

03072



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

INSTITUTO DE BIOTECNOLOGIA

5
2cy

Análisis de la interacción de Nifa de *R. meliloti* con las regiones promotoras UAS (Upstream Activator Sequences) de los genes *nif* mediante mutagénesis dirigida

T E S I S
QUE PARA OBTENER EL GRADO DE:
MAESTRIA EN CIENCIAS
P R E S E N T A :
SONIA DAVILA RAMOS

CUERNAVACA, MOR.

1998.

TESIS CON
FALLA DE ORIGEN

25 8881



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Este trabajo se realizó bajo la Dirección del Dr. J. Enrique Morett Sánchez en el Departamento de Bioestructura y Reconocimiento Molecular del Instituto de Biotecnología perteneciente a la Universidad Nacional Autónoma de México.

Durante el desarrollo del proyecto conté con el apoyo de la Beca de Maestría del Consejo Nacional de Ciencia y Tecnología.

Agradezco a los Dres. J. Enrique Morett, Lorenzo Segovia, M. Angel Cevallos, Gloria Soberón y Mario Zurita sus comentarios durante la revisión de esta tesis.

A los Dres. Xavier Soberón y Julio Collado su participación como asesores del proyecto de investigación.

A Gabriel Moreno, Katy Juárez y compañeros del laboratorio por su apoyo, consejos y discusiones.

RESUMEN

Análisis de la interacción de NifA de *R. meliloti* con las regiones promotoras UAS (Upstream Activator Sequences) de los genes *nif* mediante mutagénesis dirigida.

Los distintos elementos que participan en la activación de la transcripción de los genes *nif* nos proveen de un excelente modelo para cumplir el objetivo principal de este trabajo que es el de estudiar las interacciones moleculares entre DNA y proteína.

Durante el desarrollo del trabajo se realizaron predicciones de estructura de DNA tanto de las regiones UAS como de las posibles regiones de interacción de la proteína IHF. Para esto se utilizaron distintos modelos de predicción que toman en cuenta los ángulos que se forman entre los distintos pares de bases. Entre los principales están el de De Santis, y el de Trifonov. Lo que se pretendía era conocer si la estructura del DNA participaba directamente en la especificidad de interacción de las proteínas que interactúan con estas regiones.

En ninguna de las predicciones fue posible observar una curvatura determinada, por lo que se concluyó que muy probablemente es necesario tomar en cuenta otros parámetros que resuelvan la curvatura del DNA ya que no se descarta la posibilidad de que la estructura este involucrada directamente en la interacción y en la identificación de las proteínas involucradas.

La proteína NifA es el activador de los genes que participan en el proceso de Fijación de Nitrógeno. NifA actúa sobre estos genes uniéndose a las secuencias denominadas UAS (upstream activator sequences, TGT-N10-ACA) localizadas entre 100 y 200 pb del promotor.

El dominio carboxilo terminal de la proteína NifA contiene un motivo de Hélice-vuelta-hélice (HVH) similar a otras proteínas de unión al DNA. El residuo más conservado en las HVH es el que se encuentra al inicio de la vuelta entre las dos hélices, el cual generalmente es una glicina. Se determinó que la unión de NifA de *R. meliloti* a las UAS es muy débil y analizando las secuencias de los distintos motivos de HVH se observó que esta proteína presenta un glutámico (E) en la posición 9 sustituyendo a la glicina (G) altamente conservada, lo cual posiblemente sea una de las causas de la débil interacción. Así el residuo de glicina puede ser importante en la estructuración adecuada de la HVH para favorecer el reconocimiento específico del DNA.

Se propuso que al sustituir el residuo de glutámico por glicina podría incrementar la unión de NifA a las UAS y consecuentemente la expresión de los genes que esta proteína regula.

Se construyeron por mutagénesis dirigida un derivado de NIA con una sustitución de E por G y una mutante doble que sustituye E por G y lisina por leucina (K por L).

La actividad transcripcional de estas mutantes y de la proteína silvestre se determinó mediante la expresión de una fusión transcripcional del gene nifH-lacA.

Los resultados obtenidos muestran que la mutante sencilla en todos los ensayos cuenta con al menos el doble de actividad comparada con la actividad de la proteína silvestre, lo que indica que el residuo de glicina es importante en la estructuración correcta de la HVH para favorecer la unión al DNA.

Vo.Bo.

Dr. Juan Enrique Morett Sánchez



INDICE

INTRODUCCION	1
Interacción DNA-Proteína en sistemas relacionados con la fijación de Nitrógeno	2
Generalidades del Sistema de Regulación en la Fijación de Nitrógeno	2
Elementos principales que participan en la transcripción de los genes <i>nif</i>	3
La proteína NifA	3
La proteína IHF	4
Introducción sobre las moléculas de DNA	6
Métodos usados en el estudio de la estructura del DNA	6
Parámetros de curvatura	8
Modelos teóricos de predicción de curvatura del DNA	11
Modelos de curvatura	13
Torsimiento y curvatura	15
Interacción proteína-DNA	16
Motivos estructurales de secuencias específicas de unión	16
OBJETIVOS	22
METODOLOGIA	26
RESULTADOS Y DISCUSION	32
Predicción de estructura de las UAS	33
Predicción de estructura de las regiones identificadas por IHF	39
Motivo de unión al DNA de NifA	45
CONCLUSIONES	58
BIBLIOGRAFIA	61

INTRODUCCION

INTERACCION DNA-PROTEINA EN SISTEMAS RELACIONADOS CON LA FIJACION DE NITROGENO

Generalidades del Sistema de Regulación en la Fijación de Nitrógeno.

La Fijación de Nitrógeno es un proceso que demanda una gran cantidad de energía, por lo que la interacción que se da entre las moléculas tiene que ser específica y estar fuertemente regulada. Los genes involucrados directamente en este proceso en organismos simbiotes son los *nif* (nitrogen fixation), los *fix* (fixation) y los *nod* (nodulación) (Merrick, 1993).

La regulación de la expresión de estos genes se da tanto a nivel transcripcional como en la modificación postraduccional de las proteínas que estos codifican (David, *et al.* 1988).

El sistema de regulación de estos genes mantiene características que lo distinguen de los demás, una de ellas es que la región promotora se encuentra a -24 y -12 pb del sitio de inicio de la transcripción a diferencia de los otros promotores localizados a -35 y -10 pb (Buck, *et al.* 1985). La RNA polimerasa se une al promotor junto con un factor sigma denominado σ^{54} que es diferente a los demás factores sigma conocidos (Buck y Cannon, 1992). Otra de las características interesantes es que mantienen una región reguladora parecida a los "enhancers" en eucariotes, a estas regiones se les llama UAS (Upstream Activator Sequences) y están localizadas entre 100 y 200 pb por arriba del sitio promotor. Estas mantienen una secuencia conservada TGT-N₁₀-ACA la cual es identificada por la proteína activadora NifA que es considerada como el activador central en la regulación de los genes para la Fijación de Nitrógeno (Tuli y Merrick, 1988; Kustu, *et al.*, 1991; Better, *et al.*, 1985; Alvarez, *et al.* 1986; Buck, *et al.* 1987). El DNA al parecer requiere doblarse para que NifA unida a las UAS pueda interactuar con la RNA polimerasa- σ^{54} que se encuentra unida al promotor. Se sabe que en algunos de los genes *nif* la activación de la transcripción se da formando lo que se conoce como complejo abierto, o sea el momento en que interactúan todos los elementos y se lleva a cabo la transcripción. (Morett y Buck, 1989) (Fig. 1).

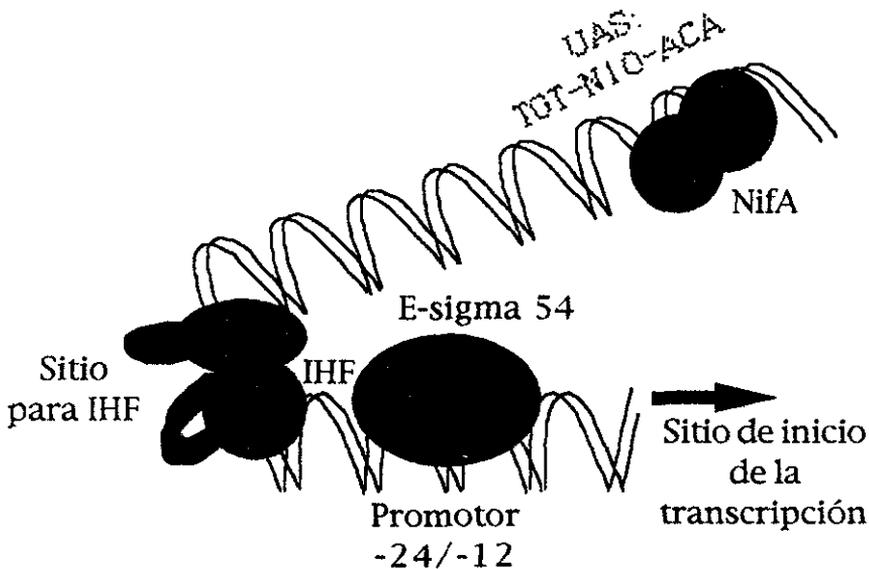


Fig. 1 Modelo de la activación de la transcripción.

La activación de la transcripción en el sistema de fijación de nitrógeno es un buen modelo para el estudio de las interacciones moleculares, desde las características que debe tener el DNA en cuanto a su estructura o secuencia para ser identificado por las proteínas y también por las proteínas que participan en esta regulación principalmente NifA e IHF.

ELEMENTOS PRINCIPALES QUE PARTICIPAN EN LA TRANSCRIPCIÓN DE LOS GENES *nif*.

El activador principal de la transcripción de los genes *nif* es la proteína NifA.

La proteína NifA pertenece a la familia de activadores de promotores σ^{54} llamada EBP's (Enhancer Binding Proteins). Las proteínas de esta familia son muy parecidas, están constituidas por 3 dominios funcionales y estructurales independientes: el dominio amino terminal no está conservado, aunque en algunas proteínas como en NtrC el dominio tiene funciones regulatorias. (Buikema, *et al.*, 1985).

Sin embargo, esta región no es absolutamente indispensable para la actividad de NifA ya que en proteínas de *B. japonicum* y de *R. meliloti* se ha deletado y la activación de los promotores de los genes *nif* no ha disminuido significativamente (Huala y Ausubel, 1984., Fisher, *et al.* 1988).

Esta región está ligada al dominio central mediante un interdominio, el Q-linker (rico en glutamina, arginina, glutamato, serina y prolina), descrito como una región flexible que une al dominio amino terminal con el central (Cannon y Drummond, 1987).

El dominio central esta altamente conservado no solo en NifA sino en todas las EBP's, esto sugiere que la región puede estar relacionada con la interacción de la proteína con el factor σ^{54} . El dominio mantiene 7 regiones altamente conservadas, las cuales han sido nombradas consecutivamente de C1 a C7. La C1 mantiene una gran similitud con proteínas que se unen o hidrolizan ATP (Morett y Segovia, 1989; Huala y Ausubel, 1989). Para las otras regiones no se conocen regiones homólogas que lleven a cabo alguna actividad.

En las Rhizobiaceas, la segunda región de interdominios que une el dominio central con el carboxilo terminal contiene 4 cisteínas que hacen a la proteína sensible a oxígeno. Se ha propuesto que en condiciones anaeróbicas NifA une un cofactor metálico y que cuando aumenta la concentración de oxígeno el cofactor se oxida y por consiguiente NifA se inactiva (Morett, *et al.*, 1991, Fisher, *et al.* 1987).

En algunas de las proteínas de la familia EBP's la región carboxilo terminal contiene alrededor de 70 residuos de aminoácidos, esta región al parecer forma un motivo estructural denominado HVH (hélice-vuelta-hélice) similar a otras proteínas que se unen a DNA. Si el dominio carboxilo de NifA de *K. pneumoniae* se muta, disminuye su actividad. (Morett y Buck., 1988). Si este dominio se expresa independientemente, se une específicamente a las UAS que se encuentran entre 100 y 200 pb del sitio promotor, aunque es incapaz de activar la transcripción. Este evento se ha probado también en experimentos de metilación *in vivo* (Morett, *et al.*, 1988).

La proteína IHF es un elemento que favorece la interacción de la proteína activadora y el complejo RNA polimerasa-factor σ , debido a la doblez que provoca en el DNA.

La proteína IHF es un heterodímero con un peso molecular de 20kDa, esta juega un papel importante en muchos procesos incluyendo recombinación sitio específica, empaquetamiento de fago, transposición, replicación de los plásmidos y transcripción (Goodrich, *et al.* 1990).

Esta proteína tiene homología con las proteínas HU las cuales son parecidas a las histonas de eucariontes, el mecanismo de acción de IHF todavía no se conoce con precisión pero se sabe que al unirse puede provocar un doblamiento en el DNA de más de 140° , por lo que es posible que el papel de IHF en la integración de lambda en *E. coli* además de otros procesos, sea el doblamiento del DNA hacia una conformación funcional. Se ha probado en el trabajo de Goodman y Nash (1989) que se puede sustituir el requerimiento de IHF en una secuencia de DNA si esta contiene una curvatura intrínseca suficiente y posicionada correctamente (Yang y Nash., 1989). Las subunidades de la proteína se expresan a partir de los genes *himA* y *hip/himD*. Mediante experimentos de "footprinting" con DNasa I se puede comprobar que IHF se une al DNA en aproximadamente 40pb. En algunos ensayos realizados mediante la migración anómala en geles de retardamiento, se observa que al pegarse IHF en distintos tramos de una secuencia de DNA esta proteína provoca una cierta curvatura en el DNA lo que ocasiona una migración diferencial en el gel.

Recientemente, experimentos de protección y de utilización de radicales hidroxilo, indican que IHF contacta al DNA en las bases del surco menor obligando a que este se doble alrededor de la proteína.

Estudios recientes han mostrado que IHF se une justo rioarriba del sitio promotor de algunos genes cuya expresión requiere de la RNA polimerasa unida al factor σ^{54} , los cuales incluyen a *nifH* y a *nifU* de *K. pneumoniae* (que son proteínas que participan en la fijación de nitrógeno), en *glnH* de *E. coli* (permeasa de glutamina) en el operon *flbG* de *C. crescentus* (proteínas flagelares), y en *xyICAB* de *P. putida* (enzimas que catabolizan xileno y tolueno), realizando un papel importante en el aumento de la actividad (Claverie-Martin y Magasanik, 1992).

El efecto de IHF en la activación del inicio de la transcripción en los promotores de *nifH* y *nifU* por la proteína NifA (Hoover, *et al.*, 1990; Cannon, *et al.*, 1990) y el promotor *glnHp2* por el regulador de nitrógeno I (NRI-fosfato) (Clavelier-Martin y Magasanik, 1992) sugieren que el doblamiento del DNA por IHF, facilita la interacción entre la proteína activadora y la RNA polimerasa- σ^{54} para unirse al promotor, y de esa manera favorecer el inicio de la transcripción. En contraste con la región reguladora de *glnH*, no existe un sitio de unión para IHF rioarriba del promotor de *glnAp2* por lo que IHF no estimula el inicio de la transcripción (Sasse-Dwight y Gralla, 1990). También la función de IHF depende de la posición de su sitio de unión pudiendo activar o reprimir el inicio de la transcripción, dependiendo de hacia donde se da el doblamiento del DNA (Claverie-Martin y Magasanik, 1992).

En conclusión, la proteína IHF es un factor importante en la transcripción ya que provoca un doblamiento en el DNA que permite la interacción de la proteína activadora unida a las UAS con la RNA polimerasa- σ^{54} unida al promotor.

Es necesario que el DNA mantenga cierta flexibilidad que permita la interacción de los distintos elementos de la transcripción, la cual se favorece al interactuar la proteína IHF. Sin embargo, se cree que existen secuencias con una curvatura intrínseca que permiten la interacción de los distintos elementos de transcripción en ausencia de la proteína IHF.

Para comprender un poco más acerca de la influencia que tiene la estructura del DNA en la regulación y funcionamiento de algunos genes y proteínas, se han desarrollado diversos métodos que permiten modelar la posible o posibles estructuras que adquiere el DNA. Los cuales se mencionan a continuación.

INTRODUCCION SOBRE LAS MOLECULAS DE DNA.

La información hereditaria de todos los organismos vivientes, con excepción de algunos virus, está contenida en moléculas de ácido desoxirribonucleico (DNA). El DNA usualmente consiste de dos cadenas complementarias enrolladas una sobre otra, formando una hélice con el giro hacia la derecha. Cada cadena esta formada por polinucleótidos lineales que consisten de cuatro nucleótidos. Hay dos purinas, adenina (A) y guanina (G) y dos pirimidinas, timina (T) y citosina (C). Las dos cadenas estan unidas por puentes de hidrógeno entre los pares de nucleótidos. Cada nucleótido en la secuencia de DNA contiene un azucar pentosa (desoxiribosa), un grupo fosfato y una base púrica o pirimídica. Las moléculas del azucar y el grupo fosfato están unidas covalentemente por uniones fosfodiester asimétricas 5'-3'. La doble hélice del DNA se encuentra arreglada de forma antiparalela.

Métodos usados en el estudio de la estructura del DNA.

