dna-protein interactions. Principles and Protocols.* Humana Press, New Jersey.

Knoepfler P.S., Calvo K.R., Chen H., Antonarakis S.E., y Kamps M.P. (1997) Meis1 and pKnox1 bind DNA cooperatively with Pbx1 utilizing an interaction surface disrupted in oncoprotein E2a-Pbx1. *Proc Natl Acad Sci USA* 94, 14553-14558.

Koizumi K., Lintas C., Nirenberg M., Maeng J., Ju J., Mack J., Gruschus J., Odenwald W., y Ferretti J. (2003) Mutations that affect the ability of the vnd/Nk-2 homeoprotein to regulate gene expression: transgenic alterations and tertiary structure. *Proc. Natl. Acad. Sci. USA* 100, 3119-3124.

Korfhage U., Trezzini G.F., Meier I., Hahlbrock K., y Somssich I.E. (1994) Plant homeodomain protein involved in transcriptional regulation of a pathogen defense-related gene. *Plant Cell* 5, 695-708.

Krusell L., Rasmussen I., y Gausing K. (1997) DNA binding sites recognised in vitro by a knotted class I homeodomain protein encoded by the *hooded* gene, *k*, in barley (*Hordeum vulgare*) *FEBS Lett.* 408, 25-29.

Kumar R., Kushalappa K., Godt D., Pidkowich M.S., Pastorelli S., Hepworth S.R., y Haughn G.W. (2007) The Arabidopsis BEL1-LIKE HOMEODOMAIN proteins SAW1 and SAW2 act redundantly to regulate KNOX expression spatially in leaf margins. *Plant Cell* (PMID: 17873098).

Kushner H. (1978) An improved method of transformation of *E. coli* with col E1-derived plasmids. En: *Genetic Engineering*. Elsevier/North-Holland. Amsterdam.

Laemmli U. (1970) Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. *Nature* 227, 680-685.

LaRonde-LeBlanc N. y Wolberger C. (2003) Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. *Genes Dev.* 17, 2060-2072.

Laughon A. (1991) DNA binding specificity of homeodomains. *Biochemistry* 30, 11357-11367.

Laughon A. y Scott M. (1984) Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. *Nature* 310, 25-31.

Levin J., Burkhoff A. y Tullius T. (1991) Using the chemistry of the hydroxyl radical to determine structural details about DNA and protein-DNA complexes. En: *A Laboratory Guide to in Vitro Studies of Protein/DNA Interactions, BioMethods Series*, (Saluz, H.P. & Jost, J.P., eds), pp.133-134, Birkhauser Verlag.

Levine M. y Hoey T. (1988) Homeobox proteins as sequence-specific transcription factors. *Cell* 55, 537-540.