Uno de los métodos más importantes para el análisis de la estructura del DNA, al menos desde el punto de vista histórico, fué la difracción de rayos X. Esta herramienta se utilizó para descubrir la forma básica de la doble hélice del DNA en 1953.

Rosalind Franklin realizó trabajos de difracción de rayos X, con fibras de DNA y obtuvo fotografías con patrones de difracción distintos, a los cuales llamó la forma A y la B. La forma A se observaba cuando las fibras estaban secas y la forma B cuando estaban húmedas.

Struther y Arnott observaron que DNA con secuencias repetidas regulares de A's (Adeninas) o con las bases A (adenina) T (timina) alternadas, pueden ser extremadamente polimórficas ya que cada uno de ellos puede producir hasta 4 diferentes patrones de difracción.

Además de los métodos de difracción de rayos X que analizan la estructura atómica que es una técnica directa con la que se obtiene un modelo tridimensional, existen los experimentos de microscopía electrónica que analizan la estructura atómica y producen imágenes que generalmente son de 2 dimensiones. Ambos métodos proveen de modelos que corresponden a la realidad física. Sin embargo, pocas personas utilizan estos métodos debido a que el equipo es muy costoso y requiere de mucha preparación para examinar los datos, además de que es muy complicado obtener una buena muestra de moléculas cristalizadas para ser analizadas en el microscopio electrónico. Por estas razones muchos científicos actualmente utilizan una gran variedad de métodos indirectos que ayudan a determinar la estructura del DNA (Shakked, 1991).

Los métodos indirectos incluyen los métodos espectroscópicos, como NMR: en donde la muestra de DNA se pone en presencia de un campo magnético fuerte, y todos los átomos de hidrógeno en la molécula de DNA se alinean en relación al campo. Después la muestra se expone a un campo electromagnético de baja energía en un rango de radiofrecuencias, los hidrógenos individuales en el DNA absorben energía en alguna frecuencia particular alineándose sobre el campo principal.

La cantidad de energía requerida para mover el hidrógeno nuevamente al campo principal es muy sensible a la localización en la molécula, también a la unión con otros átomos y a los átomos cercanos en el plano tridimensional.

La Espectroscopía de Raman: mide las frecuencias vibracionales de enlaces individuales en el DNA, permite medir sensiblemente los enlaces químicos y la estructura, estos métodos están limitados en la interpretación estructural. En comparación con la espectroscopía por rayos X o NMR, la principal ventaja de la espectroscopía de Raman es que se requiere muy poco material de la muestra.

El Dicroísmo circular mide la absorción de la luz ultravioleta polarizada en el DNA, de esto aún no hay una interpretación teórica aunque se utilizan moléculas con estructura conocida para comparar y analizar los resultados.

Para observar la macroestructura de las moléculas se han utilizado los métodos químico-enzimáticos, los cuales utilizan enzimas de restricción o químicos específicos para modificar el DNA y posteriormente observar su patrón electroforético. Algunos de estos reconocen la secuencia de las bases, y otros reconocen la estructura de doble hélice o proporcionan datos que permiten aproximar una estructura.

Por último, la migración electroforética anómala (geles de retardamiento), puede ser usada para encontrar información sobre la estructura del DNA. Sin embargo hay que tomar en cuenta que las moléculas de DNA no solo corren de acuerdo a su tamaño sino también de acuerdo a su tipo y su carga eléctrica neta. En muchos experimentos se ha mostrado que el DNA curvado se mueve más lentamente a través de los geles que el DNA lineal del mismo tamaño o largo (Calladine y Drew., 1992).

Actualmente se han desarrollado estrategias que permiten calcular de manera teórica las curvaturas más probables que adquiere cierta secuencia de DNA, para esto se toman en cuenta distintos parámetros de curvatura los cuales se describen a continuación.

Parámetros de curvatura

Los principales parámetros de curvatura son seis, tres de rotación: *twist*, *roll*, *tilt*, y tres de traslación: *rise*, *slide* y *shift* (Figura 2). Estos se identifican claramente ya que tienen ejes de hélices locales entre dos pares de bases (Dzl), o parámetros globales calculados relativamente al eje de la hélice total (Dzg). Las cantidades pueden diferir considerablemente, por ejemplo, en DNA-A el ángulo global es $Dzg=2.9^\circ$ y los locales es $Dzl=3.4^\circ$.

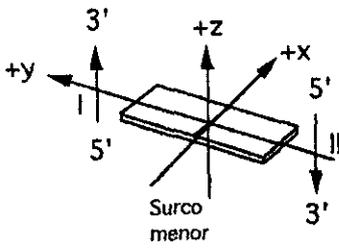
Se ha puesto cuidadosa atención al empleo del término de rotación positiva y esta siempre se refiere a un desplazamiento rotacional hacia la derecha en cualquiera de los ejes.

El *tilt* positivo es cuando el desplazamiento se dirige en dirección al surco mayor, este abre el ángulo entre los pares de bases a través de una de las cadenas del DNA denominada como banda I, visto desde el lado del surco menor. La flecha representa la rotación de los pares de bases 2 relativo a pares de bases 1 sobre el eje x.

El *roll* positivo se observa con el desplazamiento que va de la hélice de la cadena II a la de la cadena I o como la rotación de la base 2 en relación a la base 1 sobre el eje y. Como consecuencia se abre el ángulo entre las pares de bases a través del surco menor, esto se observa a lo largo del eje de las bases de la cadena o banda I a la cadena o banda II o tomando como referencia el eje y.

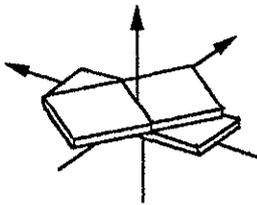
El *twist* positivo es cuando el desplazamiento es en dirección 5' a 3' de la cadena I, tomando como referencia el eje z.

El ángulo de torsión normal, es positivo visto desde abajo de lo largo del eje z de los pares de bases. El *propeller-twist* usual en DNA-A y B es negativo con esta convención, el cual se usa porque es consistente con las definiciones de los ángulos de torsión (EMBO J. 1989) (Figura 2).

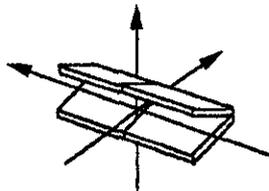


Coordenadas

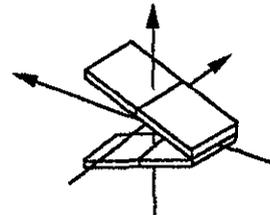
Ángulos de ROTACION



Torsimienta
(Twist)

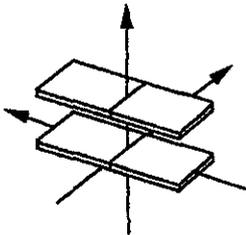


Rodaje
(Roll)

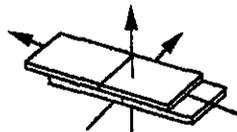


Inclinación
(Tilt)

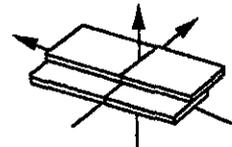
Ángulos de TRASLACION



Elevación
(Rise)



Deslizamiento
(Slide)



Cambio
(Shift)

Fig. 2 Parámetros de curvatura, ángulos de rotación y de traslación

Modelos teóricos de predicción de curvatura del DNA

El planteamiento inicial de los modelos de curvatura de DNA surgieron con la base de que el DNA puede empaquetarse en nucleosomas y por ende curvarse (Zhurkin). Al tomar en cuenta los cálculos de la energía conformacional, se observó que la curvatura en los nucleosomas puede ocurrir por "roll" cada 5 pb o alternativamente por "tilt" y esto también es válido en el caso de que la curvatura sea inducida por proteínas (Travers.,1990; Prentki, *et al.*,1987).

Trifonov y Sussman desarrollaron la idea de efectos cercanos, en donde cada dinucleótido contribuye independientemente en la estructura del DNA y presentan evidencia de repetición periódica de dinucleótidos particulares como AA, que facilitan la curvatura y por lo tanto la formación de nucleosomas en secuencias de DNA eucariótico.

Los modelos antes mencionados enfatizan la deformación "suave" del DNA por una serie de componentes pequeños e independientes de "roll" o "tilt" entre pares de bases adyacentes en un plano, estos modelos se categorizan como los modelos de cuña (*wedge*) (Ulanovsky y Trifonov., 1987).

Los estudios de construcción del modelo de Arnott, Wells y colaboradores se enfocan en un cambio abrupto en el eje de dirección de la hélice del DNA en el empalme (*junction*) de hélices de DNA-A y B. Debido a que los pares de bases en las dos formas tienen diferente orientación angular relativa al eje de la hélice, el empaquetamiento paralelo de los pares de bases en la cercanía causa que los dos ejes de las hélices no sean paralelos. El locus en el cual el eje se encuentra se llama "curvatura de empalme" y en un modelo de curvatura generalmente resulta por haber trectos repetidos de adeninas (Haran y Crothers., 1990).

Es necesario enfatizar que los modelos de "cuña y los de empalme" no son incompatibles necesariamente.

Desde el punto de vista de Crothers los modelos de cuña son más generales, partiendo de que cualquier hélice curvada puede ser descrita por la combinación de las rotaciones de hélices con una serie apropiada de ángulos de cuña. Sin embargo, el modelo de unión es más restrictivo, proporciona mayor información y por lo tanto es el más utilizado actualmente. Específicamente, el modelo de unión es apropiado cuando existen evidencias de la formación de hélices de distinta estructura obtenidas de la misma secuencia de bases y cuando existen efectos cooperativos en la nucleación de la estructura especial dentro de un segmento de B-DNA, ambas pruebas se conocen en curvaturas de DNA que ocurren en trectos de dA-dT.

La inducción cooperativa de una estructura especial debilita la aplicabilidad del modelo de cuña porque este modelo se basa en ángulos derivados de B-DNA y no pueden ser utilizados para predecir las propiedades de estructuras alteradas y por lo tanto de la curvatura de tractos largos (Crothers, *et al.*, 1990).

Los modelos de la estructura de poli dA y poli dT basados en la difracción de fibras, comparten características comunes de "roll negativo" con pares de bases adyacentes y todas producen una inclinación relativa con respecto al eje de la hélice. En el modelo de Koo y Crothers., 1988; utilizado para predecir la curvatura del DNA, coincide con datos experimentales tanto en la dirección como en la magnitud aproximada de la curvatura. Principalmente cuando los tractos de dA y dT se encuentran en el centro de la molécula sobre el surco menor o al final de la cadena 3'. De cualquier manera existe incertidumbre debido a que las estructuras cristalográficas de oligonucleótidos que contienen tractos de A's no presentan una inclinación marcada de los pares de bases inferida por los estudios de homopolímeros. Todos los colaboradores en esta área están de acuerdo con que la curvatura del DNA requiere una diferencia en los ángulos de cada base tanto de *roll* como de *tilt* (o el equivalente de la inclinación de los pares de bases o "*tip*") entre los tractos de A y los segmentos interrumpidos de B-DNA.

Se han propuesto dos modelos de secuencia dependientes de curvatura en DNA libre los cuales son: Los modelos de "*Nearest neighbor*" el cuál propone que, a larga escala, la curvatura medible puede surgir por la acumulación de pequeñas deformaciones locales en la hélice por los ángulos de *twist*, *roll*, *tilt* y *slide* entre pares de bases individuales. Los de empalme, por otro lado, proponen que la curvatura ocurre en la interfase entre dos diferentes variantes estructurales de B-DNA de doble hélice (Crothers, *et al.*, 1990; Delcourt y Blake., 1991).

Modelos de curvatura

Modelos de "Curvatura sin tractos de A"

a) Modelo de posicionamiento en un nucleosoma. Los parámetros de *roll-tilt-twist* en este modelo provienen exclusivamente de observaciones experimentales, basándose en la preferencia de localización de secuencias de trímeros (e.j. GGC o AAA) en círculos pequeños de DNA (llamados nucleosomas), sin tomar en cuenta técnicas que miden la curvatura del DNA en solución *per se*.

Con la secuencia de los tripletes y la localización en cierta cara del nucleosoma, se ha calculado la preferencia en cuanto a la posición de cada triplete, así por ejemplo la secuencia GGC tiene 45 % de preferencia por localizarse en la curva de la doble hélice donde el surco mayor está hacia fuera y el menor hacia dentro y comprimido por la curvatura (tendiendo hacia un *roll* positivo). La secuencia AAA tiene un 36% de preferencia por la orientación opuesta, donde el surco mayor se encara hacia afuera y el menor hacia adentro y comprimido (tendiendo hacia un *roll* negativo).

Después de obtener estas varianzas fraccionales los ángulos de *roll* se asignan de la siguiente manera. Los análisis de estructura de cristales por rayos X indican de manera uniforme que los pasos de adeninas continuas (AA) no tienen curvatura, por lo que se les asigna un *roll* de cero al triplete de AAA y un *roll* máximo de 10° al triplete de GGC y en escala todos los demás tripletes en forma lineal. Los ángulos de *twist* son de 34.4° de 10.5 pares de bases por vuelta y el *tilt* tiene como valor cero y como estos valores corresponden a trímeros, los valores solo están aplicados a las primeras dos bases de estos (Tabla 1).

b) Modelo de Calladine. Este modelo se formuló para explicar la migración anómala del DNA en geles. Se tomaron 25 secuencias como prueba basadas en las características estructurales observadas en el análisis de un cristal por rayos X, en particular la pérdida de *roll* o *tilt* en dodecameros como CGCAAAAAGCG que contienen tractos de A. De acuerdo con lo observado a los tractos de AA se les asignan ángulos de cero para el *roll* y el *tilt* y de 35° para el *twist*. Para las secuencias generales los valores de *roll* es de $+3.3^\circ$, el de *tilt* de cero y el *twist* uniforme de 34° . Cuando las secuencias tienen tractos de TA, se les considera anómalos y se les asigna un *roll* doble de $+6.6^\circ$ como en diversos análisis de estructura de cristales.

I Modelos de trimeros

a) Modelo de posicionamiento de nucleosoma: Adaptado de Satchwell, Drew y Travers
Tilt=0°; Twist= 34.3° (10.5 pares de bases por vuelta)

Roll	G				A				
2a base	A	G	C	T	A	G	C	T	
3a base									
1a. base.	A	3.3°	5.4°	7.5°	5.8°	0.0°	5.2°	3.7°	0.7°
	G	3.8°	6.0°	10.0°	5.4°	3.0°	5.4°	5.4°	5.3°
	C	8.3°	4.7°	7.5°	5.4°	3.3°	4.2°	6.5°	6.7°
	T	5.4°	5.4°	6.0°	5.2°	2.0°	2.2°	3.7°	2.8°

II Modelos diméricos.

Tipo de base	Purina	Purina	Purina	Pirimidina	Pirimidina	Purina	Pirimidina	Purina	Pirimidina	
Base	AA	GA	AG	GG	AT	AC	CC	TA	CA	CG
b) Modelo de Calladine, Drew y McCall.										
Roll	0°	3.3°	3.3°	3.3°	3.3°	3.3°	3.3°	6.6°	3.3°	3.3°
Tilt	0°	0°	0°	0°	0°	0°	0°	0°	0°	0°
Twist	3.5°	3.4°	3.4°	3.4°	3.4°	3.4°	3.4°	3.4°	3.4°	3.4°

c) Modelo de Bolshoi, McNamara, Harrington y Trifonov.

Roll:	-6.5°	-2.7°	8.4°	1.2°	2.6°	-0.9°	-5°	0.9°	1.6°	6.7°
Tilt:	3.2°	-4.6°	-0.3°	-1.8°	0°	-0.7°	0°	0°	3.1°	0°
Twist:	35.62°	36.9°	27.7°	33.67°	31.5°	34.4°	40°	3.6°	34.5°	29.8°

d) Modelo de Cacchione, De Santis, Foti, Palleschi y Savino.

Roll:	-5.4°	2°	+1°	1.3°	-7.3°	-2.4°	-3.7°	8°	6.7°	4.6°
Tilt:	-0.5°	-1.7°	-1.6°	-0.6°	0°	-2.7°	0°	0°	0.4°	0°
Twist:	35.9°	35.8°	35.6°	33°	35°	34.6°	33.3°	34.6°	34.5°	33.7°

e) Modelo Reverso de Calladine, Drew y McCall.

Roll:	3.3°	0°	0°	0°	0°	0°	0°	-3.3°	0°	0°
Tilt:	0°	0°	0°	0°	0°	0°	0°	0°	0°	0°
Twist:	3.5°	3.4°	3.4°	3.4°	3.4°	3.4°	3.4°	3.4°	3.4°	3.4°

f) Modelo de Koo y Crothers.

	***	***								
Roll:	0°	5.8°	0°	0°	0°	0°	0°	11.6°	5.8°	0°
Tilt:	0°	-9.6°	12°	0°	0°	12°	0°	0°	-9.6°	0°
Twist	34.3°	34.3°	34.3°	34.3°	34.3°	34.3°	34.3°	34.3°	34.3°	34.3°

En los modelos b-f Roll y Twist son de las mismas pares de base complementarias (ej. CT y AG) Tilt cambia de signo para los pares.

*** solo cuando se asocian más de 3 A's sucesivas Roll y Tilt tienen valores de cero.

+ No está explícitamente especificado en el artículo. Corresponde a 10.5 pares de bases por vuelta.

Tabla 1. Roll, Tilt y Twist, parámetros usados en varios modelos de curvatura de DNA.

Modelos de "Curvatura con trectos de A"

c) Modelo de Bolshoy. Los valores locales de *twist*, *roll* y *tilt* para este modelo se han seleccionado para explicar la movilidad anómala electroforética en geles y la circularización de distintas secuencias de oligonucleótidos.

Este modelo tiene el valor más grande de *tilt* y gran variedad de valores de *twist* (28° 40°) en comparación a los demás.

d) Modelo de Cacchione. Este modelo utiliza valores locales de *twist*, *roll* y *tilt*, que fueron obtenidos mediante cálculos de la energía conformacional para grupos de dos pares de bases. Los resultados son compatibles con los obtenidos de análisis de cristales de oligonucleótidos. Incluso son exitosos tomando en cuenta los experimentos de migración electroforética de 62 diferentes secuencias.

e) Modelo de Calladine-inverso. Una característica interesante de la serie de expresiones de Fourier utilizadas por Calladine, Drew y McCall es que los resultados de curvatura son idénticos matemáticamente si todos los valores de *roll* y *twist* se suman o se restan a cualquier constante arbitraria. El modelo de Calladine "sin trectos de A" se transforma en el de Calladine-inverso "con trectos de A" restando simplemente los tres ángulos normales de *roll* a una constante de 3.3° .

Modelo "empalme"

f) Modelo de Koo y Crothers. Este modelo se propuso específicamente para explicar la migración electroforética anómala del DNA en los geles de electroforésis. Los valores apreciables de *roll* y *tilt* se asignan a secuencias que tienen más de 3 adeninas con valores de cero en el *roll* y el *tilt*, los pasos de TA en medio de un tracto de A se considera disruptivo, pero no los AT. Un ángulo de *twist* uniforme se asume y corresponde a 10.5 pares de bases por vuelta (Goodsell y Dickerson., 1994).

Torsimiento y curvatura.

Las acciones fundamentales del DNA en la biología comprenden tanto el torsimiento como la curvatura de la doble hélice del DNA.

En *E. coli* se realizaron estudios sobre secuencias que son reconocidas por activadores transcripcionales, los cuales provocan una cierta curvatura en el DNA, estas secuencias se han deletado y sustituido por secuencias sintéticas con una curvatura definida, las cuales son capaces de restaurar la transcripción. (Bracco *et al.*, 1989).

De esta manera podemos decir que las secuencias que son reconocidas por algunas proteínas requieren de una estructura definida para que se lleve a cabo la transcripción.

INTERACCION PROTEINA-DNA.

La unión de una proteína a secuencias específicas de DNA es dependiente de dos tipos principales de interacciones. Una es la selectividad por secuencia entre los polipéptidos y los ángulos de los pares de bases localizados principalmente en el surco mayor del DNA (B-DNA), mediante puentes de hidrógeno o fuerzas de van der Waals. Y la segunda es la dependencia a la curvatura o deformación del DNA, que limita las conformaciones energéticas favorables en un sitio particular de unión (Travers, 1989;1992; Phillips y Moras., 1993).

Motivos estructurales de secuencias específicas de unión.

El reconocimiento de una secuencia de DNA específica por una cadena polipeptídica se da por un diverso pero limitado número de estructuras.

Mediante estudios estructurales, bioquímicos y de genética molecular sobre interacciones proteína-DNA, se ha identificado un pequeño número de elementos estructurales de proteínas que se unen a secuencias específicas de DNA. Algunos de los que ocurren más frecuentemente son los motivos de hélice-vuelta-hélice, que son característicos de muchas proteínas reguladoras en procariontes. El motivo de dedos de zinc, encontrado en proteínas de muchos organismos. La estructura alfa-hélice encontrada en la enzima de restricción EcoRI y una hoja beta antiparalela en el represor Met de *E. coli* (Fig. 3). En todos estos ejemplos el motivo de reconocimiento se fija en el surco mayor del DNA y muchas veces sobresale de la superficie de la proteína. Solo en el caso de la proteína IHF de *E. coli* que al parecer tiene una estructura de betas antiparalelas puede conferir una cierta selectividad que le permite unirse al surco menor del DNA (Sauer. 1995).

Además de los motivos en donde la información estructural es accesible se han identificado otras regiones de unión a DNA por similitud de secuencia, que incluyen tanto motivos conocidos como "hélice-lazo-hélice" y motivos de dimerización como los de "cierre de leucina" (Fisher, *et al*, 1992).

A continuación se describen los diferentes motivos que favorecen la unión de las proteínas al DNA.

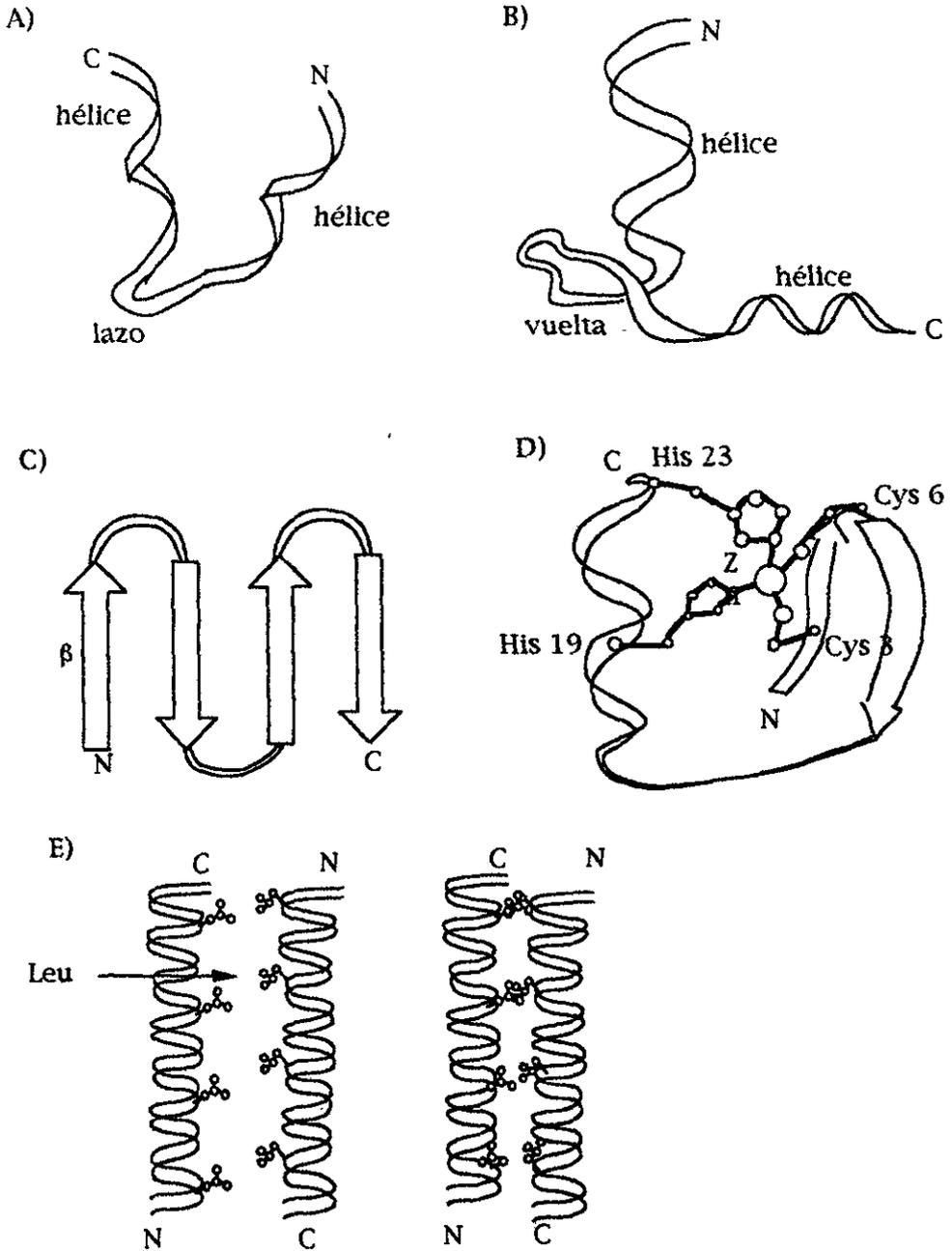


Fig. 3 Motivos de unión a DNA. A) Hélice-lazo-Hélice, B) Hélice-vuelta-Hélice, C) β plegadas, D) Dedos de Zinc y E) Cierres de Leucina.

-Motivo de Dedos de Zinc

Estos comprenden al menos tres mini dominios estructuralmente distintos, que comparten la propiedad de estabilizarse por un tetraedro coordinado por un ion de zinc. La primera de estas estructuras se reconoce como el arquetipo de dedo de zinc presente en el factor de transcripción de TFIIIA, y que se une dentro de una región codificadora para los 5S rRNA de 50 pb. La proteína tiene la capacidad de unirse tanto al DNA como al RNA. Estudios proteolíticos muestran que la proteína de 40 kDa consiste de dos regiones funcionales separables. La primera de 10kDa que se encuentra en el carboxilo terminal no esta involucrada en la unión al DNA pero sí es esencial para la transcripción eficiente. La otra región de 30kDa se une tanto al 5S RNA como a toda la región interna de control.

La estructura de dedos se estabiliza por dos cosas, la coordinación tetraédrica de zinc y por la bolsa hidrofóbica en el interior de la estructura que involucra aminoácidos altamente conservados.

El linker que separa los dedos adyacentes se encuentra conservado y se localiza entre la última histidina del dedo proximal y la primera cisteína del dedo distal, comprende normalmente de 7 a 8 aminoácidos. Sin embargo existe suficiente variación para sugerir que el camino seguido por los distintos linkers no ha sido el mismo.

Los motivos de unión a DNA que contienen Zn como elemento estructural son de tres clases, en la primera clase se encuentra el factor TFIIIA en donde el zinc es coordinado por dos cisteínas y dos histidinas. La segunda clase incluye a los receptores esteroides, en donde el zinc se une por 4 residuos de cisteína pero la unión polipeptídica entre las cisteínas coordinadoras es distinto al de TFIIIA. La tercera clase se ha encontrado en muchos factores de transcripción de levadura, esta clase de motivos contiene dos átomos de zinc con carga coordinados tetrahédricamente por 6 residuos de cisteína, dos de los cuales ligan ambos metales. En algunos casos en las proteínas que unen zinc, el ión metálico no se requiere para la interacción con el DNA aunque puede afectar la estabilidad de la proteína y aumentar la cooperatividad para la unión con una cadena sencilla de DNA. Para estos casos se ha inferido que la estructura de dedos se involucra en las interacciones proteína-proteína. Los dedos de zinc en tandem actúan como módulos estructuralmente independientes con la intervención de conexiones altamente flexibles (Harrison, 1991).

-Motivo de "cierres de leucinas" (bCC o bZIP).

Esta estructura se identifica originalmente como un motivo de dimerización en un segmento de α -hélice conocido como "cierre de leucinas".

Este segmento consta de aproximadamente 30 residuos formando la mitad carboxilo terminal. La región amino terminal está compuesta por aminoácidos básicos cargados positivamente.

El cierre de leucina es una héptada repetida en donde los residuos 1-4 son predominantemente no polares. El residuo de leucina se encuentra casi exclusivamente en la cuarta posición de las cuatro o cinco héptadas sucesivas en los genes *GCN4*, *Fos*, *Jun* y en otros miembros de la familia bCC. Algunos experimentos de protección química y de modificación indican que el contacto de estas proteínas con el DNA se hace exclusivamente en el surco mayor cubriendo aproximadamente 10pb. La ausencia de una estructura rígida en la región básica del dominio de bCC es importante para que la proteína se pueda meter alrededor de la doble hélice (Travers., 1992).

-Proteínas con motivos de β -dobladas.

Existen dos clases de proteínas procarióticas que interactúan con el DNA a través de los elementos de reconocimiento β -dobladas. En la primera clase incluyen el represor MetJ de *E. coli* y los represores Arc y Mnt de *Salmonella*. La segunda clase incluye la proteína HU que forma estructuras condensadas de nucleoproteínas en muchos procariotes, la proteína IHF de *E. coli* y el factor de transcripción TF1 de *B. subtilis*, que se unen a sitios específicos.

Algunas de estas proteínas se dimerizan para interactuar preferencialmente con el surco mayor del DNA, aunque otras como las HU contactan con el surco menor. La estructura de HU se determinó cristalográficamente y al construir un modelo de estructura no es posible distinguir si la interacción es con el surco mayor o menor. Sin embargo, estudios de protección química muestran que las interacciones son con el surco menor al igual que algunas proteínas homólogas a IHF (Van der Vliet y Verrijzer., 1993).

- Motivo de Hélice-lazo-Hélice.

El grupo de proteínas que contienen este motivo se ha caracterizado únicamente por secuencia. No existe información directa estructural. Consta de una región básica de aproximadamente 15 residuos localizada en el N-terminal con una secuencia conservada característica de las hélices anfipáticas típicas. Una región de tamaño variable entre 9 y 20 residuos separa a la hélice de otra de tamaño similar, generalmente se dimeriza y el heterodímero es una molécula que se une activamente al DNA (Harrison, 1991; Wolffe., 1994).

-Otras estructuras de unión a DNA

Las enzimas de restricción cortan la doble cadena de DNA en secuencias específicas. Enzimas como la EcoRI es un dímero de 31 kDa y la región de interacción consiste en un par de α -hélices paralelas que se insertan dentro del surco mayor. Una característica importante en la interacción es la distorsión de la estructura del DNA en el sitio de unión.

-Motivo de hélice-vuelta-hélice (HVH)

Muchas proteínas de unión a DNA que contienen especialmente el motivo de unión de HVH actúan como homodímeros con dos secuencias idénticas de dominios de reconocimiento. Este motivo se identificó como un elemento de secuencia conservada en los represores codificados por los fagos Lambda de *E. coli*. y p22 de *S. typhimurium*. Subsecuentemente este elemento se encontró en una gran cantidad de proteínas de unión a DNA, tanto en procariotes como en eucariotes y todos conservan un alto grado de conservación en estructura (Baumeister, *et al.*, 1992).

Este dominio de unión a DNA consiste de dos α -hélices cortas, entre una y otra α -hélice se encuentra en general un residuo de Glicina. Este aminoácido junto con sus vecinos actúan como una bisagra a lo largo de la cadena polipeptídica para doblarse entre las dos hélices formando casi siempre un ángulo de 120° , de manera que pueda haber contactos hidrofóbicos entre las dos hélices. Estos contactos preservan la orientación relativa y como resultado se forma un dominio estructural terciario compacto. Aunque la estructura de este motivo está altamente conservada, la orientación no siempre es la misma. Esta característica sugiere que el motivo requiere de una cierta rigidez para el reconocimiento de una secuencia específica.

Al comparar la estructura tridimensional del motivo de HVH en las proteínas Cro, el represor de lambda y en la proteína CAP de *E. coli* revelan que existen constreñimientos estéricos en varias posiciones. Entonces es posible definir algunas condiciones para que la secuencia de aminoácidos en el motivo no se llene de distorsiones estructurales (Brennan y Matthews., 1989). El residuo 9 normalmente es una glicina. La estructura en la vuelta es tal que este residuo favorece a la conformación a mano izquierda de la α -hélice. Esto es común para la glicina, pero ocurre raramente con otros aminoácidos ya que hay una interferencia estérica del átomo de carbono β con la cadena principal. Los residuos 4, 8, 10 y 15 están completamente enterrados en la proteína, son hidrofóbicos o polares débiles. En particular los residuos 4 y 15 no deben de estar cargados. Los residuos de la hélice del 3 al 7 y del 15 al 20 no son prolinas. La prolina se puede acomodar en el amino terminal de la hélice pero no dentro de ella, ya que rompe con esta estructura. El residuo 5 no debe tener una cadena lateral muy enramada. Esta cadena lateral está empalmada entre las dos hélices, y si estuviera enramada podría cambiar el alineamiento relativo de las hélices.

La HVH se ha pensado que puede ser un intermediario temprano en el plegamiento de las proteínas, quizá sea otra de las razones por la fuerte selección de los residuos de glicina en su posición.

A pesar de que la naturaleza general de las interacciones entre las HVH's y el surco mayor del DNA es muy parecida, en algunos casos como en los represores, la geometría con sus complejos a veces es firme y sin distorsión y otras contiene cierta cantidad de curvaturas que la distorsionan.

Las proteínas Cro y el represor del fago lambda están relacionadas evolutivamente, sin embargo no son tan parecidas en su estructura tridimensional, la única similitud estructural entre ellos es el motivo de 20 residuos de la hélice vuelta hélice, la situación es inusual ya que en este caso la secuencias de aminoácidos esta más conservada que la estructura tridimensional (Branden, 1991).