- Lewis E.** (1978) A gene complex controlling segmentation in *Drosophila*. *Nature* 276, 565-570.
- Li H., Ilin S., Wang W., Duncan E.M., Wysocka J., Allis C.D., y Patel D.J.** (2006) Molecular basis for site-specific readout of histone H3K4me3 by the BPTF PHD finger of NURF, *Nature* 442, 91-95.
- Li T., Stark M., Johnson A., y Wolberger C.** (1995) Crystal structure of the MATa1/MATa2 homeodomain heterodimer bound to DNA. *Science* 270, 262-269.
- Lincoln C., Long J., Yamaguchi J., Serikawa K., y Hake S.** (1994) A knotted1-like homeobox gene in *Arabidopsis* is expressed in the vegetative meristem and dramatically alters leaf morphology when overexpressed in transgenic plants. *Plant Cell* 6, 1859–1876.
- Long J.A., Moan E.I., Medford J.I., y Barton M.K.** (1996) A member of the KNOTTED class of homeodomain proteins encoded by the *SHOOTMERISTEMLESS* gene of *Arabidopsis*. *Science* 379, 66–69.
- Louret O.F., Doignon F. y Crouzet M.** (1997) Stable DNA binding yeast vector allowing high bait expression for use in the two-hybrid system. *BioTechniques*. 23, 816-819.
- Luo H., Song F., Goodman R.M., y Zheng Z.** (2005) Up-regulation of *OsBIHD1*, a rice gene encoding BELL homeodomain transcriptional factor, in disease resistance responses. *Plant Biol* 7, 459-468.
- Luo Y., Vijaychyer S., Stile J. y Zhu L.** (1996) Cloning and analysis of DNA-binding proteins by yeast one-hybrid and one-two-hybrid systems. *BioTechniques* 20, 564–568.
- Maconochie M., Nonchev S., Morrison A., y Krumlauf R.** (1996) Paralogous HOX genes: function and regulation. *Annu. Rev. Genet.* 30, 529-556.
- Malicki J., Schughart K., y McGinnis W.** (1990) Mouse Hox 2.2 specifies thoracic segmental identity in *Drosophila* embryos and larvae. *Cell* 63, 961-967.
- Mann R. y Chan S.** (1996) Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. *Trends Genet.* 12, 258-62.
- Mathias J., Zhong H., Jin Y., y Vershon A.** (2001) Altering the DNA-binding specificity of the yeast Mat α 2 homeodomain protein. *J. Biol. Chem.* 276, 32696-32703.
- Matsuoka M., Ichikawa H., Saito A., Tamda Y., Fujimura T., y Kano-Murakami Y.** (1993) Expression of a rice homeobox gene causes altered morphology of transgenic plants. *Plant Cell* 5, 1039-1048.

McGinnis N., Kuziora M., y McGinnis W. (1990) Human Hox 4.2 and Drosophila Deformed encode similar regulatory specificities in Drosophila embryos and larvae. *Cell* 63, 969-976.

McGinnis W., Garber R., Wirz J., Kuroiwa A., y Gehring W. (1984b) A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. *Cell* 37, 403-408.

McGinnis W., Levine M., Hafen E., Kuroiwa A., y Gehring W. (1984a) A conserved DNA sequence in homeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature* 308, 428-433.

Meijer A.H., Ouwerkerk P.B.F., y Hoge J.H.C. (1998) Vectors for transcription factor isolation and target gene identification by means of genetic selection in yeast. *Yeast* 14: 1407-1416.

Mele G., Ori N., Sato Y., y Hake S. (2003) The knotted1-like homeobox gene BREVIPEDICELLUS regulates cell differentiation by modulating metabolic pathways. *Genes Dev* 17, 2088-2093.

Miller J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory. Cold.

Modrusan Z., Reiser L., Feldmann K.A., Fischer R.L., y Haughn G.W. (1994) Homeotic transformation of ovules into carpel-like structures in Arabidopsis. *Plant Cell* 6, 333-349.

Müller J., Wang Y., Franzen R., Santi L., Salamini F., y Rohde W. (2001) In vitro interactions between barley TALE homeodomain proteins suggest a role for protein-protein associations in the regulation of Knox gene function. *Plant J.* 27, 13-23.

Nelson C., Albert V., Elsholtz H., Lu L., y Rosenfeld M. (1988) Activation of cell-specific expression of rat growth hormone and prolactin genes by a common transcription factor. *Science*. 239, 1400-5.

Oliphant A.R., Bryl C.J., y Struhl K. (1989) Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: analysis of yeast GCN4 protein, *Mol. Cell Biol.* 9, 2944-2949.

Otting G., Qian Y., Billeter M., Muller M., Affolter M., Gehring W., y Wuthrich K. (1990) Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO J.* 9, 3085-3092.

Pabo C. y Sauer R. (1984) Protein-DNA recognition. *Ann. Rev. Biochem.* 53, 293-321.

Palena C., González D., Guelman S., y Chan R. (1998) Expression of sunflower homeodomain containing proteins in Escherichia coli. Purification and functional studies. *Protein Expr. Purif.* 13, 97-103.

- Passner J., Ryoo H., Shen L., Mann R., y Aggarwal A.** (1999) Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex. *Nature* 397, 714-719.
- Pautot V., Dockx J., Hamant O., Kronenberger J., Gryjean O., Jublot D., y Traas J.** (2001) KNAT2: evidence for a link between knotted1-like genes and carpel development. *Plant Cell* 13, 1719–1734.
- Pellizzari L., Tell G., Fabbro D., Pucillo C., y Damante G.** (1997) Functional interference between contacting amino acids of homeodomains. *FEBS Lett.* 407, 320-324.
- Pendelton J., Nagai B., Murtha M., y Ruddle F.** (1993) Expansion of the Hox gene family and the evolution of chordates. *Proc. Natl. Acad. Sci. USA.* 90, 6300-6304.
- Penkov D., Tanaka S., Di Rocco G., Berthelsen J., Blasi F., y Ramirez F.** (2000) Cooperative interactions between PBX, PREP, and HOX proteins modulate the activity of the alpha 2(V) collagen (COL5A2) promoter. *J Biol Chem* 275, 16681-16689
- Percival-Smith A., Müller M., Affolter M., y Gehring W.** (1990) The interaction with DNA of wild-type and mutant fushi tarazu homeodomains. *EMBO J.* 9, 3967-3974.
- Plesch G., Stormann K., Torres J.T., Walden R., y Somssich I.E.** (1997) Developmental and auxin-induced expression of the Arabidopsis PRHA homeobox gene, *Plant J.* 12, 635-647.
- Pomerantz J. y Sharp P.** (1994) Homeodomain determinants of major groove recognition. *Biochemistry* 33, 10851-10858.
- Porter S. y Smith M.** (1986) Homoeo-domain homology in yeast MAT alpha 2 is essential for repressor activity. *Nature.* 320, 766-768.
- Ptashne M.** (1988) How eukaryotic transcriptional activators work. *Nature* 335, 683-689.
- Qian Y., Billeter M., Otting G., Müller M., Gehring W., y Wüthrich K.** (1989) The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. *Cell* 59, 573-580.
- Qian Y., Furukubo-Tokunaga K., Müller M., Resendez-Perez D., Gehring W., y Wüthrich K.** (1994) Nuclear magnetic resonance solution structure of the fushi tarazu homeodomain from *Drosophila* and comparison with the Antennapedia homeodomain. *J. Mol. Biol.* 238, 333-345.
- Quaedvlieg N., Dockx J., Rook F., Weisbeek P., y Smeekens S.** (1995) The homeobox gene *ATH1* of Arabidopsis is derepressed in the photomorphogenic mutants *cop1* and *det1*. *Plant Cell* 7, 117-129.

Ray A., Robinson-Beers K., Ray S., Baker S., Lang J., Preuss D., Milligan S., y Gasser C. (1994) Arabidopsis floral homeotic gene BELL (BEL1) controls ovule development through negative regulation of AGAMOUS gene (AG) Proc. Natl. Acad. Sci. USA 91, 5761–5765.

Reiser L., Modrusan Z., Margossian L., Samach A., Ohad N., Haughn G. y Fischer, R. (1995) The BELL1 gene encodes a homeodomain protein involved in pattern formation in the Arabidopsis ovule primordium. Cell 83, 735–742.

Reiser L., Sánchez-Barracaldo P., y Hake S. (2000) Knots in the family tree: evolutionary relationships and functions of knox homeobox genes. Plant Mol. Biol. 42, 151-166.

Roeder A.H., Ferrandiz C., y Yanofsky M.F. (2003) The role of the REPLUMLESS homeodomain protein in patterning the Arabidopsis fruit. Curr Biol 13, 1630–1635

Rosin F.M., Hart J.K., Horner H.T., Davies P.J., y Hannapel D.J. (2003) Overexpression of a knotted-like homeobox gene of potato alters vegetative development by decreasing gibberellin accumulation. Plant Physiol. 132, 106-17.

Ruberti I., Sessa G., Lucchetti S., y Morelli G. (1991) A novel class of plant proteins containing a homeodomain with a closely linked leucine zipper motif. EMBO J. 10, 1787-1791.