La hélice-vuelta-hélice no funciona como un motivo que siempre se alinea de la misma manera con el DNA, ya que el reconocimiento está dado por la combinación de factores, como uniones específicas de hidrógeno, contactos de van de Waals entre las cadenas laterales de la proteína y las partes de los pares de bases expuestos dentro de los surcos del DNA, interacciones que juegan un papel principal. La localización particular de las moléculas de agua que forman un puente entre la proteína y el DNA participan en el reconocimiento de secuencia específico. El DNA también puede aumentar el contacto dentro de los surcos al igual que los enlaces de fosfato.

Es así como la hélice-vuelta-hélice aparece formando un esqueleto rígido fundamental que soporta el reconocimiento de la superficie de la proteína, aunque no define como la superficie de reconocimiento se alinea con el DNA (Dodd y Egan., 1990).

OBJETIVOS

Los distintos elementos que participan en la activación de la transcripción de los genes *nif* nos proveen de un excelente modelo para cumplir el objetivo principal de este trabajo que es el de estudiar las interacciones moleculares entre DNA y proteína.

Para el estudio de la influencia de estructuras locales del DNA en la regulación, se utilizaron las regiones UAS que son identificadas por la proteína NifA, ya que si los 10 nucleótidos intermedios carecen de información adicional en cuanto a su secuencia, esta región tendría que contar con una estructura particular que le diera especificidad al sitio para que la proteína activadora pudiera identificar a las UAS verdaderas de las que no lo son. Para esto se plantearon los siguientes objetivos.

- Identificar y realizar un banco de datos con todas aquellas regiones UAS conocidas.

- Determinar si existe una ampliación en la secuencia consenso que le permita a la proteína activadora NifA identificar las UAS.

- Utilizar distintos modelos de curvatura de DNA que ayuden a predecir la estructura más probable de todas las UAS conocidas y si estas tienen una estructura característica que les permite ser identificadas por la proteína.

También se propuso predecir estructuras de segmentos de DNA mayores, para poder observar una mejor curvatura. Tomando en cuenta la participación de la región identificada por la proteína IHF en favorecer la interacción entre la proteína activadora y el complejo $E\sigma^{54}$, se eligió esta región como objeto de estudio, planteándose los siguientes objetivos.

- Obtener un banco de datos de todas las regiones del DNA rioarriba del sitio promotor de los genes *nif* hasta las UAS (aproximadamente 200pb).

- Identificar mediante un programa basado en una matriz de sitios para IHF todos los posibles sitios para IHF de las regiones regulatorias de los genes *nif* y corroborarlos con trabajos experimentales en los casos posibles.

- Predecir mediante modelos de curvatura la estructura de estas regiones, comparando las que tienen sitios para IHF de las que no y determinar en caso de no tener sitios para IHF si la curvatura de estas regiones del DNA es suficiente y característica para que se lleve a cabo la interacción entre la proteína activadora y $E\sigma^{54}$.

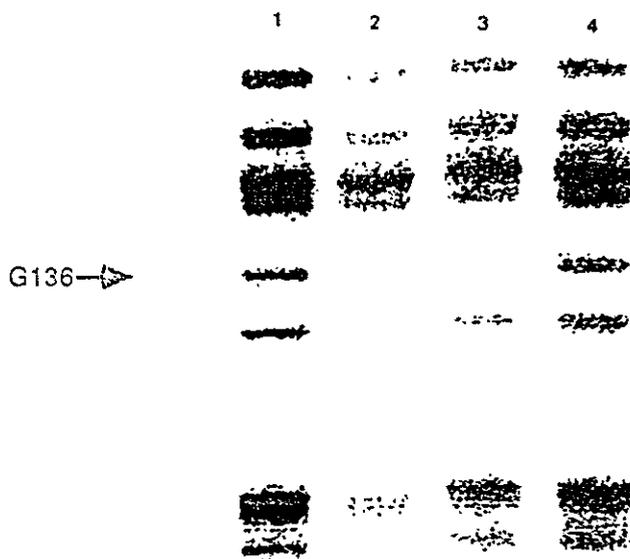
Por otro lado, es también interesante conocer más acerca del papel que juegan las proteínas en la interacción con el DNA. Utilizando el mismo sistema de regulación, se sabe que la proteína NifA interacciona con secuencias específicas mediante un motivo de unión al DNA (HVH) localizado en la región carboxilo terminal. Sin embargo, se observó que la unión de NifA de *R. meliloti* a las UAS era muy débil, aunque mantenía su actividad transcripcional comparable con otras NifA (*K. pneumoniae* y *B. japonicum*) (Tabla 2). Para conocer y determinar los principales elementos de reconocimiento en estos motivos se plantearon los siguientes objetivos.

-Realizar una predicción de estructura secundaria para determinar el sitio exacto del motivo de HVH, corroborado con predicciones anteriores.

-Efectuar un alineamiento de todas las NifA´s secuenciadas y observar si la NifA de *R. meliloti* tiene elementos de secuencia responsables de la débil interacción.

-Diseñar una mutagénesis mediante oligos dirigida al sitio para sustituir aquellos aminoácidos no conservados por aquellos que están conservados en los distintos motivos de HVH.

-Observar si mediante la mutagénesis, se restituye la actividad de unión a DNA y si la función de activación aumenta con respecto a la proteína silvestre.

A) FOOTPRINTING *In vivo* en *nifH* *K. pneumoniae*

CONDICION	ACTIVIDAD β -GALACTOSIDASA (U.M.)
nif H::lac Z	548
nif H::lacZ; NifA <i>K. pneumoniae</i>	15 500
nif H::lacZ; NifA <i>B. japonicum</i>	5 648
nif H::lacZ; NifA <i>R. meliloti</i>	20 236

Tabla 2 A) Footprinting *in vivo*. Se observa marcado con una flecha la protección de la banda correspondiente a la G-136 (UAS), del promotor *nifH* de *K. pneumoniae*, el cual esta fusionado a *lacZ*, lo que permite determinar la actividad de activación mediante la expresión de β -galactosidasa, Carril 1 NifA de *R. meliloti*, 2 NifA de *B. japonicum*, 3 NifA de *K. pneumoniae*, 4 sin NifA B) Expresión de β -galactosidasa.

METODOLOGIA

Búsqueda y alineamientos.

Para realizar la búsqueda de las distintas secuencias necesarias, se utilizaron algunos programas del paquete GCG versión 8.1 y los bancos de datos de GenEMBL, GenBank y SwissProt.

Se realizaron bancos de datos tanto de las UAS conocidas como de las regiones rioarriba del sitio promotor hasta las UAS de los genes *nif*.

Los alineamientos se realizaron con el programa de Pileup.

Predicción de estructura de DNA.

Para predecir la estructura del DNA y su posible efecto en la transcripción de algunos genes es necesario recurrir a modelos matemáticos que nos permitan calcular los ángulos de curvatura que provee cada uno de los nucleótidos.

Existen distintos programas donde se estudia la curvatura, uno de los más utilizados es el "DNA star" donde se calculan los ángulos de doblez entre pares de nucleótidos. Se han obtenido tablas de valores en donde se toman en cuenta distintos criterios, como datos de estructura por rayos X, geles de distorsión, etc. y todos coinciden con que los ángulos más importantes que alteran la curvatura del DNA son el "tilt", el "roll" y el "twist".

El programa base utilizado fue el de "DNA star", sin embargo como resultó ser un programa poco amigable gráficamente se incluyeron las tablas de valores en el paquete de "Insight", el cual permite manipular fácilmente las moléculas. De esta misma forma se pudieron incluir otras tablas de valores para poder comparar los distintos métodos de reconstrucción del DNA.

Identificación de sitios de pegado de IHF.

De algunas de las regiones rio arriba de promotores σ^{54} no se han identificado sitios para IHF por lo que se utilizó el programa "MacTargsearch", que identifica tanto sitios promotores como sitios probables de pegado de IHF. En este programa se utiliza la matriz de Goodrich que contiene la frecuencia de aparición de las bases en cierta posición.

Estas frecuencias de bases surge de la compilación y alineamientos de todos los sitios de unión conocidos de la proteína de interés. Los métodos utilizados con esta lógica han sido funcionales hasta el momento. El potencial de este programa se ha demostrado en el análisis con sitios de unión para la proteína IHF de *E. coli* (Goodrich, *et al.*, 1990).

Cepas bacterianas y plásmidos.

Se utilizó la cepa JM101 de *E. coli* y para algunos casos la GA81 (dam-, dcm-) también de *E. coli*. Los plásmidos que se utilizaron fueron el pRT22, que contiene el promotor *nifH* de *K. pneumoniae* con resistencia a Cm-20 (Cloranfenicol-20mg/ml) el plásmido pucNifA que contiene la secuencia completa que codifica para la proteína NifA silvestre de *R. meliloti* con resistencia a Amp-200 (Ampicilina-200mg/ml) el plásmido pRJ7511 que contiene la secuencia que codifica para la proteína NifA silvestre de *B. japonicum* con resistencia a Tc-15 (Tetraciclina-15mg/ml) y los plásmidos pucNifA con la sustitución de los fragmentos mutados, (NAG1, NAL1 y NAGL2).

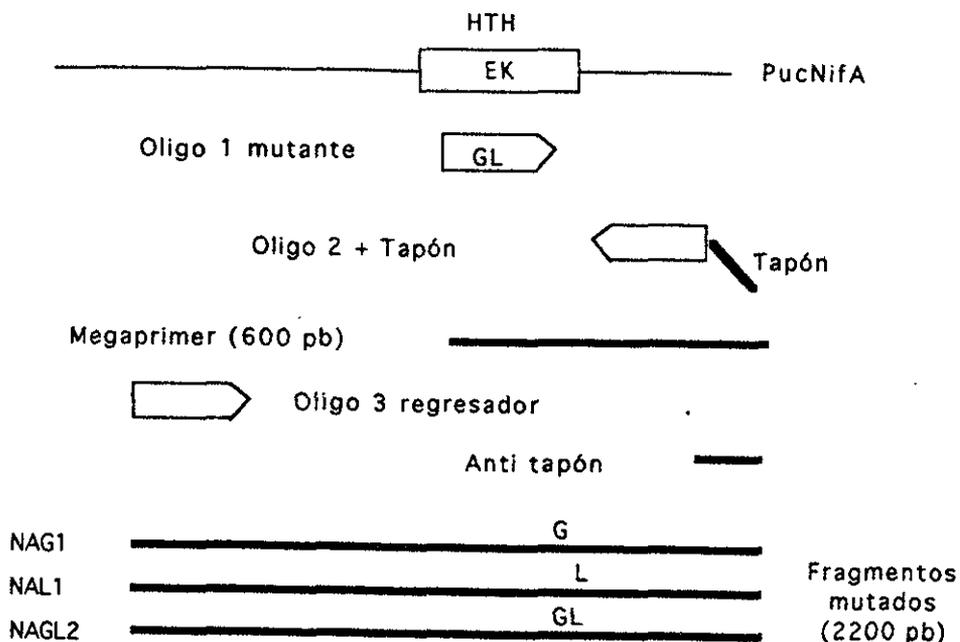
Medios de cultivo.

Todas las cepas de *E. coli* se crecieron normalmente en medio Luria a 30°C. Para los experimentos de footprinting *in vivo* las células se cultivaron en un medio NFDM modificado (Valderrama.,1996).

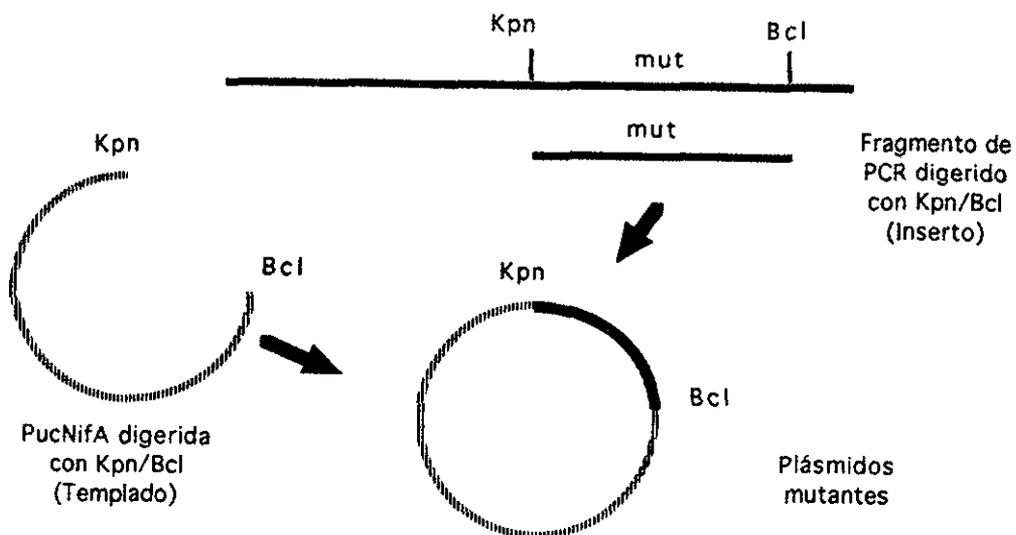
Mutagénesis oligo dirigida.

Para determinar si las mutaciones del motivo HVH de la NifA de *R. meliloti* provocan un mejor pegado al DNA y por ende una mejor activación de los genes que regula, se realizó la mutagénesis dirigida al sitio mediante la técnica de PCR cambiando el residuo de E por G con la sustitución en el codon de GAA por GGA y el de K por L de AAA por CTA ambos localizados en la HVH de la NifA de *R. meliloti*, para obtener tres mutantes, dos sencillas con cada uno de los cambios y una doble con los 2 cambios. El fragmento silvestre que se encontraba en el plásmido pucNifA fue sustituido por cada uno de los 3 fragmentos de PCR mutados, obteniendo los plásmidos NAG1, NAL1 y NAGL2 (Figura 4).

Fig.4 a) Obtención de las mutantes por PCR



b) Sustitución de los fragmentos de PCR



Ensayos de β -Galactosidasa.

Para determinar la actividad transcripcional tanto de la cepa silvestre como de las mutantes se utilizó un plásmido reportero que contiene una fusión transcripcional del gene *nifH-lacZ* (pRT22).

Se utilizaron sistemas heterólogos, es decir cepas de *E. coli* transformada con los dos plásmidos, el reportero y cada una de las mutantes o el plásmido silvestre. De cada una de las cepas transformadas se realizaron ensayos de β -galactosidasa de la manera siguiente: Se toman un cierto volumen de cultivo procurando que todos se encuentren a una densidad óptica entre 0.28-0.7 tomada a 600nm, se centrifugan en frío. las pastillas se resuspenden en 1ml. de BufferZ, a lo que se le agregan 150 μ l de cloroformo y 70 μ l de SDS al 0.2%, se agitan en el vortex. La actividad se determina como el incremento en absorbancia a 420nm después de agregar 30 μ l de ONPG (16mg/ml) y la reacción se detiene con 0.5ml de Na₂CO₃ (1M) después de que se ha desarrollado suficiente color amarillo. Una unidad de actividad específica corresponde a la cantidad de enzima que se requiere para producir 1mM de ONP por minuto por microgramo de proteína en el cultivo. Las unidades Miller se determinaron según (Miller., 1972; Ausubel, *et al.*,1989).

DMS-Footprinting.

Se cultivaron las bacterias en el medio NFDM modificado con antibióticos, estos se incuban a 30°C toda la noche en anaerobiosis. Al alcanzar una densidad optica de alrededor de 0.6, se les agrega DMS a una concentración final de 0.1%. Después de 1 minuto el cultivo se centrifuga a 10000 rpm durante 5 minutos a 4°C, la pastilla se lava dos veces con 250 ml de una solución salina de fosfato (150 mM NaCl, 40 mM K₂HPO₄, 22 mM KH₂PO₄, pH 7.2).

El DNA del plásmido se aísla por lisis alcalina, purificado por extracción fenol y precipitado con etanol. El RNA se remueve con la enzima RNasa seguida por extracción de fenol-cloroformo. El DNA plasmídico se precipita con etanol y se resuspende en 50 μ l de H₂O. Para cortar el DNA en las posiciones metiladas la preparación del plásmido se incubaba con 5 μ l de piperidina a 90°C durante 3 minutos, después el DNA se precipita y se lava extensivamente con etanol al 70%, y finalmente se resuspende en 100 μ l de H₂O.

Para la reacción de extensión del primero se necesita 0.1 pmol del oligo previamente marcado en el extremo 5' con (γ - ^{32}P)ATP, se mezcla con el DNA aproximadamente 1 μg y con 3 μl de Buffer TM 10X, se calienta 3 minutos a 95°C enfriandolo inmediatamente despues. Para la unión y la extensión, se les agrega 1.5 μl de dNTP's (10mM) y 0.5 μl de klenow (polimerasa). Se incuba a 50°C por 10 minutos, se precipita con acetato de sodio 3M y un volumen de Isopropanol, se incuba 30 minutos a -70°C, los productos se precipitan, secan y resuspenden en un buffer de corrida y separados por electroforésis en un gel de acrilamida al 6% con urea 7M. El gel se seca y se expone en una placa de Phosforimager para ser observarse al otro día (Morett y Buck., 1988).

Los oligos utilizados fueron el Kpn1 y el E58 obtenidos de Morett, E. y Buck, M. (1988).

Técnicas Inmunológicas.