Rupp H.M., Frank M., Werner T., Strnad M., y Schmulling T. (1999) Increased steady state mRNA levels of the STM and KNAT1 homeobox genes in cytokinin overproducing Arabidopsis thaliana indicate a role for cytokinins in the shoot apical meristem. Plant J. 18, 557–563.

Sakamoto T., Kamiya N., Ueguchi-Tanaka M., Iwahori S., y Matsuoka M. (2001) KNOX homeodomain protein directly suppresses the expression of a gibberellin biosynthetic gene in the tobacco shoot apical meristem. Genes Dev. 15, 581–590.

Sakamoto T., Nishimura A., Tamaoki M., Kuba M., Tanaka H., Iwahori S., y Matsuoka M. (1999) The conserved KNOX domain mediates specificity of tobacco KNOTTED1-type homeodomain proteins. Plant Cell 11, 1419-1431.

Saleh M., Rambaldi I., Yang X., y Featherstone M. (2000) Cell signaling switches HOX-PBX complexes from repressors to activators of transcription mediated by histone deacetylases and histone acetyltransferases. Mol. Cell Biol. 20, 8623-8633.

Sambrook J., Fritsch E.F. y Maniatis T. (1989) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, New York.

Sanchez-Garcia I. y Rabbitts T. (1993) Redox regulation of in vitro DNA-binding activity by the homeodomain of the Isl-1 protein. J. Mol. Biol. 231, 945-949.

Sanger F. (1981) Determination of nucleotide sequences in DNA. Science 214, 1205-1210.

- Sauer R.** (1991) *Protein-DNA Interactions. Methods in Enzymology*. Vol. 208. Academic Press, Inc.
- Scanlon M.J.** (2003) The polar auxin transport inhibitor N-1-naphthylphthalamic acid disrupts leaf initiation, KNOX protein regulation, and formation of leaf margins in maize. *Plant Physiol.* 133, 597–605.
- Schneeberger R., Tsiantis M., Freeling M., y Langdale J.A.** (1998) The *rough sheath2* gene negatively regulates homeobox gene expression during maize leaf development. *Development* 125, 2857-2865.
- Schier A. y Gehring W.** (1992) Direct homeodomain-DNA interaction in the autoregulation of the fushi tarazu gene. *Nature* 356, 804-807.
- Schindler U., Beckmann H. y Cashmore A.R.** (1993) HAT3.1, a novel Arabidopsis homeodomain protein containing a conserved cysteine-rich region. *Plant J.* 4, 137-150.
- Schwede T., Kopp J., Guex N. y Peitsch M.** (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nuc. Acids Res.* 31, 3381-3385.
- Scofield S., Dewitte W., y Murray J.A.** (2007) The KNOX gene *SHOOTMERISTEMLESS* is required for the development of reproductive meristematic tissues in Arabidopsis. *Plant J.* 50, 767-81.
- Sedmak J. y Grossberg S.** (1977) A rapid, sensitive, and versatile assay for protein using Coomassie brilliant blue G-250. *Anal. Biochem.* 79, 544-552.
- Seki M., Narusaka M., Kamiya A., Ishida J., Satou M., Sakurai T., Nakajima M., Enju A., Akiyama K., Oono, Y., Muramatsu M., Hayashizaki Y., Kawai J., Carninci P., Itoh M., Ishii Y., Arakawa T., Shibata K., Shinagawa A., y Shinozaki K.** (2002) Functional annotation of a full-length Arabidopsis cDNA collection. *Science* 296, 141-145.
- Serikawa K.A., Martinez-Laborda A. y Zambryski P.** (1996) Three knotted1-like homeobox genes in Arabidopsis. *Plant Mol. Biol.* 32, 673–683.
- Serikawa K.A., Martinez-Laborda A., Kim H.S. y Zambryski P.C.** (1997) Localization of expression of KNAT3, a class 2 knotted1-like gene. *Plant J.* 11, 853–861.
- Sessa G., Morelli G., y Ruberti I.** (1993) The Athb-1 and HD-Zip domains homodimerize forming complexes of different DNA binding specificities. *EMBO J.* 12, 3507-3517.
- Shang Z., Ebright Y., Iler N., Pendergrast P., Echelard Y., McMahon A., Ebright R., y Abate C.** (1994) DNA affinity cleaving analysis of homeodomain-DNA interaction: Identification of homeodomain consensus sites in genomic DNA. *Proc. Natl. Acad. Sci. USA.* 91, 118-122.