De los cultivos obtenidos, se realiza una extracción de proteína cruda mediante un sonicador dando 3 pulsos de 10 segundos cada 10 segundos, las muestras se corren en un gel desnaturalizante, que puede ser teñido con azul de cumassi para determinar si la extracción fué adecuada. Se corre un gel paralelamente para transferir las muestras a una membrana de nitrocelulosa y realizar el "western".

La membrana se bloquea con TBS (Tris Boratos Salino) + albúmina 3% y se lava con TBST (Tris Boratos Salino + Tween), se agrega el primer anticuerpo diluido en TBST + gelatina 1%.

Se lava con TBST se incuba con el segundo anticuerpo diluido con TBST y gelatina 1%, se lava y se le agrega el buffer de reacción con NTB (Azul de Tetrasolium) y BCIP (5-Bromo-4-Cloro-indolil-fosfato) protegiendolo de la luz y se enjuaga para parar la reacción.

Para el experimento se utilizaron 2 péptidos sintéticos, uno que corresponde a la región carboxilo de la proteína NifA de *R. meliloti* y el otro que corresponde a la región central de la proteína NifA de *B. japonicum*

*RESULTADOS Y
DISCUSION*

Predicción de la estructura de las UAS.

Como ya se mencionó, las UAS son las secuencias reconocidas por la proteína NifA por lo que es un elemento importante para la transcripción de los distintos genes que estos regulan. De los 16 nucleótidos que la forman solo 6 están conservados. Si se calcula la probabilidad de encontrar esta secuencia en un genoma como el de *E. coli*, uno podría encontrar de 1000 a 2000 UAS potenciales.

De la búsqueda que se hizo en los bancos de datos GenBank y EMBL se obtuvieron 33 secuencias UAS de los distintos genes *nif* en diferentes organismos. Estas secuencias se alinearon para observar si existían elementos de secuencia en los 10 nucleótidos intermedios que permitieran a la proteína NifA identificar las verdaderas UAS de las que no lo son (Figura 5). Las proteínas pueden identificar a las regiones mediante secuencia o por la estructura intrínseca del DNA. A partir del alineamiento se obtuvo la gráfica que mostraba el consenso de las regiones UAS. Lo que se observó es que existen otros nucleótidos como la C en la posición 4 y la T en la posición 8 que están conservados en más del 50% de las UAS, en las posiciones 5 y 13 hay purinas en todos los casos observados, esto debilita la hipótesis de que estas regiones estén identificadas por estructura y no por secuencia ya que con esto uno solo podría encontrar de 20 a 40 UAS verdaderas en un genoma como el de *E. coli* y la proteína NifA podría identificar a las regiones UAS por secuencia. Nosotros esperábamos que pudieran diferenciarse por estructura las UAS verdaderas de las que no lo son.

Para conocer la estructura del DNA se han utilizado distintos métodos de predicción en donde se han calculado los valores de los ángulos que forman los posibles pares de nucleótidos. En este caso se utilizaron tanto los valores propuestos por De Santis como los de Trifonov, para predecir la estructura de las distintas UAS obtenidas (Figura 6). Todas las secuencias se superpusieron con una secuencia determinada al azar (UAS falsa) y otra con una curvatura muy pronunciada, las secuencias se fijaron con los primeros 3 nucleótidos conservados (TGT) y se determinó si estas compartían un patrón definido de curvatura.

Lo que se obtuvo de este análisis fue que las UAS verdaderas no guardaban un patrón definido, ya que se obtenía toda una gama de grados de curvatura.

De estos resultados podemos concluir que muy probablemente las UAS no conserven una estructura intrínseca aunque no podemos descartar que estén involucrados otros factores en la predicción de estructuras de DNA.

a) Alineamiento

kp-nifh TGTTCGTITCCCACA
 kp-nife TGTCCAAAGCCACA
 kp-nifj TGTCCCTGTTCACA
 kp-nifu TGTACGACTAATACA
 kp-nifl TGTCAAACTGGGACA
 kp-nifab TGTGAGAITCAGGACA
 kp-orf1 TGTTCCTCAAGCACA
 rm-fixk TGTCCGTFAGTCCACA
 rm-fdbx TGTCCCGTGTGGAACA
 rm-nifb TGTCCGGAAGTCCGACA
 rm-p2 TGTCCAGGGCCGACA
 rm-p3 TGTCCGGCTTAACGACA
 rm-nifhdk TGTCCGGCTTAGCGACA
 rm-fixf TGTCCAGTCCGGGACA
 rs-nifh2 TGTCCGCTTGAACA
 rs-nifh1 TGTCCAGTTCGAGACA
 rl-pfixcx TGTCCAGCCCGCACA
 rl-pfixa TGTCCGGGAAGCTTACA
 rl-nifb2 TGTCAAATGCCCTGACA
 rl-nifh2 TGTCCAGATTTGCCGACA
 rl-nifh1 TGTCTGATCCGGCGACA
 bj-ndp TGTCCGACAGCTGACA
 bj-nifs TGTCCAGGAACCCGACA
 bj-nifh TGTCCAGATCCAAAACA
 bj-nifdk TGTCCGATTCGGACA
 rl-nifb TGTAGGATCAACCAACA
 rl-nifd TGTCCGGCTTCGTGACA
 rl-nifh TGTCCGACTTCGAGACA
 rs-nifh TGTCCGGCTCCCGGACA
 ab-nifh TGTCCGGGTATGTCACA
 ab-nifb TGTCCGGATTCGCCGACA
 rl-nifh TGTCCGGTACTTGACA
 rl-fixw TGTCCGGCAACCATACA

b) Gráfica de posiciones conservadas

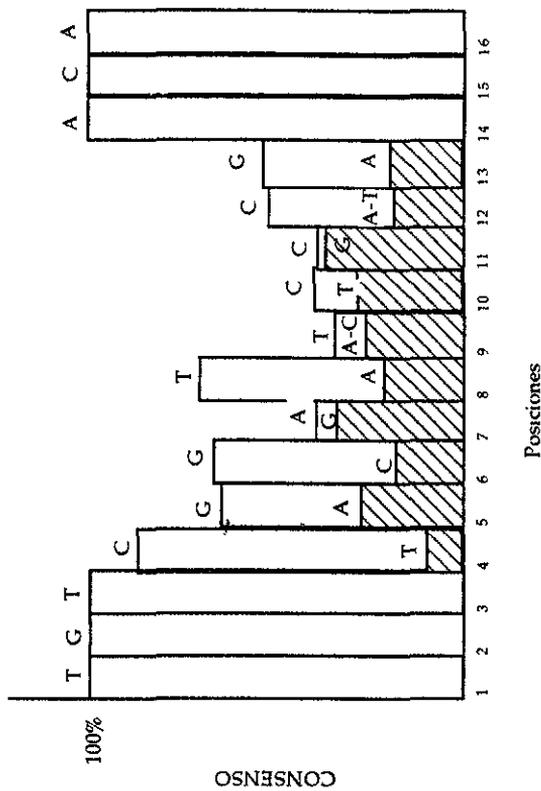


Fig. 5 a) Alineamiento de las UAS obtenidas, kp-eq. *Klebsiella pneumoniae*; rm-eq. *Rhizobium meliloti*; rs-eq. *Rhizobium sp.*; rl-eq. *Rhizobium leguminosarum*; bj-eq. *Bradyrhizobium japonicum*; ab-eq. *Azospirillum brasilense*. Las UAS que se muestran en el alineamiento regulan a los distintos genes *nif* y *fix*.

b) Gráfica de conservación de los nucleótidos de las UAS obtenidas.

abnifb TGTCGGATTGCCGACA



abnifh TGTCGGGTATGTCACA



bjndp TGTTGGACAGCTGACA



bjnifh TGTCAGATCCAAAACA



bjnifhdh TGTCGGATTGCAACA



bjnifs TGTCAGGAACCCGACA

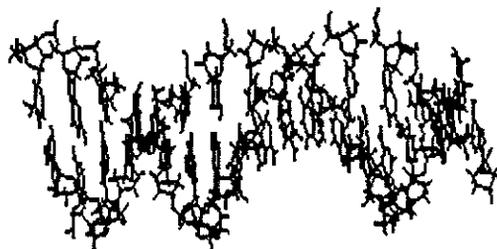


Fig. 6a Predicción de estructura de las distintas UAS.

kpnifab TGTGAGATTCAGGACA



kpnife TGTCGCAAAGCCAACA



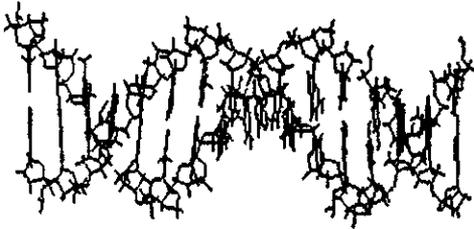
kpnifh TGTTCGTGTTTCCACA



kpnifj TGTCGCCTGTTCAACA



kpnifl TGTCATAACTGCGACA



kpnifu TGTCAGGACTAATACA

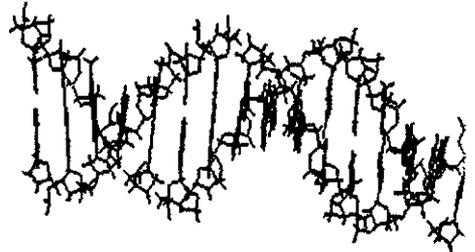
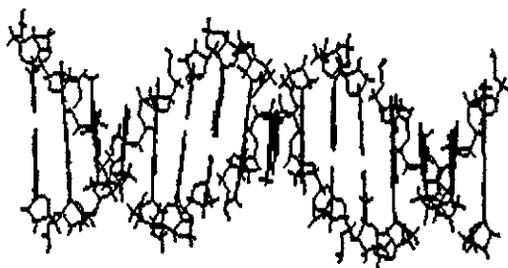


Fig. 6b Predicción de estructura de las distintas UAS.

rlnif TGTCGGTACTTGACA



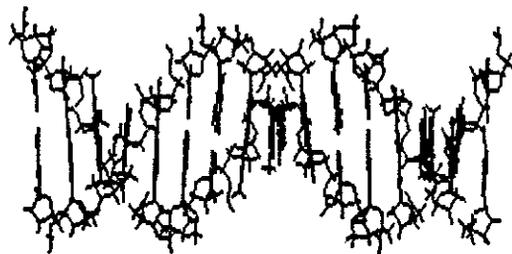
rlpfixa TGTCGGGAAGCTTACA



rlpnifh1 TGCTGATCGGCGACA



rltnifd TGTCGGCTTCGTGACA



rltnifb TGTAGGATCACCAACA



rmfixf TGTCAGGTCCGCGACA



FIG. 6c Predicción de estructura de las distintas UAS.

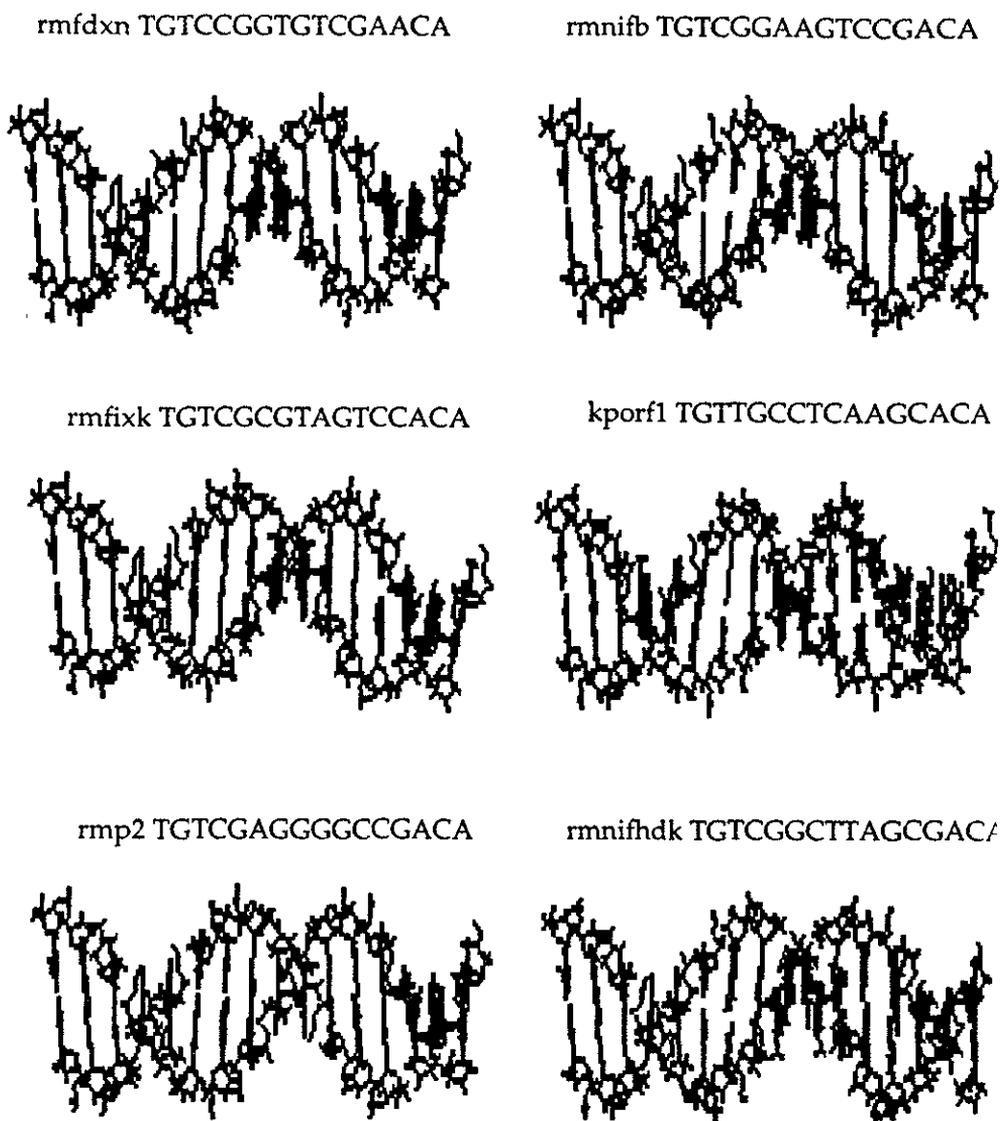


Fig. 6d Predicción de estructura de las distintas UAS.

Hasta el momento el estudio de la predicción en la estructura del DNA se ha complicado ya que existen además de las interacciones cercanas de los nucleótidos, las repercusiones en la estructura de nucleótidos alejados. Además de que al parecer, es más complicado ver curvaturas en trectos cortos de DNA en donde solo se alcanzan a apreciar dobleces y no el resultado de la suma de los distintos dobleces que permitirían ver una curvatura en un tracto grande de DNA.

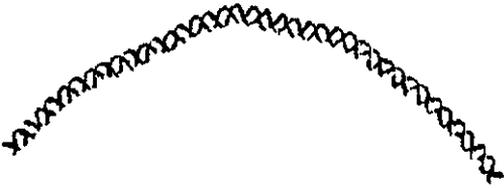
Predicción de estructura de las regiones identificadas por IHF.

De los promotores de genes que se transcriben con el factor σ^{54} identificados en el artículo de Morett y Buck, 1989, se obtuvieron las secuencias de 200 pb por arriba del sitio promotor, se determinaron las posibles secuencias de UAS, para delimitar la región donde pudiera estar interaccionando la proteína IHF.

De las regiones donde se han realizado observaciones experimentales y se conoce claramente la presencia de la proteína IHF es en el gene *glnH* de *E. coli*, también se sabe que en el gene *glnA* de *E. coli* no hay participación de la proteína IHF en el inicio y activación de la transcripción por lo que las regiones por arriba de estos dos genes nos sirvieron como control para saber si se observaba una curvatura definida utilizando las distintas tablas de ángulos.

Se utilizó la región rio arriba del promotor de *glnA* de *E. coli*, para evaluar los distintos modelos de predicción de curvatura (Trifonov, Calladine, DeSantis y Crothers) ya que se han realizado distintos experimentos donde se muestra un doblez claro en la región entre el promotor y el sitio de pegado al activador (Carmona *et al*, 1997) (Figura 7). Lo que se muestra en la figura es que tanto los ángulos utilizados por Trifonov como los de DeSantis nos muestran una curvatura congruente. Como la tabla de Trifonov además es la que se utiliza en el programa de DNASTar, se eligió esta para realizar las predicciones de estructura de las demás regiones. Lo que se esperaba era que las regiones que carecían de sitio de pegado para IHF presentaran una cierta curvatura resultado de los distintos dobleces de las pares de bases hacia el mismo sentido, que permitiera la interacción de la EBP con el complejo $E\sigma^{54}$ y por el contrario, las que tuvieran sitios para IHF que carecieran de esta curvatura ya que la proteína IHF se encarga de acercar a las distintas proteínas para la activación de la transcripción (Figura 8). En esta figura observamos que las regiones con sitios para IHF (*glnH* de *E. coli* y *nifH* de *K. pneumoniae*) no permitiría el acercamiento de las proteínas a menos que estuviera la proteína IHF, en tanto que las que carecen de sitio para IHF (*glnA* de *E. coli* y *glnA* de *K. pneumoniae*) mantenían una estructura que permitía el acercamiento de las proteínas.

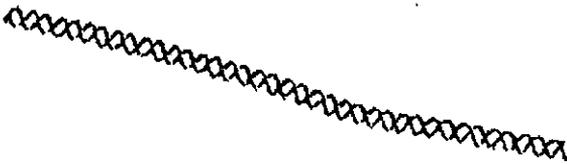
A) Modelo de Trifonov



B) Modelo de DeSantis



C) Modelo de Calladine



D) Modelo de Crothers

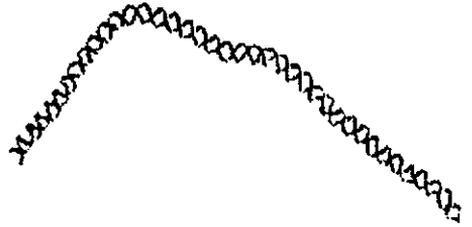


Fig. 7 Modelos de predicción de curvatura en la región por arriba del promotor de *glnA* de *E. coli*

Sin sitios para IHF

Con sitios para IHF

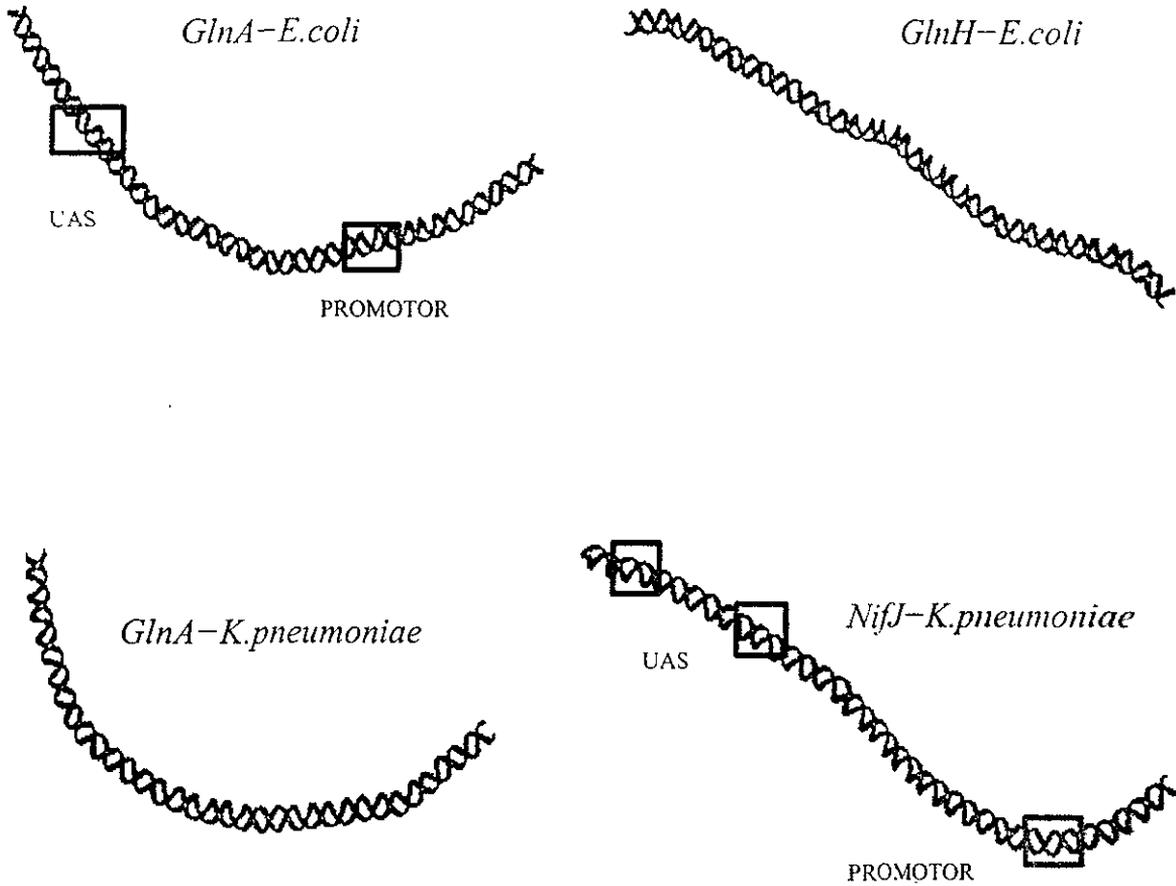


Fig. 8 Estructura de algunas regiones con y sin sitios para IHF, realizados con la tabla de Trifonov.

Esto favorecía fuertemente el planteamiento de que la estructura del DNA juega un papel importante en la activación de la transcripción, apoyado además por los trabajos realizados por Carmona *et al*, 1997, en donde evalúan el papel que juega la curvatura del DNA en el inicio de la transcripción tomando en cuenta otros factores como los de la afinidad del complejo Eo⁵⁴ por el promotor y si el DNA era lineal o superenrollado. Lo que ellos encontraron en el caso del DNA lineal fué que se requería de una buena afinidad del complejo Eo⁵⁴ por el promotor y de una curvatura ya sea intrínseca o inducida del DNA en la región localizada entre el sitio promotor y el sitio de unión a la proteína activadora. Por otro lado en el caso de DNA superenrollado se requiere de una buena afinidad al promotor o de una cierta curvatura en el DNA pero no son necesarias las dos condiciones para que se forme el complejo abierto y se lleve a cabo la transcripción como en el caso de DNA lineal (Carmona *et al*, 1997).

Posteriormente de las secuencias que se obtuvieron de los distintos genes, se identificaron los posibles sitios de IHF mediante el programa de MacTargSearch, utilizando un valor límite de corte suficiente para identificar tanto el promotor como el posible sitio para IHF, sin que aumentara la posibilidad de encontrar muchos sitios falsos. Después a todas ellas se les predijo su estructura con los ángulos de la tabla de Trifonov (Figuras 9 y 10), lo que se observó fué que no existían patrones de curvatura como los que se habían observado en los genes anteriores, tanto los genes que presentaban sitios para IHF como los que no, presentaban patrones similares. En el trabajo de Carmona *et al*, 1997 llegan a la conclusión de que la región entre el sitio promotor y el sitio de pegado a la proteína activadora es la única responsable del inicio de la transcripción en *glnA* en ausencia de IHF, sin embargo aunque muchas evidencias apuntan en ese sentido y aunque en el caso de *glnA* aparentemente pueda observarse la curvatura del DNA en esta región, no es tan evidente en regiones de otros genes. Esto sugiere que no ha sido posible determinar todos los factores que determinan la estructura del DNA.

Como ya se había mencionado con anterioridad, las repercusiones en estructura de nucleótidos no cercanos no se evalúan en este tipo de modelos, como la flexibilidad (bendability), además de la relación del DNA con el solvente o en la interacción con otras proteínas es difícil de evaluar. Por otro lado, en la regulación de la transcripción participan muchos factores que podrían no hacer necesaria la activación por estructura de los genes, entre ellos está que el promotor sea fuerte o débil, que NifA pueda interactuar en *trans*, que las UAS sean más parecidas al consenso o no y por lo tanto tengan mayor afinidad en la interacción con NifA, y por último que la misma transcripción de los genes tenga que ser la óptima y no la máxima para que se lleven a cabo los procesos a niveles normales.

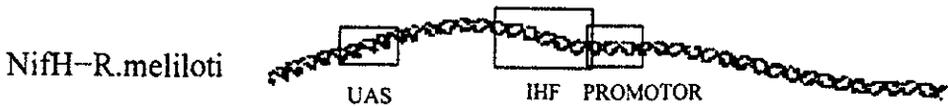
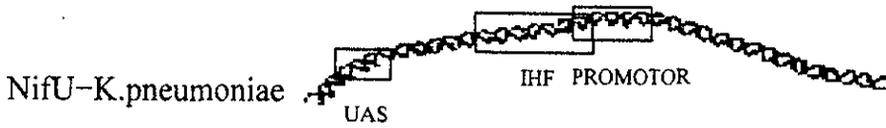
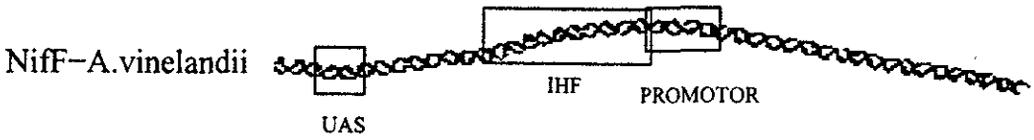
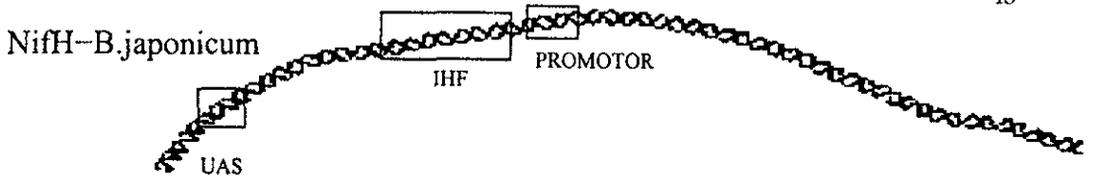


Fig. 9 Estructura de algunas regiones con sitios para IHF, realizada con la tabla de Trifonov.

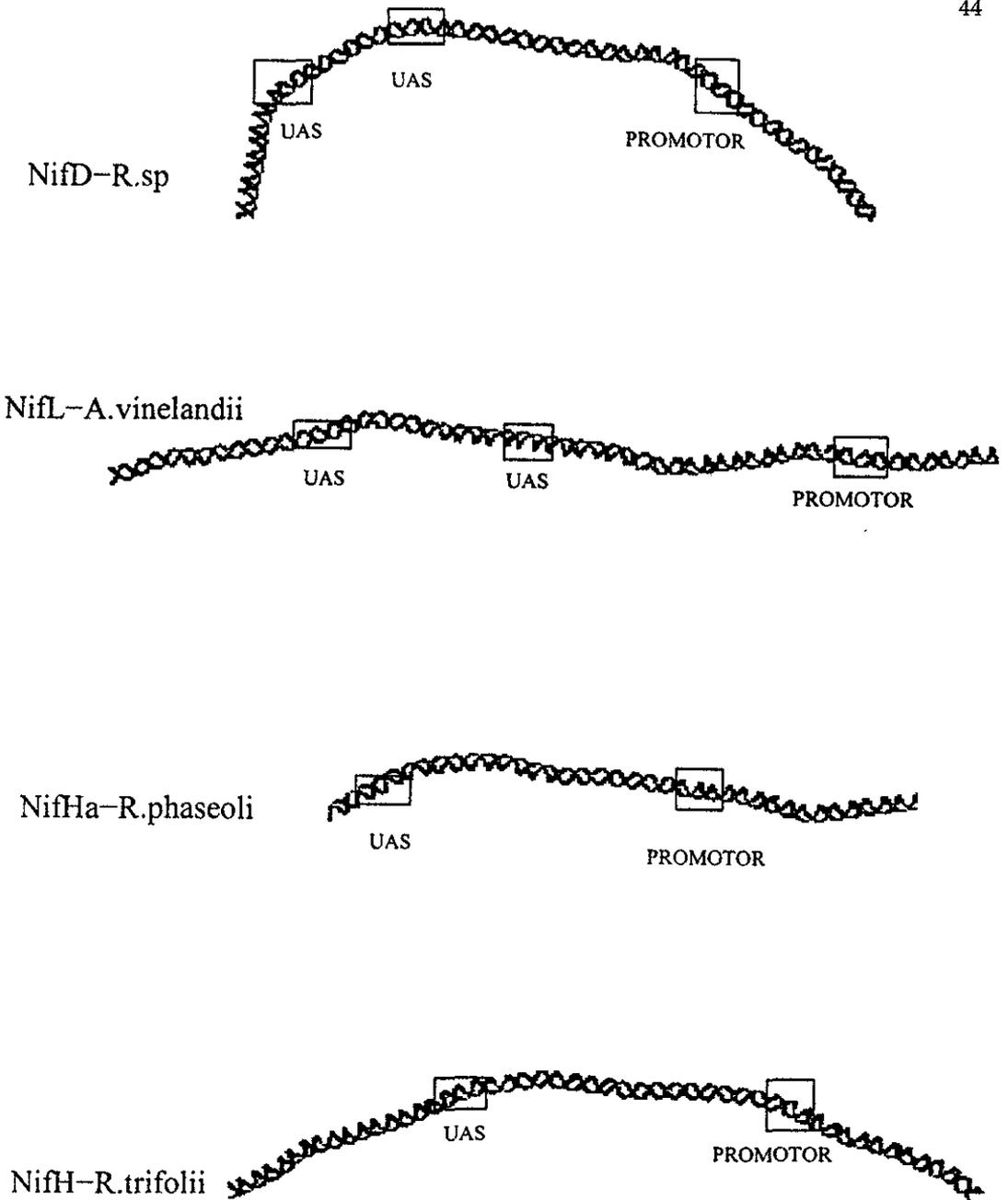


Fig. 10 Estructura de algunas regiones sin sitios para IHF, realizada con la tabla de Trifonov.

Sin embargo uno de los intereses principales es poder predecir la estructura del DNA para de esta manera avanzar en la simulación de experimentos que nos provean de información relevante sobre la interacción y funcionamiento de las moléculas.

Motivo de unión al DNA de NifA.

Como se mencionó anteriormente, la proteína NifA de *R. meliloti* es un buen activador de la transcripción a pesar de que mantiene un débil pegado al DNA, por lo que sería interesante conocer los elementos importantes dentro de el motivo de HVH que permiten la interacción entre las proteínas que lo contienen y el DNA.

Al realizar el alineamiento de todas las proteínas NifA conocidas y basándonos en la predicción de estructura secundaria obtenida mediante el banco de hssp (predicción de estructura secundaria), se observó que en casi todas ellas al igual que en la mayoría de los motivos de HVH, el residuo de glicina en la posición 9 se mantenía conservado, (Figura 11) ya que este al parecer es un elemento importante en la estructuración correcta del motivo para la interacción con el DNA. Algunos de los motivos de HVH en *H. seropedicae* y *A. caulinodans* tienen en la posición 9 asparagina. Sería interesante conocer la capacidad de unión al DNA de las proteínas que contienen estos cambios. En *R. meliloti* el residuo de glicina está sustituido por un glutámico el cual no es común en esta posición del motivo.

Debido a las diferencias en sitios importantes para la estructuración de la HVH localizadas en la secuencia de la proteína NifA de *R. meliloti*, se propuso que muy probablemente estas fueran la causa de la débil interacción.

Para esto se propuso realizar una mutagénesis dirigida al sitio, sustituyendo los residuo que no estuvieran conservados en la proteína NifA de *R. meliloti* por los residuos conservados en otros motivos de HVH de los cuales se supiera que mantenían un fuerte interacción con el DNA como la NifA de *B. japonicum*. Los residuos que se sustituyeron fueron el glutámico (E) encontrado en la NifA de *R. meliloti* por el residuo de glicina (G) conservado en la mayoría de las HVH, obteniéndose la mutante (NAG1), la segunda mutante con una sustitución de lisina (K) por leucina (L) obteniendo la mutante (NAL1) el cual es un residuo poco conservado aledaño a la G y que podría tener algún tipo de interacción que favoreciera la falta de G en la proteína de *R. meliloti*, y por último una mutante con una doble sustitución de E por G y de K por L (NAGL2). Las mutaciones se integraron por la sustitución de las regiones en un plásmido que contiene la NifA silvestre de *R. meliloti*.

<i>Rhizobium meliloti</i>	.LIERDRLISA LEEAGWNOAK AARIL@KTPR QVGYALRRHG VDVVKL
<i>Klebsiella oxytoca</i>	SLDERQRLIAA LEKAGWVOAK AARLL@MTPR QVAYRIQIMD ITMPRL
<i>Klebsiella pneumoniae</i>	SLDERQRLIAA LEKAGWVOAK AARLL@MTPR QVAYRIQIMD ITMPRL
<i>Azotobacter vinelandii</i>	TLDDRERVIAA LEQAGWVOAK AARLL@MTPR QIAYRIQTLN IHMRKI
<i>Azospirillum brasilense</i>	SGSLRDRLLWA MERTGWVOAK AARLL@MTTR QVSYALRKYN IEIKRF
<i>Azospirillum lipoferum</i>	SGSLRDRLLWA MERTGWVOAK AARLL@MTTR QVSYALRKYN IEIKRF
<i>Herbaspirillum seropedicae</i>	.TSERERLIWA MEQCGWVOAK AARALN@ISPR QMGYALQKFN IEVKKF
<i>Azorhizobium caulinodans</i>	RQSEKEQLLQA MERSGWVOAK AARLL@LTTR QVGYALRKYD IDIKRF
<i>Bradyrhizobium japonicum</i>	RMADRERVVAA MEKSGWVOAK AARLL@LTTR QVGYALRKYG IEIKRF
<i>Rhizobium leguminosarum</i>	.LTERDRLIKA MERAGWVOAK AARIL@KTPR QVGYALRRHR IDVKKE
<i>Rhodobacter capsulatus</i>	AQLSREELLRA LESAGWVOAK AARLL@MTPR QIAYALQKFE IELRKI
<i>Rhizobium etli</i>	RLTQDRRLIEA MEKAGWVOAK AARIL@LTTR QVGYALRQHR IEVKKL
Consenso	---ERERL---A ME--GWVOAK AARLL@MTPR QV-YAL----- IE---F

Fig. 11 Alineamiento de las proteínas NifA con el residuo de Glicina resaltado en la posición 9

Estos plásmidos se co-transformaron junto con un plásmido que contiene una fusión *nifH-lacZ* en una cepa de *E. coli*. De estas cepas se determinaron las actividades transcripcionales de las mutantes mediante la expresión de el gene de β -galactosidasa (Tabla 3). En las distintas cepas se observó claramente que la mutante NAG1 en todos los experimentos realizados mantiene al menos el doble de la actividad en comparación a la NifA silvestre, esto indica fuertemente que la G es un residuo importante para la estructuración del motivo de HVH y que muy probablemente favorezca su pegado al DNA. La mutante NAL1 carece de actividad, esto podría explicarse por que la carga del E no se compensa por la K que es un residuo localizado aledaño en la secuencia de la proteína y entonces se obtenga una HVH con una conformación estructural incapaz de unirse al DNA y por ende de activar la transcripción. Por último la doble mutante tiene una buena actividad pero por debajo de la silvestre. Estas actividades pueden observarse claramente en un gel de proteínas donde la expresión de β -galactosidasa es evidente.