- Shepherd J., McGinnis W., Carrasco A., De Robertis E., y Gehring W.** (1984) Fly and frog homeodomains show homologies with yeast mating type regulatory proteins. *Nature* 310, 70-71.
- Silver J., Limjoco T., y Feinstone S.** (1995) Site-specific mutagenesis using the polymerase chain reaction, 179-188. En: *PCR Strategies*. Innis M.A., Gelfand D.H., Sninsky J.J., editores. Academic Press, San Diego.
- Smeekens S., Weisbeek P. y Proveniers M.** (1998) Patent. International Publication Number WO 98/51800.
- Smith D. y Johnson K.** (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 67, 31-40.
- Smith H. y Hake S.** (2003) The interaction of two homeobox genes, *BREVIPEDICELLUS* and *PENNYWISE*, regulates internode patterning in the *Arabidopsis* inflorescence. *Plant Cell* 15, 1717-1727.
- Smith H., Boschke I., y Hake S.** (2002) Selective interaction of plant homeodomain proteins mediates high DNA-binding affinity. *Proc. Natl. Acad. Sci. USA* 99, 9579-9584.
- Smith H., Campbell B., y Hake S.** (2004) Competence to respond to floral inductive signals requires the homeobox genes *PENNYWISE* and *POUND-FOOLISH*. *Curr. Biol.* 14, 812-817
- Smith L., Greene B., Veit B., y Hake S.** (1992) A dominant mutation in the maize homeobox gene, *Knotted-1*, causes its ectopic expression in leaf cells with altered fates. *Development* 116, 21-30.
- Studier F.W. y Moffatt B.A.** (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189, 113-130.
- Tamaoki M., Kusaba S., Kano Murakami Y. y Matsuoka M.** (1997) Ectopic expression of a tobacco homeobox gene, *NTH15*, dramatically alters leaf morphology and hormone levels in transgenic tobacco. *Plant Cell Physiol.* 38: 917-927.
- Tan Q.K., e Irish V.F.** (2006) The *Arabidopsis* zinc finger-homeodomain genes encode proteins with unique biochemical properties that are coordinately expressed during floral development. *Plant Physiol.* 140, 1095-108.
- Tell G., Acquaviva R., Formisano S., Fogolari F., Pucillo C., y Damante G.** (1999) Comparative stability analysis of the thyroid transcription factor-1 and Antennapedia homeodomains: evidence for residue 54 in controlling the structural stability of the recognition helix. *Int. J. Biochem. Cell Biol.* 31, 1339-1353.