Sin embargo, para descartar que el efecto observado en la actividad se debía a los distintos niveles de expresión de las proteínas, se realizó un ensayo de inmunodetección ("western") en donde se identifica a la proteína problema y puede cuantificarse su concentración en las células (Figura 12). Para esto se utilizó un anticuerpo que identifica el dominio central de la proteína NifA y lo que se obtuvo fué que todas las proteínas se expresan a una concentración equiparable.

Para determinar de manera física como podía estar afectando el E9 en la estructuración del motivo HVH para la interacción con el DNA se realizaron unas superposiciones de las distintas proteínas NifA's (Figura 13), en este caso no se observó que existiera algún problema estérico a causa de la presencia de E en vez de la G en la posición 9, ya que el E se acomoda hacia afuera de la estructura y no parece alterar de manera importante el motivo de HVH.

También se sustituyó el motivo de HVH de la NifA silvestre de *R. meliloti* en modelos de las proteínas interaccionando con el DNA. La estructura de la proteína LCC se obtuvo por NMR, ésta contiene una Glicina en la posición 9 sirviendo así como un motivo silvestre. La proteína HCR que tiene muy buena resolución de 1.8Å se obtuvo por cristalografía y carece del residuo de glicina en la posición 9 del motivo HVH. Por último la proteína PER obtenida cristalográficamente con una resolución de 2.5 Å con glicina en la posición 9 (Figuras 14-16).

Lo que se observa de las superposiciones es que muy probablemente la glicina permita las interacciones hidrofóbicas entre los residuos de las hélices porque favorece la flexibilidad del motivo. La proteína HCR que se utiliza como modelo tampoco contiene la G9 en su motivo de HVH.

Cepas	Actividad de β -Galactosidasa
pRT22	$2 \times 10^{-1} \pm 1 \times 10^{-1}$ U.M.
pRT22+pucN1FA (silvestre <i>R. meliloti</i>)	$17 \times 10^3 \pm 2 \times 10^3$ U.M.
pRT22+NAG1	$35 \times 10^3 \pm 3 \times 10^3$ U.M.
pRT22+NAL1	$3 \times 10^{-1} \pm 1 \times 10^{-1}$ U.M.
pRT22+NAGL2	$15 \times 10^3 \pm 2 \times 10^3$ U.M.

Tabla 3 Actividad de β -gal de las cepas silvestre y la mutante

CARRIL

1 2 3 4 5 6 7 8 9 10

NifA →



Carril

Proteína

- | | |
|--------|--|
| 1 | Marcador de peso molecular |
| 2 | NifA de <i>B. japonicum</i> silvestre |
| 3 y 7 | Nif A de <i>R. meliloti</i> silvestre |
| 4 y 8 | NifA mutante G1 de <i>R. meliloti</i> |
| 5 y 9 | NifA mutante L1 de <i>R. meliloti</i> |
| 6 y 10 | NifA mutante GL2 de <i>R. meliloti</i> |

Fig. 12 Western con el anticuerpo central de *B. japonicum*, expresión de las proteínas NifA, silvestres y mutantes.

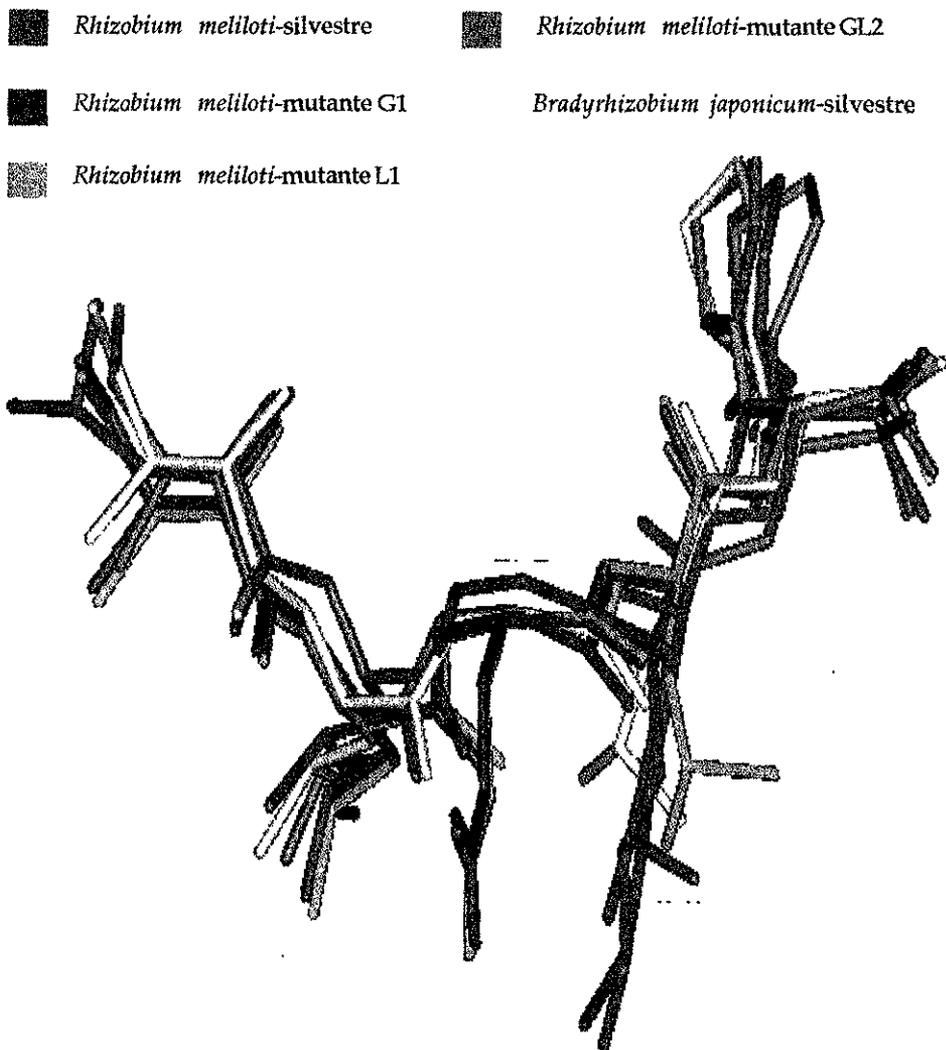


Fig. 13 Superposición del motivo HVH de las NifA mutantes y silvestres

Al superponer el motivo de HVH de la proteína silvestre de *R. meliloti*, se observa que las interacciones entre las dos hélices y el acercamiento de la 2a hélice del motivo con el DNA son más difíciles. Lo anterior resulta evidente ya que al utilizar los modelos de las HVH de las proteínas LCC y PER que si contienen el residuo de G9 se observa que son más probables tanto las interacciones hidrofóbicas entre las dos hélices como el acercamiento entre la hélice del motivo con el DNA. (Figuras 14-16).

Por último, en base al fenotipo observado y para comprobar que la G9 juega un papel importante en la interacción de NifA con las UAS y determinar el pegado de las mutantes al DNA, se realizaron ensayos que permiten observar la protección del DNA por las proteínas, mediante la metodología de "footprinting" (Figuras 17 y 18).

Se utilizaron dos oligos Kpn1 y E58 obtenidos del artículo de Morett y Buck, 1988, que permiten observar las 2 cadenas del DNA tanto la codificante como la complementaria. Mediante esta metodología se puede observar la interacción de la proteína con el DNA ya que cuando esto sucede las guaninas no se metilan con el DMS (Di-Metil-Sulfato) y por lo tanto, la piperidina no puede cortar el DNA en esos sitios; lo que resulta en la desaparición de las bandas del DNA en el gel. En estas figuras observamos que la proteína NAG1 se une fuertemente al DNA en la posición 136 observada con el oligo Kpn1, en comparación con la proteína silvestre, y en las posiciones -123, -126 y -127, en el caso de la cadena complementaria leída con el oligo E58.

El pegado de las otras NifA's mutantes no se observa, lo que es congruente con las actividades obtenidas de los ensayos de β -galactosidasa ya que al no haber un buen pegado de NifA a las UAS la actividad de transcripción tiende a decaer.

Al parecer la participación de la G tiene una fuerte influencia en la estructuración de la HVH. La proteína NifA de *R. meliloti* funciona como un fuerte activador de la transcripción, inclusive sin tener un fuerte pegado a las UAS. Debido al cambio en la vuelta del dominio HVH, era de esperarse que al restituir el E por la G en la posición 9 el pegado al DNA fuera evidente y que la activación de la transcripción aumentara, formándose así un superactivador.

Lo interesante de este fenómeno es saber de que manera NifA de *R. meliloti* dejó de ser un superactivador o si nunca lo fué. Es posible que esta proteína a lo largo de la evolución haya seguido alguno de los caminos siguientes: por un lado ser un superactivador transcripcional con su dominio HVH conservado, el cual pudiera tener como desventaja un gasto energético muy costoso para la célula, al activar fuertemente la transcripción de algunos genes que no fueran indispensables para esta o por perder la actividad de regulación fina.

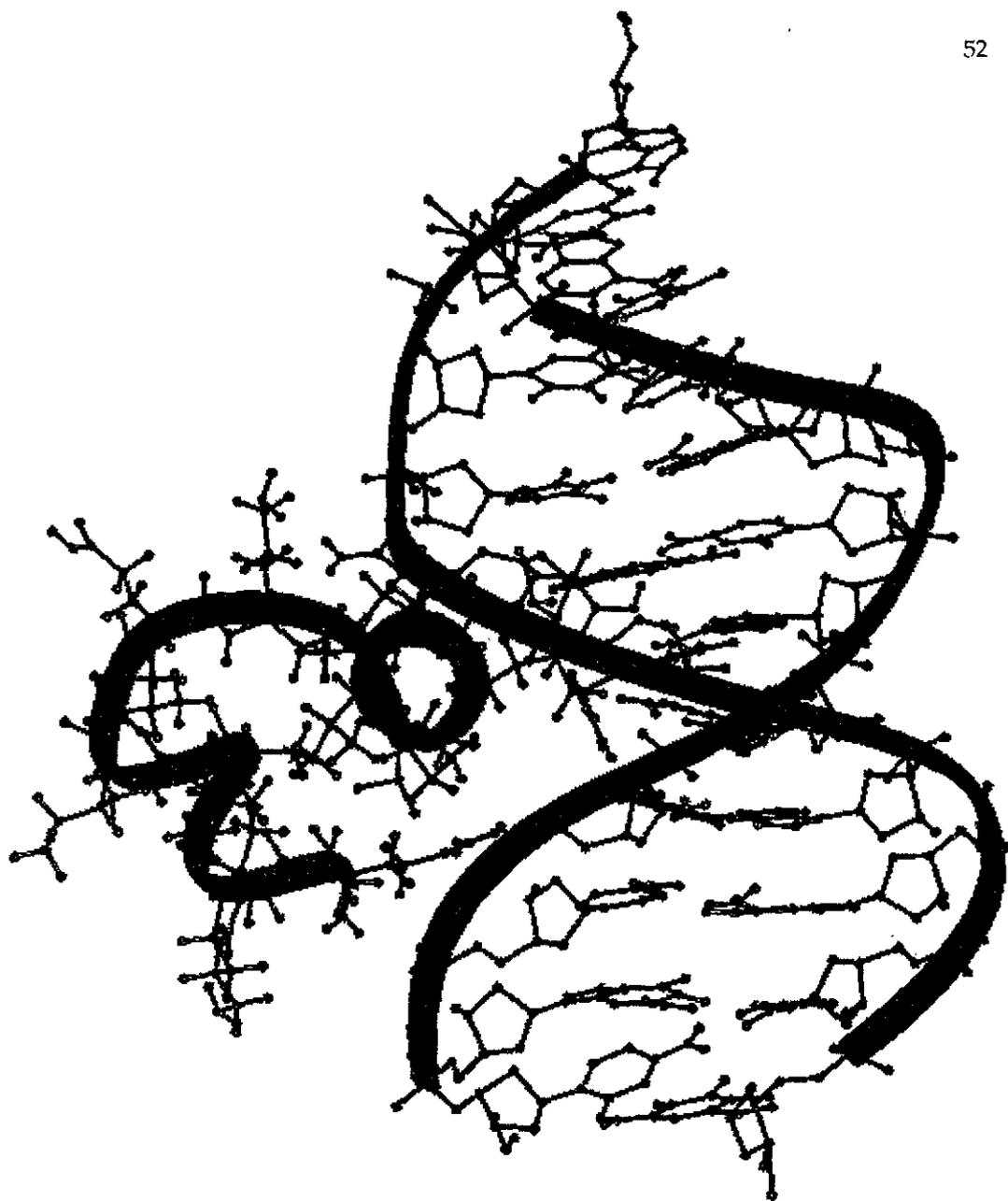


Fig. 14 Superposición del motivo HVH de la mutante NAG1 en la estructura de la proteína LCC interaccionando con el DNA

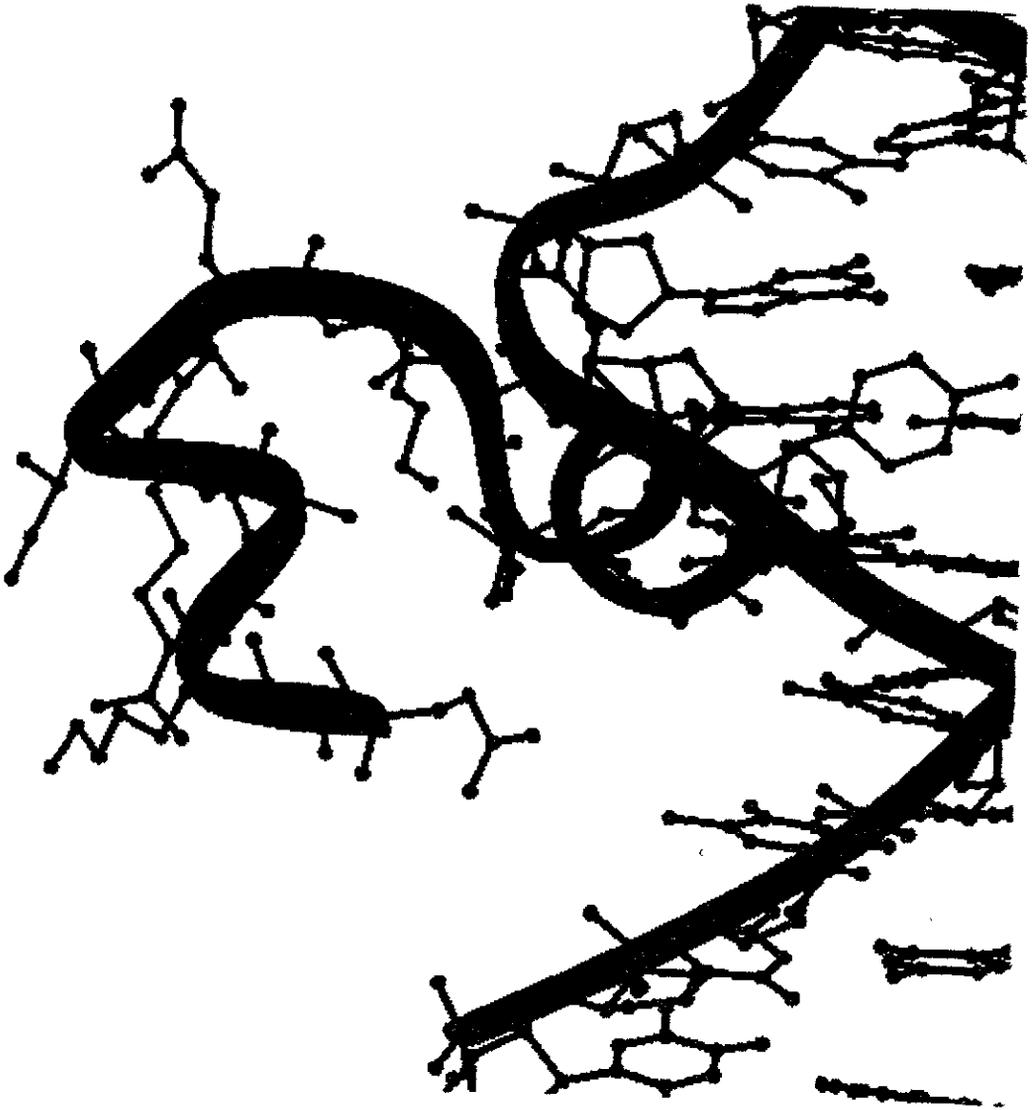


Fig. 15 Superposición del motivo HVH de la mutante NAG1 en la estructura de la proteína HCR interactuando con el DNA

Fig. 18 Footprinting realizado con el oligo E58, donde se observa la protección tanto de la mutante NAG1 como de la proteína NifA de *B. japonicum*

Así al sustituir la G por E se mantiene dentro de los niveles normales de regulación ya que pierde casi por completo la capacidad de unirse al DNA. Por otro lado puede que esta sustitución se haya dado inmediatamente después de la separación de *R. meliloti* con otras bacterias y al sufrir este cambio de G por E, perdió su capacidad de activación, la cual se ha ido compensando en los distintos dominios hasta alcanzar una actividad normal comparable con la activación de las distintas NifA's.