- Thomas B.J. y Rothstein R.** (1989) Elevated recombination rates in transcriptionally active DNA. *Cell* 56, 619-630.
- Tioni M.F., Viola I.L., Chan R.L., y Gonzalez D.H.** (2005) Site-directed mutagenesis and footprinting analysis of the interaction of the sunflower KNOX protein HAKN1 with DNA. *FEBS J.* 272, 190-202.
- Treisman J., Gonczy P., Vashishtha M., Harris E., y Desplan C.** (1989) A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell* 59, 553-562.
- Tron A.E., Bertoncini C.W., Chan R.L., y Gonzalez D.H.** (2002) Redox regulation of plant homeodomain transcription factors. *J Biol Chem.* 277, 34800-7.
- Tron, A.E., Bertoncini, C.W., Palena, C.M., Chan, R.L., y Gonzalez, D.H.** (2001) Combinatorial interactions of two amino acids with a single base pair define target site specificity in plant dimeric homeodomain proteins. *Nucleic Acids Res.* 29, 4866-4872.
- Truernit E., Siemering K.R., Hodge S., Grbic V. y Haseloff J.** (2006) A map of *KNAT* gene expression in the Arabidopsis root. *Plant Molecular Biology* 60, 1-20.
- Tsiantis M., Brown M.I., Skibinski G., y Langdale J.A.** (1999) Disruption of auxin transport is associated with aberrant leaf development in maize. *Plant Physiol* 121, 1163-1168.
- Tucker-Kellogg L., Rould M., Chambers K., Ades S., Sauer R., y Pabo C.** (1997) Engrailed (Gln50→Lys) homeodomain-DNA complex at 1.9 Å resolution: structural basis for enhanced affinity and altered specificity. *Structure* 5, 1047-1054.
- Tullius T., y Dombroski B.** (1986) Hydroxyl radical footprinting: high-resolution information about DNA-protein contacts and application to lambda repressor and Cro protein. *Proc Natl Acad Sci U S A.* 83, 5469-5473.
- Uberlacker B., Klinge B., y Werr W.** (1996) Ectopic expression of the maize homeobox genes *ZmHox1a* or *ZmHox1b* causes pleiotropic alterations in the vegetative and floral development of transgenic tobacco, *Plant Cell* 8, 349-362.
- Venglat S.P., Dumonceaux T., Rozwadowski K., Parnell L., Babic V., Keller W., Martienssen R., Selvaraj G., y Datla R.** (2002) The homeobox gene *BREVIPEDICELLUS* is a key regulator of inflorescence architecture in Arabidopsis. *Proc Natl Acad Sci USA* 99, 4730-4735.
- Vijapurkar U., Fischbach N., Shen W., Bryts C., Stokoe D., Lawrence H., y Largman C.** (2004) Protein kinase C-mediated phosphorylation of the leukemia-associated HOXA9 protein impairs its DNA binding ability and induces myeloid differentiation. *Mol Cell Biol.* 24, 3827-3837.

Vollbrecht E., Veit B., Sinha N., y Hake S. (1991) The developmental gene Knotted-1 is a member of a maize homeobox gene family. *Nature* 350, 241-243.

Wilson D., Sheng G., Jun S., y Desplan C. (1996) Conservation and diversification in homeodomain-DNA interactions: A comparative genetic analysis. *Proc. Natl. Acad. Sci. USA* 93, 6886-6891.

Wolberger C., Vershon A., Liu B., Johnson A., y Pabo C. (1991) Crystal structure of a MAT α 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interaction. *Cell* 67, 517-528.

Wolpert L. (1994) Do we understand development?. *Science* 266, 571-572.

Wüthrich K. y Gehring W. (1992) Transcriptional regulation. Ed. S.L.Mc Knight, K.R.

Yanai O., Shani E., Dolezal K., Tarkowski P., Sablowski R., Sandberg G., Samach A., y Ori N. (2005) Arabidopsis KNOX1 proteins activate cytokinin biosynthesis. *Curr Biol.* 15,1566-71.

Yang X., Ogryzko V., Nishikawa J., Howard B., y Nakatani Y. (1996) A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* 382, 319-324.

Yanisch-Perron C., Vieira J. y Messing J. (1985) Improved M13 phage cloning vectors and host strains: Nucleotide sequences of M13mp18 and pUC19 vectors. *Gene* 33, 103-119.

Zgurski J.M., Sharma R., Bolokoski D.A., y Schultz E.A. (2005) Asymmetric auxin response precedes asymmetric growth and differentiation of asymmetric leaf1 and asymmetric leaf2 Arabidopsis leaves. *Plant Cell* 17, 77-91

Zhao J., Lazzarini R., y Pick L. (1993) The mouse Hox-1.3 gene is functionally equivalent to the Drosophila sex combs reduced gene. *Gene Dev.* 7, 343-354.