CONCLUSIONES

UAS

-A partir de la ampliación del consenso obtenido de las secuencias de las UAS se observó, que la información en secuencia es más alta de lo que se había propuesto, esto aumenta la posibilidad de que las UAS sean reconocidas por secuencia. Aún así uno esperaría que la secuencia estuviera muy relacionada con la estructura y que las variaciones en esta última entre las distintas UAS fueran mínimas. Sin embargo, en este análisis se obtienen una gran variedad de estructuras que hablan de la falta de selección de una curvatura particular o simplemente que los programas no son lo suficientemente sensibles para permitirnos disociar curvaturas de trectos tan pequeños de DNA.

Para esto una posible solución sería tomar en cuenta otros parámetros como las interacciones lejanas, la interacción con proteínas que provocan curvatura etc., que permitan resolver y predecir de una forma más confiable la curvatura del DNA.

IHF

-En trectos largos del DNA no es posible tampoco observar una curvatura definida en las regiones donde pudiera o existe un sitio para IHF, aunque en algunos casos la curvatura pareciera concordar con las hipótesis planteadas, existen numerosos factores que participan en la estructura del DNA y estos aún no han sido descritos con precisión.

-Los modelos de construcción de estructura del DNA aún no nos pueden ser útiles para predecir cual es la conformación que puede tomar el DNA en algunas circunstancias, los programas tampoco nos ayudan a predecir cuál es la flexibilidad del DNA problema ya que se obtienen estructuras rígidas, sería interesante poder crear dinámicas de la estructura del DNA tomando en cuenta otros parámetros como las repercusiones en estructura que tienen nucleótidos alejados, etc.

**ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA**

NiFA

-La proteína NifA de *R.meliloti* sí es un mejor activador transcripcional ya que esta proteína *per se* mantiene buenos niveles de transcripción comparado con NifA's de otros organismos, a pesar de su débil pegado al DNA. Al corregir mediante algunas mutantes el motivo de unión, aumentan los niveles de activación al doble o más con respecto a la proteína silvestre. Esto indica además que los otros dominios de la proteína están compensando el pegado débil de la proteína silvestre para la activación de la transcripción de los genes.

-La sustitución de E9 por G9 provee a la proteína de una mayor afinidad a las UAS y por ende una mejor activación de la transcripción.

-La glicina es un residuo importante en la estructuración de la HVH para favorecer la unión al DNA.

- Si se mantiene el residuo de E9 y se cambia la K10 por L10 la proteína pierde su actividad por completo, lo que muy probablemente indique que la K es un cambio compensatorio en cuanto a estabilización de cargas en la proteína silvestre para mantener su actividad.

-La G9 es importante para que se lleven a cabo las interacciones hidrofóbicas entre las 2 hélices del motivo ya que le permite una mayor flexibilidad además de que permite según se observó en las superposiciones una mejor interacción entre la segunda hélice y el DNA.

-La proteína NifA depende del pegado al DNA para su activación aunque esta interacción sea muy débil.

-En el laboratorio se realizaron distintas mutagénesis al azar sobre el dominio central que participa en la activación de la transcripción de la NifA de *B.japonicum* para la obtención de mejores activadores, y las mutantes obtenidas en esta proteína siempre eran sustituciones que ya existían en la proteína de *R. meliloti*, lo que favorece la hipótesis de que NifA de *R.meliloti* es un mejor activador intrínseco.

BIBLIOGRAFIA

- Alvarez-Morales, A., Betancourt-Alvarez, M., Kaluza, K., y Hennecke, H. 1986. Activation of the *Bradyrhizobium japonicum* *nifH* and *nifDK* operons is dependent on promoter upstream DNA sequences. *Nucl. Acids. Res.* 14:4207-4227.
- Ausubel, F. 1984. Regulation of nitrogen fixation genes. *Cell* 37:5-6
- Baumeister, R., Müller, G., Hecht, B., y Hillen, W. 1992. Functional Roles of Amino Acid Residues Involved in Forming the α -Helix-Turn- α -Helix Operator DNA Binding Motif of Tet Repressor From Tn10. *PROTEINS: Structure, Function, and Genetics.* 14:168-177.
- Better, M., Ditta, G., y Helinski, D. 1985. Deletion analysis of *Rhizobium meliloti* symbiotic promoters. *EMBO Journal.* 4:2419-2424.
- Bracco, L., Kotlarz, D., Kolb, A., Diekmann, S., y Buc, H. 1989. Synthetic curved DNA sequences can act as transcriptional activators in *Escherichia coli*. *The EMBO J.* 8:4289-4296.
- Branden, C y Tooze, J. 1991. Introduction to protein structure. Garland Publishing, Inc. New York and London. pp 302
- Brennan, R., y Matthews, B. 1989. The Helix-Turn-Helix DNA Binding Motif. *J. of Biol. Chem.* 264:1903-1906.
- Buck, M., Khan, H., y Dixon, R. 1985. Site-directed mutagenesis of the *Klebsiella pneumoniae* *nifL* and *nifH* promoters and *in vivo* analysis of promoter activity. *Nucl. Acids. Res.* 13:7621-7638.
- Buck, M., Miller, S., Drummond, M., y Dixon, R. 1986. Upstream activator sequences are present in the promoters of nitrogen fixation genes. *Nature.* 320:374-378.
- Buck, M., Cannon, W., y Woodcock, J. 1987. Mutational analysis of upstream sequences required for transcriptional activation of the *Klebsiella pneumoniae* *nifH* promoter. *Nucl. Acids. Res.* 15:9945-9956.
- Buck, M., y Cannon, W. 1992. Specific binding of the transcription factor σ ⁻⁵⁴ to promoter DNA. *Nature.* 358:422-424.
- Buikema, W., Szeto, W., Lemley, P., Orme-Johnson, W., y Ausubel, F. 1985. Nitrogen fixation specific regulatory genes of *Klebsiella pneumoniae* and *Rhizobium meliloti* share homology with the general nitrogen regulatory gene *ntnC* of *K. pneumoniae*. *Nucl. Acids. Res.* 13:4539-4555.

- Calladine, C., Drew, H., y McCall, M. 1988. The intrinsic curvature of DNA in Solution. *J. Mol. Biol.* 201:127-137.
- Calladine, C., y Drew, H. 1992. Twisting and Curving. En *Understanding DNA The Molecule and How it Works*. Jovanovich, H (eds) Academic Press. pp 65-92.
- Calladine, C., y Drew, H. 1992. Curving in three dimensions. En *Understanding DNA The Molecule and How it Works*. Jovanovich, H (eds) Academic Press. pp 93-113.
- Calladine, C., y Drew, H. 1992. Methods Used to Study the Structure of DNA. En *Understanding DNA The Molecule and How it Works*. Jovanovich, H (eds) Academic Press. pp 169-198.
- Cannon, W., Kreutzer, R., Kent, H., Morett, E., y Buck, M. 1990. Activation of the *Klebsiella pneumoniae nifU* promoter: identification of multiple and overlapping upstream NifA binding sites. *Nucl. Acids. Res.* 18:1693-1701.
- Carmona, M., Claverie-Martin, F., y Magasanik, B. 1997. DNA bending and the initiation of transcription at σ^{54} -dependent bacterial promoters. *Proc. Natl. Acad. Sci.* 44:9568-9572.
- Claverie-Martin, F., y Magasanik, B. 1992. Positive and Negative Effects of DNA Bending on Activation of Transcription from a Distant Site. *J. Mol. Biol.* 227:996-1008.
- Crothers, D., Haran, T., y Nadeau, J. 1990. Intrinsically Bent DNA. *J. of Biol. Chem.* 265:7093-7096.
- David, M., Daveran, M-L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., Hertig, C., Boistard, P., y Kahn, D. 1988. Cascade Regulation of *nif* Gene Expression in *Rhizobium meliloti*. *Cell.* 54:671-683.
- Delcourt, S., y Blake, R. 1991. Stacking Energies in DNA. *J. of Biol. Chem.* 266:15160-15169.
- Dodd, I. y Egan, J. 1990. Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. *Nucl. Acids. Res.* 18: 5019-5026.
- Espinosa-Urgel, M., y Tormo, A. 1993. σ -dependent promoters in *Escherichia coli* are located in DNA regions with intrinsic curvature. *Nucl. Acids. Res.* 21:3667-3670.

- Fischer, H., y Hennecke, H. 1987. Direct response of *Bradyrhizobium japonicum* NifA mediated *nif* gene regulation to cellular oxygen status. *Mol. Gen. Genet.* 209:621-626
- Fischer, H., Bruderer, T., y Hennecke, H. 1988. Essential and non-essential domains in the *Bradyrhizobium japonicum* NifA: identification of indispensable cysteine residues potentially involved in redox reactivity and or metal binding. *Nucleic Acids Res.* 16:2207-2224
- Fisher, D., Parent, L., y Sharp, P. 1992. Myc/Max and other helix-loop-helix/leucine zipper proteins bend DNA toward the minor groove. *Proc. Natl. Acad. Sci.* 89:11779-11783.
- Goodrich, J., Schwartz, M., y McClure, W. 1990. Searching for and predicting the activity of sites for DNA binding proteins: compilation and analysis of the binding sites for *Escherichia coli* integration host factor (IHF). *Nucl. Acids. Res.* 18:4993-5000.
- Goodsell, D., y Dickerson, R. 1994. Bending and curvature calculations in B-DNA. *Nucl. Acids. Res.* 22:5497-5503.
- Haran, T., y Crothers, D. 1990. Cooperativity in A-tract structure and bending properties of composite TnAn Blocks. *Biochemistry.* 28:2763-2767.
- Harrison, S. 1991. A structural taxonomy of DNA-binding domains. *Nature.* 353:715-719.
- Huala, E., y Ausubel, F. 1989. The central domain of *Rhizobium meliloti* NifA is sufficient to activate transcription from the *R. meliloti* *nifH* promoter. *J. of Bacteriology.* 171:3354-3365.
- Koo, H., y Crothers, D. 1988. Calibration of DNA curvature and unified description of sequence directed bending. *Proc. Natl. Acad. Sci.* 85:1763-1767.
- Krey, R., Pühler, A., y Klipp, W. 1992. A defined amino acid exchange close to the putative nucleotide binding site is responsible for an oxygen-tolerant variant of the *Rhizobium meliloti* NifA protein. *Mol. Gen. Genet.* 234:433-441.
- Kustu, S., North, A., y Weiss, D. 1991. Prokaryotic transcriptional enhancers and enhancer-binding proteins. *TIBS.* 16:397-402.

- Lee, H-S., Berger, D., y Kustu, S. 1993. Activity of purified NifA, a transcriptional activator of nitrogen fixation genes. *Proc. Natl. Acad. Sci.* 90:2266-2270.
- Merrick, M. 1993. Organization and Regulation of Nitrogen Fixation Genes. pp.43-54. En R. Palacios, J. Mora y W. E. Newton (ed.), *New horizons in Nitrogen Fixation*, Dordrecht: Kluwer Academic Publishers.
- Miller, J. 1972. Experiments in Molecular Genetics. "Assay of β -Galactosidase". Society of Fellows, Harvard University y Département de Biologie Moléculaire Université de Genève. Cold Spring Harbor Laboratory. pp 352-355.
- Morett, E., y Buck, M. 1988. NifA-dependent *in vivo* protection demonstrates that the upstream activator sequence of *nif* promoters is a protein binding site. *Proc. Natl. Acad. Sci.* 85:9401-9405.
- Morett, E., Cannon, W., y Buck, M. 1988. The DNA-binding domain of the transcriptional activator protein NifA resides in its carboxy terminus, recognises the upstream activator sequences of *nif* promoters and can be separated from the positive control function of NifA. *Nucl. Acids. Res.* 16:11469-11488.
- Morett, E., y Buck, M. 1989. *In Vivo* Studies on the Interaction of RNA Polymerase- σ^{54} with the *Klebsiella pneumoniae* and *Rhizobium meliloti* *nifH* Promoters. The role of NifA in the formation of an Open promoter complex. *J. Mol. Biol.* 210:65-77.
- Morett, E., y Segovia, L. 1993. The σ^{54} bacterial enhancer binding protein family mechanism of action and phylogenetic relationship of their functional domains. *J. Bacteriology* 175:6067-6074
- Morett, E., Fisher, H., y Hennecke, H. 1991. Influence of Oxygen on DNA Binding, Positive Control, and Stability of the *Bradyrhizobium japonicum* NifA Regulatory Protein. *J. of Bacteriology.* 173:3478-3487.
- Phillips, S., y Moras, D. 1993. Protein-nucleic acid interactions. *Curr. Op. in Struc. Biol.* 3:1-2.
- Prentki, P., Pham, M., y Galas, D. 1987. Plasmid permutation vectors to monitor DNA bending. *Nucl. Acids. Res.* 15:10060
- Rippe, K., von Hippel, P y Langowski, J. 1995. Action at a distance: DNA-looping and initiation of transcription. *TIBS.* 20:500-506.

- Santero, E., Hoover, T., Keener, J., y Kustu, S. 1989. *In vitro* activity of the nitrogen fixation regulatory protein NifA. *Proc. Natl. Acad. Sci.* **86**:7346-7350.
- Sasse-Dwight, S., y Gralla, J. 1988. Probing the *glnALG* upstream activation mechanism *in vivo*. *Proc. Natl. Acad. Sci.* **85**:8934-8938.
- Sasse-Dwight, S., y Gralla, J. 1990. Role of eukaryotic type functional domains found in the prokaryotic enhancer receptor factor σ^{54} . *Cell.* **62**:945-954.
- Sauer, R. 1995. Minor groove DNA-recognition by σ -helices. *Structural Biology.* **2**:7-9.
- Shakked, Z. 1991. The influence of the environment on DNA structures determined by X-ray crystallography. *Curr. Op. in Struct. Biol.* **1**:446-451.
- Travers, A. 1989. DNA conformation and protein binding. *Annu. Rev. Biochem.* **58**:427-452.
- Travers, A. 1990. Why Bend DNA?. *Cell.* **60**: 177-180.
- Travers, A. 1992. DNA structure. En *DNA-Protein interactions*. Chapman and Hall. pp 1-27.
- Travers, A. 1992. DNA-protein interactions: The three-dimensional architecture of DNA-protein complexes. En *DNA-Protein interactions*. Chapman and Hall. pp 28-?
- Tuli, R., y Merrick, M. 1988. Over-production and Characterization of the *nifA* Gene Product of *Klebsiella pneumoniae*- the Transcriptional Activator of *nif* Gene Expression. *J. of General Microbiology.* **134**:425-432.
- Ulanovsky, L., y Trifonov, E. 1987. Estimation of wedge components in curved DNA. *Nature.* **326**:720-722.
- Valderrama, M. 1996. Caracterización genética de los elementos involucrados en el control transcripcional de la nitrogenasa en *R. etli*. Tesis de Doctorado en Investigación Biomédica Básica. pp 93.
- Van der Vliet, P., y Verrijzer, C. 1993. Bending of DNA by Transcription Factors. *BioEssays.* **15**:25-32.

-Virts, E., Stanfield, S., Helinski, D., y Ditta, G. 1988 Common regulatory elements control symbiotic and microaerobic induction of *nifA* in *Rhizobium meliloti*. *Proc. Natl. Acad. Sci.* **85**:3062-3065.

-Wang, X., Kolb, A., Cannon, W., y Buck, M. 1997. Nucleoprotein complex formation by enhancer binding protein NifA. *Nucl. Acids. Res.* **25**:3478-3485.

-Wolffe, A. 1994. Architectural Transcription Factors. *Science.* **264**:1100-1101.

-Wu, H., y Crothers, D. 1984. The locus of sequence directed and protein-induced DNA bending. *Nature.* **308**:509-513.

-Yang, Chien-Chin., y Nash, H. 1989. The Interaction of *E. coli* IHF Protein with Its Specific Binding Sites. *Cell.* **57**:869-880.

-1989. Definitions and nomenclature of nucleic acid structure parameters. *The EMBO J.* **8**:1-4